Contents

Wednesday, January 19, 2022

1 Real Numbers

There are two ways to introduce the real numbers. The first is to construct the natural numbers using sets. For example, we could define 0 to be the empty set $\emptyset$, then 1 to be the set $\{\emptyset\}$, 2 to be $\{\emptyset, \{\emptyset\}\}$, and 3 to be $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$. Then we construct the integer numbers $\mathbb{Z}$, the rational numbers $\mathbb{Q}$, and, finally, the real numbers $\mathbb{R}$ are constructed as "limits of rational numbers". This construction is lengthy, so we will not pursue it.

The second way to introduce the real numbers is to give them in an axiomatic way. We will use this method. The *real numbers* are a set $\mathbb{R}$ with two binary operations, *addition* and *multiplication*

$$+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \qquad \cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$
$$(x, y) \mapsto x + y \qquad (x, y) \mapsto x \cdot y$$

and a relation $\leq$ such that $(\mathbb{R}, +, ., \leq)$ is an *ordered field* satisfying the supremum property. To be precise,

(A) $(\mathbb{R}, +)$ is an *commutative group*, that is,

- (A_1) (**commutativity**) for every $a, b \in \mathbb{R}$, $a + b = b + a$,
- (A_2) (**distributivity**) for every $a, b, c \in \mathbb{R}$, $(a + b) + c = a + (b + c)$,
- (A_3) there exists a unique element in $\mathbb{R}$, called *zero* and denoted 0, such that $0 + a = a + 0 = a$ for every $a \in \mathbb{R}$,
- (A_4) for every $a \in \mathbb{R}$ there exists a unique element in $\mathbb{R}$, called the *opposite* of a and denoted $-a$, such that $(-a) + a = a + (-a) = 0$,

(M)

- (M_1) (**commutativity**) for every $a, b \in \mathbb{R}$, $a \cdot b = b \cdot a$,
- (M_2) (**distributivity**) for every $a, b, c \in \mathbb{R}$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$,
- (M_3) there exists a unique element in $\mathbb{R}$, called *one* and denoted 1, such that $1 \neq 0$ and $1 \cdot a = a \cdot 1 = a$ for every $a \in \mathbb{R}$ with $a \neq 0$,
- (M_4) for every $a \in \mathbb{R}$ with $a \neq 0$ there exists a unique element in $\mathbb{R}$, called the *inverse* of a and denoted a^{-1}, such that $a^{-1} \cdot a = a \cdot a^{-1} = 1$,

(O) $\leq$ is a *total order relation*, that is,

- (O_1) for every $a, b \in \mathbb{R}$ either $a \leq b$ or $b \leq a$,
- (O_2) for every $a, b, c \in \mathbb{R}$ if $a \leq b$ and $b \leq c$, then $a \leq c$,
- (O_3) for every $a, b \in \mathbb{R}$ if $a \leq b$ and $b \leq a$, then $a = b$,
- (O_4) for every $a \in \mathbb{R}$ we have $a \leq a$,

(AM) for every $a, b, c \in \mathbb{R}$, $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$,

(AO) for every $a, b, c \in \mathbb{R}$ if $a \leq b$, $a + c \leq b + c$,

(MO) for every $a, b \in \mathbb{R}$ if $0 \leq a$ and $0 \leq b$, then $0 \leq a \cdot b$,

(S) **(supremum property)**

If $a \leq b$ and $a \neq b$, we write $a < b$.

Exercise 1 *Using **only** the axioms (A), (M), (O), (AO), (AM) and (MO) of $\mathbb{R}$, prove the following properties of $\mathbb{R}$:*

(i) *if $a \cdot b = 0$ then either $a = 0$ or $b = 0$,*

(ii) *if $a \geq 0$ then $-a \leq 0$,*

(iii) *if $a \leq b$ and $c < 0$ then $ac \geq bc$,*

(iv) *for every $a \in \mathbb{R}$ we have $a^2 \geq 0$,*

(v) $1 > 0$.

Definition 2 *Let $E \subseteq \mathbb{R}$ be a nonempty set.*

(i) *An element $L \in \mathbb{R}$ is called an* upper bound *of E if $x \leq L$ for all $x \in E$;*

(ii) *E is said to be* bounded from above *if it has at least an upper bound;*

(iii) *if E is bounded from above, the least of all its upper bounds, if it exists, is called the* supremum *of E and is denoted* $\sup E$.

(iv) *E has a* maximum *if there exists $L \in E$ such that $x \leq L$ for all $x \in E$. We write $L = \max E$.*

We are now ready to state the supremum property.

(S) **(supremum property)** every nonempty set $E \subseteq \mathbb{R}$ bounded from above has a supremum in $\mathbb{R}$.

The supremum property says that in $\mathbb{R}$ the supremum of a nonempty set bounded from above always exists in $\mathbb{R}$. We will see that this is not the case for the rationals numbers.

Remark 3 *(i)* *Note that if a set has a maximum L, then L is also the supremum of the set.*

(ii) *If $E \subseteq \mathbb{R}$ is a set bounded from above, to prove that a number $L \in \mathbb{R}$ is the supremum of E, we need to show that L is an upper bound of E, that is, that $x \leq L$ for every $x \in E$, and that any number $s < L$ cannot be an upper bound of E, that is, that there exists $x \in E$ such that $s < x$.*

Example 4 *Let* $E := \{x \in \mathbb{R} : \ x < 1\}$*. Then* 1 *is an upper bound of the set* E *and so* E *is bounded from above. We claim that* 1 *is the supremum of the set* E*. To see this, let* $y \in \mathbb{R}$ *with* $y < 1$*. We need to prove that* y *is not an upper bound of the set* E*, that is, we need to show that there are elements in the set* E *that are larger than* y*. Take* $x := \frac{1+y}{2}$*. Since* $y < 1$*, we have that* $1 + y < 1 + 1$*, and so* $\frac{1+y}{2} < 1$*. Thus* x *belongs to* E*. On the other hand,* $x = \frac{1+y}{2} > y$*, and so* y *is not an upper bound of* E*. This shows that* $1 = \sup E$*. Note that* 1 *does not belong to the set* E *and so the set* E *has no maximum.*

Definition 5 *Let* $E \subseteq \mathbb{R}$ *be a nonempty set.*

(i) An element $\ell \in \mathbb{R}$ *is called a* lower bound *of* E *if* $\ell \le x$ *for all* $x \in E$*;*

(ii) E *is said to be* bounded from below *if it has at least a lower bound;*

(iii) if E *is bounded from below, the greatest of all its lower bounds, if it exists, is called the* infimum *of* E *and is denoted* $\inf E$*;*

(iv) E *has a* minimum *if there exists* $\ell \in E$ *such that* $\ell \le x$ *for all* $x \in E$*. We write* $\ell = \min E$*.*

Remark 6 *(i) Note that if a set has a minimum* ℓ*, then* ℓ *is also the infimum of the set.*

(ii) If $E \subseteq \mathbb{R}$ *is a set bounded from below, to prove that a number* $\ell \in \mathbb{R}$ *is the infimum of* E*, we need to show that* ℓ *is a lower bound of* E*, that is, that* $\ell \le x$ *for every* $x \in E$*, and that any number* $\ell < s$ *cannot be a lower bound of* E*, that is, that there exists* $x \in E$ *such that* $x < s$*.*

Thursday, January 20, 2022

2 Natural Numbers

Definition 7 *A set* $E \subseteq \mathbb{R}$ *is called an* inductive set *if it has the following properties*

(i) the number 1 *belongs to* E*,*

(ii) if a number x *belongs to* E*, then* $x + 1$ *also belongs to* E*.*

Example 8 *The sets* $[0, \infty) = \{x \in \mathbb{R} : \ 0 \le x\}$*,* $[1, \infty) = \{x \in \mathbb{R} : \ 1 \le x\}$*, and* $\mathbb{R}$ *are all inductive sets.*

Definition 9 *The set of the* natural numbers $\mathbb{N}$ *is defined as the intersection of all inductive sets of* $\mathbb{R}$*.*

Note that $\mathbb{N}$ is nonempty, since 1 belongs to every inductive set, and so also to $\mathbb{N}$. We also define

$$\mathbb{N}_0 = \mathbb{N} \cup \{0\} .$$

Example 10 *The number* $\frac{1}{2}$ *is not a natural number. Indeed,* $[1,\infty)$ *is an inductive set and* $\frac{1}{2}$ *does not belong to* E*, so* $\frac{1}{2}$ *cannot belong to* $\mathbb{N}$*. Also* $\frac{3}{2}$ *is not a natural number. Indeed, the set* $E = \{1\} \cup \{n \in \mathbb{N} : \ n \geq 2\}$ *is an inductive set that does not contain* $\frac{3}{2}$*. Hence,* $\frac{3}{2}$ *cannot be a natural number.*

Proposition 11 *The set* $\mathbb{N}$ *is an inductive set.*

Proof. We already know that 1 belongs to $\mathbb{N}$. If x belongs to $\mathbb{N}$, then it belongs to every inductive set E but then, since E is an inductive set, it follows that $x+1$ belongs E. Hence, $x+1$ belongs to every inductive set, and so by definition of $\mathbb{N}$, we have that $x+1$ also belongs to $\mathbb{N}$. ■

The next result is very important.

Theorem 12 (Principle of mathematical induction) *Let* $\{p_n\}$*,* $n \in \mathbb{N}$*, be a family of propositions such that*

(i) p_1 *is true,*

(ii) if p_n *is true for some* $n \in \mathbb{N}$*, then* p_{n+1} *is also true.*

Then p_n *is true for* every $n \in \mathbb{N}$.

Proof. Let $E := \{n \in \mathbb{N} \text{ such that } p_n \text{ is true}\}$. Note that $E \subseteq \mathbb{N}$. It follows by (i) and (ii) that E is an inductive set, and so E contains $\mathbb{N}$ (since $\mathbb{N}$ is the intersection of all inductive sets). Hence, $E = \mathbb{N}$. ■

If $x \in \mathbb{R}$ and $n \in \mathbb{N}$, we define

$$x^n := \underbrace{x \cdot \cdots \cdot x}_{n \text{ times}}.$$

If $x \neq 0$, we define $x^0 := 1$. We do not define 0^0.

The following will be used later on.

Exercise 13 *Let* $x \geq -1$*. Prove that*

$$(1+x)^n \geq 1 + nx \tag{1}$$

for every $n \in \mathbb{N}$*.*

Exercise 14 *Prove that*

$$1 + \cdots + n = \frac{n(n+1)}{2} \tag{2}$$

for every $n \in \mathbb{N}$

Exercise 15 *Let* $x \neq 1$*. Prove that*

$$1 + x \cdots + x^n = \frac{x^{n+1} - 1}{x - 1}$$

for all $n \in \mathbb{N}$*.*

In what follows $0! := 1$, $1! := 1$ and $n! := 1 \cdot 2 \cdot \cdots \cdot n$ for all $n \in \mathbb{N}$. The number $n!$ is called the *factorial* of n. For $n \in \mathbb{N}$ and $k \in \mathbb{N}_0$, we define

$$\binom{n}{k} := \frac{n!}{k!\,(n-k)!}.$$

Exercise 16 *Let $j, k \in \mathbb{N}$ and $a \in \mathbb{R}$. Given the function $f(x) = (x+a)^j$, prove that*

$$\frac{d^k f}{dx^k}(x) = \begin{cases} 0 & \text{if } k > j, \\ j(j-1)\cdots(j-k+1)(x+a)^{j-k} & \text{if } k < j, \\ k! & \text{if } k = j. \end{cases}$$

Exercise 17 *Let $x, y \in \mathbb{R} \setminus \{0\}$ and let $n \in \mathbb{N}$.*

(i) Prove that for every $1 \le k \le n$,

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}.$$

(ii) Prove that

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Remark 18 *If in Theorem 12 we replace property (i) with*

(i)′ if p_{n_0} is true for some $n_0 \in \mathbb{N}$,

then we can conclude that p_n is true for all $n \in \mathbb{N}$ with $n \ge n_0$. To see this, it is enough to define

$$E := \{n \in \mathbb{N} \text{ such that } p_{n+n_0-1} \text{ is true}\},$$

which is still an inductive set.

Exercise 19 *Prove that*

$$n^n > 2^n n!$$

for all $n > 6$. Hint: Use the binomial theorem.

Friday, January 21, 2022

Proposition 20 (Archimedean Property) *If $a, b \in \mathbb{R}$ with $a > 0$, then there exists $n \in \mathbb{N}$ such that $na > b$.*

Proof. If $b \le 0$, then $n = 1$ will do. Thus, assume that $b > 0$. Assume by contradiction that $na \le b$ for all $n \in \mathbb{N}$ and define the set

$$E = \{na : n \in \mathbb{N}\}.$$

Then the set E is nonempty and has an upper bound, b. By the supremum property, there exists $L = \sup E$. Hence, for every $m \in \mathbb{N}$, we have that $(m+1)\,a \leq L$, or, equivalently, $ma \leq L - a$ for all $m \in \mathbb{N}$. But this shows that $L - a$ is an upper bound of E, which contradicts the fact that L is the least upper bound. ■

In the previous section we have defined the natural numbers. Note that $(\mathbb{N}, +, \cdot, \leq)$ does not satisfy properties (A_3), (A_4), and (M_4). In particular, we cannot subtract two numbers $a, b \in \mathbb{N}$ unless, $a \geq b + 1$. For this reason, we define the *set of integers* $\mathbb{Z}$ as follows

$$\mathbb{Z} := \{\pm n : \, n \in \mathbb{N}\} \cup \{0\}\,.$$

Theorem 21 (The integer part) *Given a real number $x \in \mathbb{R}$, there exists an integer $k \in \mathbb{Z}$ such that $k \leq x < k + 1$.*

Proof. Step 1: Let $E = \{k \in \mathbb{Z} : \, k \leq x\}$. If $x \geq 0$, then $0 \in E$. If $x < 0$, let's use the Archimedean property to find $n > -x$. Then $-n < x$, and so, $-n \in E$.

Step 2: Since E is nonempty and bounded from below by x, by the supremum property, there exists $L = \sup E$. Then $L - 1$ is not an upper bound of E and so there exists $j \in E$ such that $L - 1 < j \leq L$. By adding one to both sides, we get that $L < j + 1$. Since L is the supremum of E, we have that $j + 1$ is not in E, that is, $x < j + 1$. Thus, $j \leq x < x + 1$. ■

Definition 22 *Given a real number $x \in \mathbb{R}$, the integer k given by the previous corollary is called the* integer part *of x and is denoted $\lfloor x \rfloor$. The number $x - \lfloor x \rfloor$ is called the* fractional part *of x and is denoted* $\operatorname{frac} x$ *(or $\{x\}$). Note that $0 \leq \operatorname{frac} x < 1$.*

Exercise 23 *Prove that every nonempty subset of the natural numbers has a minimum.*

3 The Rationals Numbers and the Supremum Property

Now $(\mathbb{Z}, +, \cdot, \leq)$ satisfies properties (A_3), (A_4), but not (M_4). To resolve this issue, we introduce the *set of rational numbers* $\mathbb{Q}$ defined by

$$\mathbb{Q} := \left\{ \frac{p}{q} : \, p, q \in \mathbb{Z}, \, q \neq 0 \right\},$$

where $\frac{p}{q} := p \cdot q^{-1}$. Then $(\mathbb{Q}, +, \cdot, \leq)$ satisfies properties (A), (M), (O), (AM), (AO), (MO). So, what's wrong? We will see that the rational numbers do not satisfy the supremum property.

Theorem 24 (Density of the rationals) *If $a, b \in \mathbb{R}$ with $a < b$, then there exists $r \in \mathbb{Q}$ such that $a < r < b$.*

Proof. We start by choosing the denominator. We want to find $q \in \mathbb{N}$ such that $\frac{1}{q} < b - a$. To do this, we use the Archimedean property (applied with 1 and $\frac{1}{b-a}$ in place of a and b) to find $q \in \mathbb{N}$ such that $0 < \frac{1}{b-a} < q$. So, we have $\frac{1}{q} < b - a$, or, $a < a + \frac{1}{q} < b$. Multiply by q to find

$$qa < qa + 1 < qb. \tag{3}$$

By the theorem on the integer part, there exists an integer $p \in \mathbb{Z}$ such that

$$p \leq qa < p + 1. \tag{4}$$

Since $p \leq qa$, we have $p + 1 \leq qa + 1 < qb$. Thus,

$$qa < p + 1 < qb.$$

Multiplying by $\frac{1}{q} > 0$ gives

$$a < \frac{p+1}{q} < b.$$

It suffices to define $r := \frac{p+1}{q}$. ■

Remark 25 *It follows from the previous theorem that for every $x \in \mathbb{R}$, if we consider the set*

$$E := \{r \in \mathbb{Q} : \, r < x\},$$

then

$$\sup E = x.$$

Indeed, since x is an upper bound of E, E is bounded from above, and so there exists $\sup E = L$. Moreover, $L \leq x$, since L is the least upper bound of E. We claim that $L = x$. To see this, note that if $L < x$, then by the previous theorem we can find $r \in \mathbb{Q}$ such that $L < r < x$. But then $r \in E$, and L cannot be an upper bound of E, which is a contradiction. Thus $L = x$. This property will be very useful. It says that using rational numbers we can get as close as we want to every real number.

The set $\mathbb{R} \setminus \mathbb{Q}$ is called the set of *irrational numbers*.

Exercise 26 *Prove that there does not exist a rational number r such that $r^2 = 2$*

Theorem 27 *The set of irrational numbers is nonempty.*

Monday, January 24, 2022

Proof. Take

$$E := \left\{x \in \mathbb{R} : \, 0 < x \text{ and } x^2 < 2\right\}.$$

Then E is nonempty, since $1 \in E$. Moreover, E is bounded from below, since 2 is an upper bound. Hence, by the supremum property, there exists $L \in \mathbb{R}$ such that $L = \sup E$.

We claim that $L^2 = 2$. It cannot be $L \leq 0$, since $1 \in E$ and $1 > 0$. Hence, $L > 0$. Let's prove that it cannot be $L^2 < 2$. By the Archimedean property we can choose $n \in \mathbb{N}$ so large that $n > \frac{2L+1}{2-L^2}$. Then

$$\left(L + \frac{1}{n}\right)^2 = L^2 + \frac{1}{n^2} + \frac{2L}{n} < L^2 + \frac{1}{n} + \frac{2L}{n} = L^2 + \frac{2L+1}{n} < 2,$$

by the choice of n. Hence, $L + \frac{1}{n}$ belongs to E, which contradicts the fact that L is an upper bound of E.

Let's prove that if $y \in \mathbb{R} \setminus E$ and $y > 0$, then y is an upper bound of E. Indeed, let $x \in E$. If $x > 0$, then $x^2 < 2 < y^2$, which, since $y > 0$, implies that $x < y$ (why?).

Let's prove that it cannot be $L^2 > 2$. By the archimedean property we can choose $n \in \mathbb{N}$ so large that $n > \max\left\{\frac{2L}{L^2-2}, \frac{1}{L}\right\}$. Then $L - \frac{1}{n} > 0$ and

$$\left(L - \frac{1}{n}\right)^2 = L^2 + \frac{1}{n^2} - \frac{2L}{n} > L^2 - \frac{2L}{n} > 2.$$

We claim that $L - \frac{1}{n}$ is an upper bound of E. To see this, let $x \in E$. Since $x > 0$, $L - \frac{1}{n} > 0$, and $x^2 < 2 < \left(L - \frac{1}{n}\right)^2$, we must have $x < L - \frac{1}{n}$ (why?). This shows that $L - \frac{1}{n}$ is an upper bound of E. This contradicts the fact that L is the least upper bound of E. Hence, it cannot be $L^2 > 2$, Thus, $L^2 = 2$.

The number L is denoted $\sqrt{2}$ and called *square root* of 2. ■

Exercise 28 *Prove that the rational numbers do not satisfy the supremum property, that is, it is not true that all sets $E \subseteq \mathbb{Q}$ which are nonempty and bounded from above admit a supremum in $\mathbb{Q}$.*

Corollary 29 (Density of the irrationals) *If $a, b \in \mathbb{R}$ with $a < b$, then there exists $x \in \mathbb{R} \setminus \mathbb{Q}$ such that $a < x < b$.*

Proof. Since $a < b$, we have that $\sqrt{2}a < \sqrt{2}b$. By the density of the rationals, there exists $r \in \mathbb{Q}$ such that $\sqrt{2}a < r < \sqrt{2}b$. Without loss of generality, we may assume that $r \neq 0$ (why?). Hence, $a < \frac{r}{\sqrt{2}} < b$. Since $\frac{r}{\sqrt{2}}$ is irrational (why?), the result is proved. ■

Exercise 30 *Let $(\mathbb{R}', \oplus, \odot, \preccurlyeq)$ be another ordered field satisfying the supremum property. Prove that there exists a bijection $T : \mathbb{R} \to \mathbb{R}'$ such that T is an isomorphism between the two fields, that is,*

$$T(a+b) = T(a) \oplus T(b), \qquad T(a \cdot b) = T(a) \odot T(b)$$

for all $a, b \in \mathbb{R}$, and $a \leq b$ if and only if $T(a) \preccurlyeq T(b)$.

Remark 31 *The previous exercise proves uniqueness of the real numbers. Indeed every theorem we prove for $\mathbb{R}$ would hold for $\mathbb{R}'$ because of the properties of T. Hence, for all practical purposes, we cannot distinguish $\mathbb{R}$ from $\mathbb{R}'$.*

Similarly, for every $n \in \mathbb{N}$ with n even and every $x \in \mathbb{R}$ with $x \geq 0$, we can show that there exists a unique $y \in \mathbb{R}$ with $y \geq 0$ such that $x^n = y$. On the other hand, for every $n \in \mathbb{N}$ with n odd and every $x \in \mathbb{R}$, we can show that there exists a unique $y \in \mathbb{R}$ such that $x^n = y$.

The number y is denoted $\sqrt[n]{x}$ and called *n-th root* of x.

Exercise 32 (The n-th root of a) *Given $x > 0$ and $n \in \mathbb{N}$, with $n \geq 2$, we want to define the n-th root of x.*

(i) *Prove that if $r, s \in \mathbb{Q}$ with $r < s$, then $r^n < s^n$.*

(ii) *Prove that the set*

$$E := \{r^n : r \in \mathbb{Q}, r > 1\}$$

does not have a minimum and that $\inf E = 1$.

(iii) *Given $x > 0$ consider the set*

$$F := \{y \in \mathbb{R} : y > 0, \; y^n \leq x\}.$$

Prove that F is bounded from above and nonempty. Let $\ell := \sup F$. Prove that $\ell^n = x$.

4 Powers with Real Exponents

If $x \in \mathbb{R}$ and $n \in \mathbb{N}$, then we define x^n inductively by

$$x^1 := x, \quad x^{n+1} := x^n \cdot x.$$

But what does it mean $x^{\sqrt{2}}$? Or more generally, x^a if $a \in \mathbb{R}$? To define this, we will assume that $x > 0$ (this is needed to preserve the properties of powers). If a is positive and rational, say $a = \frac{n}{m}$, where $m, n \in \mathbb{N}$, then we define

$$x^{\frac{n}{m}} := \left(\sqrt[m]{x}\right)^n.$$

Remark 33 *Note that $x^{\frac{n}{m}} = \sqrt[m]{x^n}$. Indeed, let $y = \sqrt[m]{x}$. Then*

$$(y^n)^m = (y^m)^n = x^n,$$

and so $y^n = \sqrt[m]{x^n}$, that is, $\left(\sqrt[m]{x}\right)^n = \sqrt[m]{x^n}$.

If a is rational and negative, say $a = -\frac{n}{m}$, where $m, n \in \mathbb{N}$, then we define

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

Exercise 34 *Prove that if $x > 0$ and $r, q \in \mathbb{Q}$, then*

$$\begin{aligned} x^r \cdot x^s &= x^{r+s}, \\ (x^r)^s &= (x^s)^r = x^{rs}. \end{aligned}$$

Exercise 35 *Let $x > 1$ and $r, q \in \mathbb{Q}$.*

(i) Prove that if $r > 0$, then $x^r > 1$.

(ii) Prove that if $r < s$, then $x^r < x^s$.

Define

$$\mathbb{Q}^+ := \{r \in \mathbb{Q} : \ r > 0\}.$$

We are now ready to define x^a for a real. Assume that $x > 1$ and $a > 0$. Consider the set

$$E_a := \{x^r : \ r \in \mathbb{Q}^+, \ r < a\}.$$

The set E_a is bounded from above and nonempty. We define $x^a := \sup E_a$.

Wednesday, January 26, 2022

Theorem 36 *Let $a, b \in \mathbb{R}$ with $a > 0$ and $b > 0$ and let $x \in \mathbb{R}$ with $x > 1$. Then*

$$x^a \cdot x^b = x^{a+b}$$

Proof. Consider the three sets

$$E_a := \{x^r : \ r \in \mathbb{Q}^+, \ r < a\},$$
$$E_b := \{x^s : \ s \in \mathbb{Q}^+, \ s < b\},$$
$$E_{a+b} := \{x^t : \ t \in \mathbb{Q}^+, \ t < a + b\},$$

and let $\ell_a = \sup E_a$, $\ell_b = \sup E_b$, and $\ell_{a+b} = \sup E_{a+b}$. Let's prove that

$$\ell_a \ell_b \le \ell_{a+b}.$$

If $r \in \mathbb{Q}^+$ is such that $r < a$ and $s \in \mathbb{Q}^+$ is such that $s < b$, then $r + s \in \mathbb{Q}^+$ and $r + s < a + b$. Hence,

$$x^r x^s = x^{r+s} \le \ell_{a+b}.$$

Fix $s \in \mathbb{Q}^+$ with $s < b$ and divide by x^s. Then

$$x^r \le \frac{\ell_{a+b}}{x^s}$$

for all $r \in \mathbb{Q}^+$ with $r < a$. This shows that the number $\frac{\ell_{a+b}}{x^s}$ is an upper bound for the set E_a. Hence,

$$\ell_a \le \frac{\ell_{a+b}}{x^s}.$$

Now rewrite this inequality as

$$x^s \le \frac{\ell_{a+b}}{\ell_a}.$$

Recall that $s \in \mathbb{Q}^+$ with $s < b$. Since the previous inequality is true for all such s, it shows that the number $\frac{\ell_{a+b}}{\ell_a}$ is an upper bound for the set E_a. Hence,

$$\ell_b \leq \frac{\ell_{a+b}}{\ell_a}.$$

Thus, we have proved that

$$\ell_a \ell_b \leq \ell_{a+b}.$$

Next let's prove that

$$\ell_{a+b} \leq \ell_a \ell_b.$$

Consider $t \in \mathbb{Q}^+$ with $t < a + b$. We want to find $p, q \in \mathbb{Q}^+$ with $t < p + q$, $p < a$ and $q < b$. Since $t - a < b$, by the density of the rationals there exists $q \in \mathbb{Q}$ such that $t - a < q < b$. Since $b > 0$ we can assume that $q > 0$ (if not apply the density of the rationals once more). Since $t - a < q$ we have that $t - q < a$ and so again by the density of the rationals there exists $p \in \mathbb{Q}$ such that $t - q < p < a$. Again, since $a > 0$ we can assume that $p > 0$ (if not apply the density of the rationals once more). Thus, $t < p + q$ and so by Exercises 34 and 35,

$$x^t < x^{p+q} = x^p \cdot x^q \leq \ell_a \ell_b.$$

Since this is true for all $t \in \mathbb{Q}^+$ with $t < a + b$ we have that $\ell_a \ell_b$ is an upper bound of the set E_{a+b} and so $\ell_{a+b} \leq \ell_a \ell_b$. ■

If $0 < x < 1$, we set

$$x^a := \left(x^{-1}\right)^{-a}.$$

Exercise 37 *Let $x > 0$ and $a, b \in \mathbb{R}$. Prove that*

$$\left(x^a\right)^b = \left(x^b\right)^a = x^{ab}.$$

Hint: It is enough to show $\left(x^a\right)^b = x^{ab}$. Consider first the case in which a is real and b is rational.

Given a number $x \in \mathbb{R}$, the *absolute value* of x is the number

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

The absolute value satisfies the following properties, which are left as as exercise.

Theorem 38 *Let $x, y, z \in \mathbb{R}$. Then the following properties hold.*

(i) $|x| \geq 0$ *for all* $x \in \mathbb{R}$*, with* $|x| = 0$ *if and only if* $x = 0$*,*

(ii) $|-x| = |x|$ *for all* $x \in \mathbb{R}$*,*

(iii) if $y \geq 0$ *and* $x \in \mathbb{R}$*, then* $|x| \leq y$ *if and only if* $-y \leq x \leq y$*,*

(iv) $-|x| \leq x \leq |x|$ *for all* $x \in \mathbb{R}$*,*

(iii) $|xy| = |x|\,|y|$ *for all* $x, y \in \mathbb{R}$*,*

(iv) $|x + y| \leq |x| + |y|$ *for all* $x, y \in \mathbb{R}$*.*

5 Inner Products, Norms, Distances

Definition 39 *A* vector space*, or* linear space*, over* $\mathbb{R}$ *is a nonempty set* X*, whose elements are called* vectors*, together with two operations,* addition *and* multiplication by scalars*,*

$$\begin{array}{c} X \times X \to X \\ (x, y) \mapsto x + y \end{array} \quad \textit{and} \quad \begin{array}{c} \mathbb{R} \times X \to X \\ (t, x) \mapsto tx \end{array}$$

with the properties that

(i) $(X, +)$ *is a commutative group, that is,*

- *(a)* $x + y = y + x$ *for all* $x, y \in X$ *(commutative property),*
- *(b)* $x + (y + z) = (x + y) + z$ *for all* $x, y, z \in X$ *(associative property),*
- *(c) there is a vector* $0 \in X$*, called* zero*, such that* $x + 0 = 0 + x = x$ *for all* $x \in X$*,*
- *(d) for every* $x \in X$ *there exists a vector in* X*, called the* opposite *of* x *and denoted* $-x$*, such that* $x + (-x) = 0$*,*

(ii) for all $x, y \in X$ *and* $s, t \in \mathbb{R}$*,*

- *(a)* $s\,(tx) = (st)\,x$*,*
- *(b)* $1x = x$*,*
- *(c)* $s\,(x + y) = (sx) + (sy)$*,*
- *(d)* $(s + t)x = (sx) + (tx)$*.*

Remark 40 *Instead of using real numbers, one can use a field* F*. For most or our purposes the real numbers will suffice.* From now on, whenever we don't specify, it is understood that a vector space is over $\mathbb{R}$.

Example 41 *Some important examples of vector spaces over* $\mathbb{R}$ *are the following.*

(i) The Euclidean space $\mathbb{R}^N$ *is the space of all* N*-tuples* $\boldsymbol{x} = (x_1, \ldots, x_N)$ *of real numbers. The elements of* $\mathbb{R}^N$ *are called* vectors *or* points*. The Euclidean space is a vector space with the following operations*

$$\boldsymbol{x} + \boldsymbol{y} := (x_1 + y_1, \ldots, x_N + y_N)\,, \quad t\boldsymbol{x} := (tx_1, \ldots, tx_N)$$

for every $t \in \mathbb{R}$ *and* $\boldsymbol{x} = (x_1, \ldots, x_N)$ *and* $\boldsymbol{y} = (y_1, \ldots, y_N)$ *in* $\mathbb{R}^N$*.*

(ii) The collection of all polynomials $p : \mathbb{R} \to \mathbb{R}$*.*

(iii) The space of continuous functions $f : [a, b] \to \mathbb{R}$*, where* $a < b$ *and* $[a, b] := \{x \in \mathbb{R} : \ a \leq x \leq b\}$*.*

Definition 42 *Given a set E and a function $f : E \to \mathbb{R}$, we say that f is* bounded from above *if the set*

$$f(E) := \{y \in \mathbb{R} : y = f(x), x \in E\}$$

is bounded from above. We say that f is bounded from below *if the set $f(E)$ is bounded from below. Finally, we say that f is* bounded *if the set $f(E)$ is bounded. We write*

$$\sup_E f := \sup f(E), \quad \inf_E f := \inf f(E).$$

Exercise 43 *Given a set E, consider the vector space $X := \{f : E \to \mathbb{R} \text{ bounded}\}$. Prove that X is a vector space.*

Friday, January 28, 2022

Example 44 *Consider the space $X = \{f : [a,b] \to \mathbb{R} : f \text{ is increasing in } [a,b]\}$. This is not a vector space since the difference of increasing functions is not increasing. The smallest (in the sense of inclusion) vector space that contains all increasing functions is the space of functions of pointwise bounded variation.*

Given $a < b$, consider the interval $[a,b]$. A *partition* of $[a,b]$ is a finite set $P := \{x_0, \ldots, x_n\} \subset [a,b]$, where

$$a = x_0 < x_1 < \cdots < x_n = b.$$

Given a function $f : [a,b] \to \mathbb{R}$, the *pointwise variation* of f on the interval $[a,b]$ is

$$\operatorname{Var} f := \sup \left\{ \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \right\},$$

where the supremum is taken over all partitions $P := \{x_0, \ldots, x_n\}$ of $[a,b]$, and all $n \in \mathbb{N}$. A function $f : [a,b] \to \mathbb{R}$ has *finite* or *bounded pointwise variation* if $\operatorname{Var} f < \infty$. The space of all functions $f : [a,b] \to \mathbb{R}$ of bounded pointwise variation is denoted by $BPV([a,b])$.

Exercise 45 *Prove that $BPV([a,b])$ is a vector space.*

Definition 46 *An* inner product*, or* scalar product*, on a vector space X is a function*

$$(\cdot, \cdot) : X \times X \to \mathbb{R}$$

such that

(i) $(x,x) \geq 0$ for every $x \in X$, $(x,x) = 0$ if and only if $x = 0$ (positivity);

(ii) $(x, y) = (y, x)$ *for all* $x, y \in X$ *(symmetry);*

(iii) $(sx + ty, z) = s\,(x, z) + t\,(y, z)$ *for all* $x, y, z \in X$ *and* $s, t \in \mathbb{R}$ *(bilinearity).*

An *inner product space* $(X, (\cdot, \cdot))$ is a vector space X endowed with an inner product $(\cdot, \cdot)$.

Example 47 *Some important examples of inner products are the following.*

(i) Consider the Euclidean space $\mathbb{R}^N$*, then*

$$\boldsymbol{x} \cdot \boldsymbol{y} := x_1 y_1 + \cdots + x_N y_N,$$

where $\boldsymbol{x} = (x_1, \ldots, x_N)$ *and* $\boldsymbol{y} = (y_1, \ldots, y_N)$*, is an inner product.*

(ii) Consider the space of X *of all integrable functions* $f : [a, b] \to \mathbb{R}$*. Then*

$$(f, g) := \int_a^b f(x)\, g(x)\, dx$$

is not an inner product. Indeed, if $f(x) = 0$ *for all* $x \in [a, b]$*,* $x \neq \frac{a+b}{2}$ *and* $f\left(\frac{a+b}{2}\right) = 1$*, then* $\int_a^b f^2(x)\, dx = 0$ *but* f *is not* 0*.*
To fix this problem one can take X *to be the space of all continuous functions* $f : [a, b] \to \mathbb{R}$*. Then*

$$(f, g) := \int_a^b f(x)\, g(x)\, dx$$

is an inner product.

Definition 48 *A* norm *on a vector space* X *is a map*

$$\|\cdot\| : X \to [0, \infty)$$

such that

(i) $\|x\| = 0$ *implies* $x = 0$*;*

(ii) $\|tx\| = |t|\, \|x\|$ *for all* $x \in X$ *and* $t \in \mathbb{R}$*;*

(iii) $\|x + y\| \leq \|x\| + \|y\|$ *for all* $x, y \in X$*.*

A *normed space* $(X, \|\cdot\|)$ is a vector space X endowed with a norm $\|\cdot\|$. For simplicity, we often say that X is a normed space.

Example 49 *Some important examples of norms are the following.*

(i) Consider the space $\mathbb{R}$*. By Theorem 38, the absolute value is a norm.*

(ii) Consider the Euclidean space $\mathbb{R}^N$, then

$$\|\boldsymbol{x}\| := \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}} = \sqrt{(x_1)^2 + \cdots + (x_N)^2},$$

where $\boldsymbol{x} = (x_1, \dots, x_N)$, is a norm. We will prove this below.

