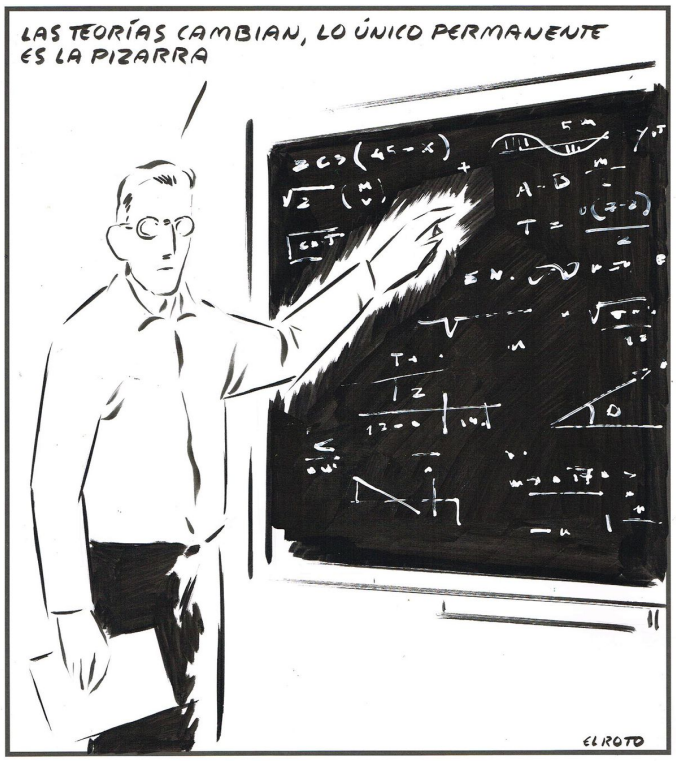
Functions of Several Real Variables[1]

Matías Raja

[1]Curso de Grado en Matemáticas en la Universidad de Murcia 2021.



I practiced every night,
now I'm ready.

L. Cohen

Contents

Preface

These lectures are based on quite a few years of teaching *Functions of Several Real Variables* for the degree in Mathematics at the University of Murcia. The purpose is to cover not only the usual topics, but a "little more" which makes the difference with other courses, notably, the inclusion of many related topics and nontrivial applications.

The corpus of Functions of Several Real Variables is often divided in two or three subjects when it comes to teach it. However, I have decided to keep some unity of the material, with the exception of the chapter on *Measure Theory and Lebesgue Integral* that is independent somehow. The contents can be grouped as follows:

1. *Topological and linear background*: chapters 1 and 2.
2. *Differential Calculus*: chapters 3, 4, and 5.
3. *Integral Calculus*: chapters 6, 7 and 8.
4. *Integration on manifolds and Vector Analysis*: chapters 9, 10 and 11.

Some comments are necessary. In case of withdrawing the Riemann integral, it is possible to adapt the change of variables theorem (chapter 7) to Lebesgue integral (chapter 8) without much trouble. As to the *Integration on manifolds*, the techniques are addressed to objects of dimension less than 3 (curves and surfaces). Therefore, the theory of differential forms is not fully developed, but all the classic applications are covered. Every chapter has a section at the end called *Rationale and remarks* where the point of view adopted in the chapter is discussed and some further developments are proposed.

I have added three appendices at the end. Two of them complement information on the most important spaces of functions that appear in the course: the space of continuous functions over a compact $C(K)$ and the spaces of integrable functions $L^p(\mu)$. The results we have included (Stone-Weierstrass theorem, relations amongst the types of convergence for integrable functions...) lie in a limbo between Real Analysis and Functional Analysis. The third appendix on Mechanics is an "experiment" based in the fact that is possible to obtain the Lagrange's equations of the movement from Newton's laws and the chain rule of Calculus.

The bibliography consists of the books I read as student to learn most of the results here, and those I have consulted during the preparation of the manuscript. I apologize for not giving precise references at all times.

An early version of this manual was presented as a *Proyecto Docente* for the promotion of the author to *Catedrático de Universidad* on July 22, 2021. The current version is updated every now and then with new examples and less mistakes, happily. However, to achieve a fully satisfactory version, as Edward E. Gibbon said, would require many years of health, of leisure and of perseverance.

Murcia, Spring 2022.

Chapter 1

Metric Spaces (for Analysis)

1.1 Generalities

The basics on metric spaces are allegedly known by the students, so we will get through this first section rather quickly.

A *metric* on a set M is a function $d : M \times M \to [0, +\infty)$ with these three properties:

a) $d(x, y) = 0$ if and only if $x = y$;

b) $d(x, y) = d(y, x)$ for any $x, y \in M$;

c) $d(x, y) \leq d(x, z) + d(z, y)$.