Exercise 50 *Given a set E, consider the vector space $X := \{f : E \to \mathbb{R} \text{ bounded}\}$. For $f \in X$, define*

$$\|f\| := \sup_E |f| .$$

Prove that $\|\cdot\|$ is a norm.

Given an inner product $(\cdot, \cdot) : X \times X \to \mathbb{R}$ on a vector space X, it turns out that the function

$$\|x\| := \sqrt{(x, x)}, \quad x \in X, \tag{5}$$

is a norm. This follows from the following result.

Proposition 51 (Cauchy–Schwarz's inequality) *Given an inner product $(\cdot, \cdot) : X \times X \to \mathbb{R}$ on a vector space X,*

$$|(x, y)| \leq \|x\| \, \|y\|$$

for all $x, y \in X$.

Proof. If $y = 0$, then both sides of the previous inequality are zeros, and so there is nothing to prove. Thus, assume that $y \neq 0$ and let $t \in \mathbb{R}$. By properties (i)-(iii),

$$0 \leq (x + ty, x + ty) = \|x\|^2 + t^2 \|y\|^2 + 2t\,(x, y) . \tag{6}$$

Taking

$$t := -\frac{(x, y)}{\|y\|^2}$$

in the previous inequality gives

$$0 \leq \|x\|^2 + \frac{(x, y)^2}{\|y\|^4} \|y\|^2 - 2\frac{(x, y)^2}{\|y\|^2},$$

or, equivalently,

$$(x, y)^2 \leq \|x\|^2 \|y\|^2 .$$

It now suffices to take the square root on both sides. ■

Monday, January 31, 2022

Remark 52 *It follows from the proof that equality holds in the Cauchy–Schwarz inequality if and only you have equality in (6), that is, if $x + ty = 0$ for some $t \in \mathbb{R}$ or $y = 0$.*

Corollary 53 *Given a scalar product* $(\cdot,\cdot) : X \times X \to \mathbb{R}$ *on a vector space* X*, the function*

$$\|x\| := \sqrt{(x,x)}, \quad x \in X,$$

is a norm.

Proof. By property (i), $\|\cdot\|$ is well-defined and $\|x\| = 0$ if and only if $x = 0$. Taking $t = 1$ in (6) and using the Cauchy–Schwarz inequality gives

$$\begin{aligned} 0 &\le \|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\,(x,y) \\ &\le \|x\|^2 + \|y\|^2 + 2\,\|x\|\,\|y\| = (\|x\| + \|y\|)^2 , \end{aligned}$$

which is the triangle inequality for the norm. Moreover, by properties (ii) and (iii) for every $t \in \mathbb{R}$,

$$\|tx\| = \sqrt{(tx,tx)} = \sqrt{t\,(x,tx)} = \sqrt{t\,(tx,x)} = \sqrt{t^2\,(x,x)} = |t|\,\|x\| .$$

Thus $\|\cdot\|$ is a norm. ■

Proposition 54 (Parallelogram law) *Given an inner product* $(\cdot,\cdot) : X \times X \to \mathbb{R}$ *on a vector space* X*,*

$$\|x+y\|^2 + \|x-y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2$$

for all $x, y \in X$*.*

Proof. Taking $t = \pm 1$ in (6), we get

$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + \|y\|^2 + 2\,(x,y) , \\ \|x-y\|^2 &= \|x\|^2 + \|y\|^2 - 2\,(x,y) . \end{aligned}$$

By adding these identities, we obtain the desired result. ■

Remark 55 *If instead of add, we subtract these two identities we get*

$$\|x+y\|^2 - \|x-y\|^2 = 4\,(x,y) ,$$

and so

$$(x,y) = \frac{1}{4}\left[\|x+y\|^2 - \|x-y\|^2\right] .$$

Exercise 56 *Prove that the following are norms in* $\mathbb{R}^N$*:*

$$\begin{aligned} \|\boldsymbol{x}\|_\infty &:= \max\{|x_1|, \ldots, |x_N|\} , \\ \|\boldsymbol{x}\|_1 &:= |x_1| + \cdots + |x_N| , \\ \|\boldsymbol{x}\|_p &:= (|x_1|^p + \cdots + |x_N|^p)^{1/p} , \end{aligned}$$

for $\boldsymbol{x} = (x_1, \ldots, x_N) \in \mathbb{R}^N$ *and where* $1 \le p < \infty$.

Example 57 *In $\mathbb{R}^N$ the norm $\|\cdot\|_\infty$ does not satisfy the parallelogram law. Take $\boldsymbol{x} = (1, 1, 0, \ldots)$, $\boldsymbol{y} = (1, -1, 0, \ldots)$. Then $\boldsymbol{x}+\boldsymbol{y} = (2, 0, \ldots)$, $\boldsymbol{x}-\boldsymbol{y} = (0, 2, 0, \ldots)$. Hence,*

$$\begin{aligned}\|\boldsymbol{x}+\boldsymbol{y}\|_\infty^2 + \|\boldsymbol{x}-\boldsymbol{y}\|_\infty^2 &= 4 + 4 = 8 \\ &\neq 2\|\boldsymbol{x}\|_\infty^2 + 2\|\boldsymbol{y}\|_\infty^2 \\ &= 2 + 2.\end{aligned}$$

Example 58 *In $\mathbb{R}^N$ the norm $\|\cdot\|_p$ for $p \neq 2$ does not satisfy the parallelogram law. Take $\boldsymbol{x} = (1, 1, 0, \ldots)$, $\boldsymbol{y} = (1, -1, 0, \ldots)$. Then $\boldsymbol{x}+\boldsymbol{y} = (2, 0, \ldots)$, $\boldsymbol{x}-\boldsymbol{y} = (0, 2, 0, \ldots)$. Hence,*

$$\begin{aligned}\|\boldsymbol{x}+\boldsymbol{y}\|_p^2 + \|\boldsymbol{x}-\boldsymbol{y}\|_p^2 &= (2^p)^{\frac{2}{p}} + (2^p)^{\frac{2}{p}} = 8 \\ &\neq 2\|\boldsymbol{x}\|_p^2 + 2\|\boldsymbol{y}\|_p^2 \\ &= 2(1^p + 1^p)^{\frac{2}{p}} + 2(1^p + 1^p)^{\frac{2}{p}} = 2^{2+\frac{2}{p}}.\end{aligned}$$

Exercise 59 *Consider the vector space $X := \{f : [0, 1] \to \mathbb{R} \text{ bounded}\}$ with*

$$\|f\| := \sup_{[0,1]} |f|.$$

Let's prove that $\|\cdot\|$ does not satisfy the parallelogram law. Take $f(x) = -x^2$ and $g(x) = x$. Then

$$\begin{aligned}\sup_{[0,1]} |f+g| &= \sup_{[0,1]} |-x^2+x| = \max_{[0,1]}(x-x^2) = \frac{1}{4}, \\ \sup_{[0,1]} |f-g| &= \sup_{[0,1]} |-x^2-x| = \max_{[0,1]}(x^2+x) = 2, \\ \sup_{[0,1]} |f| &= \sup_{[0,1]} |-x^2| = \max_{[0,1]} x^2 = 1, \\ \sup_{[0,1]} |g| &= \sup_{[0,1]} |g| = \max_{[0,1]} x = 1,\end{aligned}$$

and so

$$\begin{aligned}\left(\sup_{[0,1]} |f+g|\right)^2 + \left(\sup_{[0,1]} |f-g|\right)^2 &= \frac{1}{16} + 4 = \\ &\neq 2\left(\sup_{[0,1]} |f|\right)^2 + 2\left(\sup_{[0,1]} |g|\right)^2 \\ &= 2 + 2.\end{aligned}$$

Exercise 60 *Let $(X, \|\cdot\|)$ be a normed space. Prove that there exists an inner product $(\cdot, \cdot) : X \times X \to \mathbb{R}$ such that $\|x\| = \sqrt{(x, x)}$ for all $x \in X$ if and only if $\|\cdot\|$ satisfies the parallelogram law*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

for all $x, y \in X$.

Definition 61 *A* metric *on a set X is a map $d : X \times X \to [0, \infty)$ such that*

(i) $d(x, y) = 0$ *if and only if* $x = y$,

(ii) $d(x, y) = d(y, x)$ *for all* $x, y \in X$ *(symmetry),*

(iii) $d(x, y) \leq d(x, z) + d(z, y)$ *for all* $x, y, z \in X$ *(triangle inequality).*

A metric space (X, d) *is a set X endowed with a metric d. When there is no possibility of confusion, we abbreviate by saying that X is a metric space.*

Proposition 62 *Let $(X, \|\cdot\|)$ be a normed space. Then*

$$d(x, y) := \|x - y\|$$

is a metric.

Proof. By property (i) in Definition 48, we have that $0 = d(x, y) = \|x - y\|$ if and only if $x - y = 0$, that is, $x = y$.

By property (ii) in Definition 48, we obtain that

$$d(y, x) = \|y - x\| = \|-1(x - y)\| = |-1| \cdot \|x - y\| = \|x - y\| = d(x, y).$$

Finally, by property (ii) in Definition 48,

$$d(x, y) = \|x - y\| = \|x - z + z - y\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y).$$

■

Exercise 63 *Prove that in $\mathbb{R}$ the function*

$$d_1(x, y) := \left| \frac{x}{1 + |x|} - \frac{y}{1 + |y|} \right| \tag{7}$$

is a metric.

Wednesday, February 2, 2022

Definition 64 *Given a metric space (X, d), a point $x_0 \in X$, and $r > 0$, the* ball *centered at x_0 and of radius r is the set*

$$B(x_0, r) := \{x \in X : d(x, x_0) < r\}.$$

Definition 65 *Given a metric space (X, d), and a nonempty set $E \subseteq X$, a point $x \in E$ is called an* interior point *of E if there exists $r > 0$ such that $B(x, r) \subseteq E$. The* interior *E° of a set $E \subseteq \mathbb{R}^N$ is the union of all its interior points. A subset $U \subseteq X$ is* open *if every $x \in U$ is an interior point of U.*

Example 66 *Given a metric space (X, d), the ball $B(x_0, r)$ is open. To see this, let $x \in B(x_0, r)$. Then $B(x, r - d(x, x_0))$ is contained in $B(x_0, r)$. Indeed, if $y \in B(x_0, r - d(x, x_0))$, then*

$$d(y, x_0) \leq d(y, x) + d(x, x_0) < r - d(x, x_0) + +d(x, x_0) = r,$$

and so $y \in B(x_0, r)$.

Example 67 *Some simple examples of sets that are open and of some that are not.*

(i) The set $(a, \infty) = \{x \in \mathbb{R} : x > a\}$ is open. Indeed, if $x > a$, take $r := x - a > 0$. Then $B(x, r) \subset (a, \infty)$. Similarly, the set $(-\infty, a)$ is open.

(ii) The set $(a, b) = \{x \in \mathbb{R} : a < x < b\}$ is open. Indeed, given $a < x < b$, take $r := \min\{b - x, x - a\} > 0$. Then $B(x, r) \subseteq (a, b)$.

(iii) The set $(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$ is not open, since b belongs to the set but there is no ball $B(b, r)$ contained in $(a, b]$.

Example 68 *Consider the set $E = (0, 1) \cap \mathbb{Q}$. The interior of this set is empty. Indeed, if $x \in E$ and $r > 0$, by the density of the irrationals, we can find $y \in \mathbb{R} \setminus \mathbb{Q}$ such that $\min\{x - r, 0\} < y < \max\{x + r, 1\}$. Hence, the ball $B(x, r) = (x - r, x + r)$ contains y, which is not a point of E.*

Example 69 *Consider the set*

$$U = \mathbb{R} \setminus \left(\{0\} \cup \left\{\frac{1}{n} : n \in \mathbb{N}\right\}\right).$$

Let's prove that U is open. If $x < 0$, take $r = -x > 0$, then $B(x, r) = (-2x, 0) \subseteq U$. If $x > 1$, take $r = x - 1$, then $B(x, r) = (1, 2x - 1) \subseteq U$. If $\frac{1}{n+1} < x < \frac{1}{n}$, take $r = \min\left\{\frac{1}{n} - x, x - \frac{1}{n+1}\right\} = \frac{1}{n+1}$, then $B(x, r) \subseteq U$. Hence, U is open.

Example 70 *Consider the set*

$$E = \mathbb{R} \setminus \left(\left\{\frac{1}{n} : n \in \mathbb{N}\right\}\right).$$

Let's prove that E is not open. The point $x = 0$ belongs to E, but for every $r > 0$, by the Archimedean principle we can find $n \in \mathbb{N}$ such that $n > \frac{1}{r}$, and so $0 < \frac{1}{n} < r$, which shows that $\frac{1}{n} \in (-r, r)$. Since $\frac{1}{n}$ does not belong to E, the ball $(-r, r)$ is not contained in E for any *$r > 0$. Hence, E is not open.*

The main properties of open sets are given in the next proposition.

In what follows by an arbitrary family of sets of X we mean that there exists a set I and a function

$$f : I \to \mathcal{P}(X)$$
$$\alpha \in I \mapsto f(\alpha) = U_\alpha$$

We write $\{U_\alpha\}$ or $\{U_\alpha\}_I$ or $\{U_\alpha\}_{\alpha \in I}$ to denote the set $\{f(\alpha) : \alpha \in I\}$.

Proposition 71 *Given a metric space* (X, d)*, the following properties hold:*

(i) $\emptyset$ *and* X *are open.*

(ii) If $U_i \subseteq X$*,* $i = 1, \ldots, n$*, is a finite family of open sets of* X*, then* $U_1 \cap \cdots \cap U_n$ *is open.*

(iii) If $\{U_\alpha\}_\alpha$ *is an arbitrary collection of open sets of* X*, then* $\bigcup_\alpha U_\alpha$ *is open.*

Proof. To prove (ii), let $x \in U_1 \cap \cdots \cap U_M$. Then $x \in U_i$ for every $i = 1, \ldots, n$, and since U_i is open, there exists $r_i > 0$ such that $B(x, r_i) \subseteq U_i$. Take $r := \min\{r_1, \ldots, r_n\} > 0$. Then

$$B(x, r) \subseteq U_1 \cap \cdots \cap U_n,$$

which shows that $U_1 \cap \cdots \cap U_n$ is open.

To prove (iii), let $x \in U := \bigcup_\alpha U_\alpha$. Then there is α such that $x \in U_\alpha$ and since U_α is open, there exists $r > 0$ such that $B(x, r) \subseteq U_\alpha \subseteq U$. This shows that U is open. ■

Friday, February 4, 2022

Properties (i)–(iii) are used to define topological spaces.

Definition 72 *Let* X *be a nonempty set and let* τ *be a family of sets of* X*. The pair* (X, τ) *is called a* topological space *if the following hold.*

(i) $\emptyset, X \in \tau$*.*

(ii) If $U_i \in \tau$ *for* $i = 1, \ldots, M$*, then* $U_1 \cap \ldots \cap U_M \in \tau$*.*

(iii) If $\{U_\alpha\}_\alpha$ *is an arbitrary collection of elements of* τ*, then* $\bigcup_\alpha U_\alpha \in \tau$*.*

The elements of the family τ *are called* open sets.

Remark 73 *The intersection of infinitely many open sets is not open in general. Take* $U_n := \left(-\frac{1}{n}, \frac{1}{n}\right)$ *for* $n \in \mathbb{N}$*. Then*

$$\bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\},$$

but $\{0\}$ *is not open. Indeed, for every* $r > 0$*, the ball* $(-r, r)$ *is not contained in* $\{0\}$*.*

Remark 74 *Proposition 71 shows that the family of open sets in* $\mathbb{R}^N$ *defined in Definition 65 is a a topology, called the* Euclidean topology*. Unless specified, in* $\mathbb{R}^N$ *we will always consider the Euclidean topology.*

Example 75 *Given a nonempty set X, there are always at least two topologies on X, namely,*

$$\tau_1 = \{\emptyset, X\}$$

(so according to τ_1, the only open sets are the empty set and X) and

$$\tau_2 = \{\text{all subsets of } X\}$$

(so according to τ_2 every set $E \subseteq X$ is open).

Exercise 76 *Let $1 \leq p < \infty$. Prove that in $\mathbb{R}^N$ the norms*

$$\begin{aligned} \|\boldsymbol{x}\|_\infty &:= \max\{|x_1|, \ldots, |x_N|\}, \\ \|\boldsymbol{x}\| &:= \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}}, \\ \|\boldsymbol{x}\|_p &:= (|x_1|^p + \cdots + |x_N|^p)^{1/p}, \end{aligned}$$

generate the same topology.

The proof of following proposition is left as an exercise.

Proposition 77 *Given a metric space (X, d). Then*

(i) E° is an open subset of E,

(ii) E° is given by the union of all open subsets contained in E; that is, E° is the largest (in the sense of union) open set contained in E,

(iii) E is open if and only if $E = E^\circ$,

(iv) $(E^\circ)^\circ = E^\circ$.

Example 78 *Consider the set $E = [0, 1)$. Then 0 is not an in interior point of E, so $E^\circ \subseteq (0, 1)$. On the other hand, since $(0, 1)$ is open and contained in E, by part (ii) of the previous proposition, $E^\circ \supseteq (0, 1)$, which shows that $E^\circ = (0, 1)$.*

Exercise 79 *Some properties of the interior.*

(i) Prove that if E, F are subsets of $\mathbb{R}^N$, then

$$\begin{aligned} E^\circ \cap F^\circ &= (E \cap F)^\circ, \\ E^\circ \cup F^\circ &\subseteq (E \cup F)^\circ. \end{aligned}$$

(ii) Show that in general $E^\circ \cup F^\circ \neq (E \cup F)^\circ$.

(iii) Let $\{E_\alpha\}_\alpha$ be an arbitrary collection of sets of $\mathbb{R}^N$. What is the relation, if any, between $\bigcap_\alpha (U_\alpha)^\circ$ and $(\bigcap_\alpha U_\alpha)^\circ$? And between $\bigcup_\alpha (U_\alpha)^\circ$ and $(\bigcup_\alpha U_\alpha)^\circ$?

Definition 80 *Given a metric space (X, d), A subset $C \subseteq X$ is* closed *if its complement $X \setminus C$.*

The main properties of closed sets are given in the next proposition.

Proposition 81 *Given a metric space (X, d), the following properties hold:*

(i) $\emptyset$ *and* X *are closed.*

(ii) If $C_i \subseteq X$, $i = 1, \ldots, n$, is a finite family of closed sets of X, then $C_1 \cup \cdots \cup C_n$ is closed.

(iii) If $\{C_\alpha\}_\alpha$ is an arbitrary collection of closed sets of X, then $\bigcap_\alpha C_\alpha$ is closed.

The proof follows from Proposition 71 and De Morgan's laws. If $\{E_\alpha\}_\alpha$ is an arbitrary collection of subsets of a set $\mathbb{R}^N$, then *De Morgan's laws* are

$$X \setminus \left(\bigcup_\alpha E_\alpha \right) = \bigcap_\alpha (X \setminus E_\alpha),$$
$$X \setminus \left(\bigcap_\alpha E_\alpha \right) = \bigcup_\alpha (X \setminus E_\alpha).$$

Remark 82 *Note that the majority of sets are neither open nor closed. The set $E = (0, 1]$ is neither open nor closed.*

Definition 83 *Given a metric space (X, d) and a set $E \subseteq X$, the* closure *of E, denoted $\overline{E}$, is the intersection of all closed sets that contain E*

In other words, the closure of E is the smallest (with respect to inclusion) closed set that contains E. It follows by Proposition 81 that $\overline{E}$ is closed.

The proof of following proposition is left as an exercise.

Proposition 84 *Given a metric space (X, d), let $C \subseteq X$. Then C is closed if and only if $C = \overline{C}$.*

The previous proposition leads us to the definition of accumulation points.

Definition 85 *Given a metric space (X, d) and a set $E \subseteq X$, a point $x \in X$ is a* boundary point of E *if for every $r > 0$ the ball $B(x, r)$ contains at least one point of E and one point of $X \setminus E$. The set of all boundary points of E is denoted ∂E.*

Proposition 86 *Given a metric space (X, d), let $E \subseteq X$. Then*

$$\overline{E} = E \cup \partial E,$$

Proof. Let $x \in \overline{E}$ and assume by contradiction that $x \notin E \cup \partial E$. Since $x \notin \partial E$, there exists a ball $B(x,r)$ that either does not intersect E or does not intersect the complement of E. But since $x \notin E$, only the first possibility can occur. Hence, there exists $r > 0$ such that $B(x,r) \cap E = \emptyset$. Since $B(x,r)$ is open and $B(x,r) \cap E = \emptyset$, it follows that $X \setminus B(x,r)$ is closed and contains E. By the definition of $\overline{E}$, we have that $\overline{E} \subseteq X \setminus B(x,r)$, which contradicts the fact that $x \in \overline{E}$.

Conversely, let $x \in E \cup \partial E$ and assume that $x \notin \overline{E}$. Since $\overline{E}$ is closed, $X \setminus \overline{E}$ is open. Using the fact that $x \in X \setminus \overline{E}$, we can find $B(x,r) \subseteq X \setminus \overline{E}$, which contradicts the fact that $B(x,r) \cap E \neq \emptyset$. ■

Monday, February 7, 2022

Definition 87 *Given a metric space (X,d) and a set $E \subseteq X$, a point $x \in X$ is an* accumulation point*, or* cluster point *of E if for every $r > 0$ the ball $B(x,r)$ contains at least one point of E different from x. The set of all accumulation points of E is denoted* $\operatorname{acc} E$.

Note that x does not necessarily belong to the set E.

Remark 88 *Note take if $\boldsymbol{x} \in \mathbb{R}^N$ is an accumulation point of E, then by taking $r = \frac{1}{n}$, $n \in \mathbb{N}$, there exists a sequence $\{\boldsymbol{x}_n\} \subseteq E$ with $\boldsymbol{x}_n \neq \boldsymbol{x}$ for all $n \in \mathbb{N}$ such that $\|\boldsymbol{x}_n - \boldsymbol{x}\| < \frac{1}{n} \to 0$. Thus $\{x_n\}$ converges to x. Conversely, if there exists $\{\boldsymbol{x}_n\} \subseteq E$ with $\boldsymbol{x}_n \neq \boldsymbol{x}$ for all $n \in \mathbb{N}$ such that $\|\boldsymbol{x}_n - \boldsymbol{x}\| \to 0$, then $\boldsymbol{x}$ is an accumulation point of E.*

It turns out that the closure of a set is given by the set and all its accumulations points.

Proposition 89 *Given a metric space (X,d) and a set $E \subseteq X$, then*

$$\overline{E} = E \cup \operatorname{acc} E.$$

In particular, a set $C \subseteq X$ is closed if and only if C contains all its accumulation points.

Proof. Exercise. ■

Exercise 90 *(i) Prove that if $E_1, \dots, E_n$ are subsets of $\mathbb{R}^N$, then*

$$\overline{E_1} \cap \cdots \cap \overline{E_n} \supseteq \overline{E_1 \cap \cdots \cap E_n},$$
$$\overline{E_1} \cup \cdots \cup \overline{E_n} = \overline{E_1 \cup \cdots \cup E_n}.$$

(ii) Show that in general $\overline{E_1} \cap \cdots \cap \overline{E_n} \neq \overline{E_1 \cap \cdots \cap E_n}$.

(iii) Let $\{E_\alpha\}_\alpha$ be an arbitrary collection of sets of $\mathbb{R}^N$. What is the relation, if any, between between $\bigcap_\alpha \overline{E_\alpha}$ and $\overline{\bigcap_\alpha E_\alpha}$? And between $\bigcup_\alpha \overline{E_\alpha}$ and $\overline{\bigcup_\alpha E_\alpha}$?

Definition 91 *Given a metric space* (X,d)*, a set* $E \subseteq X$ *is* bounded *if it is contained in a ball.*

Theorem 92 (Bolzano–Weierstrass) *Every bounded set* $E \subseteq \mathbb{R}^N$ *with infinitely many elements has at least one accumulation point.*

The proof relies on a few preliminary results, which are of interest in themselves.

Lemma 93 *Let* $\{[a_n, b_n]\}_n$ *be a sequence of closed bounded intervals such that* $[a_n, b_n] \supseteq [a_{n+1}, b_{n+1}]$ *for all* $n \in \mathbb{N}$*. Then the intersection*

$$\bigcap_{n=1}^{\infty} [a_n, b_n]$$

is nonempty.

Proof. Since

$$\cdots \subseteq [a_{n+1}, b_{n+1}] \subseteq [a_n, b_n] \subseteq \cdots \subseteq [a_1, b_1],$$

we have that

$$a_1 \leq \cdots \leq a_n \leq a_{n+1} \leq \cdots, \tag{8}$$
$$b_1 \geq \cdots \geq b_n \geq b_{n+1} \geq \cdots. \tag{9}$$

Let

$$A := \{a_1, \ldots, a_n, \ldots\}.$$

By (8) and (9), for $n \in \mathbb{N}$,

$$a_n \leq b_n \leq b_1.$$

Hence, A is bounded from above, and so by the supremum property, there exists $x := \sup A \in \mathbb{R}$ and

$$a_n \leq x$$

for all $n \in \mathbb{N}$. We claim that $x \leq b_n$ for all $n \in \mathbb{N}$. If not, then there exists $m \in \mathbb{N}$ such that $b_m < x$. Since x is the least upper bound of A, there exists $n \in \mathbb{N}$ such that $b_m < a_n \leq x$. Find $k \geq m, n$. Then by (8) and (9),

$$b_m < a_n \leq a_k \leq b_k \leq b_m,$$

which is a contradiction. This proves the claim. Hence, $x \in [a_n, b_n]$ for all $n \in \mathbb{N}$, and so $x \in \bigcap_{n=1}^{\infty} [a_n, b_n]$. ■

Given N bounded intervals $I_1, \ldots, I_N \subset \mathbb{R}$, a *rectangle* in $\mathbb{R}^N$ is a set of the form

$$R := I_1 \times \cdots \times I_N.$$

If all the intervals have the same length, we call R a cube.

Lemma 94 *Let $\{R_n\}_n$ be a sequence of closed bounded rectangles in $\mathbb{R}^N$ such that $R_n \supseteq R_{n+1}$ for all $n \in \mathbb{N}$. Then the intersection*

$$\bigcap_{n=1}^{\infty} R_n$$

is nonempty.

Proof. Each rectangle R_n has the form

$$R_n = [a_{n,1}, b_{n,1}] \times \cdots \times [a_{n,N}, b_{n,N}] \,.$$

Since $R_n \supseteq R_{n+1}$ for all $n \in \mathbb{N}$, for every fixed $k = 1, \ldots, N$, we have that $[a_{n,k}, b_{n;k}] \supseteq [a_{n+1,k}, b_{n+1,k}]$ for all $n \in \mathbb{N}$, and so by the previous lemma there exists $x_k \in \bigcap_{n=1}^{\infty} [a_{n,k}, b_{n;k}]$. Define $\boldsymbol{x} = (x_1, \ldots, x_N)$. Then $\boldsymbol{x} = (x_1, \ldots, x_N) \in [a_{n,1}, b_{n,1}] \times \cdots \times [a_{n,N}, b_{n,N}] = R_n$ for every $n \in \mathbb{N}$, and so $\boldsymbol{x} \in \bigcap_{n=1}^{\infty} R_n$. ■

Wednesday, February 9, 2022

We are now ready to prove the Bolzano–Weierstrass theorem.

Proof of the Bolzano–Weierstrass theorem. Since E is bounded, it is contained in ball, and in turn a ball is contained in a cube Q_1 of side-length ℓ. Divide Q_1 into 2^N two closed cubes of side-length $\frac{\ell}{2}$. Since E has infinitely many elements, at least one of these 2^N closed cubes contains infinitely many elements of E. Let's call this closed interval Q_2. Then $Q_2 \subset Q_1$, and Q_2 contains infinitely many elements of E.

Divide Q_2 into into 2^N two closed cubes of side-length $\frac{\ell}{2^2}$. Since E has infinitely many elements, at least one of these 2^N closed cubes contains infinitely many elements of E. Let's call this closed interval Q_3. By induction, we construct a sequence of closed cubes Q_n, $n \in \mathbb{N}$, with $Q_n \supseteq Q_{n+1}$, such that the side-length of Q_n is $\frac{\ell}{2^{n-1}}$ and Q_n contains infinitely many elements of E. By the previous lemma, there exists $\boldsymbol{x} \in \bigcap_{n=1}^{\infty} Q_n$. We claim that $\boldsymbol{x}$ is an accumulation point of E.

Fix $r > 0$ and consider the ball $B(\boldsymbol{x}, r)$. We claim that for n sufficiently large, $Q_n \subset B(\boldsymbol{x}, r)$. To see this, let $\boldsymbol{y} \in Q_n$. Then

$$\|\boldsymbol{y} - \boldsymbol{x}\| = \sqrt{(y_1 - x_1)^2 + \cdots + (y_1 - x_1)^2} < \sqrt{N \left(\frac{\ell}{2^{n-1}}\right)^2} = \frac{2\ell}{2^n}\sqrt{N}$$

By the Archimedean property, there exists $n \in \mathbb{N}$ such that

$$\frac{2\ell\sqrt{N}}{r} < 1 + n \leq 2^n,$$

and so $r > \frac{2\ell}{2^n}\sqrt{N}$, which proves the claim. Since Q_n contains infinitely many elements of E, the same holds for $B(\boldsymbol{x}, r)$ and so $\boldsymbol{x}$ is an accumulation point of E. ■

Definition 95 *Given a metric space (X, d) and a set $E \subseteq X$, a point $x \in$ is a* boundary point *of E if for every $r > 0$ the ball $B(x, r)$ contains at least one point of E and one point of $X \setminus E$. The set of boundary points of E is denoted ∂E.*

The following theorem is left as an exercise.

Theorem 96 *Let $E \subseteq \mathbb{R}^N$. Then*

(i) $\overline{E} = E \cup \partial E$,

(ii) *E is closed if and only if it contains all its boundary points,*

(iii) $\partial E = \partial \left(\mathbb{R}^N \setminus E\right)$,

(iv) $\partial E = \overline{(\mathbb{R}^N \setminus E)} \cap \overline{E}$.

6 Compactness

Exercise 97 *Let $K \subseteq \mathbb{R}^N$ be closed and bounded. Prove that if $E \subseteq K$ has infinitely many elements, then E has an accumulation point that belongs to K.*

Exercise 98 *Let $K_n \subset \mathbb{R}^N$ be nonempty, bounded, and closed. Assume that $K_n \supseteq K_{n+1}$ for all $n \in \mathbb{N}$. Prove that $\bigcap_{n=1}^{\infty} K_n$ is nonempty.*

Definition 99 *Given a metric space (X, d), a set $K \subseteq X$ is* compact *if for every open cover of K, i.e., for every collection $\{U_\alpha\}_\alpha$ of open sets such that $\bigcup_\alpha U_\alpha \supseteq K$, there exists a finite subcover (i.e., a finite subcollection of $\{U_\alpha\}_\alpha$ whose union still contains K).*

Example 100 *The set $(0, 1]$ is not compact, since taking $U_n := \left(\frac{1}{n}, 2\right)$, a finite number of U_n does not cover $(0, 1]$.*

Here the problem is that 0 does not belong to E. But what if E is closed?

Example 101 *The set $[0, \infty)$ is not compact, since taking $U_n := (-1, n)$, a finite number of U_n does not cover $[0, \infty)$.*

Here E is closed but the problem is that E is not bounded.

Theorem 102 *Given a metric space (X, d), a compact set $K \subseteq X$ is closed and bounded.*

Proof. To prove that K is closed, we show that $X \setminus K$ is open. Fix $x \in X \setminus K$. For every $y \in K$ consider the balls $B(y, r_y)$ and $B(x, r_y)$, where $r := \frac{d(x,y)}{4}$. These two balls do not intersect each other (why?). Then $\{B(y, r_y)\}_{y \in K}$ is an open cover of K, and so there exist $y_1, \ldots, y_m \in K$ such that

$$K \subseteq \bigcup_{i=1}^{m} B(y_i, r_{y_i}).$$

Let $r := \min\{r_{y_1}, \ldots, r_{y_m}\} > 0$. Then $x \in B(x, r)$ and the ball $B(x, r)$ does not intersect $B(y_i, r_{y_i})$ for any $i = 1, \ldots, m$. Hence, $B(x, r)$ is contained in $X \setminus K$.

This shows that every point x of $X \setminus K$ is an interior point, and so $X \setminus K$ is open.

To prove that K is bounded, consider a point $x_0 \in X$ and $B(x_0, n)$. The family of balls $\{B(x_0, n)\}_{n\in\mathbb{N}}$ covers the entire space X and in particular K. By compactness K is contained in a finite number of balls. Since the balls are one contained into the other, we have that K is contained in the ball of largest radius. Hence, K is bounded. ■

Remark 103 *For a topological space (X, τ) we can still prove that a compact set $K \subseteq X$ is closed, provided the topological space X is a* Hausdorff space*, that is, for every x and $y \in X$, with $x \neq y$, there exist disjoint neighborhoods of x and y.*

A very simple example of a space that is not Hausdorff can be obtained by considering a nonempty set X and taking as topology $\tau := \{\emptyset, X\}$. If X has at least two elements, then any singleton $\{x\}$ is compact but not closed.

There is a way to define a notion of boundedness for special topological spaces, called topological vector spaces.

Friday, February 11, 2022

Theorem 104 *A closed and bounded set $K \subset \mathbb{R}^N$ is compact.*

Proof. Let $\{U_\alpha\}_\alpha$ be a family of open sets such that $\bigcup_\alpha U_\alpha \supseteq K$ and assume by contradiction that no finite subcover covers K. Since K is bounded, it is contained in ball, and in turn a ball is contained in a cube Q_1 of side-length ℓ. Divide Q_1 into 2^N two closed cubes of side-length $\frac{\ell}{2}$. If $K \cap Q'$ is contained in a finite subcover for every such subcube, then K would be contained in a finite subcover. Hence, there exists at least one subcube Q_1 such that $K \cap Q_1$ is not contained in a finite subcover of $\{U_\alpha\}_\alpha$. Note that this imply, in particular, that $K \cap Q_1$ has infinitely many distinct elements.

By induction, we construct a sequence of closed cubes Q_n, $n \in \mathbb{N}$, with $Q_n \supseteq Q_{n+1}$, such that the side-length of Q_n is $\frac{\ell}{2^{n-1}}$ and $K \cap Q_n$ is not contained in a finite subcover of $\{U_\alpha\}_\alpha$. Again, this implies that $K \cap Q_n$ has infinitely many distinct elements. As in the proof of the Bolzano–Weierstrass theorem, there exists $\boldsymbol{x} \in \bigcap_{n=1}^\infty Q_n$ and $\boldsymbol{x}$ is an accumulation point of K. Since K is closed, K contains all its accumulation points (exercise). Hence, $\boldsymbol{x} \in K$. Since $\{U_\alpha\}_\alpha$ covers K, there exists β such that $\boldsymbol{x} \in U_\beta$. On the other hand, U_β is open, and so, there is a ball $B(\boldsymbol{x}, r)$ contained in U_β. As in the proof of the Bolzano–Weierstrass theorem, we have that for n sufficiently large, $Q_n \subseteq B(\boldsymbol{x}, r) \subseteq U_\beta$, which contradicts the fact that $K \cap Q_n$ is not contained in a finite subcover of $\{U_\alpha\}_\alpha$. ■

Remark 105 *The previous theorem fails for infinite dimensional normed spaces, and so, in general, for infinite dimensional metric spaces.*

7 Functions

Given two sets X and Y consider a function $f : E \to Y$, where $E \subseteq X$. The set E is called the *domain* of f. When $X = \mathbb{R}^M$, if E is not specified, then E should be taken to be the largest set of x for which $f(x)$ makes sense. This means that:

If there are even roots, their arguments should be nonnegative. If there are logarithms, their arguments should be strictly positive. Denominators should be different from zero. If a function is raised to an irrational number, then the function should be nonnegative.

Given a set $F \subseteq E$, the set $f(F) = \{y \in Y : y = f(x) \text{ for some } x \in F\}$ is called the *image* of F through f.