A pair (M, d) consisting on a set and a metric on it is called *metric space*. The open and closed balls with center $x \in M$ and radius $r > 0$ are defined in this way

$$B(x, r) = \{y \in M : d(x, y) < r\}$$
$$B[x, r] = \{y \in M : d(x, y) \leq r\}$$

A set $A \subset M$ is said *open* if for every $x \in A$ there is some $r > 0$ such that $B(x, r) \subset A$. The family of all the open sets of M is called its *topology* and has the following properties:

a) $\emptyset$ and M are open;

b) arbitrary unions of open sets give open sets;

c) finite intersections of open sets give open sets.

Statements on metric spaces that can be formulated in terms of open sets (the topology) are called topological. For instance, we may define convergence of sequences in a metric space as follows $(x_n) \subset M$ is converging to $x \in M$ if $\lim_n d(x_n, x) = 0$. Apparently, the definition strongly uses the metric, however convergence of sequences is a topological notion actually, because it can be equivalently formulated as: for every $U \ni x$ open there is n_U such that $x_n \in U$ whenever $n \geq n_U$. We say that x is a *cluster point* of a sequence (x_n) if there is a subsequence (x_{n_k}) with limit x. A cluster point of an infinite set is the limit of a sequence of different point from the set.

A set $A \subset M$ is said *closed* if its complement is an open set. It follows that $\emptyset$ and M are closed as well, closed sets are stable by finite unions and arbitrary intersections. The interior of a set A denoted A° is the largest open contained in a set A, and the closure of A denoted $\overline{A}$ is the smaller closed set containing A. Of course, we have the duality formula $M \setminus A^\circ = \overline{M \setminus A}$. Moreover, we have this useful characterization in terms of convergent sequences.

Proposition 1.1.1. *The closure $\overline{A}$ of a set $A \subset M$ is the set of limits of all the sequences contained in A which are convergent. In particular, a set A is closed if and only if the limits of sequences of points from A remain in A.*

Given two metric spaces (M_1, d_1) and (M_2, d_2) and a map $f : M_1 \to M_2$ we say that f is *continuous* at $x \in M_1$ is for every $\varepsilon > 0$ there is $\delta > 0$ such that for $y \in M_1$ and $d_1(x, y) < \delta$ then $d_2(f(y), f(x)) < \varepsilon$. It is not difficult to check that continuity is a topological property. Note that continuity at one point is characterized by means of sequences: f is continuous at x if $\lim_n d_2(f(x_n), f(x)) = 0$ for every sequence $(x_n) \subset M_1$ converging to x. A map $f : M_1 \to M_2$ is said continuous if it is continuous at every point of M_1. Note that continuity of f is equivalent to say that $f^{-1}(U)$ is open whenever $U \in M_2$ is open. With the same notation, the map f is said uniformly continuous if for every $\varepsilon > 0$ there is $\delta > 0$ such that for $x, y \in M_1$ then $d_2(f(y), f(x)) < \varepsilon$ (note the change of position of a quantifier). A particular important case of uniformly continuous maps those satisfying the *Lipschitz* condition. We say that f is Lipschitz if there is some $\lambda > 0$ such that $d_2(f(y), f(x)) \leq \lambda\, d_1(x, y)$.

Examples of real functions are provided by the distance functions to sets: for $A \subset M$ set $d(x, A) = \inf\{d(x, y) : y \in A\}$. Indeed, continuity follows easily from $|d(x, A) - d(y, A)| \leq d(x, y)$. Given two disjoint closed sets $A, B \subset M$,

the continuous function

$$f(x) = \frac{d(x, A)}{d(x, A) + d(x, B)}$$

satisfies that $f(x) \in [0, 1]$, $A = f^{-1}(0)$ and $B = f^{-1}(1)$.

Given two metrics d_1 and d_2 on the same set M, we say that d_1 is *finer* than d_2 (equivalently, d_2 is *coarser* than d_1) if any open set with respect to d_2 is also open with respect to d_1. Note that this is equivalent to the continuity of the identity map $Id : (M, d_1) \to (M, d_2)$. The two metrics on M are said *equivalent* if they produce the same topology, that is, the identity map is continuous forth and back (topological homeomorphism). Given a metric space (M, d), we may always suppose that the metric is bounded just taking the equivalent metric $d_1(x, y) = \min\{1, d(x, y)\}$.