Given a set $G \subseteq \mathbb{R}$, the set $f^{-1}(G) = \{x \in E : f(x) \in G\}$ is called the *inverse image* or *preimage* of F through f. It has NOTHING to do with the inverse function. It is just one of those unfortunate cases in which we use the same symbol for two different objects.

The *graph* of a function is the set of $X \times Y$ defined by

$$\operatorname{gr} f = \{(x, f(x)) : x \in E\}.$$

A function f is said to be

- *one-to-one* or *injective* if $f(x) \neq f(z)$ for all $x, z \in E$ with $x \neq z$.
- *onto* or *surjective* if $f(E) = F$,
- *bijective* or *invertible* if it is one-to-one and onto. The function $f^{-1} : F \to E$, which assigns to each $y \in F = f(E)$ the unique $x \in E$ such that $f(x) = y$, is called the *inverse* function of f.

8 Limits of Functions

Definition 106 *If (X, d_X) and (Y, d_Y) are two metric spaces, $E \subseteq X$, $x_0 \in X$ is an accumulation point of E and $f : E \to Y$, we say that $\ell \in Y$ is the* limit *of $f(x)$ as x approaches x_0 if for every $\varepsilon > 0$ there exists a real number $\delta = \delta(\varepsilon, x_0) > 0$ with the property that*

$$d_Y(f(x), \ell) < \varepsilon$$

for all $x \in E$ with $0 < d_X(x, x_0) < \delta$. We write

$$\lim_{x \to x_0} f(x) = \ell \quad \text{or} \quad f(x) \to \ell \text{ as } x \to x_0.$$

Remark 107 *Note that even when $x \in E$, we cannot take $x = x_0$ since in the definition we require $0 < d_X(x, x_0)$.*

Remark 108 *Let $E \subseteq \mathbb{R}^N$, let $\boldsymbol{x}_0 \in \mathbb{R}^N$ be an accumulation point of E, and let $\boldsymbol{f} : E \to \mathbb{R}^N$. We say that a number $\boldsymbol{\ell} \in \mathbb{R}^M$ is the* limit *of $\boldsymbol{f}(\boldsymbol{x})$ as $\boldsymbol{x}$ approaches $\boldsymbol{x}_0$ if for every $\varepsilon > 0$ there exists a real number $\delta = \delta(\varepsilon, \boldsymbol{x}_0) > 0$ with the property that*

$$\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{\ell}\| < \varepsilon$$

for all $\boldsymbol{x} \in E$ with $0 < \|\boldsymbol{x} - \boldsymbol{x}_0\| < \delta$. We write

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{\ell} \quad \text{or} \quad \boldsymbol{f}(\boldsymbol{x}) \to \boldsymbol{\ell} \text{ as } \boldsymbol{x} \to \boldsymbol{x}_0.$$

Remark 109 *If (X, τ_X) and (Y, τ_Y) are two topological spaces, $E \subseteq X$, $x_0 \in X$ is an accumulation point of E and $f : E \to Y$, we say that $\ell \in Y$ is the* limit *of $f(x)$ as x approaches x_0 if for every neighborhood V of ℓ there exists a neighborhood U of x_0 with the property that*

$$f(x) \in V$$

for all $x \in E$ with $x \in U \setminus \{x_0\}$. We write

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

Note that unless the space Y is Hausdorff, the limit may not be unique.

Theorem 110 *Let (X, d_X) and (Y, d_Y) be two metric spaces, let $E \subseteq X$, let $x_0 \in X$ be an accumulation point of E and $f : E \to Y$. If the limit*

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

exists, it is unique.

Proof. Assume by contradiction that there exist

$$\lim_{x \to x_0} f(x) = \ell \quad \text{and} \quad \lim_{x \to x_0} f(x) = L$$

with $\ell \neq L$. Then $d_Y(\ell, L) > 0$. Fix $0 < \varepsilon = \frac{1}{2} d_Y(\ell, L)$. Since $\lim_{x \to x_0} f(x) = \ell$, there exists $\delta_1 > 0$ with the property that

$$d_Y(f(x), \ell) < \varepsilon$$

for all $x \in E$ with $0 < d_X(x, x_0) < \delta_1$, while, since $\lim_{x \to x_0} f(x) = L$, there exists $\delta_2 > 0$ with the property that

$$d_Y(f(x), L) < \varepsilon$$

for all $\boldsymbol{x} \in E$ with $0 < d_X(x, x_0) < \delta_2$.

Take $\delta = \min\{\delta_1, \delta_2\} > 0$ and take $x \in E$ with $0 < d_X(x, x_0) < \delta$. Note that such x exists because x_0 is an accumulation point of E. Then by the properties of the distance,

$$\begin{aligned} d_Y(\ell, L) &\le d_Y(f(x), \ell) + d_Y(f(x), L) \\ &< \varepsilon + \varepsilon = d_Y(\ell, L), \end{aligned}$$

which implies that $d_Y(\ell, L) < d_Y(\ell, L)$. This contradiction proves the theorem. ∎

Remark 111 *For topological spaces in general the limit is not unique. Given* (X, τ_X) *and* (Y, τ_Y) *are two topological spaces,* $E \subseteq X$*,* $x_0 \in X$ *is an accumulation point of* E *and* $f : E \to Y$*, it can be shown that the limit is unique if the space* Y *is Hausdorff. A topological space* Y *is a* Hausdorff space*, if for every* x *and* $y \in Y$*, with* $x \neq y$*, there exist disjoint neighborhoods of* x *and* y*.*

Monday, February 14, 2022

Definition 112 *If* (X, d_X) *and* (Y, d_Y) *are two metric spaces,* $E \subseteq X$*, and* $f : E \to Y$*, given a subset* $F \subseteq E$ *we denote by* $f|_F$ *the* restriction *of the function* f *to the set* F*, that is, the function* $f : F \to Y$*.*

Remark 113 *Let* (X, d_X) *and* (Y, d_Y) *be two metric spaces, let* $E \subseteq X$*, let* $x_0 \in X$ *be an accumulation point of* E *and* $f : E \to Y$*. Assume that there exists*

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

Then for every $\varepsilon > 0$ *there exists a real number* $\delta = \delta(\varepsilon, x_0) > 0$ *with the property that*

$$d_Y(f(x), \ell) < \varepsilon \tag{10}$$

for all $x \in E$ *with* $0 < d_X(x, x_0) < \delta$*. if* $F \subseteq E$ *is a subset such that* x_0 *is an accumulation point of* F*, then by restricting (10) we have that*

$$d_Y(f(x), \ell) < \varepsilon$$

for all $x \in F$ *with* $0 < d_X(x, x_0) < \delta$*. Hence, there exists*

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

It follows that if we can find two sets $F \subseteq E$ *and* $G \subseteq E$ *such that* $x_0 \in \operatorname{acc} F$ *and* $x_0 \in \operatorname{acc} G$

$$\lim_{x \to x_0} f|_F(x) = \ell_1 \neq \ell_2 = \lim_{x \to x_0} f|_G(x),$$

then by the uniqueness of the limit (which we will prove later), it follows that the limit over E *cannot exist.*

Example 114 *Let's study the limit*

$$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2},$$

where $m \in \mathbb{N}$. In this case $f(x,y) = \frac{xy}{x^2+y^2}$ and the domain is $E = \mathbb{R}^2 \setminus \{(0,0)\}$. Note that $(0,0)$ is an accumulation point of E.

Taking $F = \{(x,x) : x \in \mathbb{R} \setminus \{0\}\}$, we have that $(0,0)$ is an accumulation point of F. For $(x,x) \in F$ we have

$$f(x,x) = \frac{x^2}{x^2+x^2} = \frac{1}{2} \to \frac{1}{2}$$

as $x \to 0$, while taking $G = \{(x,0) : x \in \mathbb{R} \setminus \{0\}\}$, we have that $(0,0)$ is an accumulation point of G. For $(x,0) \in F$ we have

$$f(x,0) = \frac{0}{x^2+0} = 0 \to 0,$$

and so the limit does not exist.

Remark 115 *Note that the degree of the numerator is 2 and the degree of the numerator is 2, so that in this particular example the limit does not exist if the degree of the numerator is the same as the degree of the numerator.*

Example 116 *Let's study the limit*

$$\lim_{(x,y)\to(0,0)} \frac{x^2y}{x^2+y^2}.$$

In this case $f(x,y) = \frac{x^2y}{x^2+y^2}$ and the domain is $\mathbb{R}^2 \setminus \{(0,0)\}$. To try to guess what the limit should be, let's consider the restriction $x = 0$. For $y \neq 0$, we have

$$f(0,y) = \frac{0y}{0+y^2} = \frac{0}{y^2} = 0 \to 0$$

as $y \to 0$. This says that if the limit exists, then it must be zero. To prove that the limit exist, we use the fact that $x^2 \leq x^2+y^2$ and that $|y| = \sqrt{y^2} \leq \sqrt{x^2+y^2}$ to estimate

$$\left|\frac{x^2y}{x^2+y^2} - 0\right| = \frac{x^2\,|y|}{x^2+y^2} \leq \frac{(x^2+y^2)\sqrt{x^2+y^2}}{x^2+y^2} = \sqrt{x^2+y^2} < \varepsilon$$

for all $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$ with $0 < \sqrt{x^2+y^2} < \delta$, provided we take $\delta = \varepsilon$.

Remark 117 *Note that the degree of the numerator is 3 and the degree of the numerator is 2, so that in this particular example the limit exists if the degree of the numerator is strictly bigger that the degree of the numerator.*

Example 118 *Let's study the limit*

$$\lim_{(x,y)\to(0,0)} \frac{x^{100}y}{x-y}.$$

In this case $f(x,y) = \frac{x^{100}y}{x-y}$ *and the domain is* $\mathbb{R}^2 \setminus \{(x,x) : x \in \mathbb{R}\}$. *Taking* $y = 0$*, we have that*

$$f(x,0) = \frac{0}{x-0} = 0 \to 0.$$

Let us take $y = x + x^a$*, where* $a > 1$ *has to be chosen. Then*

$$f(x, x + x^a) = \frac{x^{100}(x + x^a)}{x - (x + x^a)} = \frac{x^{101} + x^{100+a}}{-x^a}.$$

Take $a = 101$. *Then*

$$\begin{aligned} f\left(x, x + x^{101}\right) &= -\frac{x^{101} + x^{201}}{x^{101}} = -\frac{x^{101}\left(1 + x^{100}\right)}{x^{101}} \\ &= -\frac{1 + x^{100}}{1} \nrightarrow 0. \end{aligned}$$

Hence the limit does not exist.

Remark 119 *Note that the degree of the numerator is* 101 *and the degree of the numerator is* 2*, but in this case the limit never exists no matter how high is the degree of the numerator. The problem is that the domain is* $\mathbb{R}^2$ *minus a curve passing through the origin.*

Exercise 120 *Study the limit*

$$\lim_{(x,y)\to(0,0)} \frac{x^4 y}{y - \sin x}.$$

Hint: Try $y = x^m + \sin x$*, where* m *has to be chosen.*

We list some important limits.

$$\lim_{x\to 0} \frac{\sin x}{x} = 1, \qquad \lim_{x\to 0} \frac{1-\cos x}{x^2} = \frac{1}{2}, \qquad \lim_{x\to 0} \frac{\log(1+x)}{x} = 1,$$
$$\lim_{x\to 0} \frac{(1+x)^a - 1}{x} = a \quad \text{for } a \in \mathbb{R}, \qquad \lim_{x\to 0} \frac{e^x - 1}{x} = 1.$$

Example 121 *Let's study the limit*

$$\lim_{(x,y)\to(0,0)} \frac{\sin(x^m y)}{x^2 + y^2},$$

where $m \in \mathbb{N}$. *In this case* $f(x,y) = \frac{\sin(x^m y)}{x^2+y^2}$ *and the domain is* $\mathbb{R}^2 \setminus \{(0,0)\}$. *We want to use the limit*

$$\lim_{t\to 0} \frac{\sin t}{t} = 1.$$

Consider $f(x,0) = \frac{0}{x^2+0} = 0 \to 0$ *as* $x \to 0$ *and* $f(0,y) = \frac{0}{0+y^2} = 0 \to 0$ *as* $y \to 0$. *If* $xy \neq 0$, *then we can divide by* $x^m y$ *to see that*

$$\frac{\sin(x^m y)}{x^2+y^2} = \frac{\sin(x^m y)}{x^m y}\frac{x^m y}{x^2+y^2}.$$

Then

$$\lim_{(x,y)\to(0,0)} \frac{\sin(x^m y)}{x^m y} = 1$$

and so (using the theorem on product of limits which we will prove later) it remains to study

$$\lim_{(x,y)\to(0,0)} \frac{x^m y}{x^2+y^2}.$$

If $m = 1$ *we have seen in the previous example that the limit does not exist.*

For $m \geq 2$, *we have that the limit is* 0. *Indeed, using the facts that* $|x| = \sqrt{x^2} \leq \sqrt{x^2+y^2}$ *and* $|y| = \sqrt{y^2} \leq \sqrt{x^2+y^2}$ *we have*

$$\left|\frac{x^m y}{x^2+y^2} - 0\right| = \frac{|x|^m |y|}{x^2+y^2} \leq \frac{\left(x^2+y^2\right)^{m/2}\left(x^2+y^2\right)^{1/2}}{x^2+y^2} = \left(x^2+y^2\right)^{(m-1)/2} \to 0$$

as $(x,y) \to (0,0)$.

Remark 122 *Note that the degree of the numerator is* $m+1$ *and the degree of the denominator is* 2, *so that in this particular example the limit exists if the degree of the numerator is higher than the degree of the denominator, that is, if* $m+1 > 2$.

Wednesday, February 16, 2022

The next example shows that checking the limit on every line passing through $\boldsymbol{x}_0$ is not enough to guarantee the existence of the limit.

Example 123 *Let*

$$f(x,y) := \begin{cases} 1 & \text{if } y = x^2,\ x \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Given the line $y = mx$, *the line intersects the parabola* $y = x^2$ *only in* $\mathbf{0}$ *and in at most one point. Hence, if* x *is very small,*

$$f(x, mx) = 0 \to 0$$

as $x \to 0$. *However, since* $f\left(x, x^2\right) = 1 \to 1$ *as* $x \to 0$, *the limit does not exists.*

Exercise 124 *Study the limit*

$$\lim_{(x,y)\to(0,0)} \frac{x^2 y}{x^4+y^2}.$$

Exercise 125 (Important) *Let (X, d_X) and (Y, d_Y) be two metric spaces, let $E \subseteq X$, let $x_0 \in X$ be an accumulation point of E and let $f : E \to Y$. Assume that there exists $\ell \in Y$ and a function $g : [0, \infty) \to (0, \infty)$ with*

$$\lim_{s \to 0^+} g(s) = 0,$$

such that for every $\varepsilon > 0$ there exists a real number $\delta = \delta(\varepsilon, x_0) > 0$ with the property that

$$d_Y(f(x), \ell) < g(\varepsilon)$$

for all $x \in E$ with $0 < d_X(x, x_0) < \delta$. Prove that there exists

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

Remark 126 (Important) *The previous exercise says that we do not have to be very precise when applying the definition of limit, in the sense that, if we can prove that for every $\varepsilon > 0$ there exists a real number $\delta = \delta(\varepsilon, x_0) > 0$ with the property that*

$$d_Y(f(x), \ell) < 4\varepsilon^{1/2}$$

for all $x \in E$ with $0 < d_X(x, x_0) < \delta$ or

$$d_Y(f(x), \ell) < 4\varepsilon^3 + 16\varepsilon$$

for all $x \in E$ with $0 < d_X(x, x_0) < \delta$ or anything like that, then we know that we can conclude that there exists

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

This is very useful when proving theorems about limits.

Example 127 *Let's compute the limit*

$$\lim_{(x,y) \to (x_0, y_0)} \frac{x}{y},$$

where $x_0 \in \mathbb{R}$ and $y_0 \in \mathbb{R}$, $y_0 \neq 0$. The limit is $\frac{x_0}{y_0}$. To prove it, let's write

$$\begin{aligned} \left| \frac{x}{y} - \frac{x_0}{y_0} \right| &= \left| \frac{xy_0 - yx_0}{yy_0} \right| = \left| \frac{xy_0 - x_0y_0 + x_0y_0 - yx_0}{yy_0} \right| \\ &= \left| \frac{(x - x_0)y_0 + x_0(y_0 - y)}{yy_0} \right| \\ &\leq \frac{|x - x_0||y_0| + |x_0||y_0 - y|}{|y||y_0|} \\ &\leq \frac{|x - x_0|}{|y|} + \frac{|x_0||y_0 - y|}{|y||y_0|}. \end{aligned}$$

The problem is when y gets too closed to zero. The idea is that since y is close to y_0 and $y_0 \neq 0$, we can make sure that y stays away from zero. Assume that $|y - y_0| < \frac{1}{2}|y_0|$. Then

$$|y| \geq |y_0| - |y - y_0| \geq |y_0| - \frac{1}{2}|y_0| = \frac{|y_0|}{2}$$

and so,

$$\frac{1}{|y|} \leq \frac{2}{|y_0|}.$$

Therefore, given $\varepsilon > 0$ for $\sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta$ we have

$$\begin{aligned}\left|\frac{x}{y} - \frac{x_0}{y_0}\right| &\leq \frac{|x - x_0|}{|y|} + \frac{|x_0||y_0 - y|}{|y||y_0|} \\ &\leq \frac{2}{|y_0|}|x - x_0| + \frac{2|x_0|}{y_0^2}|y_0 - y| \\ &\leq \frac{2}{|y_0|}\varepsilon + \frac{2|x_0|}{y_0^2}\varepsilon\end{aligned}$$

provided we take $\delta = \min\{\varepsilon, \frac{1}{2}|y_0|\}$.

Exercise 128 *Prove that the limit*

$$\lim_{(x,y)\to(x_0,y_0)} x + y = x_0 + y_0,$$

where $x_0 \in \mathbb{R}$ and $y_0 \in \mathbb{R}$.

Example 129 *Prove that the limit*

$$\lim_{(x,y)\to(x_0,y_0)} xy = x_0 y_0,$$

where $x_0 \in \mathbb{R}$ and $y_0 \in \mathbb{R}$.

Friday, February 18, 2022

Theorem 130 *Let (X, d_X), (Y, d_Y), (Z, d_Z) be three metric spaces, let $E \subseteq X$, let $x_0 \in E$ be an accumulation point of E, and let $F \subseteq Y$. Given two functions $f : E \to F$ and $g : F \to Z$ assume that there exist*

$$\lim_{x\to x_0} f(x) = \ell \in Y,$$

that ℓ is an accumulation point of F and that there exists

$$\lim_{y\to \ell} g(y) = L \in Z.$$

Suppose also that either $f(x) \neq \ell$ for all $x \in E \setminus \{x_0\}$, or that $\ell \in F$ and $L = g(\ell)$. Then there exists $\lim_{x\to x_0} g(f(x)) = L$.

Proof. For every $\varepsilon > 0$ there exists $\eta = \eta(\varepsilon, \ell) > 0$ such that

$$d_Z(g(y), L) < \varepsilon \tag{11}$$

for all $y \in F$ with $0 < d_Y(y, \ell) < \eta$.

Since $\lim_{x \to x_0} f(x) = \ell$, there exists $\delta = \delta(x_0, \eta) > 0$ such that

$$d_Y(f(x), \ell) < \eta$$

for all $x \in E$ with $0 < d_X(x, x_0) < \delta$.

Case 1: Assume that $f(x) \neq \ell$ for all $x \in E \setminus \{x_0\}$. Then for all $x \in E$ with $0 < d_X(x, x_0) < \delta$, we have that $d_Y(f(x), \ell) < \eta$, and so we can take $y = f(x)$ in (11) to get that

$$d_Z(g(f(x)), L) < \varepsilon,$$

which implies that there exists $\lim\limits_{x \to x_0} g(f(x)) = L$.

Case 2: Assume that $\ell \in F$ and $L = g(\ell)$. For every $\varepsilon > 0$ there exists $\eta = \eta(\varepsilon, \ell) > 0$ such that

$$d_Z(g(y), g(\ell)) < \varepsilon$$

for all $y \in F$ with $0 < d_Y(y, \ell) < \eta$. Note that, if we take $y = \ell$, we have that

$$d_Z(g(\ell), g(\ell)) = 0 < \varepsilon.$$

Thus,

$$d_Z(g(y), g(\ell)) < \varepsilon \tag{12}$$

for all $y \in F$ with $d_Y(y, \ell) < \eta$ (so we can take $y = \ell$). Then for all $x \in E$ with $0 < d_X(x, x_0) < \delta$, we have that $d_Y(f(x), \ell) < \eta$, and so we can take $y = f(x)$ in (12) to get that

$$d_Z(g(f(x)), L) < \varepsilon,$$

which implies that there exists $\lim\limits_{x \to x_0} g(f(x)) = L$. ■

Example 131 *Let's prove that the previous theorem fails without the hypotheses that either $\boldsymbol{f}(\boldsymbol{x}) \neq \boldsymbol{\ell}$ for all $\boldsymbol{x} \in E$ near $\boldsymbol{x}_0$. Consider the function*

$$g(y) := \begin{cases} 1 & \text{if } y \neq 0, \\ 2 & \text{if } y = 0. \end{cases}$$

Then there exists

$$\lim_{y \to 0} g(y) = 1.$$

So $L = 1$. Consider the function $f(x) := 0$ for all $x \in \mathbb{R}$. Then for every $x_0 \in \mathbb{R}$, we have that

$$\lim_{x \to x_0} f(x) = 0.$$

So $\ell = 0$. However, $g(f(x)) = g(0) = 2$ for all $x \in \mathbb{R}$. Hence,

$$\lim_{x \to x_0} g(f(x)) = \lim_{x \to x_0} 2 = 2 \neq 1,$$

which shows that the conclusion of the theorem is violated..

Theorem 132 *Let (X,d) be a metric space, let $E \subseteq X$, let $x_0 \in X$ be an accumulation point of E. Given two functions $f,g : E \to \mathbb{R}$, assume that there exist*

$$\lim_{x\to x_0} f(x) = \ell_1 \in \mathbb{R}, \qquad \in \lim_{x\to x_0} g(x) = \ell_2\mathbb{R}.$$

Then

(i) there exists $\lim_{x\to x_0} (f+g)(x) = \ell_1 + \ell_2$,

(ii) there exists $\lim_{x\to x_0} (f\cdot g)(x) = \ell_1 \cdot \ell_2$,

(iii) if $\ell_2 \neq 0$, then $g(x) \neq 0$ for all x close to x_0 and there exists $\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{\ell_1}{\ell_2}$.

Proof. One can use Theorem 130 to prove Theorem 132. Indeed, $f+g$ is the composition of the function $h_1 : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ given by

$$h_1(s,t) := s+t$$

with the function $\boldsymbol{P} : E \to \mathbb{R}^2$ given by $\boldsymbol{P}(x) = (f(x), f(x))$, while $f\cdot g$ is the composition of the function $h_2 : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ given by

$$h_2(s,t) := s\cdot t$$

with the function $\boldsymbol{P}$, while $\frac{f(x)}{g(x)}$ is the composition of the function $h_3 : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ given by

$$h_3(s,t) := \frac{s}{t}$$

with the function $\boldsymbol{P}$. By Exercises 128 and 129 and Example 127, the functions h_1, h_2, and h_3 are continuous.

In item (iii), to prove that if $\ell_2 \neq 0$, then $g(x) \neq 0$ for all x close to x_0, take $\varepsilon = \frac{|\ell_2|}{2} > 0$. Since $\lim_{x\to x_0} g(x) = \ell_2$, we can find $\varepsilon > 0$ such that

$$|g(x) - \ell_2| < \frac{|\ell_2|}{2}$$

for all $x \in E$ with $0 < d(x,x_0) < \delta$. Hence,

$$|g(x)| = |\ell_2 + g(x) - \ell_2| \geq |\ell_2| - |g(x) - \ell_2| \geq |\ell_2| - \frac{|\ell_2|}{2} = \frac{|\ell_2|}{2} > 0$$

for all $x \in E$ with $0 < d(x,x_0) < \delta$, which implies that $g(x) \neq 0$ for all x close to x_0. ■

Remark 133 *The previous theorem continues to hold if $\ell_1, \ell_2 \in [-\infty,\infty]$, provided we avoid the cases $\infty - \infty$, 0∞, $\frac{0}{0}$, $\frac{\infty}{\infty}$.*

Monday, February 21, 2022

Theorem 134 (Squeeze Theorem) *Let (X, d) be a metric space, let $E \subseteq X$, let $x_0 \in X$ be an accumulation point of E. Given three functions $f, g, h : E \to \mathbb{R}$, assume that there exist*

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

and that $f(x) \leq h(x) \leq g(x)$ for every $x \in E$. Then there exists $\lim_{x \to x_0} h(x) = \ell$.

Proof. Given $\varepsilon > 0$ there exist $\delta_1 > 0$ such that

$$|f(x) - \ell| \leq \varepsilon$$

for all $x \in E$ with $0 < d_X(x, x_0) < \delta_1$ and $\delta_2 > 0$ such that

$$|g(x) - \ell| \leq \varepsilon$$

for all $x \in E$ with $0 < d_X(x, x_0) < \delta_2$. Then for all $x \in E$ with $0 < d_X(x, x_0) < \delta = \min\{\delta_1, \delta_2\}$, we have that

$$\ell - \varepsilon \leq f(x) \leq h(x) \leq g(x) \leq \ell + \varepsilon.$$

Hence,

$$|h(x) - \ell| \leq \varepsilon$$

for all $x \in E$ with $0 < d_X(x, x_0) < \delta$, which shows that $\lim_{x \to x_0} h(x) = \ell$. ■

Example 135 *The previous theorem can be used for example to show that for $a > 0$*

$$\lim_{x \to 0} |x|^a \sin\frac{1}{x} = 0.$$

Indeed,

$$0 \leq \left| |x|^a \sin\frac{1}{x} \right| = |x|^a \left| \sin\frac{1}{x} \right| \leq |x|^a$$

and since $|x|^a \to 0$ as $x \to 0$ we can apply the squeeze theorem. We could also used the following Exercise.

Exercise 136 *Let (X, d) be a metric space, let $E \subseteq X$, let $x_0 \in X$ be an accumulation point of E. Given two functions $f, g : E \to \mathbb{R}$, assume that there exists*

$$\lim_{x \to x_0} f(x) = 0,$$

and that g is bounded, that is, $|g(x)| \leq L$ for all $x \in E$ and for some $L > 0$. Prove that there exists $\lim_{x \to x_0} (fg)(x) = 0$.

9 Limits of Monotone Functions

Let $E \subseteq \mathbb{R}$ and let $f : E \to \mathbb{R}$. Then f is said to be

- *increasing* if $f(x) \leq f(y)$ for all $x, y \in E$ with $x < y$,
- *strictly increasing* if $f(x) < f(y)$ for all $x, y \in E$ with $x < y$,
- *decreasing* if $f(x) \geq f(y)$ for all $x, y \in E$ with $x < y$,
- *strictly decreasing* if $f(x) > f(y)$ for all $x, y \in E$ with $x < y$,
- *monotone* if one of the four property above holds.

Given $E \subseteq \mathbb{R}$ and $f : E \to \mathbb{R}$, if $x_0 \in \mathbb{R}$ is an accumulation point of $E \cap (-\infty, x_0)$, we define the left limit of f as x approaches x_0 as

$$\lim_{x \to x_0^-} f(x) := \lim_{x \to x_0} f|_{E \cap (-\infty, x_0)}(x),$$

provided the limit $\lim_{x \to x_0} f|_{E \cap (-\infty, x_0)}(x)$ exists. Similarly, if $x_0 \in \mathbb{R}$ is an accumulation point of $E \cap (x_0, \infty)$, we define the right limit of f as x approaches x_0 as

$$\lim_{x \to x_0^+} f(x) := \lim_{x \to x_0} f|_{E \cap (x_0, \infty)}(x),$$

provided the limit $\lim_{x \to x_0} f|_{E \cap (x_0, \infty)}(x)$ exists.

In what follows if a nonempty set $F \subseteq \mathbb{R}$ is not bounded from above, we set $\sup F := \infty$. Similarly, if a nonempty set $F \subseteq \mathbb{R}$ is not bounded from below, we set $\inf F := -\infty$.

Theorem 137 *Let $E \subseteq \mathbb{R}$ and let $f : E \to \mathbb{R}$ be a monotone function. If $x_0 \in \mathbb{R}$ is an accumulation point of $E \cap (-\infty, x_0)$, then there exists*

$$\lim_{x \to x_0^-} f(x) = \begin{cases} \sup_{E \cap (-\infty, x_0)} f & \text{if } f \text{ is increasing,} \\ \inf_{E \cap (-\infty, x_0)} f & \text{if } f \text{ is decreasing,} \end{cases}$$

while if $x_0 \in \mathbb{R}$ is an accumulation point of $E \cap (x_0, \infty)$ then there exists

$$\lim_{x \to x_0^+} f(x) = \begin{cases} \sup_{E \cap (x_0, \infty)} f & \text{if } f \text{ is decreasing,} \\ \inf_{E \cap (x_0, \infty)} f & \text{if } f \text{ is increasing.} \end{cases}$$

Proof. Assume that $x_0 \in \mathbb{R}$ is an accumulation point of $E \cap (-\infty, x_0)$ and that f is increasing (the other cases are similar). There are two cases.

Case 1: The function f is bounded from above in $E \cap (-\infty, x_0)$. Hence, there exists

$$\sup_{E \cap (-\infty, x_0)} f = \ell \in \mathbb{R}.$$

We need to prove that there exists

$$\lim_{x \to x_0^-} f(x) = \ell. \tag{13}$$

Let $\varepsilon > 0$. Since ℓ is the supremum of $f(E \cap (-\infty, x_0))$, we have that $f(x) \leq \ell$ for all $x \in E$ with $x < x_0$. On the other hand, since $\ell - \varepsilon$ is not an upper bound for the set $f(E \cap (-\infty, x_0))$ there exists $x_1 \in E \cap (-\infty, x_0)$ such that $\ell - \varepsilon < f(x_1)$. But since f is increasing, for all $x \in E$ with $x_1 < x < x_0$ we have that $\ell - \varepsilon < f(x_1) \leq f(x)$. Thus,

$$\ell - \varepsilon < f(x) \leq \ell < \ell + \varepsilon$$

for all $x \in E \cap (-\infty, x_0)$ with $x_1 < x < x_0$. Take $\delta := x_0 - x_1 > 0$. Then $|f(x) - \ell| < \varepsilon$ for all $x \in E \cap (-\infty, x_0)$ with $0 < |x - x_0| < \delta$. This proves (13).

Case 2: The function f is not bounded from above in $E \cap (-\infty, x_0)$. We need to prove that there exists

$$\lim_{x \to x_0^-} f(x) = \infty. \tag{14}$$

Let $M > 0$. Since the set $f(E \cap (-\infty, x_0))$ is not bounded from above there exists $x_1 \in E \cap (-\infty, x_0)$ such that $f(x_1) > M$. But since f is increasing, for all $x \in E$ with $x_1 < x < x_0$ we have that $M < f(x_1) \leq f(x)$. Thus,

$$M < f(x)$$

for all $x \in E \cap (-\infty, x_0)$ with $x_1 < x < x_0$. Take $\delta := x_0 - x_1 > 0$. Then $f(x) > M$ for all $x \in E \cap (-\infty, x_0)$ with $0 < |x - x_0| < \delta$. This proves (14). ■

Remark 138 *A similar result holds if $E \subseteq \overline{\mathbb{R}}$ and if $f : E \to \overline{\mathbb{R}}$, where we recall that $\overline{\mathbb{R}} = [-\infty, \infty]$ is the extended real line.*

Definition 139 *A set $E \subseteq \mathbb{R}^N$ is* countable *if there exists a one-to-one function $f : E \to \mathbb{N}$.*

Remark 140 *It can be shown that $\mathbb{Q}$ is countable and that if $E_n \subseteq \mathbb{R}$, $n \in \mathbb{N}$, is countable, then*

$$E = \bigcup_{n=1}^{\infty} E_n$$

is countable. Using Cantor's diagonal argument one can show that $\mathbb{R}$ and the irrationals are NOT countable.

Definition 141 *A set $I \subseteq \mathbb{R}$ is an* interval *if for every $x, y \in I$, with $x < y$, we have that the interval $[x, y]$ is contained in I.*

Definition 142 *Given a set X and a function $f : X \to [0, \infty]$ the* infinite sum

$$\sum_{x \in X} f(x)$$

is defined as

$$\sum_{x \in X} f(x) := \sup \left\{ \sum_{x \in Y} f(x) : Y \subset X,\ Y \text{ finite} \right\}.$$

Proposition 143 *Given a set X and a function $f : X \to [0, \infty]$, if*

$$\sum_{x \in X} f(x) < \infty,$$

then the set $\{x \in X : f(x) > 0\}$ is countable. Moreover, f does not take the value ∞.

Proof. Define

$$M := \sum_{x \in X} f(x) < \infty.$$

For $k \in \mathbb{N}$ set $X_k := \left\{x \in X : f(x) > \frac{1}{k}\right\}$ and let Y be a finite subset of X_k. Then

$$\frac{1}{k}\text{number of elements of } Y \le \sum_{x \in Y} f(x) \le M,$$

which shows that Y cannot have more than $\lfloor kM \rfloor$ elements, where $\lfloor \cdot \rfloor$ is the integer part. In turn, X_k has a finite number of elements, and so

$$\{x \in X : f(x) > 0\} = \bigcup_{k=1}^{\infty} X_k$$

is countable. ■

Exercise 144 *Given a nonempty set X and two functions $f, g : X \to [0, \infty]$.*

(i) Prove that

$$\sum_{x \in X} (f(x) + g(x)) \le \sum_{x \in X} f(x) + \sum_{x \in X} g(x).$$

(ii) If $f \le g$, then

$$\sum_{x \in X} f(x) \le \sum_{x \in X} g(x).$$

Wednesday, February 23, 2022

Theorem 145 *Let $I \subseteq \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$ be a monotone function. Then there exists*

$$\lim_{y \to x} f(y) = f(x)$$

for all $x \in I$ except at most for countably many.

Proof. Step 1: Assume that $I = [a, b]$ and, without loss of generality, that f is increasing. For every $x \in (a, b)$ there exist

$$\lim_{y \to x^+} f(y) =: f_+(x), \quad \lim_{y \to x^-} f(y) =: f_-(x).$$

Let $S(x) := f_+(x) - f_-(x) \geq 0$ be the jump of f at x. Then

$$\lim_{y \to x} f(y) = f(x)$$

if and only if $S(x) = 0$. Let $J \subseteq [a, b]$ be any finite subset, and write

$$J = \{x_1, \dots, x_k\}, \quad \text{where } x_1 < \cdots < x_k.$$

Since f is increasing, we have that

$$\begin{aligned} f(a) &\leq f_-(x_1) \leq f_+(x_1) \leq f_-(x_2) \leq f_+(x_2) \\ &\leq \cdots \leq f_-(x_k) \leq f_+(x_k) \leq f(b), \end{aligned}$$

and so,

$$\sum_{x \in J} S(x) = \sum_{i=1}^{k} (f_+(x_i) - f_-(x_i)) \leq f(b) - f(a),$$

which implies that

$$\sum_{x \in (a,b)} S(x) \leq f(b) - f(a).$$

By the previous proposition, it follows that the set of discontinuity points of f is at most countable.

Step 2: If I is an arbitrary interval, construct an increasing sequence of intervals $[a_n, b_n]$ such that

$$a_n \searrow \inf I, \quad b_n \nearrow \sup I.$$

Since the union of countable sets is countable and on each interval $[a_n, b_n]$ the set of discontinuity points of f is at most countable, by the previous step it follows that the set of discontinuity points of f in I is at most countable. ■

Conversely, given a countable set $E = \{r_n : n \in \mathbb{N}\}$, we can construct an increasing function $f : \mathbb{R} \to \mathbb{R}$ such that

$$\lim_{y \to x} f(y) = f(x)$$

for all $x \in \mathbb{R} \setminus E$ and f jumps at every point of E. Consider

$$f_n(x) = \begin{cases} \frac{1}{2^n} & \text{if } x \geq r_n, \\ 0 & \text{if } x < r_n. \end{cases}$$

Then the function $f : \mathbb{R} \to \mathbb{R}$, defined by

$$f(x) := \sum_{n=1}^{\infty} f_n(x), \quad x \in \mathbb{R},$$

is increasing and has the desired properties. We will see this later.

Exercise 146 *Let $E \subseteq \mathbb{R}$ be not bounded from above and let $f : E \to \mathbb{R}$ be an increasing function. Prove that there exists*

$$\lim_{x \to \infty} f(x) = \sup_{x \in E} f(x).$$

10 Series

Definition 147 *Given a normed space X and a sequence $\{x_n\}_n$ of vectors in X, we call the n-th* partial sum *the vector*

$$s_n = x_1 + \cdots + x_n.$$

The sequence $\{s_n\}_n$ of partial sums is called infinite series *or* series *and is denoted*

$$\sum_{n=1}^{\infty} x_n.$$

If there exists $\lim_{n\to\infty} s_n = S \in X$, we say that the series $\sum_{n=1}^{\infty} x_n$ is convergent*. The number S is called* sum *of the series. while if the $\lim_{n\to\infty} s_n$ does not exist, we say that the series $\sum_{n=1}^{\infty} x_n$* oscillates.

If $X = \mathbb{R}$ and $\lim_{n\to\infty} s_n = \infty$ or $\lim_{n\to\infty} s_n = -\infty$, we say that series $\sum_{n=1}^{\infty} x_n$ is *divergent.*

Friday, February 25, 2022

First interim exam.

Monday, March 01, 2022

Solutions, first interim exam.

Wednesday, March 03, 2022

Remark 148 *There is nothing special about 1, we will also consider series of the type $\sum_{n=0}^{\infty} x_n$ or $\sum_{n=n_0}^{\infty} x_n$, where $n_0 \in \mathbb{N}$. The only change is that in the partial sums, one should consider $s_n = x_0 + \cdots + x_n$ and $s_n = x_{n_0} + \cdots + x_n$, respectively.*

Theorem 149 *Given a normed space X and a sequence $\{x_n\}_n$ of vectors in X, if the series $\sum_{n=1}^{\infty} x_n$ converges, then there exists*

$$\lim_{n\to\infty} x_n = 0.$$

Proof. Since the series $\sum_{n=1}^{\infty} x_n$ converges, there exists $\lim_{n\to\infty} s_n = S \in X$. Hence,

$$a_n = s_{n+1} - s_n \to S - S = 0$$

as $n \to \infty$. Note that here it is important that $S \in X$. ■

Corollary 150 *Given a series $\sum_{n=1}^{\infty} x_n$, if either the limit $\lim_{n\to\infty} x_n$ does not exist or it exists but is different from zero, then the series $\sum_{n=1}^{\infty} x_n$ cannot converge.*

Example 151 (Geometric series) *The series*

$$\sum_{n=0}^{\infty} x^n,$$

where $x \in \mathbb{R}$, is called a geometric series *with* ratio x. *Since*

$$\lim_{n\to\infty} x^n = 0 \quad \textit{if and only if } -1 < x < 1,$$

by the previous theorem for $|x| \geq 1$, the series cannot converge. It remains to study what happens when $-1 < x < 1$. By Exercise 15,

$$s_n = 1 + x \cdots + x^n = \frac{1 - x^{n+1}}{1 - x},$$

and so

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} \frac{1 - x^{n+1}}{1 - x} = \frac{1}{1 - x}.$$

Thus for $-1 < x < 1$, the series converges and its sum is $\frac{1}{1-x}$.

10.1 Series of Nonnegative Terms

A series $\sum_{n=1}^{\infty} x_n$ is called a series of nonnegative terms if $x_n \geq 0$ for all $n \in \mathbb{N}$. These series have the important property that they cannot oscillate.

Theorem 152 *Given a series $\sum_{n=1}^{\infty} x_n$ with $x_n \geq 0$ for all $n \in \mathbb{N}$, then the series either converges or it diverges to ∞.*

Proof. For all $n \in \mathbb{N}$, we have that $s_{n+1} = s_n + a_n \geq s_n$, and so the sequence $\{s_n\}$ is increasing. Thus, by Exercise 146, there exists $\lim_{n\to\infty} s_n = S \in [0, \infty]$. Hence, the series either converges or it diverges to ∞. ■

Remark 153 *The same proof continues to work if we only assume that there exists $n_0 \in \mathbb{N}$ such that $x_n \geq 0$ for all $n \geq n_0$. In this case, we have that $s_{n+1} \geq s_n$ for all $n \geq n_0$, which still implies that the limit* $\lim_{n\to\infty} s_n$ *exists, although this time it can also be negative.*

Next we study some texts for convergence of series of nonnegative terms.

Theorem 154 (Comparison Test) *Given two series $\sum_{n=1}^{\infty} x_n$ and $\sum_{n=1}^{\infty} y_n$ with $x_n \geq 0$ and $y_n \geq 0$ for all $n \in \mathbb{N}$. Assume that there exists $N \in \mathbb{N}$ such that*

$$x_n \leq y_n$$

for all $n \geq N$.

(i) If the series $\sum_{n=1}^{\infty} y_n$ converges, then so does the series $\sum_{n=1}^{\infty} x_n$.

(ii) If the series $\sum_{n=1}^{\infty} x_n$ diverges to ∞, then so does the series $\sum_{n=1}^{\infty} y_n$.

Proof. (i) Let $t_n = y_1 + \cdots + y_n$. By hypothesis, there exists $\lim_{n\to\infty} t_n = T \in \mathbb{R}$. It follows that $\{t_n\}$ is bounded by T. Hence, $0 \le t_n \le T$ for all $n \in \mathbb{N}$. For $n \ge N$, we have that

$$\begin{aligned} s_n &= x_1 + \cdots + x_{N-1} + x_N + \cdots + y_n \le (x_1 + \cdots + x_{N-1}) + y_N + \cdots + y_n \\ &\le (x_1 + \cdots + x_{N-1}) + y_1 + \cdots + y_n = (x_1 + \cdots + x_{N-1}) + t_n \\ &\le (x_1 + \cdots + x_{N-1}) + T. \end{aligned}$$

Thus, the sequence $\{s_n\}$ is bounded. Since it is increasing, it follows that it converges.

(ii) By hypothesis, there exists $\lim_{n\to\infty} s_n = \infty$. As before, for $n \ge N$, we have that

$$s_n \le (x_1 + \cdots + x_{N-1}) + t_n.$$

Letting $n \to \infty$, we conclude that $\lim_{n\to\infty} t_n = \infty$. ■

Example 155 *Consider the series*

$$\sum_{n=1}^{\infty} \left(\frac{1 + \cos n}{3} \right)^n .$$

Note that $1 + \cos n \ge 0$. *Moreover,*

$$0 \le \left(\frac{1 + \cos n}{3} \right)^n \le \left(\frac{2}{3} \right)^n .$$

Hence, by the comparison test, the series $\sum_{n=1}^{\infty} \left(\frac{1+\cos n}{3} \right)^n$ *converges.*

Given a real number $t \in \mathbb{R}$, the *positive part of* t is defined as $t_+ := \max\{t, 0\}$, while the *negrative part of* t is $t_- := \max\{-t, 0\}$. Observe that

$$t = t_+ - t_-, \quad |t| = t_+ + t_-.$$

Corollary 156 *Let* $\{\boldsymbol{x}_n\}_n$ *be a sequence of vectors in* $\mathbb{R}^N$ *such that* $\sum_{n=1}^{\infty} \|\boldsymbol{x}_n\|$ *converges. Then* $\sum_{n=1}^{\infty} \boldsymbol{x}_n$ *converges.*

Proof. Step 1. Assume first that $N = 1$. Then $0 \le (x_n)_+ \le |x_n|$ and $0 \le (x_n)_- \le |x_n|$. Since $\sum_{n=1}^{\infty} |x_n|$ converges, by the comparison principle, so do $\sum_{n=1}^{\infty} (x_n)_-$ and $\sum_{n=1}^{\infty} (x_n)_+$. In turn, by the theorem on the sum of limits, there exist

$$\lim_{n\to\infty} (x_1 + \cdots + x_n) = \lim_{n\to\infty} ((x_1)_+ + \cdots + (x_n)_+) - \lim_{n\to\infty} (((x_1)_- + \cdots + ((x_n)_-),$$

which implies that the series $\sum_{n=1}^{\infty} x_n$ converges.

Step 2. If $N \ge 2$, write $\boldsymbol{x}_n = (x_n^{(1)}, \ldots, x_n^{(N)})$. Then for each $i = 1, \ldots, N$, $|x_n^{(i)}| \le \sqrt{(x_n^{(1)})^2 + \cdots + (x_n^{(N)})^2} = \|\boldsymbol{x}_n\|$. Hence, by the previous step, the series $\sum_{n=1}^{\infty} x_n^{(i)}$ converges for every $i = 1, \ldots, N$. In turn, $\sum_{n=1}^{\infty} \boldsymbol{x}_n$ converges. ■

Remark 157 *You will see in other courses, that if X is a complete normed space, then the previous corollary continues to hold.*

Remark 158 *In step 2 of the previous proof, we used the fact that if $\boldsymbol{f} : E \to \mathbb{R}^M$, where $E \subseteq X$ and $x_0 \in X$ is an accumulation point of E, then there exists*

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

if and only for each $i = 1, \ldots, M$, there exists $\lim_{x \to x_0} f_i(x) = \ell_i$, where $\boldsymbol{f} = (f_1, \ldots, f_M)$. We leave this fact as an exercise.

Monday, March 14, 2022

Theorem 159 (Root Test) *Given a series $\sum_{n=1}^{\infty} x_n$ with $x_n \geq 0$ for all $n \in \mathbb{N}$, if*

$$\limsup_{n \to \infty} \sqrt[n]{x_n} < 1,$$

then the series converges. If

$$\limsup_{n \to \infty} \sqrt[n]{x_n} > 1,$$

then the series diverges to ∞.

Proof. Let $\ell = \limsup_{n \to \infty} \sqrt[n]{x_n}$. Assume first that $\ell < 1$ and fix $\varepsilon > 0$ so small that $\ell + \varepsilon < 1$. By an exercise in your homework, there exists $N \in \mathbb{N}$ such that

$$\sqrt[n]{x_n} \leq \ell + \varepsilon$$

for all $n \geq N$, and so

$$x_n \leq (\ell + \varepsilon)^n$$

for all $n \geq N$. Since $\ell + \varepsilon < 1$, the geometric series $\sum_{n=1} (\ell + \varepsilon)^n$ converges. Hence, so does $\sum_{n=1}^{\infty} x_n$ by the comparison test.

On the other hand, if $\ell > 1$, fix $\varepsilon > 0$ so small that $\ell - \varepsilon > 1$. Again by your homework,

$$\sqrt[n]{x_n} \geq \ell - \varepsilon$$

for infinitely many n, and so

$$x_n \geq (\ell - \varepsilon)^n$$

for infinitely many n. Thus,

$$\limsup_{n \to \infty} x_n \geq \limsup_{n \to \infty} (\ell - \varepsilon)^n = \infty,$$

since $\ell - \varepsilon > 1$. It follows by Theorem 149, that the series $\sum_{n=1}^{\infty} x_n$ cannot converge. In turn, by Theorem 152, $\sum_{n=1}^{\infty} x_n$ diverges to ∞. ■

Remark 160 *If* $\limsup_{n\to\infty} \sqrt[n]{x_n} = 1$*, then the test is inconclusive and one should try a different test. We will see that the series* $\sum_{n=1}^{\infty} \frac{1}{n}$ *diverges, while the series* $\sum_{n=1}^{\infty} \frac{1}{n^2}$*. However, in both cases,* $\limsup_{n\to\infty} \sqrt[n]{x_n} = 1$*. Indeed,*

$$\sqrt[n]{\frac{1}{n}} = \left(\frac{1}{n}\right)^{\frac{1}{n}} = e^{\log\left(\frac{1}{n}\right)^{\frac{1}{n}}} = e^{\frac{1}{n}\log\left(\frac{1}{n}\right)} = e^{\frac{-\log n}{n}} \to e^0 = 1,$$

$$\sqrt[n]{\frac{1}{n^2}} = \left(\frac{1}{n^2}\right)^{\frac{1}{n}} = \left(\frac{1}{n}\right)^{\frac{2}{n}} = e^{\log\left(\frac{1}{n}\right)^{\frac{2}{n}}} = e^{\frac{2}{n}\log\left(\frac{1}{n}\right)} = e^{\frac{-2\log n}{n}} \to e^0 = 1.$$

Example 161 *Consider the series of functions*

$$\sum_{n=1}^{\infty} \frac{e^{nx}}{n}.$$

Note that $\frac{e^{nx}}{n} \geq 0$*. Then*

$$\lim_{n\to\infty} \frac{e^{nx}}{n} = \begin{cases} \infty & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases}$$

Thus, by Theorems 149 and 152 the series diverges to ∞ *for* $x > 0$*. It remains to study what happens for* $x \leq 0$*. We have that*

$$\sqrt[n]{\frac{e^{nx}}{n}} = e^x \left(\frac{1}{n}\right)^{\frac{1}{n}} = e^x e^{\frac{1}{n}\log\left(\frac{1}{n}\right)} = e^x e^{\frac{-\log n}{n}} \to e^x e^0 < 1$$

for $x < 0$*. Thus, for* $x < 0$*, the series converges by the root test, while for* $x = 0$ *we have* $\sum_{n=1}^{\infty} \frac{1}{n}$*, which diverges, as you will see later.*

To study the uniform convergence of the series, consider $\varepsilon > 0$ *and a set* $E = (-\infty, -\varepsilon]$*. Since* $\frac{d}{dx}(\frac{e^{nx}}{n}) = e^{nx} > 0$*, the function* $f_n(x) = \frac{e^{nx}}{n}$ *is increasing. Therefore,*

$$\sup_{(-\infty,\varepsilon]} f_n(x) = f_n(-\varepsilon) = \frac{e^{-n\varepsilon}}{n}.$$

Hence,

$$\sum_{n=1}^{\infty} \sup_{(-\infty,\varepsilon]} f_n(x) = \sum_{n=1}^{\infty} \frac{e^{-n\varepsilon}}{n}.$$

We have already seen that $\sum_{n=1}^{\infty} \frac{e^{-n\varepsilon}}{n}$ *converges. Thus,* $\sum_{n=1}^{\infty} \sup_{(-\infty,\varepsilon]} f_n(x)$ *converges, so by your homework,* $\sum_{n=1}^{\infty} \frac{e^{nx}}{n}$ *converges uniformly in* $(-\infty, -\varepsilon]$*. It remains to show that if* $E \subseteq (-\infty, 0)$ *is such that* $\sup E = 0$*, then the series does not converge uniformly in* E*. I will skip this because is very similar to your homework.*

Theorem 162 (Ratio Test) *Given a series* $\sum_{n=1}^{\infty} x_n$ *with* $x_n > 0$ *for all* $n \in \mathbb{N}$*, if*

$$\limsup_{n\to\infty} \frac{x_{n+1}}{x_n} < 1,$$

then the series converges. If

$$\liminf_{n\to\infty}\frac{x_{n+1}}{x_n} > 1,$$

then the series diverges to ∞.

Proof. If $\limsup_{n\to\infty}\frac{x_{n+1}}{x_n} < 1$, then by your homework, $\limsup_{n\to\infty}\sqrt[n]{x_n} < 1$, and so by the root test, the series converges. On the other hand, if $\ell = \liminf_{n\to\infty}\frac{x_{n+1}}{x_n} > 1$, fix $\varepsilon > 0$ so small that $\ell - \varepsilon > 1$. By an exercise in your homework, there exists $N \in \mathbb{N}$ such that

$$\frac{x_{n+1}}{x_n} \geq \ell - \varepsilon$$

for all $n \geq N$, and so $x_{n+1} \geq (\ell - \varepsilon)\, x_n \geq x_n$ for all $n \geq N$, which implies the sequence $\{x_n\}_n$ is increasing for $n \geq N$. Hence, there exists $\lim_{n\to\infty} x_n = \sup_{n\geq N} x_n > 0$. It follows by Theorem 149, that the series $\sum_{n=1}^{\infty} x_n$ cannot converge. In turn, by Theorem 152, $\sum_{n=1}^{\infty} x_n$ diverges to ∞. ■

Remark 163 *It follows from the second part of the proof, that if there exists* $N \in \mathbb{N}$ *such that* $\frac{x_{n+1}}{x_n} \geq 1$ *for all* $n \geq N$, *then the series* $\sum_{n=1}^{\infty} x_n$ *diverges to* ∞.

Remark 164 *In view of your homework, the ratio test is worse than the root test.*

Example 165 *Consider the series*

$$\sum_{n=1}^{\infty}\frac{n!x^n}{n^n},$$

where $x > 0$. *Note that* $\frac{n!x^n}{n^n}$. *By Stirling's formula*

$$\lim_{n\to\infty}\frac{n!x^n}{n^n} = \lim_{n\to\infty}\frac{n^n e^{-n}\sqrt{2n\pi}x^n}{n^n} = \lim_{n\to\infty}\left(\frac{x}{e}\right)^n\sqrt{2n\pi} = \begin{cases}\infty & \text{if } x \geq e,\\ 0 & \text{if } x < e.\end{cases}$$

Thus, by Theorems 149 and 152 the series diverges to ∞ *for* $x \geq e$. *It remains to study what happens for* $x < e$. *We have that*

$$\frac{x_{n+1}}{x_n} = \frac{\frac{(n+1)!x^{n+1}}{(n+1)^{n+1}}}{\frac{n!x^n}{n^n}} = \frac{\frac{n!(n+1)x^n x}{(n+1)^n(n+1)}}{\frac{n!x^n}{n^n}} = \frac{xn^n}{(n+1)^n} = \frac{x}{\left(\frac{n+1}{n}\right)^n} = \frac{x}{\left(1+\frac{1}{n}\right)^n} \to \frac{x}{e} < 1$$

for $x < e$. *Thus, for* $x < e$, *the series converges by the ratio test.*

Let's prove that we have uniform convergence in the set $E = (0, e-\varepsilon]$, *where* $0 < \varepsilon < e$. *Since* x^n *is increasing for* $x > 0$,

$$\sup_{(0,e-\varepsilon]} f_n(x) = f_n(e-\varepsilon) = \frac{n^n e^{-n}\sqrt{2n\pi}(e-\varepsilon)^n}{n^n}.$$

Hence,

$$\sum_{n=1}^{\infty} \sup_{(0,e-\varepsilon]} f_n(x) = \sum_{n=1}^{\infty} \frac{n^n e^{-n}\sqrt{2n\pi}(e-\varepsilon)^n}{n^n}.$$

We have already seen that $\sum_{n=1}^{\infty} \frac{n^n e^{-n}\sqrt{2n\pi}(e-\varepsilon)^n}{n^n}$ *converges. Thus,* $\sum_{n=1}^{\infty} \sup_{(0,e-\varepsilon]} f_n(x)$ *converges, so by your homework,* $\sum_{n=1}^{\infty} \frac{n!x^n}{n^n}$ *converges uniformly in* E*. It remains to show that if* $E \subseteq (0, e)$ *is such that* $\sup E = e$*, then the series does not converge uniformly in* E*. I will skip this because is very similar to your homework.*

Wednesday, March 16, 2022

11 Continuity

We recall that

Definition 166 *Let* (X, d_X)*,* (Y, d_Y) *be metric spaces, let* $E \subseteq X$*, and let* $f : E \to Y$*. We say that* f *is* continuous *at* $x_0 \in E \cap \operatorname{acc} E$ *if there exists*

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

If f *is continuous at every point of* $E \cap \operatorname{acc} E$ *we say that* f *is* continuous *on* E *and we write* $f \in C(E)$ *or* $f \in C^0(E)$*.*

Remark 167 *If* (X, τ_X) *and* (Y, τ_Y) *are two topological spaces,* $E \subseteq X$*,* $x_0 \in E \cap \operatorname{acc} E$*, and* $f : E \to Y$*, we say that* f *is* continuous *at* x_0 *if for every neighborhood* V *of* $f(x_0)$ *there exists a neighborhood* U *of* x_0 *with the property that*

$$f(x) \in V$$

for all $x \in E$ *with* $x \in U$*.*

Exercise 168 *Prove that the functions* $\sin x$*,* $\cos x$*,* x^n*, where* $n \in \mathbb{N}$*, are continuous.*

The following theorems follows from the analogous results for limits.

Theorem 169 *Let* (X, d)*, be a metric space, let* $E \subseteq X$*, and let* $x_0 \in E$*. Given two functions* f*,* $g : E \to \mathbb{R}$ *assume that* f *and* g *are continuous at* x_0*. Then*

(i) $f + g$ *and* fg *are continuous at* x_0*;*

(ii) if $g(x) \neq 0$ *for all* $x \in E$*, then* $\dfrac{f}{g}$ *is continuous at* x_0*.*

Example 170 *In view of Exercise 168 and the previous theorem, the functions* $\tan x = \frac{\sin x}{\cos x}$ *and* $\cot x = \frac{\cos x}{\sin x}$ *are continuous in their domain of definition.*

Theorem 171 *Let (X, d_X), (Y, d_Y), (Z, d_Z) be metric spaces, let $E \subseteq X$, let $F \subseteq Y$, and let $f : E \to F$ and $g : F \to Z$. Assume that f is continuous at x_0 and that g is continuous at $f(x_0)$. Then $g \circ f : E \to Z$ is continuous at x_0.*

Let $E \subseteq \mathbb{R}^N$ and let $\boldsymbol{f} : E \to \mathbb{R}^M$. Given $\boldsymbol{x}_0 \in E$, what happens when $\boldsymbol{f}$ is discontinuous at $\boldsymbol{x}_0$? Then $\boldsymbol{x}_0$ is an accumulation point of E. The following situations can arise. It can happen that there exists

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{\ell} \in \mathbb{R}^M$$

but $\boldsymbol{\ell} \neq \boldsymbol{f}(\boldsymbol{x}_0)$. In this case, we say that $\boldsymbol{x}_0$ is a *removable discontinuity.* Indeed, consider the function $g : E \to \mathbb{R}^M$ defined by

$$g(\boldsymbol{x}) := \begin{cases} \boldsymbol{f}(\boldsymbol{x}) & \text{if } \boldsymbol{x} \neq \boldsymbol{x}_0, \\ \boldsymbol{\ell} & \text{if } \boldsymbol{x} = \boldsymbol{x}_0. \end{cases}$$

Then there exists

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} g(\boldsymbol{x}) = \boldsymbol{\ell} = g(\boldsymbol{x}_0),$$

and so the new function g is continuous at $\boldsymbol{x}_0$.

Another type of discontinuity is when x_0 is an accumulation point of $E^- := E \cap (-\infty, x_0]$ and of $E^+ := E \cap (x_0, \infty)$ and there exist

$$\lim_{x \to x_0^-} \boldsymbol{f}(x) = \boldsymbol{\ell} \in \mathbb{R}^M, \quad \lim_{x \to x_0^+} \boldsymbol{f}(x) = L \in \mathbb{R}^M$$

but $\boldsymbol{\ell} \neq L$. In this case the point x_0 is called a *jump discontinuity* of $\boldsymbol{f}$.

Example 172 *The integer and fractional part of x have jump discontinuity at every integer.*

Finally, the last type of discontinuity is when at least one of the limits $\lim_{x \to x_0^-} \boldsymbol{f}(x)$ and $\lim_{x \to x_0^+} \boldsymbol{f}(x)$ is not finite or does not exist. In this case, the point x_0 is called an *essential discontinuity* of $\boldsymbol{f}$.

Example 173 *The function*

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

and

$$g(x) := \begin{cases} \log x & \text{if } x > 0, \\ 1 & \text{if } x = 0, \end{cases}$$

have an essential discontinuity at $x = 0$.

Theorem 174 *Let (X, d_X), (Y, d_Y) be metric spaces, let $E \subseteq X$, and let $f : E \to Y$.*

(i) Then f is continuous if and only if $f^{-1}(V)$ is relatively open for every open set $V \subseteq Y$.

(ii) Then f is continuous if and only if $f^{-1}(C)$ is relatively closed for every closed set $C \subseteq Y$.

Proof. (i) **Step 1:** Let $V \subseteq Y$ be open. Assume that f is continuous. If $f^{-1}(V)$ is empty, then there is nothing to prove. Otherwise, let $x_0 \in f^{-1}(V)$. Since V is open and $f(x_0) \in V$, there exists $\varepsilon > 0$ such that $B_Y(f(x_0), \varepsilon) \subseteq V$. Since f is continuous at x_0 there exists $\delta_{x_0} > 0$ such that for all $x \in E$ with $d_X(x, x_0) < \delta_{x_0}$, we have

$$d_Y(f(x), f(x_0)) < \varepsilon.$$

Hence, for all $x \in E$ with $d_X(x, x_0) < \delta_{x_0}$,

$$f(x) \in B_Y(f(x_0), \varepsilon) \subseteq V,$$

and so $B_X(x_0, \delta_{x_0}) \cap E \subseteq f^{-1}(V)$.

Take

$$U := \bigcup_{x \in f^{-1}(V)} B_X(x, \delta_x).$$

Then U is open and $f^{-1}(V) \subseteq U$. Hence,

$$U \cap E = f^{-1}(V),$$

which shows that $f^{-1}(V)$ is relatively open.

Step 2: Assume that $f^{-1}(V)$ is relatively open for every open set $V \subseteq Y$. Let $x_0 \in E \cap \operatorname{acc} E$ and let $\varepsilon > 0$. Consider the open set $V = B_Y(f(x_0), \varepsilon)$. Then $f^{-1}(V)$ is relatively open and so there exists an open set $U \subseteq X$ such that $U \cap E = f^{-1}(V)$. Since $x_0 \in f^{-1}(U)$, we have $x_0 \in U$. Hence there exists $B_X(x_0, \delta) \subseteq U$. It follows that for every $x \in U \cap E$ with $0 < d_X(x, x_0) < \delta$, then x belongs to $U \cap E = f^{-1}(V)$ and so $f(x) \in V = B_Y(f(x_0), \varepsilon)$, that is,

$$d_Y(f(x), f(x_0)) < \varepsilon,$$

which shows that there exists

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

(ii) Exercise. ■

As a corollary, we get.

Corollary 175 *Let (X, d_X), (Y, d_Y) be metric spaces, let $E \subseteq X$, and let $f : E \to Y$.*

(i) If E is open, then f is continuous if and only if $f^{-1}(V)$ is open for every open set $V \subseteq Y$.

(ii) If E is closed, then f is continuous if and only if $f^{-1}(C)$ is relatively closed for every closed set $C \subseteq Y$.

Remark 176 *The previous characterization of continuous functions is useful to define continuity in a topological space.*

Example 177 *The previous theorem implies in particular that sets of the form*

$$\{x \in \mathbb{R} : \ 4 \sin x - \log (1 + |x|) > 0\}$$

are open. We used this in the exercises.

Next we show that continuous functions preserve compactness.

Proposition 178 *Let (X, d_X), (Y, d_Y) be metric spaces, let $E \subseteq X$, and let $f : E \to Y$ be continuous. Then $f(K)$ is compact for every compact set $K \subseteq E$.*

Proof. Let $\{U_\alpha\}_{\alpha \in \Lambda}$ be an open cover of $f(K)$. By continuity, $f^{-1}(U_\alpha)$ is relatively open for every $\alpha \in \Lambda$, and so there exists W_α open such that $f^{-1}(U_\alpha) = E \cap W_\alpha$. The family $\{W_\alpha\}_{\alpha \in \Lambda}$ is an open cover of K. Since K is compact, we may find $U_{\alpha_1}, \dots, U_{\alpha_l}$ such that $\{W_{\alpha_i}\}_{i=1}^l$ cover K. In turn, $U_{\alpha_1}, \dots, U_{\alpha_l}$ cover $f(K)$. Indeed, if $y \in f(K)$, then there exists $x \in K$ such that $f(x) = y$. Let $i = 1, \dots, l$ be such that $x \in f^{-1}(U_{\alpha_i}) = E \cap W_{\alpha_i}$. Then $y = f(x) \in U_{\alpha_i}$. ■

An important consequence of the previous theorem is the following result.

Theorem 179 (Weierstrass) *Let (X, d) be a metric space, let $K \subseteq X$ be compact and let $f : K \to \mathbb{R}$ be continuous. Then there exist $x_0, x_1 \in K$ such that*

$$f(x_0) = \min_{x \in K} f(x), \quad f(x_1) = \max_{x \in K} f(x)$$

Proof. By the previous theorem $f(K)$ is compact in $\mathbb{R}$. It follows that $f(K)$ is closed and bounded. Hence, there exist $L = \sup f(K)$. There are now two cases. Either $L \in f(K)$ or $L \notin f(K)$. In the first case, there exists $x_1 \in K$ such that $f(x_1) = L = \sup f(K)$, that is, $f(x_1) \geq f(x)$ for all $x \in K$. On the other hand, if $L \notin f(K)$, then L would be an accumulation point of the set $f(K)$, but a closed set contains all its accumulation points. Hence, the case $L \notin f(K)$ cannot happen. This shows that f admits a maximum.

Similarly, taking $\ell = \inf f(K)$, we can show that $\ell \in f(K)$. ■

We now discuss the continuity of inverse functions and of composite functions. If a continuous function f is invertible its inverse function f^{-1} may not be continuous.

Example 180 *Let*

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

Then $f^{-1} : [0, 2] \to R$ is given by

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

which is not continuous at $x = 1$.

Friday, March 18, 2022

We will see that this cannot happen if $f : I \to \mathbb{R}$ with I an interval and when E is a compact set.

Exercise 181 *Consider the function* $\boldsymbol{f} : (-\pi, \pi) \to \mathbb{R}^2$ *given by* $\boldsymbol{f}(t) = (\sin(2t), \sin t)$. *Prove that* f *is injective, continuous but that the inverse is not continuous.*

Remark 182 *However, if* $U \subseteq \mathbb{R}^N$ *and* $\boldsymbol{f} : U \to \mathbb{R}^N$ *is continuous and injective, then* $\boldsymbol{f}(U)$ *is open and* $\boldsymbol{f}^{-1} : \boldsymbol{f}(U) \to \mathbb{R}^N$ *is continuous. This is a deep theorem known as invariance of the domain. We will not prove it in this course.*

Theorem 183 *Let* (X, d_X), (Y, d_Y) *be metric spaces, let* $K \subseteq X$ *be compact, and let* $f : K \to Y$ *be one-to-one and continuous. Then the inverse function* $f^{-1} : f(K) \to X$ *is continuous.*

Lemma 184 *Let* (X, d_X), *let* $K \subset X$ *be a compact set, and let* $C \subseteq K$ *be a closed set. Then* C *is compact.*

Proof. Let $\{U_\alpha\}_{\alpha\in\Lambda}$ be an open cover of C. Since C is closed, the set $U := X \setminus C$ is open. Note that

$$K = (K \setminus C) \cup C \subseteq U \cup \bigcup_\alpha U_\alpha.$$

Since K is compact, there exist $U_{\alpha_1}, \ldots, U_{\alpha_l}$ such that

$$K \subseteq U \cup \bigcup_{i=1}^{l} U_{\alpha_i}.$$

But since $U = X \setminus C$, it follows that

$$C \subseteq \bigcup_{i=1}^{l} U_{\alpha_i},$$

which shows that C is compact. ■

Proof of Theorem 183. Let $C \subseteq X$ be a closed set. By the previous lemma $K \cap C$ is compact. By Proposition 178 we have that $f(K \cap C)$ is compact. In particular, $f(K \cap C)$ is closed by Theorem 102. Let $g := f^{-1}$. Then

$$f(K \cap C) = g^{-1}(C),$$

which shows that $g^{-1}(C)$ is closed for every closed set $C \subseteq X$. Thus, by Theorem 174, g is continuous. ■

Remark 185 *Here we used the fact that a compact set is closed, so to extend this to a function* $f : K \to Y$, *where* $K \subseteq X$ *and* X *and* Y *are topological spaces, we need* Y *to be a Hausdorff topological space (see Remark 103).*

Example 186 *In view of the previous theorem and Exercise 168, the functions* $\arccos x$, $\arcsin x$, $\arctan x$ *are continuous.*

Given $a > 0$*, the function* $\log_a x$ *is continuous for* $x > 0$*, since it is the inverse of* a^x*.*

Given $n \in \mathbb{N}$*, the function* $\sqrt[2n+1]{x}$, $x \in \mathbb{R}$*, is continuous, since it is the inverse of* x^{2n+1}*. The function* $\sqrt[2n]{x}$, $x \in [0, \infty)$*, is continuous, since it is the inverse of* x^{2n}*.*

Given $a > 0$*, since* e^x *and* $\log x$ *are continuous in* $(0, \infty)$*, by writing*

$$x^a = e^{\log x^a} = e^{a \log x},$$
$$x^x = e^{\log x^x} = e^{x \log x},$$

it follows from Theorems 169 and 171, that x^a *and* x^x *are continuous in* $(0, \infty)$*.*

12 Directional Derivatives and Differentiability

Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be normed spaces, let $E \subseteq X$, let $f : E \to Y$ and let $x_0 \in E$. Given a direction $v \in X$ with $\|v\|_X = 1$, let L be the line through x_0 in the direction v, that is,

$$L := \{x \in X : x = x_0 + tx, t \in \mathbb{R}\},$$

and assume that x_0 is an accumulation point of the set $E \cap L$. The *directional derivative* of f at x_0 in the direction v is defined as

$$\frac{\partial f}{\partial v}(x_0) := \lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t},$$

provided the limit exists in Y.

If $X = \mathbb{R}^N$ and $\boldsymbol{v} = \boldsymbol{e}_i$, where $\boldsymbol{e}_i$ is a vector of the canonical basis, the directional derivative $\frac{\partial f}{\partial \boldsymbol{e}_i}(\boldsymbol{x}_0)$, if it exists, is called the *partial derivative* of f with respect to x_i and is denoted $\frac{\partial f}{\partial x_i}(\boldsymbol{x}_0)$ or $f_{x_i}(\boldsymbol{x}_0)$ or $D_i f(\boldsymbol{x}_0)$.

Remark 187 *When* $X = \mathbb{R}$*, taking* $v = 1$*, we get that*

$$\lim_{t \to 0} \frac{f(x_0 + t) - f(x_0)}{t},$$

which is the standard definition of derivative $f'(x_0)$*. It can be shown that if* $f'(x_0)$ *exists in* $\mathbb{R}$*, then* f *is continuous at* x_0*.*

In view of the previous remark, one would be tempted to say that if the directional derivatives at x_0 exist and are finite in every direction, then f is continuous at x_0. This is false in general, as the following example shows.