A very useful operation with metric spaces (and more general topological spaces) is the *product*. Let (M_1, d_1) and (M_2, d_2) be metric spaces. We can endow $M_1 \times M_2$ with the metric defined by

$$d((x_1, x_2), (y_1, y_2)) = d_1(x_1, y_1) + d_2(x_2, y_2).$$

That operation can be extended to more finitely many factors. In the particular case of the product of "copies" of $\mathbb{R}$, it is not difficult to check that the product metric is equivalent to the Euclidean distance. We may even consider countable many factors $\{(M_n, d_n)\}_{n\in\mathbb{N}}$. In that case, define the metric by a series

$$d((x_n), (y_n)) = \sum_{n=1}^{\infty} 2^{-n} d_n(x_n, y_n)$$

where d_n is an equivalent metric on M_n bounded by 1.

1.2 Separability

A subset $A \subset M$ is said to *dense* if $\overline{A} = M$. A metric space is said to be *separable* if it contains a countable dense set $\{x_n : n \in \mathbb{N}\}$. Note that in such a case, the collection of balls $\{B(x_n, 1/m) : n, m \in \mathbb{N}\}$ is a countable *base* of the topology, that is, every open set can be expressed as a union of balls from that collection. A metric (or more generally, topological) space is said *Lindelöf* if every cover of the space by open sets has a countable subcover. With all these definitions we have the following.

Proposition 1.2.1. *A metric space is separable if and only if it is Lindelöf.*

Proof. If M is separable, there is a countable base (B_n). For every $x \in M$ there is an open set U form the cover such that $x \in U$, and by the definition of base, there is $n \in \mathbb{N}$ such that $x \in B_n \subset U$. Doing this operation for all the points in M involves only a countable number of sets from (B_n). The collection of open supersets is a countable cover of M. On the other hand, assume M is Lindelöf, and for $m \in \mathbb{N}$ take the cover $\{B(x, 1/m) : x \in M\}$, which turns out to have a countable subfamily covering M. Let $(x_{n,m})$ be the collection of centres of the balls for the countable subcovering. By construction, $\{x_{n,m} : n, m \in \mathbb{N}\}$ is a dense set. ∎

Let $\varepsilon > 0$. A set $A \subset M$ is said *ε-discrete* if $d(x, y) \geq \varepsilon$ for any $x, y \in A$ with $x \neq y$. The set is said (metrically) *discrete* if it is ε-discrete for some $\varepsilon > 0$. We have the following.

Proposition 1.2.2. *A metric space is separable if and only if it not contains an uncountable discrete set.*

Proof. If M contains an ε-discrete uncountable set A, then $\{B(x, \varepsilon/3) : x \in A\}$ is a disjoint uncountable collection of balls. A countable subset of M cannot meet all those balls by cardinality, so that set cannot be dense. In the other hand, assume that discrete sets are countable. Given $\varepsilon > 0$ there is a maximal ε-discrete set A_ε. Such a set is countable and has the property that for any $x \in M$ there is $y \in A_\varepsilon$ such that $d(x, y) < \varepsilon$. Taking $\varepsilon = 1/n$ we get that $\bigcup_{n=1}^{\infty} A_{1/n}$ is dense. ∎

Separability implies that ε-discrete sets are countable. We say that the metric space M is *totally bounded* if all the discrete sets are finite. We will call *ε-net* to a maximal ε-discrete set.

1.3 Completeness

A sequence (x_n) is said Cauchy if for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that $d(x_n, x_m) < \varepsilon$ whenever $n, m \geq N$ (equivalently, $\lim_{n,m} d(x_n, x_m) = 0$). The reader could establish these easy facts: every convergent sequence is Cauchy; a Cauchy sequence with a cluster point must be convergent. A metric space is said complete if every Cauchy sequence is convergent. The notion of completeness is non topological. Observe that $(-\pi/2, \pi/2)$ is not complete with the

usual metric on $\mathbb{R}$ but the metric $d(x,y) = |\tan x - \tan y|$ makes its complete.

A useful observation is that completeness is inherited by closed subsets and products (with the standard product metric). On the other hand, a subset of a metric space which is complete with respect when endowed with the restricted metric must be closed in the overspace.

Proposition 1.3.1 (Cantor)**.** *A metric space M is complete if and only if $\bigcap_{n=1}^{\infty} F_n \neq \emptyset$ whenever (F_n) is a decreasing sequence of nonempty closed sets of M such that $\lim_n \operatorname{diam}(F_n) = 0$. In such a case, we have $\bigcap_{n=1}^{\infty} F_n = \{x\}$ for some $x \in M$.*

Proof. Observe that the hypothesis implies that (x_n) is a Cauchy sequence for any choice of $x_n \in F_n$. If the space is complete then $\lim_n x_n = x$ and clearly $\{x\} = \bigcap_{n=1}^{\infty} F_n$. On the other hand, if M is not complete then there is a Cauchy sequence (x_n) with no limit. Since (x_n) cannot have cluster points, the sets $F_n = \{x_k : k \geq n\}$ are closed. Cauchy property implies that $\lim_n \operatorname{diam}(F_n) = 0$, but we have $\bigcap_{n=1}^{\infty} F_n = \emptyset$. ∎

The following is the celebrated Baire's theorem.