Example 188 *Let*

$$f(x, y) := \begin{cases} 1 & \textit{if } y = x^2,\ x \neq 0, \\ 0 & \textit{otherwise.} \end{cases}$$

Given a direction $\boldsymbol{v} = (v_1, v_2)$*, the line* L *through* $(0,0)$ *in the direction* $\boldsymbol{v}$ *intersects the parabola* $y = x^2$ *only in* $(0,0)$ *and in at most one point. Hence, if* t *is very small,*

$$f(0 + tv_1, 0 + tv_2) = 0.$$

It follows that

$$\frac{\partial f}{\partial \boldsymbol{v}}(0,0) = \lim_{t\to 0} \frac{f(0 + tv_1, 0 + tv_2) - f(0,0)}{t} = \lim_{t\to 0} \frac{0 - 0}{t} = 0.$$

However, f *is not continuous in* $(0,0)$*, since* $f\left(x, x^2\right) = 1 \to 1$ *as* $x \to 0$*, while* $f(x,0) = 0 \to 0$ *as* $x \to 0$*.*

Example 189 *Let*

$$f(x,y) := \begin{cases} \frac{x^2 y}{x^4 + y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Let's find the directional derivatives of f *at* $(0,0)$*. Given a direction* $\boldsymbol{v} = (v_1, v_2)$*, with* $v_1^2 + v_2^2 = 1$*, we have*

$$f(0 + tv_1, 0 + tv_2) = 0.$$

It follows that

$$\begin{aligned} \frac{f(0 + tv_1, 0 + tv_2) - f(0,0)}{t} &= \frac{\frac{(tv_1)^2 tv_2}{(tv_1)^4 + (tv_2)^2} - 0}{t} \\ &= \frac{t^3 v_1^2 v_2}{t^5 v_1^4 + t^3 v_2^2}. \end{aligned}$$

If $v_2 = 0$ *then*

$$\frac{f(0 + tv_1, 0 + tv_2) - f(0,0)}{t} = \frac{0}{t^5 v_1^4 + 0} = 0 \to 0$$

as $t \to 0$*, so* $\frac{\partial f}{\partial x}(0,0) = 0$*. If* $v_2 \neq 0$*, then,*

$$\frac{f(0 + tv_1, 0 + tv_2) - f(0,0)}{t} = \frac{v_1^2 v_2}{t^2 v_1^4 + v_2^2} \to \frac{v_1^2 v_2}{0 + v_2^2} = \frac{v_1^2}{v_2},$$

so

$$\frac{\partial f}{\partial \boldsymbol{v}}(0,0) = \frac{v_1^2}{v_2}.$$

In particular, $\frac{\partial f}{\partial y}(0,0) = \frac{0}{1} = 0$*. Now let's prove that* f *is not continuous at* $(0,0)$*. We have*

$$f(x,0) = \frac{0}{0 + y^2} = 0 \to 0$$

as $x \to 0$*, while*

$$f\left(x, x^2\right) = \frac{x^2 x^2}{x^4 + x^4} = \frac{1}{2} \to \frac{1}{2}$$

as $x \to 0$*. Hence, the limit* $\lim_{(x,y)\to(0,0)} f(x, y)$ *does not exists and so* f *is not continuous at* $(0, 0)$*. Note that* f *is continuous at all other points* $(x, y) \neq (0, 0)$ *by Theorem 169, since* $h(x, y) = x$ *and* $g(x, y) = y$ *are continuous functions in* $\mathbb{R}^2$*.*

Monday, March 21, 2022

The previous examples show that in dimension $N \geq 2$ partial derivatives do not give the same kind of results as in the case $N = 1$. To solve this problem, we introduce a stronger notion of derivative, namely, the notion of differentiability.

Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be normed spaces. We recall that a function $L : X \to Y$ is *linear* if

$$L(x_1 + x_2) = L(x_1) + L(x_2)$$

for all $x_1, x_2 \in X$ and

$$L(sx) = sL(x)$$

for all $s \in \mathbb{R}$ and $x \in X$.

Remark 190 *If* $X = \mathbb{R}^N$ *and* $Y = \mathbb{R}^M$*, then every linear function* $\boldsymbol{L} : \mathbb{R}^N \to \mathbb{R}^M$ *is continuous. Indeed, Write* $\boldsymbol{x} = \sum_{i=1}^N x_i \boldsymbol{e}_i$*. Then by the linearity of* $\boldsymbol{L}$*,*

$$\boldsymbol{L}(\boldsymbol{x}) = \boldsymbol{L}\left(\sum_{i=1}^N x_i \boldsymbol{e}_i\right) = \sum_{i=1}^N x_i \boldsymbol{L}(\boldsymbol{e}_i).$$

Define $\boldsymbol{b}_i := \boldsymbol{L}(\boldsymbol{e}_i) \in \mathbb{R}^M$*. Then the previous calculation shows that*

$$\boldsymbol{L}(\boldsymbol{x}) = \sum_{i=1}^N x_i \boldsymbol{b}_i \qquad \text{for all } \boldsymbol{x} \in \mathbb{R}^N,$$

which is continuous by Theorem 169.

The following example shows that when X is infinite-dimensional there exist linear functions which are not continuous.

Example 191 *Let* $X := \{f : [-1, 1] \to \mathbb{R} :$ *there exists* $f'(x) \in \mathbb{R}$ *for all* $x \in [-1, 1]\}$*. The vector space* X *is a normed space with the norm* $\|f\| := \max_{x\in[-1,1]} |f(x)|$*. Note that since* f *has a finite derivative at every* x*, it follows that* f *is continuous at every* x*. By the theorem on composition of continuous functions, the function* $|f(x)|$ *is also continuous. Since* $[-1, 1]$ *is compact, by the Weierstrass theorem, there exists* $\max_{x\in[-1,1]} |f(x)|$*. Hence,* $\|f\|$ *is well-defined. We have already seen in Exercise 50 that it is a norm.*

Consider the linear function $L : X \to \mathbb{R}$ *defined by*

$$L(f) := f'(0).$$

Then L is linear. To prove that L is not continuous, consider

$$f_n(x) := \frac{1}{n}\sin\left(n^2 x\right).$$

Then

$$\|f_n - 0\| \leq \frac{1}{n} \to 0$$

but

$$f_n'(x) = n\cos\left(n^2 x\right)$$

so that

$$L(f_n) = f_n'(0) = n \to \infty$$

and so L is not continuous, since $L(f_n) \nrightarrow L(0) = 0$.

Definition 192 *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be normed spaces, let $E \subseteq X$, let $f : E \to Y$, and let $x_0 \in E$ be an accumulation point of E. The function f is* differentiable at x_0 *if there exists a continuous linear function $L : X \to Y$ (depending on f and x_0) such that*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|_X} = 0. \tag{15}$$

provided the limit exists. The function L, if it exists, is called the differential *of f at x_0 and is denoted $df(x_0)$ or df_{x_0}.*

Remark 193 *Since f takes values in Y the limit (15) is equivalent to*

$$\lim_{x \to x_0} \left\| \frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|_X} \right\|_Y = 0.$$

Exercise 194 *Prove that if $N = 1$, then f is differentiable at x_0 if and only there exists the derivative $f'(x_0) \in \mathbb{R}$.*

The next theorem shows that differentiability in dimension $N \geq 2$ plays the same role of the derivative in dimension $N = 1$.

Theorem 195 *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be normed spaces, let $E \subseteq X$, let $f : E \to Y$, and let $x_0 \in E$ be an accumulation point of E. If f is differentiable at x_0, then f is continuous at x_0.*

Proof. Let L be the differential of f at x_0. We have

$$\begin{aligned} f(x) - f(x_0) &= f(x) - f(x_0) - L(x - x_0) + L(x - x_0) \\ &= \frac{f(x) - f(x) - L(x - x_0)}{\|x - x_0\|_X} \|x - x_0\| + L(x - x_0). \end{aligned}$$

Hence, by the properties of the norm for $x \in E$, $x \neq x_0$,

$$0 \leq \|f(x) - f(x_0)\|_Y \leq \left\| \frac{f(x) - f(x) - L(x - x_0)}{\|x - x_0\|_X} \right\|_Y \|x - x_0\|_X + \|L(x - x_0)\|_Y$$
$$\to \|0\|_Y \|0\|_X + \|L(0)\| = 0$$

as $x \to x_0$. It follows that f is continuous at x_0. ■

Next we study the relation between directional derivatives and differentiability. Here we need x_0 to be an interior point of E.

Theorem 196 *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be normed spaces, let $E \subseteq X$, let $f : E \to Y$, and let $x_0 \in E^\circ$. If f is differentiable at x_0, then all the directional derivatives of f at x_0 exist and*

$$\frac{\partial f}{\partial v}(x_0) = L(v),$$

where L is the differential of f at x_0. In particular, the function

$$v \mapsto \frac{\partial f}{\partial v}(x_0)$$

is linear.

Proof. Since x_0 is an interior point, there exists $B(x_0, r) \subseteq E$. Let $v \in X$ be a direction, so that $\|v\|_X = 1$, and take $x = x_0 + tv$. Note that for $|t| < r$, we have that

$$\|x - x_0\|_X = \|x_0 + tv - x_0\|_X = \|tv\|_X = |t| \|v\|_X = |t| < r$$

and so $x_0 + tv \in B(x_0, r) \subseteq E$. Moreover, $x \to x_0$ as $t \to 0$ and so, since f is differentiable at x_0,

$$0 = \lim_{x \to x_0} \frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|_X} = \lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0) - L(x_0 + tv - x_0)}{\|x_0 + tv - x_0\|}$$
$$= \lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0) - tL(v)}{|t|}.$$

By considering the left and right limits we get

$$0 = \lim_{t \to 0^+} \frac{f(x_0 + tv) - f(x_0) - tL(v)}{t} = 0,$$
$$0 = \lim_{t \to 0^-} \frac{f(x_0 + tv) - f(x_0) - tL(v)}{-t} = -\lim_{t \to 0^-} \frac{f(x_0 + tv) - f(x_0) - tL(v)}{t}$$

and so

$$0 = \lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0) - tL(v)}{t} = \lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t} - L(v).$$

This shows that there exists $\frac{\partial f}{\partial v}(x_0) = L(v)$. ■

Remark 197 *If in the previous theorem x_0 is not an interior point but for some direction $v \in X$, the point x_0 is an accumulation point of the set $E \cap L$, where L is the line through x_0 in the direction v, then as in the first part of the proof we can show that there exists the directional derivative $\frac{\partial f}{\partial v}(x_0)$ and*

$$\frac{\partial f}{\partial v}(x_0) = T(v).$$

Remark 198 *In particular, if $X = \mathbb{R}^N$, then by the previous theorem*

$$L(\boldsymbol{e}_i) = \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0),$$

and so, writing $\boldsymbol{v} = \sum_{i=1}^N v_i \boldsymbol{e}_i$, by the linearity of L we have

$$L(\boldsymbol{v}) = L\left(\sum_{i=1}^N v_i \boldsymbol{e}_i\right) = \sum_{i=1}^N v_i L(\boldsymbol{e}_i) = \sum_{i=1}^N \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0) v_i.$$

Thus, only at interior points *of E, to check differentiability it is enough to prove that*

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{f(\boldsymbol{x}) - f(\boldsymbol{x}_0) - \sum_{i=1}^N \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0)(x_i - x_{0i})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0. \tag{16}$$

Wednesday, March 23, 2022

If all the partial derivatives of f at $\boldsymbol{x}_0$ exist, the vector

$$\left(\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0), \ldots, \frac{\partial f}{\partial x_N}(\boldsymbol{x}_0)\right) \in \mathbb{R}^N$$

is called the *gradient* of f at $\boldsymbol{x}_0$ and is denoted by $\nabla f(\boldsymbol{x}_0)$ or $\operatorname{grad} f(\boldsymbol{x}_0)$ or $Df(\boldsymbol{x}_0)$. Note the previous theorem shows that

$$df(\boldsymbol{x}_0)(\boldsymbol{v}) = L(\boldsymbol{v}) = \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{v} = \sum_{i=1}^N \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0) v_i. \tag{17}$$

for all directions $\boldsymbol{v}$.

Exercise 199 *Let*

$$f(x,y) := \begin{cases} \frac{x^2 y}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Prove that f is continuous at 0, that all directional derivatives of f at 0 exist but that the formula

$$\frac{\partial f}{\partial \boldsymbol{v}}(0,0) = \frac{\partial f}{\partial x}(0,0) v_1 + \frac{\partial f}{\partial y}(0,0) v_2$$

fails.

Exercise 200 *Let*

$$f(x,y) := \begin{cases} \frac{x^2 y}{x^2+y^4} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Find all directional derivatives of f at $\mathbf{0}$. Study the continuity and the differentiability of f at $\mathbf{0}$.

Exercise 201 *Let $f : E \to \mathbb{R}$ be Lipschitz and let $\boldsymbol{x}_0 \in E^\circ$.*

(i) Assume that all the directional derivatives of f at $\boldsymbol{x}_0$ exist and that $\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0) v_i$ for every direction $\boldsymbol{v}$. Prove that f is differentiable at $\boldsymbol{x}_0$.

(ii) Assume that all the partial derivatives of f at $\boldsymbol{x}_0$ exist, that the directional derivatives $\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0)$ exist for all $\boldsymbol{v} \in S$, where S is dense in the unit sphere, and that $\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0) v_i$ for every direction $\boldsymbol{v} \in S$. Prove that f is differentiable at $\boldsymbol{x}_0$.

The next theorem gives an important sufficient condition for differentiability at a point $\boldsymbol{x}_0$.

Theorem 202 *Let $E \subseteq \mathbb{R}^N$, let $f : E \to \mathbb{R}$, let $\boldsymbol{x}_0 \in E^\circ$, and let $i \in \{1, \ldots, N\}$. Assume that there exists $r > 0$ such that $B(\boldsymbol{x}_0, r) \subseteq E$ and for all $j \neq i$ and for all $\boldsymbol{x} \in B(\boldsymbol{x}_0, r)$ the partial derivative $\frac{\partial f}{\partial x_j}$ exists at $\boldsymbol{x}$ and is continuous at $\boldsymbol{x}_0$. Assume also that $\frac{\partial f}{\partial x_i}(\boldsymbol{x}_0)$ exists. Then f is differentiable at $\boldsymbol{x}_0$.*

The proof makes use of the following theorem, which was proved in recitations.

Theorem 203 (Lagrange or Mean Value Theorem) *Assume that $f : [a,b] \to \mathbb{R}$ is continuous in $[a,b]$ and has a derivative in (a,b). Then there exists $c \in (a,b)$ such that*

$$f(b) - f(a) = f'(c)(b-a).$$

We turn to the proof of Theorem 202.

We are now ready to prove Theorem 202

Proof of Theorem 202. Without loss of generality, we may assume that $i = N$. Let $\boldsymbol{x} \in B(\boldsymbol{x}_0, r)$. Write $\boldsymbol{x} = (x_1, \ldots, x_N)$ and $\boldsymbol{x}_0 = (y_1, \ldots, y_N)$. Then

$$\begin{aligned} f(\boldsymbol{x}) - f(\boldsymbol{x}_0) &= (f(x_1, \ldots, x_N) - f(y_1, x_2, \ldots, x_N)) \\ &\quad + \cdots + (f(y_1, \ldots, y_{N-1}, x_N) - f(y_1, \ldots, y_N)). \end{aligned}$$

By the mean value theorem applied to the function of one variable $f(\cdot, x_2, \ldots, x_N)$,

$$f(x_1, \ldots, x_N) - f(y_1, x_2, \ldots, x_N) = \frac{\partial f}{\partial x_1}(\boldsymbol{z}_1)(x_1 - y_1),$$

where $\boldsymbol{z}_1 := (\theta_1 x_1 + (1-\theta_1) y_1, x_2, \dots, x_N)$ for some $\theta_1 \in (0,1)$. Note that

$$\|\boldsymbol{z}_1 - \boldsymbol{x}_0\| \le \|\boldsymbol{x} - \boldsymbol{x}_0\|.$$

Similarly, for $i = 2, \dots, N-1$,

$$f(y_1, \dots, y_{i-1}, x_i, \dots, x_N) - f(y_1, \dots, y_{i-1}, y_i, \dots, x_N) = \frac{\partial f}{\partial x_i}(\boldsymbol{z}_i)(x_i - y_i),$$

where $\boldsymbol{z}_i := (y_1, \dots, y_{i-1}, \theta_i x_i + (1-\theta_i) y_i, x_{i+1}, \dots, x_N)$ for some $\theta_i \in (0,1)$ and

$$\|\boldsymbol{z}_i - \boldsymbol{x}_0\| \le \|\boldsymbol{x} - \boldsymbol{x}_0\|.$$

Write

$$\begin{aligned} &f(\boldsymbol{x}) - f(\boldsymbol{x}_0) - \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) \\ &\quad = \sum_{i=1}^{N-1} \left(\frac{\partial f}{\partial x_i}(\boldsymbol{z}_i) - \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0) \right)(x_i - y_i) \\ &\qquad + \left(\frac{f(y_1, \dots, y_{N-1}, x_N) - f(y_1, \dots, y_N)}{x_N - y_N} - \frac{\partial f}{\partial x_N}(\boldsymbol{x}_0) \right)(x_N - y_N). \end{aligned}$$

Then

$$\begin{aligned} \frac{|f(\boldsymbol{x}) - f(\boldsymbol{x}_0) - \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0)|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} &\le \sum_{i=1}^{N-1} \left| \frac{\partial f}{\partial x_i}(\boldsymbol{z}_i) - \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0) \right| \frac{|x_i - y_i|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} \\ + \left| \frac{f(y_1, \dots, y_{N-1}, y_N + (x_N - y_N)) - f(y_1, \dots, y_N)}{x_N - y_N} \right. &- \left. \frac{\partial f}{\partial x_N}(\boldsymbol{x}_0) \right| \frac{|x_N - y_N|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|}. \end{aligned}$$

Since $\frac{|x_i - y_i|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} \le 1$, we have that

$$\begin{aligned} 0 \le \frac{|f(\boldsymbol{x}) - f(\boldsymbol{x}_0) - \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0)|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} &\le \sum_{i=1}^{N-1} \left| \frac{\partial f}{\partial x_i}(\boldsymbol{z}_i) - \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0) \right| \qquad (18) \\ + \left| \frac{f(y_1, \dots, y_{N-1}, y_N + (x_N - y_N)) - f(y_1, \dots, y_N)}{x_N - y_N} - \frac{\partial f}{\partial x_N}(\boldsymbol{x}_0) \right|&. \end{aligned}$$

Using the fact that $\|\boldsymbol{z}_i - \boldsymbol{x}_0\| \le \|\boldsymbol{x} - \boldsymbol{x}_0\| \to 0$ as $\boldsymbol{x} \to \boldsymbol{x}_0$, together with the continuity of $\frac{\partial f}{\partial x_i}$ at $\boldsymbol{x}_0$, gives

$$\left| \frac{\partial f}{\partial x_i}(\boldsymbol{z}_i) - \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0) \right| \to 0,$$

while, since $t := x_N - y_N \to 0$ as $\boldsymbol{x} \to \boldsymbol{x}_0$, we have that

$$\begin{aligned} &\frac{f(y_1, \dots, y_{N-1}, y_N + (x_N - y_N)) - f(y_1, \dots, y_N)}{x_N - y_N} \\ &= \frac{f(y_1, \dots, y_{N-1}, y_N + t) - f(y_1, \dots, y_N)}{t} \to \frac{\partial f}{\partial x_N}(\boldsymbol{x}_0), \end{aligned}$$

and so the right-hand side of (18) goes to zero as $\boldsymbol{x} \to \boldsymbol{x}_0$. ■

Friday, March 25, 2022

Example 204 *Let*

$$f(x,y) := \begin{cases} \frac{x^2|y|}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Let's study continuity, partial derivatives and differentiability. For $(x,y) \neq (0,0)$, we have that f is continuous by Theorem 169, while for $(x,y) = (0,0)$, we need to check that

$$\lim_{(x,y)\to(0,0)} f(x,y) = f(0,0).$$

We have

$$0 \leq |f(x,y) - f(0,0)| = \left|\frac{x^2|y|}{x^2+y^2} - 0\right| = \frac{x^2|y|}{x^2+y^2} \leq \frac{\left(x^2+y^2\right)|y|}{x^2+y^2} = |y| \to 0$$

as $(x,y) \to (0,0)$. Hence, f is continuous at $(0,0)$.

Next, let's study partial derivatives. For $(x,y) \neq (0,0)$, by the quotient rule, we have

$$\frac{\partial f}{\partial x}(x,y) = \frac{2x|y|\left(x^2+y^2\right) - x^2|y|(2x+0)}{\left(x^2+y^2\right)^2}, \tag{19}$$

while for $(x,y) = (0,0)$,

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t\to 0} \frac{f(0+t1, 0+t0) - f(0,0)}{t} = \lim_{t\to 0} \frac{\frac{t^2|0|}{t^2+0} - 0}{t} = \lim_{t\to 0} \frac{0}{t^3} = 0.$$

For $y \neq 0$, by the quotient rule, we have

$$\frac{\partial f}{\partial y}(x,y) = \frac{x^2\frac{y}{|y|}\left(x^2+y^2\right) - x^2|y|(0+2y)}{\left(x^2+y^2\right)^2}, \tag{20}$$

while at a point $(x_0, 0)$,

$$\frac{\partial f}{\partial y}(x_0,0) = \lim_{t\to 0} \frac{f(x_0+t0, 0+t1) - f(x_0,0)}{t} = \lim_{t\to 0} \frac{\frac{x_0^2|t|}{x_0^2+t^2} - 0}{t} = \lim_{t\to 0} \frac{|t|}{t}\frac{x_0^2}{x_0^2+t^2}.$$

If $x_0 = 0$, then $\frac{|t|}{t}\frac{x_0^2}{x_0^2+t^2} = \frac{|t|}{t}\frac{0}{0+t^2} = 0 \to 0$ as $t \to 0$, so $\frac{\partial f}{\partial y}(0,0) = 0$, while if $x_0 \neq 0$, we have

$$\lim_{t\to 0^+} \frac{|t|}{t}\frac{x_0^2}{x_0^2+t^2} = \lim_{t\to 0^+} \frac{t}{t}\frac{x_0^2}{x_0^2+t^2} = \lim_{t\to 0^+} \frac{x_0^2}{x_0^2+t^2} = \frac{x_0^2}{x_0^2+0} = 1,$$
$$\lim_{t\to 0^-} \frac{|t|}{t}\frac{x_0^2}{x_0^2+t^2} = \lim_{t\to 0^+} \frac{-t}{t}\frac{x_0^2}{x_0^2+t^2} = -\lim_{t\to 0^+} \frac{x_0^2}{x_0^2+t^2} = -\frac{x_0^2}{x_0^2+0} = -1.$$

Hence, $\frac{\partial f}{\partial y}(x_0,0)$ does not exist at $(x_0,0)$ for $x_0 \neq 0$, and so by Theorem 196, f is not differentiable at $(x_0,0)$ for $x_0 \neq 0$.

On the other hand, at points (x, y) with $y \neq 0$, we have that $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist in a small ball centered at (x, y) (see (19) and (20)) and they are continuous by Theorem 169. Hence, we can apply Theorem 202 below to conclude that f is differentiable at all points (x, y) with $y \neq 0$.

It remains to study differentiability at $(0, 0)$. By 16, we need to prove that

$$\lim_{(x,y)\to(0,0)} \frac{f(x, y) - f(0, 0) - \nabla f(0, 0) \cdot ((x, y) - (0, 0))}{\|(x, y) - (0, 0)\|} = 0.$$

We have

$$\begin{aligned} \frac{f(x, y) - f(0, 0) - \nabla f(0, 0) \cdot ((x, y) - (0, 0))}{\|(x, y) - (0, 0)\|} &= \frac{\frac{x^2|y|}{x^2+y^2} - 0 - (0, 0) \cdot ((x, y) - (0, 0))}{\sqrt{x^2 + y^2}} \\ &= \frac{x^2 |y|}{(x^2 + y^2)^{3/2}}. \end{aligned}$$

Taking $y = x$, with $x > 0$, we get

$$\frac{x^2 |x|}{(x^2 + x^2)^{3/2}} = \frac{x^2 x^3}{(x^2 + x^2)^{3/2}} = \frac{1}{(2)^{3/2}} \nrightarrow 0.$$

Hence, f is not differentiable at $(0, 0)$.

Exercise 205 *Study the differentiability of the function*

$$f(x, y) = |x|y, \quad (x, y) \in \mathbb{R}^2.$$

Exercise 206 *Given the function*

$$f(x, y) = \sqrt{(y - x^2)(y - 2x^2)},$$

defined in $E = \{(x, y) \in \mathbb{R}^2 : (y - x^2)(y - 2x^2) \geq 0\}$, study the differentiability of f in E.

Remark 207 *In the previous exercise, at points on ∂E we cannot use (16) since we only proved it for interior points.*

We study the differentiability of composite functions.

Theorem 208 (Chain Rule) *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, $(Z, \|\cdot\|_Z)$ be three normed spaces, let $E \subseteq X$, let $x_0 \in E$ be an accumulation point of E, let $F \subseteq Y$, and let $f : E \to F$ and $g : F \to Z$. Assume that there exists the directional derivative $\frac{\partial f}{\partial v}(x_0)$, that $f(x_0) \in F$ and that g is differentiable at $f(x_0)$. Then there exists the directional derivative*

$$\frac{\partial (g \circ f)}{\partial v}(x_0) = dg(f(x_0)) \left(\frac{\partial f}{\partial v}(x_0) \right). \tag{21}$$

Moreover, if f is differentiable at x_0, then $g \circ f$ is differentiable at x_0 with

$$d(g \circ f)(x_0) = dg(f(x_0)) \circ df(x_0).$$

Remark 209 *Assume that* $Y = \mathbb{R}^M$ *and* $Z = \mathbb{R}$*. Then* $\boldsymbol{f} : E \to \mathbb{R}^M$*. Let* $\boldsymbol{y}_0 := \boldsymbol{f}(x_0)$*. If* $\boldsymbol{y}_0 \in F^\circ$*, then by (16),*

$$dg(\boldsymbol{y}_0)(\boldsymbol{v}) = \nabla g(\boldsymbol{y}_0) \cdot \boldsymbol{v} = \sum_{i=1}^{M} \frac{\partial g}{\partial y_i}(\boldsymbol{y}_0) v_i.$$

Hence, (21)becomes

$$\begin{aligned}\frac{\partial (g \circ \boldsymbol{f})}{\partial v}(x_0) &= \sum_{i=1}^{M} \frac{\partial g}{\partial y_i}(\boldsymbol{f}(x_0)) \frac{\partial f_i}{\partial v}(x_0) \\ &= \nabla g(\boldsymbol{f}(x_0)) \cdot \frac{\partial \boldsymbol{f}}{\partial v}(x_0).\end{aligned}$$

Monday, March 28, 2022

Proof. Since g is differentiable at $f(x_0)$, there exists $L : Y \to Z$ linear and continuous such that

$$\lim_{y \to f(x_0)} \frac{g(y) - g(f(x_0)) - L(y - f(x_0))}{\|y - f(x_0)\|_Y} = 0, \tag{22}$$

where L is the differential of g at $f(x_0)$, so $L = dg(f(x_0))$.

Since there exists the directional derivative $\frac{\partial f}{\partial v}(x_0)$, we have that there exists

$$\lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t} = \frac{\partial f}{\partial v}(x_0). \tag{23}$$

It follows that the function of one variable

$$t \mapsto f(x_0 + tv)$$

is continuous at $t = 0$. Hence, if we take $y = f(x_0 + tv)$, we have that

$$y = f(x_0 + tv) \to f(x_0) \quad \text{as } t \to 0. \tag{24}$$

Case 1: Assume that $f(x_0 + tv) \neq f(x_0)$ for all t small. Then by (22), (23), (24),

$$\begin{aligned}&\frac{g(f(x_0 + tv)) - g(f(x_0))}{t} - L\left(\frac{\partial f}{\partial v}(x_0)\right) \\ &= \frac{g(f(x_0 + tv)) - g(f(x_0)) - L(f(x_0 + tv) - f(x_0))}{t} \\ &\quad + L\left(\frac{f(x_0 + tv) - f(x_0)}{t} - \frac{\partial f}{\partial v}(x_0)\right) \\ &= \frac{g(f(x_0 + tv)) - g(f(x_0)) - L(f(x_0 + tv) - f(x_0))}{\|f(x_0 + tv) - f(x_0)\|_Y} \left\|\frac{f(x_0 + tv) - f(x_0)}{t}\right\|_Y \frac{|t|}{t} \\ &\quad + L\left(\frac{f(x_0 + tv) - f(x_0)}{t} - \frac{\partial f}{\partial v}(x_0)\right) \\ &\to 0 \left\|\frac{\partial f}{\partial v}(x_0)\right\|_Y (\pm 1) + L(0) = 0.\end{aligned}$$

This shows that there exists

$$\frac{\partial(g\circ f)}{\partial v}(x_0)=L\left(\frac{\partial f}{\partial v}(x_0)\right).$$

Case 2: There exists countably many t approaching zero such that $f(x_0+tv)=f(x_0)$. Hence, for these t,

$$\frac{f(x_0+tv)-f(x_0)}{t}=0\to 0$$

as $t\to 0$, which implies that $\frac{\partial f}{\partial v}(x_0)=0$.

Let $F:=\{t\in\mathbb{R}: f(x_0+tv)=f(x_0)\}$. For $t\in F$,

$$\frac{g(f(x_0+tv))-g(f(x_0))}{t}=\frac{0}{t}\to 0.$$

On the other hand, if $t\notin F$, then $f(x_0+tv)\neq f(x_0)$ and so

$$\begin{aligned}
&\frac{g(f(x_0+tv))-g(f(x_0))}{t}\\
&=\frac{g(f(x_0+tv))-g(f(x_0))-L(f(x_0+tv)-f(x_0))}{t}\\
&\quad+L\left(\frac{f(x_0+tv)-f(x_0)}{t}\right)\\
&=\frac{g(f(x_0+tv))-g(f(x_0))-L(f(x_0+tv)-f(x_0))}{\|f(x_0+tv)-f(x_0)\|_Y}\left\|\frac{f(x_0+tv)-f(x_0)}{t}\right\|_Y\frac{|t|}{t}\\
&\quad+L\left(\frac{f(x_0+tv)-f(x_0)}{t}\right)\\
&\to 0\left\|\frac{\partial f}{\partial v}(x_0)\right\|_Y(\pm 1)+L(0)=0.
\end{aligned}$$

by (22), (23), (24). This proves the first part of the statement.

The second part of the statement is left as an exercise. ■

Exercise 210 *Prove the second part of the theorem.*

Example 211 (Quotient Rule) *Let $(X,\|\cdot\|_X)$ be a normed space, let $E\subseteq X$ and let $\boldsymbol{f}:E\to\mathbb{R}^2$, with $\boldsymbol{f}(x)=(f_1(x),f_2(x))$, be such that $f_2(x)\neq 0$ for all $x\in E$. Let $g:\mathbb{R}^2\to\mathbb{R}$ be defined by $g(y_1,y_2)=\frac{y_1}{y_2}$. Then (exercise)*

$$\frac{\partial g}{\partial y_1}(y_1,y_2)=\frac{1}{y_2},\quad \frac{\partial g}{\partial y_2}(y_1,y_2)=-\frac{y_1}{(y_2)^2}.$$

If there exists $\frac{\partial \boldsymbol{f}}{\partial v}(x_0)$*, then by Remark 209,*

$$\begin{aligned}
\frac{\partial}{\partial v}\left(\frac{f_1}{f_2}\right)(x_0) &= \frac{\partial (g \circ \boldsymbol{f})}{\partial v}(x_0) \\
&= \frac{\partial g}{\partial y_1}\left(\boldsymbol{f}\,(x_0)\right)\frac{\partial f_1}{\partial v}(x_0) + \frac{\partial g}{\partial y_2}\left(\boldsymbol{f}\,(x_0)\right)\frac{\partial f_2}{\partial v}\,(x_0) \\
&= \frac{1}{f_2(x_0)}\frac{\partial f_1}{\partial v}(x_0) - \frac{f_1(x_0)}{(f_2(x_0))^2}\frac{\partial f_2}{\partial v}(x_0) \\
&= \frac{\frac{\partial f_1}{\partial v}(x_0) f_2(x_0) - f_1(x_0)\frac{\partial f_1}{\partial v}(x_0)}{(f_2(x_0))^2},
\end{aligned}$$

which is the quotient rule.

Example 212 (Product Rule) *Let* $(X, \|\cdot\|_X)$ *be a normed space, let* $E \subseteq X$ *and let* $\boldsymbol{f} : E \to \mathbb{R}^2$*. Let* $g : \mathbb{R}^2 \to \mathbb{R}$ *be defined by* $g(y_1, y_2) = y_1 y_2$*. Then (exercise)*

$$\frac{\partial g}{\partial y_1}(y_1, y_2) = y_2, \quad \frac{\partial g}{\partial y_2}(y_1, y_2) = y_1.$$

If there exists $\frac{\partial \boldsymbol{f}}{\partial v}(x_0)$*, then by Remark 209,*

$$\begin{aligned}
\frac{\partial}{\partial v}\left(f_1 f_2\right)(x_0) &= \frac{\partial (g \circ \boldsymbol{f})}{\partial v}(x_0) \\
&= \frac{\partial g}{\partial y_1}\left(\boldsymbol{f}\,(x_0)\right)\frac{\partial f_1}{\partial v}(x_0) + \frac{\partial g}{\partial y_2}\left(\boldsymbol{f}\,(x_0)\right)\frac{\partial f_2}{\partial v}\,(x_0) \\
&= f_2(x_0)\frac{\partial f_1}{\partial v}(x_0) + f_1(x_0)\frac{\partial f_2}{\partial v}(x_0),
\end{aligned}$$

which is the product rule.

Example 213 *Consider the function*

$$h\left(\boldsymbol{x}\right) := g\left(\|\boldsymbol{x}\|\right) = g\left(\sqrt{x_1^2 + x_2^2 + \cdots + x_N^2}\right),$$

where $g : [0, \infty) \to \mathbb{R}$ *is differentiable. Since the norm is differentiable at all* $\boldsymbol{x} \neq \boldsymbol{0}$*, by Theorem 208, we have that* h *is differentiable at* $\boldsymbol{x} \neq \boldsymbol{0}$*,*

$$\frac{\partial h}{\partial x_i}\left(\boldsymbol{x}\right) = g'\left(\sqrt{x_1^2 + x_2^2 + \cdots + x_N^2}\right)\frac{2x_i}{2\left(\sqrt{x_1^2 + x_2^2 + \cdots + x_N^2}\right)}.$$

On the other hand, the Euclidean norm is not differentiable at $\boldsymbol{x} = \boldsymbol{0}$ *and so we cannot apply the previous results. Hence, we use the definition to get*

$$\frac{h(\boldsymbol{0} + t\boldsymbol{e}_i) - h(\boldsymbol{0})}{t} = \frac{g(\|t\boldsymbol{e}_i\|) - g(\|\boldsymbol{0}\|)}{t} = \frac{g(|t|1) - g(0)}{t}.$$

We have

$$\lim_{t\to 0^+}\frac{g(|t|1)-g(0)}{t}=\lim_{t\to 0^+}\frac{g(t)-g(0)}{t}=g'(0),$$
$$\lim_{t\to 0^-}\frac{g(|t|1)-g(0)}{t}=\lim_{t\to 0^-}\frac{g(-t)-g(0)}{t}=-\lim_{t\to 0^-}\frac{g(-t)-g(0)}{-t}=-g'(0).$$

Hence, $\frac{\partial h}{\partial x_i}(\mathbf{0})$ *exists if and only if* $g'(0)=0$*. Next, assume that* $g'(0)=0$ *and let's study differentiability at* $\boldsymbol{x}=\mathbf{0}$*. We have*

$$\begin{aligned}\frac{h(\boldsymbol{x})-h(\mathbf{0})-\nabla h(\mathbf{0})\cdot(\boldsymbol{x}-\mathbf{0})}{\|\boldsymbol{x}-\mathbf{0}\|}&=\frac{g(\|\boldsymbol{x}\|)-g(\|\mathbf{0}\|)-\mathbf{0}\cdot\boldsymbol{x}}{\|\boldsymbol{x}\|}\\&=\frac{g(\|\boldsymbol{x}\|)-g(0)}{\|\boldsymbol{x}\|}\to g'(0)=0\end{aligned}$$

as $\boldsymbol{x}\to\mathbf{0}$*.*

Next we define the Jacobian of a vectorial function $\boldsymbol{f}:E\to\mathbb{R}^M$.