Theorem 1.3.2 (Baire)**.** *Let M be a complete metric space. If (U_n) a sequence of dense open sets of M, then $\bigcap_{n=1}^{\infty} U_n$ is dense.*

Proof. Denseness of $\bigcap_{n=1}^{\infty} U_n$ is equivalent to $U \cap \bigcap_{n=1}^{\infty} U_n \neq \emptyset$ for every nonempty open U. Fix $U \subset M$ a nonempty open set. Since U_1 is dense, $U \cap U_1 \neq \emptyset$ and there are $x_1 \in M$ and $r_1 \leq 1$ such that $B[x_1, r_1] \subset U \cap U_1$. Again $B(x_1, r_1) \cap U_2 \neq \emptyset$ so there are $x_2 \in M$ and $r_2 \leq 1/2$ such that $B[x_2, r_2] \subset B(x_1, r_1) \cap U_1$. Proceeding in this way we may have sequences $(x_n) \subset M$ and $(r_n) \subset \mathbb{R}^+$ such that $B[x_n, r_n] \subset U \cap U_n$ and

$$B[x_1, r_1] \supset B[x_2, r_2] \supset \cdots \supset B[x_n, r_n] \supset \ldots$$

thus by the previous proposition $\bigcap_{n=1}^{\infty} B[x_n, r_n] = \{x\}$. By construction $x \in U \cap \bigcap_{n=1}^{\infty} U_n$ and so $U \cap \bigcap_{n=1}^{\infty} U_n \neq \emptyset$. ∎

Baire's theorem is sometimes preferred in this equivalent form.

Corollary 1.3.3. *Let M be a complete metric space and (F_n) a sequence of closed subsets of M such that $M = \bigcup_{n=1}^{\infty} F_n$. There there is $n \in \mathbb{N}$ such that F_n has nonempty interior.*

The following is Banach's fixed point for contractive mappings.

Theorem 1.3.4 (Banach)**.** *Let M be a complete metric space and $f : M \to M$ a map such that there is $\lambda < 1$ such that*

$$d(f(x), f(y)) \leq \lambda\, d(x, y)$$

Then there is unique point $x \in M$ such that $f(x) = x$ (a fixed point for f). Moreover, whenever $x_1 \in M$ is chosen, the sequence defined inductively by $x_n = f(x_{n-1})$ for $n \geq 2$ is converging to x.

Proof. If $y \in M$ is another fixed point for f and $y \neq x$, then we have

$$d(x, y) = d(f(x), f(y)) \leq \lambda\, d(x, y) < d(x, y)$$

which is a contradiction proving the uniqueness.
Now, for an arbitrary chosen $x_1 \in M$, consider the sequence (x_n) recursively generated as in the statement. Note that in case (x_n) converges to some point $x \in M$, then it is a fixed point. Indeed

$$f(x) = \lim_n f(x_n) = \lim_n x_{n+1} = x.$$

It just remains to prove that (x_n) is converging, and this will be done checking that (x_n) is Cauchy. For that aim, firstly observe that

$$d(x_n, x_{n-1}) = d(f(x_{n-1}), f(x_{n-2})) \leq \lambda\, d(x_{n-1}, x_{n-2}).$$

Recursively we have

$$d(x_n, x_{n-1}) \leq \lambda^{n-2} d(x_2, x_1).$$

Triangle inequality gives us for $n > m \geq 1$ that

$$d(x_n, x_m) \leq d(x_n, x_{n-1}) + \cdots + d(x_{m+1}, x_m)$$

$$\leq (\lambda^{n-2} + \cdots + \lambda^{m-1})\, d(x_2, x_1) \leq \frac{d(x_2, x_1)}{1 - \lambda}\, \lambda^{m-1}.$$

The inequality clearly implies that (x_n) is Cauchy. ∎

1.4 Compactness

The compactness is one of the most important properties for Analysis.

Definition 1.4.1. *A topological space is said to be compact if any open cover has a finite subcover.*

Passing to complement sets compactness is equivalent to the following property: *a family of closed sets has nonempty intersection whenever it has the finite intersection property*, that is, if any of its finite subfamilies have nonempty intersection. Note as well that compactness is preserved by continuous maps.