Definition 214 *Given a set* $E\subseteq\mathbb{R}^N$ *and a function* $\boldsymbol{f}:E\to\mathbb{R}^M$*, the* Jacobian matrix *of* $\boldsymbol{f}=(f_1,\ldots,f_M)$ *at some point* $\boldsymbol{x}_0\in E$*, whenever it exists, is the* $M\times N$ *matrix*

$$J_{\boldsymbol{f}}(\boldsymbol{x}_0):=\begin{pmatrix}\nabla f_1(\boldsymbol{x}_0)\\ \vdots\\ \nabla f_M(\boldsymbol{x}_0)\end{pmatrix}=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(\boldsymbol{x}_0)&\cdots&\frac{\partial f_1}{\partial x_N}(\boldsymbol{x}_0)\\ \vdots&&\vdots\\ \frac{\partial f_M}{\partial x_1}(\boldsymbol{x}_0)&\cdots&\frac{\partial f_M}{\partial x_N}(\boldsymbol{x}_0)\end{pmatrix}.$$

It is also denoted

$$\frac{\partial(f_1,\ldots,f_M)}{\partial(x_1,\ldots,x_N)}(\boldsymbol{x}_0).$$

When $M=N$*,* $J_{\boldsymbol{f}}(\boldsymbol{x}_0)$ *is an* $N\times N$ *square matrix and its determinant is called the* Jacobian determinant *of* $\boldsymbol{f}$ *at* $\boldsymbol{x}_0$*. Thus,*

$$\det J_{\boldsymbol{f}}(\boldsymbol{x}_0)=\det\left(\frac{\partial f_j}{\partial x_i}(\boldsymbol{x}_0)\right)_{i,j=1,\ldots,N}.$$

Remark 215 *Let* $E\subseteq\mathbb{R}^N$*, let* $\boldsymbol{f}:E\to\mathbb{R}^M$*, and let* $\boldsymbol{x}_0\in E^\circ$*. Assume that* $\boldsymbol{f}$ *is differentiable at* $\boldsymbol{x}_0$*. Then all its components* f_j*,* $j=1,\ldots,M$*, are differentiable at* $\boldsymbol{x}_0$ *with*

$$d\boldsymbol{f}(\boldsymbol{x}_0)=(df_1(\boldsymbol{x}_0),\ldots,df_M(\boldsymbol{x}_0)).$$

Since $\boldsymbol{x}_0$ *is an interior point, it follows from (17) that for every direction* $\boldsymbol{v}$*,*

$$df_j(\boldsymbol{x}_0)(\boldsymbol{v})=\nabla f_j(\boldsymbol{x}_0)\cdot\boldsymbol{v}=\sum_{i=1}^N\frac{\partial f_j}{\partial x_i}(\boldsymbol{x}_0)v_i.$$

Hence,

$$\begin{aligned}d\boldsymbol{f}(\boldsymbol{x}_0)(\boldsymbol{v})&=(df_1(\boldsymbol{x}_0)(\boldsymbol{v}),\ldots,df_M(\boldsymbol{x}_0)(\boldsymbol{v}))\\&=J_{\boldsymbol{f}}(\boldsymbol{x}_0)\boldsymbol{v}^T.\end{aligned}$$

As a corollary of Theorem 208, we have the following result.

Corollary 216 *Let $E \subseteq \mathbb{R}^N$, $F \subseteq \mathbb{R}^M$, let $\boldsymbol{f} : E \to F$ and let $\boldsymbol{g} : F \to \mathbb{R}^P$. Assume that $\boldsymbol{f}$ is differentiable at some point $\boldsymbol{x}_0 \in E^\circ$ and that $\boldsymbol{g}$ is differentiable at the point $\boldsymbol{f}(\boldsymbol{x}_0)$ and that $\boldsymbol{f}(\boldsymbol{x}_0) \in F^\circ$. Then the composite function $\boldsymbol{g} \circ \boldsymbol{f}$ is differentiable at $\boldsymbol{x}_0$ and*

$$J_{\boldsymbol{g} \circ \boldsymbol{f}}(\boldsymbol{x}_0) = J_{\boldsymbol{g}}(\boldsymbol{f}(\boldsymbol{x}_0)) J_{\boldsymbol{f}}(\boldsymbol{x}_0).$$

Wednesday, March 30, 2022

13 Higher Order Derivatives

Let $E \subseteq \mathbb{R}^N$, let $f : E \to \mathbb{R}$ and let $\boldsymbol{x}_0 \in E$. Let $i \in \{1, \ldots, N\}$ and assume that there exists the partial derivatives $\frac{\partial f}{\partial x_i}(\boldsymbol{x})$ for all $\boldsymbol{x} \in E$. If $j \in \{1, \ldots, N\}$ and $\boldsymbol{x}_0$ is an accumulation point of $E \cap L$, where L is the line through $\boldsymbol{x}_0$ in the direction $\boldsymbol{e}_j$, then we can consider the partial derivative of the function $\frac{\partial f}{\partial x_i}$ with respect to x_j, that is,

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) =: \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Note that in general the order in which we take derivatives is important.

Example 217 *Let*

$$f(x, y) := \begin{cases} \frac{x^3 y - x y^3}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

If $(x, y) \neq (0, 0)$, then by Examples 212 and 211,

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial}{\partial x}\left(\frac{x^3 y - x y^3}{x^2 + y^2}\right) = \frac{(3x^2 y - 1y^3)(x^2 + y^2) - (x^3 y - x y^3)(2x + 0)}{(x^2 + y^2)^2},$$

and

$$\frac{\partial f}{\partial y}(x, y) = \frac{\partial}{\partial y}\left(\frac{x^3 y - x y^3}{x^2 + y^2}\right) = \frac{(x^3 1 - x 3 y^2)(x^2 + y^2) - (x^3 y - x y^3)(0 + 2y)}{(x^2 + y^2)^2},$$

while at $(0, 0)$ we have:

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{t \to 0} \frac{f(t, 0) - f(0, 0)}{t} = \lim_{t \to 0} \frac{\frac{0}{t^2 + 0} - 0}{t} = 0,$$
$$\frac{\partial f}{\partial y}(0, 0) = \lim_{t \to 0} \frac{f(0, 0 + t) - f(0, 0)}{t} = \lim_{t \to 0} \frac{\frac{0}{0 + t^2} - 0}{t} = 0.$$

Thus,

$$\frac{\partial f}{\partial x}(x,y)=\begin{cases}\frac{(3x^2y-1y^3)(x^2+y^2)-(x^3y-xy^3)(2x+0)}{(x^2+y^2)^2} & \text{if } (x,y)\neq(0,0),\\ 0 & \text{if } (x,y)=(0,0),\end{cases}$$
$$\frac{\partial f}{\partial y}(x,y)=\begin{cases}\frac{(x^31-x3y^2)(x^2+y^2)-(x^3y-xy^3)(0+2y)}{(x^2+y^2)^2} & \text{if } (x,y)\neq(0,0),\\ 0 & \text{if } (x,y)=(0,0).\end{cases}$$

To find $\frac{\partial^2 f}{\partial y\partial x}(0,0)$*, we calculate*

$$\frac{\partial^2 f}{\partial y\partial x}(0,0)=\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(0,0)=\lim_{t\to0}\frac{\frac{\partial f}{\partial x}(0,0+t)-\frac{\partial f}{\partial x}(0,0)}{t}$$
$$=\lim_{t\to0}\frac{\frac{(0-1t^3)(0+t^2)-0}{(0+t^2)^2}-0}{t}=\lim_{t\to0}-1=-1,$$

while

$$\frac{\partial^2 f}{\partial x\partial y}(0,0)=\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(0,0)=\lim_{t\to0}\frac{\frac{\partial f}{\partial y}(0+t,0)-\frac{\partial f}{\partial y}(0,0)}{t}$$
$$=\lim_{t\to0}\frac{\frac{(t^31-0)(t^2+0)-0}{(t^2+0)^2}-0}{t}=\lim_{t\to0}1=1.$$

Hence, $\frac{\partial^2 f}{\partial x\partial y}(0,0)\neq\frac{\partial^2 f}{\partial y\partial x}(0,0)$.

Exercise 218 *Let*

$$f(x,y):=\begin{cases}y^2\arctan\frac{x}{y} & \text{if } y\neq0,\\ 0 & \text{if } y=0.\end{cases}$$

Prove that $\frac{\partial^2 f}{\partial x\partial y}(0,0)\neq\frac{\partial^2 f}{\partial y\partial x}(0,0)$.

We present an improved version of the Schwartz theorem.

Theorem 219 (Schwartz) *Let* $E\subseteq\mathbb{R}^N$*, let* $f:E\to\mathbb{R}$*, let* $\boldsymbol{x}_0\in E^\circ$*, and let* $i,j\in\{1,\dots,N\}$*. Assume that there exists* $r>0$ *such that* $B(\boldsymbol{x}_0,r)\subseteq E$ *and for all* $\boldsymbol{x}\in B(\boldsymbol{x}_0,r)$*, the partial derivatives* $\frac{\partial f}{\partial x_i}(\boldsymbol{x})$*,* $\frac{\partial f}{\partial x_j}(\boldsymbol{x})$*, and* $\frac{\partial^2 f}{\partial x_j\partial x_i}(\boldsymbol{x})$ *exist. Assume also that* $\frac{\partial^2 f}{\partial x_j\partial x_i}$ *is continuous at* $\boldsymbol{x}_0$*. Then there exists* $\frac{\partial^2 f}{\partial x_i\partial x_j}(\boldsymbol{x}_0)$ *and*

$$\frac{\partial^2 f}{\partial x_i\partial x_j}(\boldsymbol{x}_0)=\frac{\partial^2 f}{\partial x_j\partial x_i}(\boldsymbol{x}_0).$$

Lemma 220 *Let* $A:((-r,r)\setminus\{0\})\times((-r,r)\setminus\{0\})\to\mathbb{R}$*. Assume that the double limit* $\lim_{(s,t)\to(0,0)}A(s,t)$ *exists in* $\mathbb{R}$ *and that the limit* $\lim_{t\to0}A(s,t)$ *exists in* $\mathbb{R}$ *for all* $s\in(-r,r)\setminus\{0\}$*. Then the iterated limit* $\lim_{s\to0}\lim_{t\to0}A(s,t)$ *exists and*

$$\lim_{s\to0}\lim_{t\to0}A(s,t)=\lim_{(s,t)\to(0,0)}A(s,t).$$

Taking the lemma for granted for the time being, let's prove the theorem.

Proof of Theorem 219. Step 1: Assume that $N = 2$. Let $0 < |t|, |s| < \frac{r}{\sqrt{2}}$. Then the points $(x_0 + s, y_0)$, $(x_0 + s, y_0 + t)$, and $(x_0, y_0 + t)$ belong to $B((x_0, y_0), r)$. Define

$$A(s,t) := \frac{f(x_0 + s, y_0 + t) - f(x_0 + s, y_0) - f(x_0, y_0 + t) + f(x_0, y_0)}{st}, \quad 0 < |t|, |s| < \frac{r}{\sqrt{2}}.$$

Fix $0 < |t| < \frac{r}{\sqrt{2}}$ and consider the function

$$g(x) := f(x, y_0 + t) - f(x, y_0)$$

By the mean value theorem

$$A(s,t) = \frac{g(x_0 + s) - g(x_0)}{st} = \frac{g'(\xi)}{t} = \frac{\frac{\partial f}{\partial x}(\xi_t, y_0 + t) - \frac{\partial f}{\partial x}(\xi_t, y_0)}{t}$$

where ξ is between x_0 and $x_0 + t$. Consider the function

$$h(y) := \frac{\partial f}{\partial x}(\xi_t, y).$$

By the mean value theorem,

$$h(y_0 + t) - h(y_0) = h'(\eta_t)(y_0 + t - y_0) = \frac{\partial^2 f}{\partial y \partial x}(\xi_t, \eta_t)\, t$$

for some η_t is between y_0 and $y_0 + t$. This gives

$$\frac{\partial f}{\partial x}(\xi_t, y_0 + t) - \frac{\partial f}{\partial x}(\xi_t, y_0) = \frac{\partial^2 f}{\partial y \partial x}(\xi_t, \eta_t)\, t.$$

Hence,

$$A(s,t) = \frac{\partial^2 f}{\partial y \partial x}(\xi_t, \eta_t) \to \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0),$$

where we have used the fact that $(\xi, \eta) \to (x_0, y_0)$ as $(s,t) \to (0,0)$ together with the continuity of $\frac{\partial^2 f}{\partial y \partial x}$ at (x_0, y_0). Note that this shows that there exists the limit

$$\lim_{(s,t) \to (0,0)} A(s,t) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0).$$

On the other hand, for all $s \neq 0$,

$$\begin{aligned}\lim_{t \to 0} A(s,t) &= \frac{1}{s} \lim_{t \to 0} \left[\frac{f(x_0 + s, y_0 + t) - f(x_0 + s, y_0)}{t} - \frac{f(x_0, y_0 + t) - f(x_0, y_0)}{t} \right] \\ &= \frac{\frac{\partial f}{\partial y}(x_0 + s, y_0) - \frac{\partial f}{\partial y}(x_0, y_0)}{s}.\end{aligned}$$

Hence, we are in a position to apply the previous lemma to obtain

$$\begin{aligned}\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) &= \lim_{(s,t)\to(0,0)} A(s,t) = \lim_{s\to 0}\lim_{t\to 0} A(s,t) \\ &= \lim_{s\to 0} \frac{\frac{\partial f}{\partial y}(x_0 + s, y_0) - \frac{\partial f}{\partial y}(x_0, y_0)}{s} = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0).\end{aligned}$$

Step 2: In the case $N \geq 2$ let $\boldsymbol{x} = (x_1, \ldots, x_N)$, $\boldsymbol{x}_0 = (c_1, \ldots, c_N)$. Assume that $1 < i < j < N$ (the cases $i = 1$ and $j = N$ are similar) and apply Step 1 to the function of two variables

$$F(x_i, x_j) := f(c_1, \ldots, c_{i-1}, x_i, c_{i+1}, \ldots, c_{j-1}, x_j, c_{j+1}, \ldots, c_N)$$

■

Friday, April,1 2022

Let's now prove the lemma.

Proof. Let $\ell = \lim_{(s,t)\to(0,0)} A(s,t)$. Then for every $\varepsilon > 0$ there exists $\delta = \delta((0,0), \varepsilon) > 0$ such that

$$|A(s,t) - \ell| \leq \varepsilon$$

for all $(s,t) \in ((-r,r) \setminus \{0\}) \times ((-r,r) \setminus \{0\})$, with $\sqrt{|s-0|^2 + |t-0|^2} \leq \delta$.

Fix $s \in \left(-\frac{\delta}{2}, \frac{\delta}{2}\right) \setminus \{0\}$. Then for all $t \in \left(-\frac{\delta}{2}, \frac{\delta}{2}\right) \setminus \{0\}$,

$$|A(s,t) - \ell| \leq \varepsilon$$

and so letting $t \to 0$ in the previous inequality (and using the fact that the limit $\lim_{t\to 0} A(s,t)$ exists), we get

$$\left|\lim_{t\to 0} A(s,t) - \ell\right| \leq \varepsilon$$

for all $s \in \left(-\frac{\delta}{2}, \frac{\delta}{2}\right) \setminus \{0\}$. But this implies that there exists $\lim_{s\to 0}\lim_{t\to 0} A(s,t) = \ell$. ■

Next we prove Taylor's formula in higher dimensions. We recall that $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. A *multi-index* $\boldsymbol{\alpha}$ is a vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N) \in \mathbb{N}_0^N$. The *length* of a multi-index is defined as

$$|\boldsymbol{\alpha}| := \alpha_1 + \cdots + \alpha_N.$$

Given a multi-index $\boldsymbol{\alpha}$, the partial derivative $\frac{\partial^{\boldsymbol{\alpha}}}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}$ is defined as

$$\frac{\partial^{\boldsymbol{\alpha}}}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}} := \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial x_1^{\alpha_1} \cdots \partial x_N^{\alpha_N}},$$

where $\boldsymbol{x} = (x_1, \ldots, x_N)$. If $\boldsymbol{\alpha} = \boldsymbol{0}$, we set $\frac{\partial^{\boldsymbol{0}} f}{\partial \boldsymbol{x}^{\boldsymbol{0}}} := f$.

Example 221 *If $N = 3$ and $\boldsymbol{\alpha} = (2, 1, 0)$, then*

$$\frac{\partial^{(2,1,0)}}{\partial (x, y, z)^{(2,1,0)}} = \frac{\partial^3}{\partial x^2 \partial y}.$$

Given a multi-index $\boldsymbol{\alpha}$ and $\boldsymbol{x} \in \mathbb{R}^N$, we set

$$\boldsymbol{\alpha}! := \alpha_1! \cdots \alpha_N!, \quad \boldsymbol{x}^{\boldsymbol{\alpha}} := x_1^{\alpha_1} \cdots x_N^{\alpha_N}.$$

If $\boldsymbol{\alpha} = \mathbf{0}$, we set $\boldsymbol{x}^{\mathbf{0}} := 1$.

Definition 222 *Given an open set $U \subseteq \mathbb{R}^N$, for every nonnegative integer $m \in \mathbb{N}_0$, we denote by $C^m(U)$ the space of all functions that are continuous in U together with their partial derivatives up to order m (included). We set $C^\infty(U) := \bigcap_{m=0}^{\infty} C^m(U)$. When $N = 1$ we also define $C^m([a, b])$ the space of all functions that are continuous in $[a, b]$ together with their derivatives up to order m (included).*

Theorem 223 (Taylor's Formula) *Let $U \subseteq \mathbb{R}^N$ be an open set, let $f \in C^m(U)$, $m \in \mathbb{N}$, and let $\boldsymbol{x}_0 \in U$. Then for every $\boldsymbol{x} \in U$,*

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \sum_{\boldsymbol{\alpha} \text{ multi-index, } 1 \le |\boldsymbol{\alpha}| \le m} \frac{1}{\boldsymbol{\alpha}!} \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0)^{\boldsymbol{\alpha}} + R_m(\boldsymbol{x}),$$

where

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{R_m(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|^m} = 0.$$

Definition 224 *Given a metric space (X, d_X), a set $E \subseteq X$, and two functions $f : E \to \mathbb{R}$ and $g : E \to \mathbb{R}$ and a point $x_0 \in \operatorname{acc} E$, we say that the function f is a* little o *of g as $x \to x_0$, and we write $f = \mathrm{o}(g)$, if $g \neq 0$ in E and*

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0.$$

Hence, a little o of g is simply a function that goes to zero faster than g as $x \to x_0$. Therefore, Taylor's formula can be written as

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \sum_{\boldsymbol{\alpha} \text{ multi-index, } 1 \le |\boldsymbol{\alpha}| \le m} \frac{1}{\boldsymbol{\alpha}!} \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0)^{\boldsymbol{\alpha}} + o(\|\boldsymbol{x} - \boldsymbol{x}_0\|^m)$$

as $\boldsymbol{x} \to \boldsymbol{x}_0$.

Thursday, March 31, 2022

Recitation

Theorem 225 (Taylor's Formula) *Let $f \in C^{(m)}((a,b))$ and let $x_0 \in (a,b)$. Then for every $x \in (a,b)$,*

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2$$
$$+ \cdots + \frac{f^{(m)}(x_0)}{m!}(x - x_0)^m + R_m(x),$$

where the remainder $R_m(x, x_0)$ satisfies

$$\lim_{x \to x_0} \frac{R_m(x)}{(x - x_0)^m} = 0.$$

Lemma 226 *Let $g \in C^{(m)}((a,b))$ and let $x_0 \in (a,b)$. Then*

$$\lim_{x \to x_0} \frac{g(x)}{(x - x_0)^m} = 0 \tag{25}$$

if and only if

$$g(x_0) = g'(x_0) = \cdots = g^{(m)}(x_0) = 0. \tag{26}$$

Proof. Assume that (26) holds. By applying De l'Hôpital's theorem several times we get

$$\lim_{x \to x_0} \frac{g(x)}{(x - x_0)^m} = \lim_{x \to x_0} \frac{g'(x)}{m(x - x_0)^{m-1}} = \lim_{x \to x_0} \frac{g^{(2)}(x)}{m(m-1)(x - x_0)^{m-2}}$$
$$= \cdots = \lim_{x \to x_0} \frac{g^{(m-1)}(x)}{m!(x - x_0)} = \lim_{x \to x_0} \frac{g^{(m)}(x)}{m!1} = \frac{g^{(m)}(x_0)}{m!} = 0.$$

Conversely, assume (25). If $g^{(k)}(x_0) \neq 0$ for some $0 \leq k < m$, then by what we just proved (with k in place of m)

$$\lim_{x \to x_0} \frac{g(x)}{(x - x_0)^k} = \frac{g^{(k)}(x_0)}{k!} \neq 0.$$

On the other hand,

$$\frac{g(x)}{(x - x_0)^k} = \frac{g(x)}{(x - x_0)^k} \frac{(x - x_0)^{m-k}}{(x - x_0)^{m-k}} = \frac{g(x)}{(x - x_0)^m}(x - x_0)^{m-k} \to 0$$

as $x \to x_0$, which is a contradiction. ■

We now turn to the proof of Theorem 225.

Proof of Theorem 225. Note that given a polynomial of degree m,

$$p(x) = a_0 + a_1(x - x_0) + \cdots + a_m(x - x_0)^m = \sum_{i=0}^{m} a_i (x - x_0)^i,$$

we have that

$$p^{(k)}(x) = \sum_{i=k}^{m} i(i-1)\cdots(i-k+1)\,a_i\,(x-x_0)^{i-k},$$

so that

$$p^{(k)}(x_0) = k!a_k.$$

We apply the lemma to the function

$$g(x) := f(x) - p(x)$$

to conclude that

$$\lim_{x\to x_0} \frac{f(x)-p(x)}{(x-x_0)^m} = 0$$

if and only if for all $k = 0, \ldots, m$,

$$0 = g^{(k)}(x_0) = f^{(k)}(x_0) - p^{(k)}(x_0) = f^{(k)}(x_0) - k!a_k,$$

that is

$$a_k = \frac{f^{(k)}(x_0)}{k!}.$$

Thus

$$g(x) = R_m(x) = f(x) - \left[f(x_0) + f'(x_0)(x-x_0) + \cdots + \frac{f^{(m)}(x_0)}{m!}(x-x_0)^m\right].$$

■

Exercise 227 *Let $g : [a,b] \to \mathbb{R}$ and $m \in \mathbb{N}$. Assume g, g', $\ldots$, $g^{(m-1)}$ exist and are continuous in $[a,b]$ and that $g^{(m-1)}$ is differentiable in (a,b). Prove that if*

$$g(a) = g'(a) = \cdots = g^{(m-1)}(a) = 0, \quad g(b) = 0,$$

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

Exercise 228 *Let $m \in \mathbb{N}$, $f \in C^{(m)}((a,b))$, and $x_0 \in (a,b)$. Prove that for every $x \in (a,b)$,*

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(m)}(x_0)}{m!}(x-x_0)^m + R_m(x),$$

where

$$R_m(x) = \frac{1}{m!}[f^{(m)}(c) - f^{(m)}(x_0)](x-x_0)^m$$

for some c between x_0 and x. Deduce that R_m satisfies

$$\lim_{x\to x_0} \frac{R_m(x)}{(x-x_0)^m} = 0.$$

We prove Theorem 223.

Exercise 229 (Multinomial theorem) *Let* $\boldsymbol{x} = (x_1, \ldots, x_N) \in \mathbb{R}^N$ *and let* $n \in \mathbb{N}$. *Prove that*

$$(x_1 + \cdots + x_N)^n = \sum_{\boldsymbol{\alpha} \text{ multi-index, } |\boldsymbol{\alpha}|=n} \frac{n!}{\boldsymbol{\alpha}!} \boldsymbol{x}^{\boldsymbol{\alpha}}.$$

We are now ready to prove Taylor's formula.

Proof of Theorem 223. Since $\boldsymbol{x}_0 \in U$ and U is open, there exists $B(\boldsymbol{x}_0, r) \subseteq U$. Fix $\boldsymbol{x} \in B(\boldsymbol{x}_0, r)$, $\boldsymbol{x} \neq \boldsymbol{x}_0$ and let $\boldsymbol{v} := \frac{\boldsymbol{x}-\boldsymbol{x}_0}{\|\boldsymbol{x}-\boldsymbol{x}_0\|}$ and consider the function $g(t) := f(\boldsymbol{x}_0 + t\boldsymbol{v})$ defined for $t \in [0, r]$. By Theorem 208, we have that

$$\frac{dg}{dt}(t) = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0 + t\boldsymbol{v})\, v_i = (\boldsymbol{v} \cdot \nabla) f(\boldsymbol{x}_0 + t\boldsymbol{v})$$

with for all $t \in [0, r]$. By repeated applications of Theorem 208, we get that

$$\frac{d^{(n)}g}{dt^n}(t) = (\boldsymbol{v} \cdot \nabla)^n f(\boldsymbol{x}_0 + t\boldsymbol{v})$$

for all $n = 1, \ldots, m$, where $(\boldsymbol{v} \cdot \nabla)^n$ means that we apply the operator

$$\boldsymbol{v} \cdot \nabla = v_1 \frac{\partial}{\partial x_1} + \cdots + v_N \frac{\partial}{\partial x_N}$$

n times to f. By the multinomial theorem, and the fact that for functions in C^m partial derivatives commute,

$$\begin{aligned}(\boldsymbol{v} \cdot \nabla)^n &= \left(v_1 \frac{\partial}{\partial x_1} + \cdots + v_N \frac{\partial}{\partial x_N}\right)^n \\ &= \sum_{\boldsymbol{\alpha} \text{ multi-index, } |\boldsymbol{\alpha}|=n} \frac{n!}{\boldsymbol{\alpha}!} \boldsymbol{v}^{\boldsymbol{\alpha}} \frac{\partial^{\boldsymbol{\alpha}}}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}},\end{aligned}$$

and so

$$\frac{d^{(n)}g}{dt^n}(t) = \sum_{\boldsymbol{\alpha} \text{ multi-index, } |\boldsymbol{\alpha}|=n} \frac{n!}{\boldsymbol{\alpha}!} \boldsymbol{v}^{\boldsymbol{\alpha}} \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x}_0 + t\boldsymbol{v}).$$

Using Taylor's formula for g (see Exercise 228) we get

$$g(t) = g(0) + \sum_{n=1}^{m} \frac{1}{n!} \frac{d^{(n)}g}{dt^n}(0)(t-0)^n + R_m(t),$$

where

$$R_m(t) = \frac{1}{m!}[g^{(m)}(c) - g^{(m)}(0)]t^m \tag{27}$$

for some $0 < |c| < |t|$. Substituting, we obtain

$$\begin{aligned} f(\boldsymbol{x}_0 + t\boldsymbol{v}) &= f(\boldsymbol{x}_0) + \sum_{n=1}^{m} \frac{t^n}{n!} \sum_{\boldsymbol{\alpha} \text{ multi-index, } |\boldsymbol{\alpha}|=n} \frac{n!}{\boldsymbol{\alpha}!} \boldsymbol{v}^{\boldsymbol{\alpha}} \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x}_0) + R_m(t) \\ &= f(\boldsymbol{x}_0) + \sum_{\boldsymbol{\alpha} \text{ multi-index, } 1 \le |\boldsymbol{\alpha}| \le m} \frac{t^{|\boldsymbol{\alpha}|}}{\boldsymbol{\alpha}!} \boldsymbol{v}^{\boldsymbol{\alpha}} \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x}_0) + R_m(t). \end{aligned}$$

Take $t = \|\boldsymbol{x} - \boldsymbol{x}_0\|$. Then

$$\boldsymbol{x}_0 + t\boldsymbol{v} = \boldsymbol{x}_0 + \|\boldsymbol{x} - \boldsymbol{x}_0\| \frac{\boldsymbol{x} - \boldsymbol{x}_0}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = \boldsymbol{x}$$

and

$$t^{|\boldsymbol{\alpha}|} \boldsymbol{v}^{\boldsymbol{\alpha}} = \|\boldsymbol{x} - \boldsymbol{x}_0\|^{|\boldsymbol{\alpha}|} \frac{(\boldsymbol{x} - \boldsymbol{x}_0)^{\boldsymbol{\alpha}}}{\|\boldsymbol{x} - \boldsymbol{x}_0\|^{|\boldsymbol{\alpha}|}} = (\boldsymbol{x} - \boldsymbol{x}_0)^{\boldsymbol{\alpha}}$$

and so

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \sum_{\boldsymbol{\alpha} \text{ multi-index, } 1 \le |\boldsymbol{\alpha}| \le m} \frac{(\boldsymbol{x} - \boldsymbol{x}_0)^{\boldsymbol{\alpha}}}{\boldsymbol{\alpha}!} \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x}_0) + R_m(\|\boldsymbol{x} - \boldsymbol{x}_0\|).$$

Similarly, by (27),

$$R_m(\|\boldsymbol{x} - \boldsymbol{x}_0\|) = \sum_{\boldsymbol{\alpha} \text{ multi-index, } |\boldsymbol{\alpha}|=m} \frac{(\boldsymbol{x} - \boldsymbol{x}_0)^{\boldsymbol{\alpha}}}{\boldsymbol{\alpha}!} \left[\frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x}_0 + c\boldsymbol{v}) - \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x}_0) \right].$$

Hence,

$$\begin{aligned} \frac{|R_m(\|\boldsymbol{x} - \boldsymbol{x}_0\|)|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|^m} &\le \sum_{\boldsymbol{\alpha} \text{ multi-index, } |\boldsymbol{\alpha}|=n} \frac{\|\boldsymbol{x} - \boldsymbol{x}_0\|^m}{\|\boldsymbol{x} - \boldsymbol{x}_0\|^m \boldsymbol{\alpha}!} \left| \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x}_0 + c\boldsymbol{v}) - \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x}_0) \right| \\ &\le \sum_{\boldsymbol{\alpha} \text{ multi-index, } |\boldsymbol{\alpha}|=n} \left| \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x}_0 + c\boldsymbol{v}) - \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x}_0) \right| \to 0 \end{aligned}$$

as $\boldsymbol{x} \to \boldsymbol{x}_0$ since $0 < |c| < \|\boldsymbol{x} - \boldsymbol{x}_0\| \to 0$. ■

Wednesday, April 6, 2022

Properties of little o

- $f(x)\, o(g) = o(fg)$

Example: as $x \to 0$
$x^6 o\left(x^3\right) = o\left(x^{6+3}\right) = o\left(x^9\right)$

- $o(f) + o(g) = o(f+g) = o$ (the slower between f and g)

Example: as $x \to 0$
$o\left(x^5\right) + o\left(x^3\right) = o\left(x^5 + x^3\right) = o\left(x^3\right)$

- $(o(f))^a = o(f^a)$ where $a > 0$

Example: as $x \to 0$
$\left(o\left(x^5\right)\right)^{1/3} = o\left(\left(x^5\right)^{1/3}\right) = o\left(x^{5/3}\right)$

- $c\, o(f) = o(f)$ where c is any number different from zero

Examples: as $x \to 0$
$3o\left(x^4\right) = o\left(x^4\right)$
$-o\left(x^5\right) = o\left(x^5\right)$

- $f + o(g) = o(g)$ if f is faster than g

Example: as $x \to 0$
$x^5 + 2x^4 + o\left(x^4\right) - x^2 + x^8 + x = 2x^4 + o\left(x^4\right) - x^2 + x$ (hence the little o $o\left(x^4\right)$ absorbs all the powers of degree strictly bigger than 4, while all the powers of degree 4 or smaller than 4 remain)

- $o(o(g)) = o(g)$

- $\dfrac{o(f)}{o(g)}$ does not make sense in general.

Example 230 *Let's calculate the limit*

$$\lim_{(x,y)\to(0,0)} \frac{(1+x)^y - 1}{x^2 + y^2}.$$

First Method: *Let's use Taylor's formula of order $m = 2$ at $(0,0)$,*

$$\begin{aligned} f(x,y) &= f(0,0) + \frac{\partial f}{\partial x}(0,0)(x-0) + \frac{\partial f}{\partial y}(0,0)(y-0) \\ &\quad + \frac{1}{(2,0)!}\frac{\partial^2 f}{\partial x^2}(0,0)(x-0)^2 + \frac{1}{(1,1)!}\frac{\partial^2 f}{\partial x \partial y}(0,0)(x-0)(y-0) \\ &\quad + \frac{1}{(0,2)!}\frac{\partial^2 f}{\partial y^2}(0,0)(y-0)^2 + o\left(x^2+y^2\right). \end{aligned}$$

There are too many derivatives to compute so we will skip this.

Second Method: *A simpler method is to use the Taylor's formulas for e^t and $\log(1+x)$. We have*

$$\log(1+x) = x - \frac{1}{2}x^2 + o\left(x^2\right),$$
$$e^t = 1 + t + \frac{1}{2!}t^2 + o\left(t^2\right),$$

and so

$$
\begin{aligned}
f(x,y) &= e^{y\log(1+x)} - 1 = e^{y\left(x-\frac{1}{2}x^2+o\left(x^2\right)\right)} - 1 \\
&= 1 + \left(xy - \frac{1}{2}x^2y + o\left(x^2y\right)\right) + \frac{1}{2!}\left(xy - \frac{1}{2}x^2y + o\left(x^2y\right)\right)^2 + o\left(\left(xy - \frac{1}{2}x^2y + o\left(x^2y\right)\right)^2\right) - 1 \\
&= xy - \frac{1}{2}x^2y + o\left(x^2y\right) + \frac{1}{2!}\left(xy - \frac{1}{2}x^2y + o\left(x^2y\right)\right)^2 + o\left(\left(xy - \frac{1}{2}x^2y + o\left(x^2y\right)\right)^2\right) \\
&= xy + o(x^2+y^2)
\end{aligned}
$$

and so

$$\frac{f(x,y)}{x^2+y^2} = \frac{xy+o(x^2+y^2)}{x^2+y^2} = \frac{xy}{x^2+y^2} + \frac{o(x^2+y^2)}{x^2+y^2} = \frac{xy}{x^2+y^2} + o(1).$$

Taking $x = y$ *we have*

$$\frac{f(x,x)}{x^2+x^2} = \frac{x^2}{x^2+x^2} + o(1) \to \frac{1}{2} + 0 \quad \textit{as } x \to 0.$$

while taking $x = 0$ *we have*

$$\frac{f(0,y)}{0+y^2} = \frac{0}{x^2+y^2} + o(1) \to 0 \quad \textit{as } y \to 0.$$

Hence, the limit does not exist.

Third Method: (This method in general does not work for $m \geq 2$***)***
If either $x = 0$ *or* $y = 0$*, we get*

$$\frac{(1+x)^y - 1}{x^2+y^2} = \frac{0}{x^2+y^2} = 0.$$

If $x \neq 0$ *and* $y \neq 0$*, then*

$$\frac{(1+x)^y - 1}{x^2+y^2} = \frac{e^{y\log(1+x)} - 1}{y\log(1+x)} \frac{\log(1+x)}{x} \frac{xy}{x^2+y^2}.$$

Now, using the limits $\lim_{t\to 0}\frac{e^t-1}{t} = 1$ *and* $\lim_{x\to 0}\frac{\log(1+x)}{x} = 1$*, we have*

$$\lim_{(x,y)\to(0,0)} \frac{e^{y\log(1+x)} - 1}{y\log(1+x)} = 1, \quad \lim_{x\to 0}\frac{\log(1+x)}{x} = 1,$$

while if $g(x,y) = \frac{xy}{x^2+y^2}$ *and we take* $x = y$ *we get that* $g(x,x) = \frac{x^2}{x^2+x^2} = \frac{1}{2}$ *and so*

$$\frac{f(x,x)}{x^2+x^2} \to 1 \times 1 \times \frac{1}{2}$$

as $x \to 0$*. Hence, the limit does not exist.*

Important Taylor's formulas with center $x = 0$

- $e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \cdots + \frac{1}{n!}x^n + o\left(x^n\right)$ hence the first order formula is
$$e^x = 1 + x + o\left(x\right)$$
while the second order formula is
$$e^x = 1 + x + \frac{1}{2!}x^2 + o\left(x^2\right)$$

- $\log\left(1 + x\right) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \cdots + \left(-1\right)^{n+1}\frac{1}{n}x^n + o\left(x^n\right)$ hence the first order formula is
$$\log\left(1 + x\right) = x + o\left(x\right)$$
while the second order formula is
$$\log\left(1 + x\right) = x - \frac{1}{2}x^2 + o\left(x^2\right)$$

- $\left(1 + x\right)^a = 1 + ax + \frac{1}{2}a\left(a - 1\right)x^2 + \frac{1}{6}a\left(a - 1\right)\left(a - 2\right)x^3 + \frac{1}{4!}a\left(a - 1\right)\left(a - 2\right)\left(a - 3\right)x^4 + \cdots + \frac{1}{n!}a\left(a - 1\right)\left(a - 2\right)\left(a - 3\right)\cdots\left(a - n + 1\right)x^n + o\left(x^n\right)$ hence the first order formula is
$$\left(1 + x\right)^a = 1 + ax + o\left(x\right)$$
while the second order formula is
$$\left(1 + x\right)^a = 1 + ax + \frac{1}{2}a\left(a - 1\right)x^2 + o\left(x^2\right)$$

- $\dfrac{1}{1 + x} = \left(1 + x\right)^{-1} = 1 - x + x^2 - x^3 + x^4 + \cdots + \left(-1\right)^{n+1}x^n + o\left(x^n\right)$ hence the first order formula is
$$\frac{1}{1 + x} = \left(1 + x\right)^{-1} = 1 - x + o\left(x\right)$$
while the second order formula is
$$\frac{1}{1 + x} = \left(1 + x\right)^{-1} = 1 - x + x^2 + o\left(x^2\right)$$

- $\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \frac{1}{8!}x^8 \cdots + \left(-1\right)^k\frac{1}{2k!}x^{2k} + o\left(x^{2k+1}\right)$ hence the third order formula is
$$\cos x = 1 - \frac{1}{2!}x^2 + o\left(x^3\right)$$
while the fifth order formula is
$$\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + o\left(x^5\right)$$

- $\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \frac{1}{9!}x^9 \cdots + (-1)^k \frac{1}{(2k+1)!}x^{2k+1} + o\left(x^{2k+2}\right)$
  hence the second order formula is

$$\sin x = x + o\left(x^2\right)$$

  while the fourth order formula is

$$\sin x = x - \frac{1}{3!}x^3 + o\left(x^4\right)$$

- $\arctan x = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7 + \frac{1}{9}x^9 \cdots + (-1)^k \frac{1}{(2k+1)}x^{2k+1} + o\left(x^{2k+2}\right)$
  hence the second order formula is

$$\arctan x = x + o\left(x^2\right)$$

  while the fourth order formula is

$$\arctan x = x - \frac{1}{3}x^3 + o\left(x^4\right).$$

Example 231 *Let's calculate*

$$\lim_{(x,y)\to(0,0)} \frac{\log\left(1+\sin^2(xy)\right) - x^2y^2}{\left(x^2+y^2\right)^2}$$

Taylor's formula of $\sin t$ *of order one is given by*

$$\sin t = t + o\left(t^2\right)$$

and so

$$\begin{aligned}\sin^2 t &= \left(t + o\left(t^2\right)\right)^2 = t^2 + \left(o\left(t^2\right)\right)^2 + 2to\left(t^2\right)\\ &= t^2 + o\left(t^3\right)\end{aligned}$$

where we have used the properties of the little o. Hence

$$\log\left(1+\sin^2 t\right) = \log\left(1 + t^2 + o\left(t^3\right)\right),$$

Let's use now Taylor's formula

$$\log(1+s) = s + o(s),$$

where for us $s = \sin^2 t = t^2 + o\left(t^3\right)$*. We get*

$$\begin{aligned}&\log\left(1+\sin^2 t\right) = \log\left(1 + t^2 + o\left(t^3\right)\right)\\ &= \left(t^2 + o\left(t^3\right)\right) + o\left(t^2 + o\left(t^3\right)\right) = t^2 + o\left(t^2\right).\end{aligned}$$

Hence,

$$\frac{\log\left(1+\sin^2(xy)\right)-x^2y^2}{\left(x^2+y^2\right)^2} = \frac{x^2y^2+o\left(x^2y^2\right)-x^2y^2}{\left(x^2+y^2\right)^2}$$
$$= \frac{o\left(x^2y^2\right)}{\left(x^2+y^2\right)^2} = \frac{x^2y^2}{\left(x^2+y^2\right)^2}\frac{o\left(x^2y^2\right)}{x^2y^2}$$

if $x \neq 0$ and $y \neq 0$. Now

$$0 \leq \frac{x^2y^2}{\left(x^2+y^2\right)^2} \leq \frac{1}{2},$$

and so

$$\frac{x^2y^2}{\left(x^2+y^2\right)^2}\frac{o\left(x^2y^2\right)}{x^2y^2} \to 0$$

by Theorem while if either $x = 0$ or $y = 0$, we get

$$\frac{\log\left(1+\sin^2(xy)\right)-x^2y^2}{\left(x^2+y^2\right)^2} = \frac{0}{\left(x^2+y^2\right)^2} = 0.$$

Hence,

$$\lim_{(x,y)\to(0,0)} \frac{\log\left(1+\sin^2(xy)\right)-x^2y^2}{\left(x^2+y^2\right)^2} = 0$$

Exercise 232 *Calculate the limit*

$$\lim_{(x,y)\to(0,0)} \frac{\log\left(1+\sin^2(xy)\right)-x^2y^2}{\left(x^2+y^2\right)^4}.$$

Monday, April 11, 2022

14 Local Minima and Maxima

We recall that

Definition 233 *Let (X,d) be a metric space, let $E \subseteq X$, let $f : E \to \mathbb{R}$, and let $x_0 \in E$. We say that*

(i) f attains a local minimum *at x_0 if there exists $r > 0$ such that $f(x) \geq f(x_0)$ for all $x \in E \cap B(x_0, r)$,*

(ii) f attains a global minimum *at x_0 if $f(x) \geq f(x_0)$ for all $x \in E$,*

(iii) f attains a local maximum *at x_0 if there exists $r > 0$ such that $f(x) \leq f(x_0)$ for all $x \in E \cap B(x_0, r)$,*

(iv) f attains a global maximum *at x_0 if $f(x) \leq f(x_0)$ for all $x \in E$.*

Theorem 234 *Let* $(X, \|\cdot\|)$ *be a normed space, let* $E \subseteq X$*, let* $f : E \to \mathbb{R}$*, and let* $x_0 \in E$*. Assume that* f *attains a local minimum (or maximum) at* x_0 *and that there exists the directional derivative* $\frac{\partial f}{\partial v}(x_0)$*. If* x_0 *is an accumulation point for both sets* $E \cap \{x_0 + tv : \ t > 0\}$ *and* $E \cap \{x_0 + tv : \ t < 0\}$*, then necessarily,* $\frac{\partial f}{\partial v}(x_0) = 0$*. In particular, if* x_0 *is an interior point of* E *and* f *is differentiable at* x_0*, then all the directional derivatives of* f *at* x_0 *are zero.*

Proof. Assume that f attains a local minimum (the case of a local maximum is similar). Then there exists $r > 0$ such that $f(x) \geq f(x_0)$ for all $x \in E \cap B(x_0, r)$. Take $x = x_0 + tv$, where $|t| < r / \|v\|$. Then

$$\|x_0 + tv - x_0\| = \|tv\| = |t| \, \|v\| < r,$$

and so $f(x_0 + tv) \geq f(x_0)$. If $t > 0$, then

$$\frac{f(x_0 + tv) - f(x_0)}{t} \geq 0.$$

Since x_0 is an accumulation point for the set $E \cap \{x_0 + tv : \ t > 0\}$, there are infinitely many $t > 0$ approaching zero. Hence, letting $t \to 0^+$ and using the fact that there exists $\frac{\partial f}{\partial v}(x_0)$, we get that $\frac{\partial f}{\partial v}(x_0) \geq 0$.

If $t < 0$, then $f(x_0 + tv) \geq f(x_0)$ and

$$\frac{f(x_0 + tv) - f(x_0)}{t} \leq 0.$$

Since x_0 is an accumulation point for the set $E \cap \{x_0 + tv : \ t < 0\}$, there are infinitely many $t < 0$ approaching zero. Hence, letting $t \to 0^-$ and using the fact that there exists $\frac{\partial f}{\partial v}(x_0)$, we get that $\frac{\partial f}{\partial v}(x_0) \leq 0$.

This shows that $\frac{\partial f}{\partial v}(x_0) = 0$. ∎

Remark 235 *If if* x_0 *is a point of local minimum and* $\frac{\partial f}{\partial v}(x_0)$ *exists, then* x_0 *is an accumulation point for the set* $E \cap \{x_0 + tv : \ t \in \mathbb{R}\}$*, so* x_0 *is an accumulation point for* $E \cap \{x_0 + tv : \ t > 0\}$*, in which case* $\frac{\partial f}{\partial v}(x_0) \geq 0$*, or* x_0 *is an accumulation point for* $E \cap \{x_0 + tv : \ t < 0\}$*, in which case* $\frac{\partial f}{\partial v}(x_0) \leq 0$*.*

Remark 236 *In view of Theorem 234, when looking for local minima and maxima, we have to search among the following:*

- *Interior points at which* f *is differentiable and* $\nabla f(\boldsymbol{x}) = \mathbf{0}$*, these are called* critical points*. Note that if* $\nabla f(\boldsymbol{x}_0) = \mathbf{0}$*, the function* f *may not attain a local minimum or maximum at* $\boldsymbol{x}_0$*. Indeed, consider the function* $f(x) = x^3$*. Then* $f'(0) = 0$*, but* f *is strictly increasing, and so* f *does not attain a local minimum or maximum at* 0*.*

- *Interior points at which* f *is not differentiable. The function* $f(x) = |x|$ *attains a global minimum at* $x = 0$*, but* f *is not differentiable at* $x = 0$*.*

- *Boundary points.*

To find sufficient conditions for a critical point to be a point of local minimum or local maximum, we study the second order derivatives of f.

Definition 237 *Let $f : E \to \mathbb{R}$, where $E \subseteq \mathbb{R}^N$, and let $\boldsymbol{x}_0 \in E$. The Hessian matrix of f at $\boldsymbol{x}_0$ is the $N \times N$ matrix*

$$\begin{aligned} H_f(\boldsymbol{x}_0) &:= \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_N \partial x_1}(\boldsymbol{x}_0) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_N}(\boldsymbol{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_N^2}(\boldsymbol{x}_0) \end{pmatrix} \\ &= \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}_0) \right)_{i,j=1}^{N}, \end{aligned}$$

whenever it is defined.

Remark 238 *If the hypotheses of Schwartz's theorem are satisfied for all $i, j = 1, \ldots, N$, then*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}_0) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}_0),$$

which means that the Hessian matrix $H_f(\boldsymbol{x}_0)$ is symmetric.

Given an $N \times N$ matrix H, the *characteristic polynomial* of H is the polynomial

$$p(t) := \det(tI_N - H), \quad t \in \mathbb{R}.$$

Theorem 239 *Let H be an $N \times N$ matrix. If H is symmetric, then all roots of the characteristic polynomial are real.*

Theorem 240 *Given a polynomial of the form*

$$p(t) = t^N + a_{N-1}t^{N-1} + a_{N-2}t^{N-2} + \cdots + a_1 t + a_0, \quad t \in \mathbb{R},$$

where the coefficients a_i are real for every $i = 0, \ldots, N-1$, assume that all roots of p are real. Then

(i) all roots of p are positive if and only if the coefficients alternate sign, that is, $a_{N-1} < 0$, $a_{N-2} > 0$, $a_{N-3} < 0$, etc.

(ii) all roots of p are negative if and only $a_i > 0$ for every $i = 0, \ldots, N-1$.

Remark 241 *Given a vector $\boldsymbol{x} = (x_1, \ldots, x_N) \in \mathbb{R}^N$ we can identify $\boldsymbol{x}$ with the $1 \times N$ matrix $\begin{pmatrix} x_1 & \cdots & x_N \end{pmatrix}$. In turn, its transpose $\boldsymbol{x}^T$ becomes the $N \times 1$ matrix*

$$\begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}.$$

Hence, given an $N \times N$ matrix

$$H = \begin{pmatrix} h_{1,1} & \cdots & h_{1,N} \\ \vdots & & \\ h_{N,1} & & h_{N,N} \end{pmatrix},$$

we have that

$$H\boldsymbol{x}^T = \begin{pmatrix} h_{1,1} & \cdots & h_{1,N} \\ \vdots & & \\ h_{N,1} & & h_{N,N} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} h_{1,1}x_1 + \cdots + h_{1,N}x_N \\ \vdots \\ h_{N,1}x_1 + \cdots + h_{N,N}x_N \end{pmatrix}.$$

In turn, $\boldsymbol{x}H\boldsymbol{x}^T$ becomes

$$\begin{aligned} \boldsymbol{x}H\boldsymbol{x}^T &= \begin{pmatrix} h_{1,1}x_1 + \cdots + h_{1,N}x_N \\ \vdots \\ h_{N,1}x_1 + \cdots + h_{N,N}x_N \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} \\ &= \sum_{j=1}^{N} \sum_{i=1}^{N} h_{i,j} x_j x_i. \end{aligned} \tag{28}$$

If H is symmetric, then its eigenvalues $\lambda_1, \ldots, \lambda_N$ are real. Moreover, we can find corresponding eigenvectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_1$ which forms an orthonormal basis. Since

$$H\boldsymbol{v}_i^T = \lambda_i \boldsymbol{v}_i^T$$

for every $i = 1, \ldots, N$, we get

$$\boldsymbol{v}_i H \boldsymbol{v}_i^T = \lambda_i \boldsymbol{v}_i \cdot \boldsymbol{v}_i = \lambda_i \left\| \boldsymbol{v}_i \right\|^2 = \lambda_i 1. \tag{29}$$

Since $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_1\}$ forms a basis in $\mathbb{R}^N$, we can write

$$\boldsymbol{x} = c_1 \boldsymbol{v}_1 + \cdots + c_N \boldsymbol{v}_N.$$

Then

$$\begin{aligned} \boldsymbol{x}H\boldsymbol{x}^T &= \left(\sum_{j=1}^{N} c_j \boldsymbol{v}_j \right) H \left(\sum_{i=1}^{N} c_i \boldsymbol{v}_i^T \right) \\ &= \left(\sum_{j=1}^{N} c_j \boldsymbol{v}_j \right) \left(\sum_{i=1}^{N} c_i H \boldsymbol{v}_i^T \right) = \sum_{j=1}^{N} \sum_{i=1}^{N} \lambda_j c_i c_j \boldsymbol{v}_i \cdot \boldsymbol{v}_j \\ &= \sum_{j=1}^{N} \lambda_j c_j^2, \end{aligned}$$

where we used the fact that $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 1$ if $i = j$ and 0 otherwise. In particular, if we let $m := \min\{\lambda_1, \ldots, \lambda_N\}$, we have that

$$\boldsymbol{x}H\boldsymbol{x}^T = \sum_{j=1}^{N} \lambda_j c_j^2 \geq m \sum_{j=1}^{N} c_j^2 = m \left\| \boldsymbol{x} \right\|^2.$$

Definition 242 *Given an $N \times N$ matrix H, we say that*

(i) H *is* positive definite *if* $\boldsymbol{x}H\boldsymbol{x}^T > 0$ *for all* $\boldsymbol{x} \in \mathbb{R}^N \setminus \{\mathbf{0}\}$,

(ii) H *is* positive semidefinite *if* $\boldsymbol{x}H\boldsymbol{x}^T \geq 0$ *for all* $\boldsymbol{x} \in \mathbb{R}^N$,

(iii) H *is* negative definite *if* $\boldsymbol{x}H\boldsymbol{x}^T < 0$ *for all* $\boldsymbol{x} \in \mathbb{R}^N \setminus \{\mathbf{0}\}$,

(iii) H *is* negative semidefinite *if* $\boldsymbol{x}H\boldsymbol{x}^T \leq 0$ *for all* $\boldsymbol{x} \in \mathbb{R}^N \setminus \{\mathbf{0}\}$.

Exercise 243 *Let H be an $N \times N$ symmetric matrix.*

(i) Prove that H is positive definite if and only if all its eigenvalues are positive.

(ii) Prove that H is positive semidefinite if and only if all its eigenvalues are nonnegative.

(iii) Prove that H is negative definite if and only if all its eigenvalues are negative.

(iv) Prove that H is negative semidefinite if and only if all its eigenvalues are nonpositive.

The next theorem gives necessary and sufficient conditions for a point to be of local minimum or maximum.

Theorem 244 *Let $U \subseteq \mathbb{R}^N$ be open, let $f : U \to \mathbb{R}$ be of class $C^2(U)$ and let $\boldsymbol{x}_0 \in U$ be a critical point of f.*

(i) If $H_f(\boldsymbol{x}_0)$ is positive definite, then f attains a local minimum at $\boldsymbol{x}_0$,

(ii) if f attains a local minimum at $\boldsymbol{x}_0$, then $H_f(\boldsymbol{x}_0)$ is positive semidefinite,

(iii) if $H_f(\boldsymbol{x}_0)$ is negative definite, then f attains a local maximum at $\boldsymbol{x}_0$,

(iv) if f attains a local maximum at $\boldsymbol{x}_0$, then $H_f(\boldsymbol{x}_0)$ is negative semidefinite.

Proof. (i) Assume that $H_f(\boldsymbol{x}_0)$ is positive definite. Then by Remark 241,

$$\sum_{j=1}^{N}\sum_{i=1}^{N}\frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}_0)\, y_i y_j = \boldsymbol{y}H_f(\boldsymbol{x}_0)\,\boldsymbol{y}^Y \geq m\,\|\boldsymbol{y}\|^2 \tag{30}$$

for all $\boldsymbol{y} \in \mathbb{R}^N$ and for some $m > 0$.

We now apply Taylor's formula of order two to obtain

$$\begin{aligned} f(\boldsymbol{x}) &= f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)\cdot(\boldsymbol{x}-\boldsymbol{x}_0) + \sum_{|\boldsymbol{\alpha}|=2}\frac{1}{\boldsymbol{\alpha}!}\frac{\partial^{\boldsymbol{\alpha}} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x}_0)\cdot(\boldsymbol{x}-\boldsymbol{x}_0)^{\boldsymbol{\alpha}} + R_2(\boldsymbol{x}) \\ &= f(\boldsymbol{x}_0) + 0 + \frac{1}{2}\sum_{j=1}^{N}\sum_{i=1}^{N}\frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)_i(\boldsymbol{x}-\boldsymbol{x}_0)_j + R_2(\boldsymbol{x}) \\ &= f(\boldsymbol{x}_0) + 0 + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_0)H_f(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^Y + R_2(\boldsymbol{x}) \end{aligned}$$

where we have used the fact that $\boldsymbol{x}_0$ is a critical point and where

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \frac{R_2(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{x}_0\|^2} = 0.$$

■

Wednesday, April 12, 2022

Proof. Using the definition of limit with $\varepsilon = \frac{m}{2}$, we can find $\delta > 0$ such that

$$\left| \frac{R_2(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{x}_0\|^2} \right| \leq \frac{m}{2}$$

for all $\boldsymbol{x} \in E$ with $\|\boldsymbol{x}-\boldsymbol{x}_0\| \leq \delta$. Using this property and (30), we get

$$\begin{aligned} f(\boldsymbol{x}) &\geq f(\boldsymbol{x}_0) + m\|\boldsymbol{x}-\boldsymbol{x}_0\|^2 + R_2(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \|\boldsymbol{x}-\boldsymbol{x}_0\|^2 \left(m + \frac{R_2(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{x}_0\|^2} \right) \\ &\geq f(\boldsymbol{x}_0) + \|\boldsymbol{x}-\boldsymbol{x}_0\|^2 \left(m - \frac{m}{2} \right) = f(\boldsymbol{x}_0) + \|\boldsymbol{x}-\boldsymbol{x}_0\|^2 \frac{m}{2} > f(\boldsymbol{x}_0) \end{aligned}$$

for all $\boldsymbol{x} \in E$ with $0 < \|\boldsymbol{x}-\boldsymbol{x}_0\| \leq \delta$. This shows that f attains a (strict) local minimum at $\boldsymbol{x}_0$.

(ii) Assume that if f attains a local minimum at $\boldsymbol{x}_0$. Then there exists $B(\boldsymbol{x}_0, r) \subseteq U$ such that

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}_0)$$

for all $\boldsymbol{x} \in B(\boldsymbol{x}_0, r)$. Assume by contradiction that $H_f(\boldsymbol{x}_0)$ is not positive semidefinite. This means that there exists an eigenvalue $\lambda_i < 0$. Let $\boldsymbol{v}$ be an eigenvector of norm 1 for λ_i. Then

$$\boldsymbol{v} H_f(\boldsymbol{x}_0) \boldsymbol{v}^T = \lambda_i$$

As in the previous step

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + 0 + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_0) H_f(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^Y + R_2(\boldsymbol{x}).$$

Take $\boldsymbol{x} = \boldsymbol{x}_0 + t\boldsymbol{v}_i$, Then

$$\begin{aligned} f(\boldsymbol{x}_0 + t\boldsymbol{v}_i) &= f(\boldsymbol{x}_0) + 0 + \frac{1}{2} t^2 \lambda_i + R_2(\boldsymbol{x}_0 + t\boldsymbol{v}_i) \\ &= f(\boldsymbol{x}_0) + \frac{\lambda_i}{2} t^2 + o(t^2) \\ &= f(\boldsymbol{x}_0) + t^2 \left(\frac{\lambda_i}{2} + \frac{o(t^2)}{t^2} \right). \end{aligned}$$

Take $\varepsilon = -\frac{\lambda_i}{4} > 0$, as in the previous step we have that

$$\begin{aligned} f(\boldsymbol{x}_0 + t\boldsymbol{v}_i) &= f(\boldsymbol{x}_0) + t^2\left(\frac{\lambda_i}{2} + \frac{o(t^2)}{t^2}\right) \\ &\le f(\boldsymbol{x}_0) + t^2\left(\frac{\lambda_i}{2} - \frac{\lambda_i}{4}\right) \\ &= f(\boldsymbol{x}_0) + t^2\frac{\lambda_i}{4} < f(\boldsymbol{x}_0), \end{aligned}$$

since $\lambda_i < 0$. This contradicts the fact that f has a local minimum at $\boldsymbol{x}_0$. ■

Remark 245 *Note that in view of the previous theorem, if at a critical point $\boldsymbol{x}_0$ the characteristic polynomial of $H_f(\boldsymbol{x}_0)$ has one positive root and one negative root, then f does not admit a local minimum or a local maximum at $\boldsymbol{x}_0$.*

Example 246 *Let $f(x,y,z) := x^2 + y^4 + y^2 + z^3 - 2xz$. We have*

$$\begin{cases} \frac{\partial f}{\partial x} = 2x - 2z = 0 \\ \frac{\partial f}{\partial y} = 4y^3 + 2y = 0 \\ \frac{\partial f}{\partial z} = 3z^2 - 2x = 0 \end{cases} \iff \begin{cases} x - z = 0 \\ y\left(2y^2 + 1\right) = 0 \\ 3z^2 - 2x = 0 \end{cases} \iff \begin{cases} x - z = 0 \\ y = 0 \\ 3z^2 - 2z = 0 \end{cases}$$

$$\iff \begin{cases} x - z = 0 \\ y = 0 \\ z(3z - 2) = 0 \end{cases}$$

and so the critical points are $(0,0,0)$ and $\left(\frac{2}{3}, 0, \frac{2}{3}\right)$. Note that $(0,0,0)$ is not a point of local minimum or maximum, since $f(0,0,z) = z^3$ which changes sign near 0. Let's study the point $\left(\frac{2}{3}, 0, \frac{2}{3}\right)$. We have

$$H_f = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial z \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial z \partial y} \\ \frac{\partial^2 f}{\partial x \partial z} & \frac{\partial^2 f}{\partial y \partial z} & \frac{\partial^2 f}{\partial z^2} \end{pmatrix} = \begin{pmatrix} 2 & 0 & -2 \\ 0 & 12y^2 + 2 & 0 \\ -2 & 0 & 6z \end{pmatrix}$$

and so

$$H_f\left(\frac{2}{3}, 0, \frac{2}{3}\right) = \begin{pmatrix} 2 & 0 & -2 \\ 0 & 2 & 0 \\ -2 & 0 & 6 \end{pmatrix}.$$

We have

$$\begin{aligned} 0 &= \det\left(t\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 2 & 0 & -2 \\ 0 & 2 & 0 \\ -2 & 0 & 4 \end{pmatrix}\right) \\ &= \det\begin{pmatrix} t-2 & 0 & 2 \\ 0 & t-2 & 0 \\ 2 & 0 & t-4 \end{pmatrix} = t^3 - 8t^2 + 16t - 8. \end{aligned}$$

The eigenvalues are all positive by Theorem 240. Hence, at $\left(\frac{2}{3}, 0, \frac{2}{3}\right)$ we have a local minimum.

15 Lagrange Multipliers

In Section 14 (see Theorem 244) we have seen how to find points of local minima and maxima of a function $f : E \to \mathbb{R}$ in the interior E° of E. Now we are ready to find points of local minima and maxima of a function $f : E \to \mathbb{R}$ on the boundary ∂E of E. We assume that the boundary of E has a special form, that is, it is given by a set of the form

$$\left\{ \boldsymbol{x} \in \mathbb{R}^N : \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0} \right\}.$$

Definition 247 *Let $f : E \to \mathbb{R}$, where $E \subseteq \mathbb{R}^N$, let $F \subseteq E$ and let $\boldsymbol{x}_0 \in F$. We say that*

(i) f attains a constrained local minimum *at $\boldsymbol{x}_0$ if there exists $r > 0$ such that $f(\boldsymbol{x}) \geq f(\boldsymbol{x}_0)$ for all $\boldsymbol{x} \in F \cap B(\boldsymbol{x}_0, r)$,*

(ii) f attains a constrained local maximum *at $\boldsymbol{x}_0$ if there exists $r > 0$ such that $f(\boldsymbol{x}) \leq f(\boldsymbol{x}_0)$ for all $\boldsymbol{x} \in F \cap B(\boldsymbol{x}_0, r)$.*

The set F is called the *constraint*.

Theorem 248 (Lagrange multipliers) *Let $U \subseteq \mathbb{R}^N$ be an open set, let $f : U \to \mathbb{R}$ be a function of class C^1 and let $\boldsymbol{g} : U \to \mathbb{R}^M$ be a class of function C^1, where $M \leq N$, and let*

$$F := \{\boldsymbol{x} \in U : \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0}\}.$$

Let $\boldsymbol{x}_0 \in F$ and assume that f attains a constrained local minimum (or maximum) at $\boldsymbol{x}_0$. If the vectors $\nabla g_i(\boldsymbol{x}_0)$, $i = 1, \ldots, M$ are linearly independent, then there exist $\lambda_1, \ldots, \lambda_M \in \mathbb{R}$ such that

$$\nabla f(\boldsymbol{x}_0) = \lambda_1 \nabla g_1(\boldsymbol{x}_0) + \cdots + \lambda_M \nabla g_M(\boldsymbol{x}_0).$$

We will prove this theorem using manifolds in MS.

Example 249 *Given a point $\boldsymbol{x}_0 \in \mathbb{R}^N$, find*

$$\operatorname{dist}(\boldsymbol{x}_0, S_{N-1}),$$

where $S_{N-1} := \partial B(\boldsymbol{0}, 1)$ is the unit sphere in $\mathbb{R}^N$. Note that

$$\begin{aligned} \operatorname{dist}(\boldsymbol{x}_0, S_{N-1}) &= \inf\{\|\boldsymbol{x}_0 - \boldsymbol{x}\| : \boldsymbol{x} \in S_{N-1}\} \\ &= \inf\{\|\boldsymbol{x}_0 - \boldsymbol{x}\| : \|\boldsymbol{x}\| = 1\}. \end{aligned}$$

To simplify our life, we can square everything, so we are looking for the minimum of the function

$$f(\boldsymbol{x}) = \|\boldsymbol{x}_0 - \boldsymbol{x}\|^2$$

subject to the constraint $\|\boldsymbol{x}\|^2 = 1$. Take $g(\boldsymbol{x}) := \|\boldsymbol{x}\|^2 - 1$. Since S_{N-1} is closed and bounded and f is continuous, by the Weierstrass theorem f has a global *minimum and a global maximum in S_{N-1}. Thus, we can apply the theorem on Lagrange multipliers. We are looking for a solution of the following system*

$$\nabla f = \lambda \nabla g,$$

subject to $g = 0$, that is

$$\begin{cases} \frac{\partial f}{\partial x_i}(\boldsymbol{x}) - \lambda \frac{\partial g}{\partial x_i}(\boldsymbol{x}) = 0, \quad i = 1, \ldots, N, \\ g(\boldsymbol{x}) = 0. \end{cases}$$

We have

$$\begin{cases} 2(x_i - x_{0,i}) - 2\lambda x_i = 0, \quad i = 1, \ldots, N, \\ \|\boldsymbol{x}\|^2 = 1, \end{cases}$$

that is,

$$\begin{cases} (1-\lambda) x_i = x_{0,i}, \quad i = 1, \ldots, N, \\ \|\boldsymbol{x}\|^2 = 1, \end{cases} \quad \Leftrightarrow \quad \begin{cases} (1-\lambda)\boldsymbol{x} = \boldsymbol{x}_0, \\ \|\boldsymbol{x}\|^2 = 1. \end{cases}$$

If $\boldsymbol{x}_0 = \mathbf{0}$, then $\lambda = 1$, and so every point on the sphere is at maximum distance from $\mathbf{0}$. If $\boldsymbol{x}_0 \neq \mathbf{0}$, then $\lambda \neq 1$, and so $\boldsymbol{x} = \frac{1}{1-\lambda}\boldsymbol{x}_0$. Plugging this into $\|\boldsymbol{x}\|^2 = 1$, we get

$$1 = \frac{1}{(1-\lambda)^2} \|\boldsymbol{x}_0\|^2 \quad \Leftrightarrow \quad (1-\lambda)^2 = \|\boldsymbol{x}_0\|^2$$

which gives $\lambda = 1 \pm \|\boldsymbol{x}_0\|$, and in turn

$$\boldsymbol{x} = \pm \frac{1}{\|\boldsymbol{x}_0\|} \boldsymbol{x}_0.$$

The closest point to $\boldsymbol{x}_0$ will be $\frac{1}{\|\boldsymbol{x}_0\|}\boldsymbol{x}_0$ and the furthest $-\frac{1}{\|\boldsymbol{x}_0\|}\boldsymbol{x}_0$, as we expected.

16 Implicit and Inverse Function

Definition 250 *Given an open set $U \subseteq \mathbb{R}^N$ and a function $\boldsymbol{f} : U \to \mathbb{R}^M$, we say that $\boldsymbol{f}$ is of class C^m for some nonnegative integer $m \in \mathbb{N}_0$, if all its components f_i, $i = 1, \ldots, M$, are of class C^m. The space of all functions $\boldsymbol{f} : U \to \mathbb{R}^M$ of class C^m is denoted $C^m\left(U; \mathbb{R}^M\right)$. We set $C^\infty\left(U; \mathbb{R}^M\right) := \bigcap_{m=0}^{\infty} C^m\left(U; \mathbb{R}^M\right)$.*

Theorem 251 (Inverse Function) *Let $U \subseteq \mathbb{R}^N$ be open, let $\boldsymbol{f} : U \to \mathbb{R}^N$, and let $\boldsymbol{a} \in U$. Assume that $\boldsymbol{f} \in C^m\left(U; \mathbb{R}^N\right)$ for some $m \in \mathbb{N}$ and that*

$$\det J_{\boldsymbol{f}}(\boldsymbol{a}) \neq 0.$$

Then there exists $B(\boldsymbol{a}, r) \subseteq U$ such that $\boldsymbol{f}(B(\boldsymbol{a}, r))$ is open, the function

$$\boldsymbol{f} : B(\boldsymbol{a}, r) \to \boldsymbol{f}(B(\boldsymbol{a}, r))$$

is invertible and $\boldsymbol{f}^{-1} \in C^m\left(\boldsymbol{f}(B(\boldsymbol{a}, r)); \mathbb{R}^N\right)$. Moreover,

$$J_{\boldsymbol{f}^{-1}}(\boldsymbol{y}) = \left(J_{\boldsymbol{f}}\left(\boldsymbol{f}^{-1}(\boldsymbol{y})\right)\right)^{-1}.$$

We will prove this theorem using a fixed point theorem in MS.

The next exercise shows that differentiability is not enough for the inverse function theorem.

Exercise 252 *Consider the function* $\boldsymbol{f} : \mathbb{R}^2 \to \mathbb{R}^2$ *defined by*

$$f_1(x, y) = \begin{cases} 0 & \text{if } x = 0, \\ x + 2x^2 \sin \frac{1}{x} & \text{if } x \neq 0, \end{cases}$$
$$f_2(x, y) = y.$$

Prove that $\boldsymbol{f} = (f_1, f_2)$ *is differentiable in* $(0,0)$ *and* $J_{\boldsymbol{f}}(0,0) = 1$*. Prove that* $\boldsymbol{f}$ *is not one-to-one in any neighborhood of* $(0,0)$*.*

The next exercise shows that the existence of a local inverse at every point does not imply the existence of a global inverse.

Exercise 253 *Consider the function* $\boldsymbol{f} : \mathbb{R}^2 \to \mathbb{R}^2$ *defined by*

$$\boldsymbol{f}(x, y) = (e^x \cos y, e^x \sin y).$$

Prove that $\det J_{\boldsymbol{f}}(x, y) \neq 0$ *for all* $(x, y) \in \mathbb{R}^2$ *but that* $\boldsymbol{f}$ *is not injective.*

Given a function f of two variables $(x, y) \in \mathbb{R}^2$, consider the equation

$$f(x, y) = 0.$$

We want to solve for y, that is, we are interested in finding a function $y = g(x)$ such that

$$f(x, g(x)) = 0.$$

We will see under which conditions we can do this. The result is going to be local.

In what follows given $\boldsymbol{x} \in \mathbb{R}^N$ and $\boldsymbol{y} \in \mathbb{R}^M$ and $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y})$, we write

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}(\boldsymbol{x}, \boldsymbol{y}) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}, \boldsymbol{y}) & \cdots & \frac{\partial f_1}{\partial x_N}(\boldsymbol{x}, \boldsymbol{y}) \\ \vdots & & \vdots \\ \frac{\partial f_M}{\partial x_1}(\boldsymbol{x}, \boldsymbol{y}) & \cdots & \frac{\partial f_M}{\partial x_N}(\boldsymbol{x}, \boldsymbol{y}) \end{pmatrix}$$

and

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}}(\boldsymbol{x}, \boldsymbol{y}) := \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(\boldsymbol{x}, \boldsymbol{y}) & \cdots & \frac{\partial f_1}{\partial y_M}(\boldsymbol{x}, \boldsymbol{y}) \\ \vdots & & \vdots \\ \frac{\partial f_M}{\partial y_1}(\boldsymbol{x}, \boldsymbol{y}) & \cdots & \frac{\partial f_M}{\partial y_M}(\boldsymbol{x}, \boldsymbol{y}) \end{pmatrix}.$$

Theorem 254 (Implicit Function) *Let* $U \subseteq \mathbb{R}^N \times \mathbb{R}^M$ *be open, let* $\boldsymbol{f} : U \to \mathbb{R}^M$*, and let* $(\boldsymbol{a}, \boldsymbol{b}) \in U$*. Assume that* $\boldsymbol{f} \in C^m(U; \mathbb{R}^M)$ *for some* $m \in \mathbb{N}$*, that*

$$\boldsymbol{f}(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{0} \quad \text{and} \quad \det \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b}) \neq 0.$$

Then there exist $B_N(\boldsymbol{a}, r_0) \subset \mathbb{R}^N$ and $B_M(\boldsymbol{b}, r_1) \subset \mathbb{R}^M$, with $B_N(\boldsymbol{a}, r_0) \times B_M(\boldsymbol{b}, r_1) \subseteq U$ and a unique function

$$\boldsymbol{g} : B_N(\boldsymbol{a}, r_0) \to B_M(\boldsymbol{b}, r_1)$$

such that $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{g}(\boldsymbol{x})) = \boldsymbol{0}$ for all $\boldsymbol{x} \in B_N(\boldsymbol{a}, r_0)$. Moreover, $\boldsymbol{g}$ is of class C^m and $\boldsymbol{g}(\boldsymbol{a}) = \boldsymbol{b}$.

Remark 255 *When we say "unique function" we mean that for every $\boldsymbol{x} \in B_N(\boldsymbol{a}, r_0)$ there exists a unique $\boldsymbol{y}_{\boldsymbol{x}} \in B_M(\boldsymbol{b}, r_1)$ (depending on $\boldsymbol{x}$) such that $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}_{\boldsymbol{x}}) = \boldsymbol{0}$. The function $\boldsymbol{g}$ is defined by $\boldsymbol{g}(\boldsymbol{x}) := \boldsymbol{y}_{\boldsymbol{x}}$. Hence, we are saying that in the set $B_N(\boldsymbol{a}, r_0) \times B_M(\boldsymbol{b}, r_1)$ the only solutions to the equation*

$$\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0}$$

are given by $(\boldsymbol{x}, \boldsymbol{g}(\boldsymbol{x}))$, $\boldsymbol{x} \in B_N(\boldsymbol{a}, r_0)$.

Proof. We apply the inverse function theorem to the function $\boldsymbol{h} : U \to \mathbb{R}^N \times \mathbb{R}^M$ defined by

$$\boldsymbol{h}(\boldsymbol{x}, \boldsymbol{y}) := (\boldsymbol{x}, \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y})).$$

We have

$$\det J\boldsymbol{h}(\boldsymbol{a}, \boldsymbol{b}) = \det \begin{pmatrix} I_{N\times N} & 0_{N\times M} \\ \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}(\boldsymbol{a}, \boldsymbol{b}) & \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b}) \end{pmatrix} = \det \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b}) \neq 0.$$

Hence, by the inverse function theorem there exists $B((\boldsymbol{a}, \boldsymbol{b}), r) \subseteq U$ such that $\boldsymbol{h}(B((\boldsymbol{a}, \boldsymbol{b}), r))$ is open, the function

$$\boldsymbol{h} : B((\boldsymbol{a}, \boldsymbol{b}), r) \to \boldsymbol{h}(B((\boldsymbol{a}, \boldsymbol{b}), r))$$

is invertible and $\boldsymbol{h}^{-1} \in C^m\left(\boldsymbol{h}(B((\boldsymbol{a}, \boldsymbol{b}), r)); \mathbb{R}^N\right)$. Note that $\boldsymbol{h}(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{a}, \boldsymbol{0}) \in \boldsymbol{h}(B((\boldsymbol{a}, \boldsymbol{b}), r))$. Since $\boldsymbol{h}(B((\boldsymbol{a}, \boldsymbol{b}), r))$ is open, we can find $r_0 > 0$ and $r_1 > 0$ such that $B(\boldsymbol{a}, r_0) \times B(\boldsymbol{0}, r_1) \subseteq \boldsymbol{h}(B((\boldsymbol{a}, \boldsymbol{b}), r))$.

Write $\boldsymbol{h}^{-1}(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{p}(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{q}(\boldsymbol{x}, \boldsymbol{y})) \in \mathbb{R}^N \times \mathbb{R}^M$. Then for $(\boldsymbol{x}, \boldsymbol{y}) \in B(\boldsymbol{a}, r_0) \times B(\boldsymbol{0}, r_1)$,

$$\begin{aligned}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{h}(\boldsymbol{h}^{-1}(\boldsymbol{x}, \boldsymbol{y})) &= \boldsymbol{h}((\boldsymbol{p}(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{q}(\boldsymbol{x}, \boldsymbol{y})) \\ &= (\boldsymbol{p}(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{f}(\boldsymbol{p}(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{q}(\boldsymbol{x}, \boldsymbol{y}))),\end{aligned}$$

so

$$\begin{aligned}\boldsymbol{x} &= \boldsymbol{p}(\boldsymbol{x}, \boldsymbol{y}), \\ \boldsymbol{y} &= \boldsymbol{f}(\boldsymbol{p}(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{q}(\boldsymbol{x}, \boldsymbol{y})).\end{aligned}$$

Substituting the first identity into the second, we obtain

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{q}(\boldsymbol{x}, \boldsymbol{y}))$$

for $(\boldsymbol{x}, \boldsymbol{y}) \in B(\boldsymbol{a}, r_0) \times B(\boldsymbol{0}, r_1)$. In particular, taking $\boldsymbol{y} = \boldsymbol{0}$ gives

$$\boldsymbol{0} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{q}(\boldsymbol{x}, \boldsymbol{0}))$$

for $\boldsymbol{x} \in B(\boldsymbol{a}, r_0)$. So we can define $\boldsymbol{g}(\boldsymbol{x}) := \boldsymbol{q}(\boldsymbol{x}, \boldsymbol{0})$ for $\boldsymbol{x} \in B(\boldsymbol{a}, r_0)$. ■

The next examples show that when $\det \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{0}$, then anything can happen.

Example 256 *In all these examples $N = M = 1$ and $\frac{\partial f}{\partial y}(x_0, y_0) = 0$.*

(i) Consider the function

$$f(x, y) := (y - x)^2.$$

Then $f(0,0) = 0$, $\frac{\partial f}{\partial y}(0,0) = 0$ and $g(x) = x$ satisfies $f(x, g(x)) = 0$.

(ii) Consider the function

$$f(x, y) := x^2 + y^2.$$

Then $f(0,0) = 0$, $\frac{\partial f}{\partial y}(0,0) = 0$ but there is no function g defined near $x = 0$ such that $f(x, g(x)) = 0$.

(iii) Consider the function

$$f(x, y) := (xy - 1)\left(x^2 + y^2\right).$$

Then $f(0,0) = 0$, $\frac{\partial f}{\partial y}(0,0) = 0$ but

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

which is discontinuous.

Friday, April 22, 2022

Next we give an example on how to apply the implicit function theorem.

Example 257 *Consider the function*

$$\boldsymbol{f}(x, y, z) = \left(y \cos(xz) - x^2 + 1, y \sin(xz) - x\right).$$

Let's prove that there exist $r > 0$ and $\boldsymbol{g} : (1 - r, 1 + r) \to \mathbb{R}^2$ of class C^∞ such that $\boldsymbol{g}(1) = \left(1, \frac{\pi}{2}\right)$ and $\boldsymbol{f}(x, \boldsymbol{g}(x)) = 0$. Note that $\boldsymbol{f}$ is of class C^∞. Here the point is $\left(1, 1, \frac{\pi}{2}\right)$ and

$$\boldsymbol{f}\left(1, 1, \frac{\pi}{2}\right) = \left(1 \cos\left(1\frac{\pi}{2}\right) - 1 + 1, 1 \sin\left(1\frac{\pi}{2}\right) - 1\right) = (0, 0).$$

Moreover,

$$\begin{aligned} \frac{\partial \boldsymbol{f}}{\partial (y, z)}(x, y, z) &= \begin{pmatrix} \frac{\partial f_1}{\partial y}(x, y, z) & \frac{\partial f_1}{\partial z}(x, y, z) \\ \frac{\partial f_2}{\partial y}(x, y, z) & \frac{\partial f_1}{\partial z}(x, y, z) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial}{\partial y}\left(y \cos(xz) - x^2 + 1\right) & \frac{\partial}{\partial z}\left(y \cos(xz) - x^2 + 1\right) \\ \frac{\partial}{\partial y}\left(y \sin(xz) - x\right) & \frac{\partial}{\partial z}\left(y \sin(xz) - x\right) \end{pmatrix} \\ &= \begin{pmatrix} 1 \cos(xz) - 0 - 0 & -xy \sin(xz) - 0 - 0 \\ 1 \sin(xz) - 0 & xy \cos(xz) - 0 \end{pmatrix} \end{aligned}$$

and so

$$\det \frac{\partial \boldsymbol{f}}{\partial (y,z)} \left(1, 1, \frac{\pi}{2}\right) = \det \begin{pmatrix} \cos\left(1\frac{\pi}{2}\right) & -1\sin\left(1\frac{\pi}{2}\right) \\ \sin\left(1\frac{\pi}{2}\right) & 1\cos\left(1\frac{\pi}{2}\right) \end{pmatrix} = 1 \neq 0.$$

Hence, by the implicit function theorem there exist $r > 0$ and $\boldsymbol{g} : (1-r, 1+r) \to \mathbb{R}^2$ of class C^∞ such that $\boldsymbol{g}(1) = \left(1, \frac{\pi}{2}\right)$ and $\boldsymbol{f}(x, \boldsymbol{g}(x)) = 0$ for all $x \in (1-r, 1+r)$, that is,

$$\begin{cases} g_1(x)\cos(xg_2(x)) - x^2 + 1 = 0, \\ g_1(x)\sin(xg_2(x)) - x = 0. \end{cases}$$

Reasoning as before, we can use Taylor's formula to find the behavior of g_1 and g_2 near $x = 1$, that is,

$$\begin{aligned} g_1(x) &= g_1(1) + g_1'(1)(x-1) + o((x-1)), \\ g_2(x) &= g_2(1) + g_2'(1)(x-1) + o((x-1)). \end{aligned}$$

Let's differentiate the two equations. We get

$$\begin{cases} g_1'(x)\cos(xg_2(x)) - g_1(x)(1g_2(x) + xg_2'(x))\sin(xg_2(x)) - 2x + 0 = 0, \\ g_1'(x)\sin(xg_2(x)) + g_1(x)(1g_2(x) + xg_2'(x))\cos(xg_2(x)) - 1 = 0. \end{cases}$$

Taking $x = 1$ and using the fact that $g_1(1) = 1$ and $g_2(1) = \frac{\pi}{2}$, we obtain

$$\begin{cases} g_1'(1)\cos\left(1\frac{\pi}{2}\right) - 1\left(\frac{\pi}{2} + 1g_2'(1)\right)\sin\left(1\frac{\pi}{2}\right) - 2 = 0, \\ g_1'(1)\sin\left(1\frac{\pi}{2}\right) + 1\left(\frac{\pi}{2} + 1g_2'(1)\right)\cos\left(1\frac{\pi}{2}\right) - 1 = 0, \end{cases}$$

that is,

$$\begin{cases} 0 - 1\left(\frac{\pi}{2} + g_2'(1)\right)1 - 2 = 0, \\ g_1'(1)\,1 + 0 - 1 = 0, \end{cases}$$

and so $g_1'(1) = 1$ and $g_2'(1) = -2 - \frac{\pi}{2}$. Hence,

$$\begin{aligned} g_1(x) &= 1 + 1(x-1) + o((x-1)), \\ g_2(x) &= \frac{\pi}{2} + \left(-2 - \frac{\pi}{2}\right)(x-1) + o((x-1)). \end{aligned}$$

17 Lebesgue Measure

Given a bounded interval $I \subseteq \mathbb{R}$, the *length* of I is defined as

$$\operatorname{length} I := \sup I - \inf I.$$

Given N bounded intervals $I_1, \ldots, I_N \subset \mathbb{R}$, a *rectangle* in $\mathbb{R}^N$ is a set of the form

$$R := I_1 \times \cdots \times I_N.$$

The *elementary measure* of a rectangle R as

$$\operatorname{meas} R := \operatorname{length} I_1 \cdot \cdots \cdot \operatorname{length} I_N.$$

Given a set $E \subseteq \mathbb{R}^N$, we recall that the *Lebesgue outer measure of* E is defined by

$$\mathcal{L}_o^N(E) := \inf \left\{ \sum_{i=1}^{\infty} \operatorname{meas} R_i : \ R_i \text{ rectangles}, \ \bigcup_{i=1}^{\infty} R_i \supseteq E \right\}. \tag{31}$$

Proposition 258 *Let $R \subset \mathbb{R}^N$ be a rectangle. Then $\mathcal{L}_o^N(R) = \operatorname{meas} R$.*

Monday, April 25, 2022

Exercise 259 *Let $R \subset \mathbb{R}^N$ be a rectangle and assume that*

$$R = \bigcup_{i=1}^{n} R_i,$$

with R_i pairwise disjoint rectangles. Prove that

$$\operatorname{meas} R = \sum_{i=1}^{n} \operatorname{meas} R_i.$$

Exercise 260 *Let $R \subset \mathbb{R}^N$ be a rectangle and assume that*

$$R \subseteq \bigcup_{i=1}^{n} R_i,$$

with R_i rectangles (not necessarily disjoint). Prove that

$$\operatorname{meas} R \leq \sum_{i=1}^{n} \operatorname{meas} R_i.$$

Exercise 261 *Let $R \subset \mathbb{R}^N$ be a rectangle. Prove that $\mathcal{L}_o^N(\partial R) = 0$.*

Proposition 262 *The following properties hold.*

(i) If $E \subseteq F \subseteq \mathbb{R}^N$, then $\mathcal{L}_o^N(E) \leq \mathcal{L}_o^N(F)$.

(ii) If $E \subseteq \bigcup_{n=1}^{\infty} E_n$, then $\mathcal{L}_o^N(E) \leq \sum_{n=1}^{\infty} \mathcal{L}_o^N(E_n)$.

We now show that if $E \cap F = \emptyset$, then it can happen that

$$\mathcal{L}_o^N(E \cup F) \neq \mathcal{L}_o^N(E) + \mathcal{L}_o^N(F).$$

Exercise 263 *Let $E \subseteq \mathbb{R}^N$ and let $\boldsymbol{x}_0 \in \mathbb{R}^N$. Prove that*

$$\mathcal{L}_o^N(E) = \mathcal{L}_o^N(\boldsymbol{x}_0 + E)$$

Example 264 *On the real line we consider the equivalence relation $x \sim y$ if $x - y \in \mathbb{Q}$. By the axiom of choice we may construct a set $E \subset (0,1)$ that contains exactly one element from each equivalence class. The following properties are satisfied:*

(i) *If $x \in (0,1)$, then $x \in r+E$ for some $r \in (-1,1)\cap\mathbb{Q}$. To see this, observe that by construction of E, for any $x \in (0,1)$ there exists $y \in E$ such that that $x \sim y$, that is, $x - y = r \in (-1,1)\cap\mathbb{Q}$.*

(ii) *If r, $q \in \mathbb{Q}$, with $r \neq q$, then $(r+E)\cap(q+E) = \emptyset$. Indeed, if not, then we may write $r+x = q+y$ for some x, $y \in E$. But then $x-y = q-r \in \mathbb{Q}\setminus\{0\}$, which implies that $x \sim y$. By the construction of E this is possible only if $x = y$, which is impossible.*

Define

$$F := \bigcup_{r\in(-1,1)\cap\mathbb{Q}} (r+E) \subset (-1,2)\,.$$

Observ that $F \supset (0,1)$ by property (a)

Assume by contradiction that

$$\mathcal{L}_o^1(E_1 \cup E_2) = \mathcal{L}_o^1(E_1) + \mathcal{L}_o^1(E_2) \tag{32}$$

for all $E_1, E_2 \subseteq \mathbb{R}$ with $E_1 \cap E_2 = \emptyset$. Consider an enumeration $\{r_n : n \in \mathbb{N}\}$ of $(-1,1)\cap\mathbb{Q}$ and define

$$E_n := r_n + E.$$

Then the sets E_n are pairwise disjoint and so if (32) were to hold, then by induction

$$3 \geq \mathcal{L}_o^1\Big(\bigcup_{n=1}^m E_n\Big) = \sum_{n=1}^m \mathcal{L}_o^1(E_n) = \sum_{n=1}^m \mathcal{L}_o^1(E) = m\mathcal{L}_o^1(E).$$

Taking m large enough we have a contradiction unless $\mathcal{L}_o^1(E) = 0$, but if $\mathcal{L}_o^1(E) = 0$, then $\mathcal{L}_o^1(F) = 0$, which contradicts the fact that $F \supset (0,1)$.

To recover property (32), we need to introduce the notion of Lebesgue measurability.

Definition 265 *Given a set $E \subseteq \mathbb{R}^N$, we say that E is* Lebesgue measurable *if for every $\varepsilon > 0$ there exists an open set $U \supseteq E$ such that*

$$\mathcal{L}_o^N(U \setminus E) \leq \varepsilon.$$

Proposition 266 *The following properties hold.*

(i) *Open sets are Lebesgue measurable.*

(ii) *If $E \subseteq \mathbb{R}^N$ has Lebesgue outer measure zero, then E and its subsets are Lebesgue measurable.*

(iii) *If $E = \bigcup_{n=1}^\infty E_n$, and each E_n is Lebesgue measurable, then E is Lebesgue measurable.*

(iv) *Compact sets are Lebesgue measurable.*

(v) Closed sets are Lebesgue measurable.

(vi) If $E \subseteq \mathbb{R}^N$ is Lebesgue measurable, then $\mathbb{R}^N \setminus E$ is Lebesgue measurable.

(vii) If $E = \bigcap_{n=1}^{\infty} E_n$, and each E_n is Lebesgue measurable, then E is Lebesgue measurable.

Let

$$\mathfrak{M} = \{E \subseteq \mathbb{R}^N : E \text{ is Lebesgue measurable}\}.$$

For every $E \in \mathfrak{M}$, the *Lebesgue measure of* E is defined to be $\mathcal{L}^N(E) := \mathcal{L}_o^N(E)$.

Proposition 267 *Let $E_n \subseteq \mathbb{R}^N$, $n \in \mathbb{N}$, be Lebesgue measurable.*

(i) If the sets E_n are disjoint, then

$$\mathcal{L}^N\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mathcal{L}^N(E_n).$$

(ii) If $E_n \subseteq E_{n+1}$ for all n, then

$$\mathcal{L}^N\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mathcal{L}^N(E_n).$$

(iii) If $E_n \supseteq E_{n+1}$ for all n and $\mathcal{L}^N(E_n) < \infty$ for some n, then

$$\mathcal{L}^N\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mathcal{L}^N(E_n).$$

17.1 Integrable Functions

We are now in a position to introduce the notion of integral. Given a set $F \subseteq \mathbb{R}^N$ the *characteristic function of* F is the function χ_f, defined by

$$\chi_F(\boldsymbol{x}) := \begin{cases} 1 & \text{if } \boldsymbol{x} \in F, \\ 0 & \text{otherwise.} \end{cases}$$

Let $E \subseteq \mathbb{R}^N$ be a Lebesgue measurable set and let $F \subseteq E$ be a Lebesgue measurable set. We define the *Lebesgue integral* of χ_F over F as

$$\int_E \chi_F \, d\boldsymbol{x} := \mathcal{L}^N(F).$$

Wednesday, April 27, 2022

Definition 268 *Let $E \subseteq \mathbb{R}^N$ be a Lebesgue measurable set. A* simple function *is a function $s : E \to \mathbb{R}$ that can be written as*

$$s = \sum_{n=1}^{\ell} c_n \chi_{E_n},$$

where $c_1, \dots, c_\ell \in \mathbb{R}$ and the sets E_n are Lebesgue measurable.

Let $E \subseteq \mathbb{R}^N$ be a Lebesgue measurable set and let $s : E \to [0, \infty)$ be a nonnegative simple function. If $s \neq 0$, we can write

$$s = \sum_{n=1}^{\ell} c_n \chi_{E_n},$$

where the sets $E_n \subseteq E$ are pairwise disjoint, $E_n \cap E_k = \emptyset$ if $n \neq k$, and $c_n > 0$ for all $n = 1, \ldots, \ell$. We define the *Lebesgue integral* of s over E as

$$\int_E s \, d\boldsymbol{x} := \sum_{n=1}^{\ell} c_n \mathcal{L}^N (E_n). \tag{33}$$

Exercise 269 *Let $E \subseteq \mathbb{R}^N$ be a Lebesgue measurable set. Let $s, t : E \to [0, \infty)$ be simple functions. Prove that for every Lebesgue measurable set $G \subseteq E$,*

$$\int_G (s + t) \, d\boldsymbol{x} = \int_G s \, d\boldsymbol{x} + \int_G t \, d\boldsymbol{x}.$$

Definition 270 *Let $E \subseteq \mathbb{R}^N$ be a Lebesgue measurable set and let $f : E \to [0, \infty]$. We say that f is* Lebesgue measurable *if there exists a sequence of simple functions $s_n : E \to [0, \infty)$ such that $s_n \leq f$ for every n and $s_n \to f$ pointwise in E.*

Theorem 271 *Let $E \subseteq \mathbb{R}^N$ be a Lebesgue measurable set and let $f : E \to [0, \infty)$ be a continuous function. Then f is Lebesgue measurable.*

Let $E \subseteq \mathbb{R}^N$ be a Lebesgue measurable set and let $f : E \to [0, \infty]$ be a measurable function. The *Lebesgue integral* of f over E is defined as

$$\int_E f \, d\boldsymbol{x} := \sup \left\{ \int_E s \, d\boldsymbol{x} : \, s \text{ simple}, \, 0 \leq s \leq f \text{ in } E \right\}.$$

We list below some basic properties of Lebesgue integration for nonnegative functions.

Proposition 272 *Let $E \subseteq \mathbb{R}^N$ be a Lebesgue measurable set, let $f, g : E \to [0, \infty]$ be two Lebesgue measurable functions.*

(i) If $0 \leq f \leq g$ in E, then $\int_E f \, d\boldsymbol{x} \leq \int_E g \, d\boldsymbol{x}$.

(ii) If $c \in [0, \infty)$, then $\int_E cf \, d\boldsymbol{x} = c \int_E f \, d\boldsymbol{x}$ (here we set $0\infty := 0$).

(iii) If $\int_E f \, d\boldsymbol{x} = 0$ then there exists a Lebesgue measurable set $G \subseteq E$ with $\mathcal{L}^N (G) = 0$ such that $f = 0$ in $E \setminus G$.

(iv) If $\mathcal{L}^N (E) = 0$, then $\int_E f \, d\boldsymbol{x} = 0$, even if $f \equiv \infty$ in E.

(v) $\int_F f \, d\boldsymbol{x} = \int_E \chi_F f \, d\boldsymbol{x}$ for every Lebesgue measurable set $F \subseteq E$.

17.2 Lebesgue Integration of Functions of Arbitrary Sign

Let $E \subseteq \mathbb{R}^N$ be a Lebesgue measurable set. In order to extend the notion of integral to functions of arbitrary sign, consider $f : E \to [-\infty, \infty]$ and set

$$f^+ := \max\{f, 0\}, \quad f^- := \max\{-f, 0\}.$$

Note that $f = f^+ - f^-$, $|f| = f^+ + f^-$.We say that f is *Lebesgue measurable* if f^+ and f^- are Lebesgue measurable.

Definition 273 *Let $E \subseteq \mathbb{R}^N$ be a Lebesgue measurable set and let $f : E \to [-\infty, \infty]$ be a measurable function. If at least one of the two integrals $\int_E f^+ \, d\boldsymbol{x}$ and $\int_E f^- \, d\boldsymbol{x}$ is finite, then we define the* Lebesgue integral of f *over E by*

$$\int_E f \, d\boldsymbol{x} := \int_E f^+ \, d\boldsymbol{x} - \int_E f^- \, d\boldsymbol{x}.$$

If both $\int_E f^+ \, d\boldsymbol{x}$ and $\int_E f^- \, d\boldsymbol{x}$ are finite, then f is said to be Lebesgue integrable *over E.*

Example 274 *Consider the function*

$$f(x) := \frac{\sin x}{x}, \quad x \geq \pi.$$

Let's prove that the limit

$$\lim_{\ell \to \infty} \int_\pi^\ell \frac{\sin x}{x} \, dx \in \mathbb{R},$$

exists so that f is Riemann integrable in $[\pi, \infty)$. Integrating by parts, we have

$$\begin{aligned} \int_\pi^\ell \frac{\sin x}{x} \, dx &= \left[-\frac{1}{x} \cos x \right]_{x=\pi}^{x=\ell} - \int_\pi^\ell \frac{1}{x^2} \cos x \, dx \\ &= -\frac{1}{\ell} \cos \ell - \frac{1}{\pi} - \int_\pi^\ell \frac{1}{x^2} \cos x \, dx. \end{aligned}$$

Since

$$\int_\pi^\ell \left| \frac{1}{x^2} \cos x \right| dx \leq \int_\pi^\infty \frac{1}{x^2} \, dx = \frac{1}{\pi} < \infty,$$

we have that there exists the limit

$$\lim_{\ell \to \infty} \int_\pi^\ell \frac{1}{x^2} \cos x \, dx = \ell \in \mathbb{R}.$$

Hence,

$$\lim_{\ell \to \infty} \int_\pi^\ell \frac{\sin x}{x} \, dx = \lim_{\ell \to \infty} -\frac{1}{\ell} \cos \ell - \frac{1}{\pi} - \lim_{\ell \to \infty} \int_\pi^\ell \frac{1}{x^2} \cos x \, dx = -\frac{1}{\pi} - \ell.$$

On the other hand,

$$\int_{\pi}^{\infty}\left(\frac{\sin x}{x}\right)^{+}dx=\int_{\pi}^{\infty}\left(\frac{\sin x}{x}\right)^{-}dx=\infty,$$

so that the Lebesgue integral of f is not defined. To see this, observe that

$$\int_{\pi}^{\infty}\left(\frac{\sin x}{x}\right)^{+}dx\geq\sum_{n=1}^{\infty}\int_{(2n-1)\pi}^{2n\pi}\left(\frac{\sin x}{x}\right)^{+}dx\geq\sum_{n=1}^{\infty}\frac{1}{2n\pi}\int_{(2n-1)\pi}^{2n\pi}(\sin x)^{+}dx$$
$$=\sum_{n=1}^{\infty}\frac{1}{2n\pi}\int_{\pi}^{2\pi}(\sin x)^{+}dx=\infty.$$

The other integral can be estimated in a similar way.

Proposition 275 *Let $E\subseteq\mathbb{R}^N$ be a Lebesgue measurable set and let $f,\,g:E\to[-\infty,\infty]$ be two measurable functions.*

(i) If f and g are integrable and $\alpha,\,\beta\in\mathbb{R}$, then $\alpha f+\beta g$ is integrable and

$$\int_E(\alpha f+\beta g)\;d\boldsymbol{x}=\alpha\int_E f\,d\boldsymbol{x}+\beta\int_E g\,d\boldsymbol{x}.$$

(ii) $\left|\int_E f\,d\boldsymbol{x}\right|\leq\int_E|f|\;d\boldsymbol{x}$.

(iii) If f is Lebesgue integrable, then the set $\{\boldsymbol{x}\in E:\,|f(\boldsymbol{x})|=\infty\}$ has measure zero.

(iv) If $f(\boldsymbol{x})=g(\boldsymbol{x})$ for $\mathcal{L}^N$ a.e. $\boldsymbol{x}\in E$, then $\int_E f^{\pm}\,d\boldsymbol{x}=\int_E g^{\pm}\,d\boldsymbol{x}$, so that $\int_E f\,d\boldsymbol{x}$ is well-defined if and only if $\int_E g\,d\boldsymbol{x}$ is well-defined, and in this case we have

$$\int_E f\,d\boldsymbol{x}=\int_E g\,d\boldsymbol{x}.\tag{34}$$

Given a set $E\subseteq\mathbb{R}^N\times\mathbb{R}^M$, for every $\boldsymbol{x}\in\mathbb{R}^N$ consider the section

$$E_{\boldsymbol{x}}:=\{\boldsymbol{y}\in\mathbb{R}^M:\,(\boldsymbol{x},\boldsymbol{y})\in E\},$$

and for $\boldsymbol{y}\in\mathbb{R}^M$ consider the section

$$E_{\boldsymbol{y}}:=\{\boldsymbol{x}\in\mathbb{R}^N:\,(\boldsymbol{x},\boldsymbol{y})\in E\}.$$

Let

$$G:=\{\boldsymbol{x}\in\mathbb{R}^N:\,E_{\boldsymbol{x}}\neq\emptyset\},\quad H:=\{\boldsymbol{y}\in\mathbb{R}^M:\,E_{\boldsymbol{y}}\neq\emptyset\}.$$

Theorem 276 *Let $E\subseteq\mathbb{R}^N\times\mathbb{R}^M$ be a Lebesgue measurable set. Then G is Lebesgue measurable, for $\mathcal{L}^N$-a.e. $\boldsymbol{x}\in G$ the section $E_{\boldsymbol{x}}$ is Lebesgue measurable, and the function $\boldsymbol{x}\in G\mapsto\mathcal{L}^M(E_{\boldsymbol{x}})$ is measurable. Similarly, H is Lebesgue measurable, for $\mathcal{L}^M$-a.e. $\boldsymbol{y}\in H$ the section $E_{\boldsymbol{y}}$ is Lebesgue measurable, and the function $\boldsymbol{y}\in H\mapsto\mathcal{L}^N(E_{\boldsymbol{y}})$ is Lebesgue measurable. Moreover,*

$$\mathcal{L}^{N+M}(E)=\int_G\mathcal{L}^M(E_{\boldsymbol{x}})\,d\boldsymbol{x}=\int_H\mathcal{L}^M(E_{\boldsymbol{y}})\,d\boldsymbol{y}.$$

Friday, April 27, 2022

By applying the previous theorem first to χ_E, then to simple functions, then to pointwise limits of simple functions we obtain the following theorem.

Theorem 277 (Tonelli) *Let $E \subseteq \mathbb{R}^N \times \mathbb{R}^M$ be a Lebesgue measurable set, and let $f : E \to [0, \infty]$ be a Lebesgue measurable function. Then for $\mathcal{L}^N$-a.e. $\boldsymbol{x} \in G$ the section $E_{\boldsymbol{x}}$ is Lebesgue measurable, the function $\boldsymbol{y} \in E_{\boldsymbol{x}} \mapsto f(\boldsymbol{x}, \boldsymbol{y})$ is Lebesgue measurable, and the function $\boldsymbol{x} \in G \mapsto \int_{E_{\boldsymbol{x}}} f(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{y}$ is measurable. Similarly, for $\mathcal{L}^M$-a.e. $\boldsymbol{y} \in H$ the section $E_{\boldsymbol{y}}$ is Lebesgue measurable, the function $\boldsymbol{x} \in E_{\boldsymbol{y}} \mapsto f(\boldsymbol{x}, \boldsymbol{y})$ is Lebesgue measurable, and the function $\boldsymbol{y} \in H \mapsto \int_{E_{\boldsymbol{y}}} f(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{x}$ is measurable. Moreover,*

$$\begin{aligned}\int_E f(\boldsymbol{x}, \boldsymbol{y})\, d(\boldsymbol{x}, \boldsymbol{y}) &= \int_G \left(\int_{E_{\boldsymbol{x}}} f(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{y} \right) d\boldsymbol{x} \\ &= \int_H \left(\int_{E_{\boldsymbol{y}}} f(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{x} \right) d\boldsymbol{y}.\end{aligned}$$

The version of Tonelli's theorem for integrable functions of arbitrary sign is the well–known Fubini's theorem:

Theorem 278 (Fubini) *Let $E \subseteq \mathbb{R}^N \times \mathbb{R}^M$ be a Lebesgue measurable set, and let $f : E \to [-\infty, \infty]$ be Lebesgue integrable. Then for $\mathcal{L}^N$-a.e. $\boldsymbol{x} \in G$ the section $E_{\boldsymbol{x}}$ is Lebesgue measurable, the function $\boldsymbol{y} \in E_{\boldsymbol{x}} \mapsto f(\boldsymbol{x}, \boldsymbol{y})$ is Lebesgue integrable, and the function $\boldsymbol{x} \in G \mapsto \int_{E_{\boldsymbol{x}}} f(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{y}$ is Lebesgue integrable. Similarly, for $\mathcal{L}^M$-a.e. $\boldsymbol{y} \in H$ the section $E_{\boldsymbol{y}}$ is Lebesgue measurable, the function $\boldsymbol{x} \in E_{\boldsymbol{y}} \mapsto f(\boldsymbol{x}, \boldsymbol{y})$ is Lebesgue integrable, and the function $\boldsymbol{y} \in H \mapsto \int_{E_{\boldsymbol{y}}} f(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{x}$ is Lebesgue integrable. Moreover,*

$$\begin{aligned}\int_E f(\boldsymbol{x}, \boldsymbol{y})\, d(\boldsymbol{x}, \boldsymbol{y}) &= \int_G \left(\int_{E_{\boldsymbol{x}}} f(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{y} \right) d\boldsymbol{x} \\ &= \int_H \left(\int_{E_{\boldsymbol{y}}} f(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{x} \right) d\boldsymbol{y}.\end{aligned}$$

Exercise 279 *The next example shows that Fubini's theorem fails without assuming the integrability of the function f. Consider the function*

$$f(x, y) := \frac{x^2 - y^2}{(x^2 + y^2)^2}, \quad (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}.$$

Prove that the Lebesgue integral of f is not defined over $[0, 1]^2 \setminus \{(0, 0\})$ is not defined and that the iterated integrals are different.

$$\int_0^1 \left(\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dx \right) dy = \int_0^1 \left[-\frac{x}{x^2 + y^2} \right]_{x=0}^{x=1} dy = \int_0^1 -\frac{1}{y^2 + 1} dy = -[\arctan y]_{y=0}^{y=1} = -\frac{1}{4}\pi,$$

while

$$\int_0^1 \left(\int_0^1 \int \frac{x^2 - y^2}{(x^2+y^2)^2} dy \right) dx = \int_0^1 \left[\frac{y}{x^2+y^2} \right]_{y=0}^{y=1} dx = \int_0^1 \frac{1}{x^2+1} dx = [\arctan x]_{x=0}^{x=1} = \frac{1}{4}\pi,$$

On the other hand

$$\int_0^1 \int_0^1 \frac{(x^2-y^2)^+}{(x^2+y^2)^2} dxdy = \int_0^1 \left(\int_y^1 \frac{x^2-y^2}{(x^2+y^2)^2} dx \right) dy$$
$$= \int_0^1 \left[-\frac{x}{x^2+y^2} \right]_{x=y}^{x=1} dy = \int_0^1 \frac{1}{2y} - \frac{1}{y^2+1} dy = \infty$$

while

$$\int_0^1 \int_0^1 \frac{(x^2-y^2)^-}{(x^2+y^2)^2} dxdy = \int_0^1 \left(\int_x^1 \frac{y^2-x^2}{(x^2+y^2)^2} dy \right) dx$$
$$= \int_0^1 \left[-\frac{y}{x^2+y^2} \right]_{y=x}^{y=1} dx = \int_0^1 \frac{1}{2x} - \frac{1}{x^2+1} dx = \infty.$$

Corollary 280 *Let $E \subseteq \mathbb{R}^N$ be a Lebesgue measurable set, let $\alpha : E \to \mathbb{R}$ and $\beta : E \to \mathbb{R}$ be two Lebesgue measurable functions, with $\alpha(\boldsymbol{x}) \le \beta(\boldsymbol{x})$ for all $\boldsymbol{x} \in E$, and let*

$$F := \{(\boldsymbol{x}, y) \in E \times \mathbb{R} : \alpha(\boldsymbol{x}) \le y \le \beta(\boldsymbol{x})\}.$$

Then F is Lebesgue measurable. Moreover, if $f : F \to \mathbb{R}$ is Lebesgue integrable or $f : F \to [0, \infty)$ is Lebesgue measurable, then

$$\int_F f(\boldsymbol{x}, y)\, d(\boldsymbol{x}, y) = \int_E \left(\int_{\alpha(\boldsymbol{x})}^{\beta(\boldsymbol{x})} f(\boldsymbol{x}, y)\, dy \right) d\boldsymbol{x}.$$

Example 281 *Let's calculate the integral*

$$\iint_E x(1-y)\, dxdy,$$

where

$$E := \left\{(x, y) \in \mathbb{R}^2 : y \le x,\, x^2 + y^2 \le 1,\, x \ge 0,\, y \ge 0\right\}.$$

We can rewrite E as follows,

$$E = \left\{ (x, y) \in \mathbb{R}^2 : 0 \le y \le \frac{\sqrt{2}}{2},\, y \le x \le \sqrt{1-y^2} \right\}$$

and since the function $f(x, y) := x(1-y)$ is continuous in E and the functions $\alpha(y) := y$ and $\beta(y) := \sqrt{1-y^2}$ are continuous, we can apply the previous *corollary to conclude that*

$$\iint_E x(1-y)\,dxdy = \int_0^{\frac{\sqrt{2}}{2}} \left(\int_y^{\sqrt{1-y^2}} x(1-y)\,dx \right) dy = \int_0^{\frac{\sqrt{2}}{2}} (1-y) \left(\left[\frac{x^2}{2} \right]_{x=y}^{x=\sqrt{1-y^2}} \right) dy$$

$$= \int_0^{\frac{\sqrt{2}}{2}} (1-y) \left(\frac{1-y^2}{2} - \frac{y^2}{2} \right) dy$$

$$= \frac{1}{6}\sqrt{2} - \frac{1}{16}.$$

Exercise 282 *Calculate the integral*

$$\iiint_E (x+z)\,dxdydz,$$

where

$$E := \left\{ (x,y,z) \in \mathbb{R}^3 :\, x+y+z \leq 1,\, x \geq 0,\, y \geq 0,\, z \geq 0 \right\}.$$