Proposition 1.4.2. *"Compactness versus countable compactness".*

1. *In a compact topological space any infinite subset has a cluster point.*
2. *If topological space satisfies that any infinite subset has a cluster point, then any countable open cover has a finite subcover.*

Proof. Suppose that the infinite set $A \subset X$ has not cluster points. Then, for any subset $B \subset A$ the set $A \setminus B$ is closed. Now note that

$$\{A \setminus B : A \supset B \text{ is finite}\}$$

is a family of closed subsets with the finite intersection property and empty intersection.
For the second statement it is enough to show that a decreasing sequence (F_n) of nonempty closed subsets of X has nonempty intersection. Indeed, take a point $x_n \in F_n$. If (x_n) is finite, then $x = x_n$ for infinitely many $n \in \mathbb{N}$ and so $x \in \bigcap_{n=1}^{\infty} F_n$. Otherwise, (x_n) is infinite and thus it has a cluster point x which is a cluster point of any set $\{x_k : k \geq n\} \subset F_n$. Therefore $x \in \bigcap_{n=1}^{\infty} F_n$. In any case, the intersection $\bigcap_{n=1}^{\infty} F_n$ is nonempty. ∎

Proposition 1.4.3. *For a metric space M, the following are equivalent:*

(i) M is compact;

(ii) any sequence in M has a convergent subsequence;

(iii) M is complete and totally bounded.

Proof. (i)⇒(ii) Clearly we may assume that the sequence has infinitely many points, and so it has a cluster point as infinite set. In a metric space, a cluster point of a sequence is the limit of some of its subsequences.
(ii)⇒(i) Note that any infinite subset of M has a cluster point. Note as well that the metric space M must be separable since otherwise it would contain an uncountable metrically discrete subset, and any sequence of different points made from that set has no convergent subsequence. Now M is Lindelöf, any open cover has a countable subcover. By previous result, this countable cover has further a finite subcover.
(i)+(ii)⇒(iii) Given $\varepsilon > 0$, then $\{B(x,\varepsilon) : x \in M\}$ is an open cover of M, which has a finite subcover of the form $\{B(x_k,\varepsilon) : 1 \leq k \leq n\}$. Clearly, $\{x_k : 1 \leq k \leq n\}$ is a finite ε-net of M. Given a Cauchy sequence (x_n), it has a convergent subsequence with limit $x \in M$. That implies that the whole sequence (x_n) is converging to x.
(iii)⇒(ii) Suppose we are given a sequence $(x_n) \subset M$. Since M is covered by finitely many balls of radius $1/2$, there is at least one that contains infinitely many terms of the sequence (x_n), that is, there is $A_1 \subset \mathbb{N}$ infinite such that $d(x_n, x_m) \leq 1$ for $n, m \in A_1$. With the same argument, we can find $A_2 \subset A_1$ infinite such that $d(x_n, x_m) \leq 1/2$ for $n, m \in A_2$. Proceeding in this way, we will have $A_1 \supset A_2 \supset \cdots \supset A_k \supset \ldots$ all infinite such that if $x_n, x_m \in A_k$ then $d(x_n, x_m) \leq 1/2^k$. Now, we may take inductively $n_1 < n_2 < \cdots < n_k < \ldots$ such that $x_{n_k} \in A_k$, the construction shows that (x_{n_k}) is a Cauchy sequence, which should be convergent by the completeness of M. ∎

The characterization of compactness in $\mathbb{R}^n$ follows straight. For the time being, $\mathbb{R}^n$ is endowed with the Euclidean metric.

Corollary 1.4.4 (Heine-Borel)**.** *A subset of $\mathbb{R}^n$ is compact if and only if it is bounded and closed.*

Proof. Note that being $\mathbb{R}^n$ complete, the crux of the proof is to prove that in $\mathbb{R}^n$ boundedness and total boundedness is the same, which is essentially reduced to the Archimedean property of $\mathbb{R}$. ∎

Let us put together two classic results.

Theorem 1.4.5 (Heine - Weierstrass)**.** *A real function defined on a compact metric space is bounded, uniformly continuous and attains its maximum and its minimum.*

Proof. The readers should be able to prove that by theirselves.

Next results provides the so called Lebesgue's number of a covering.

Proposition 1.4.6 (Lebesgue)**.** *Let (M, d) be a metric compact space and let $\{U_i\}_{i\in I}$ be an open cover of M. Then there is $\xi > 0$ such that for any $x \in M$ there is some $i \in I$ such that $B(x, \xi) \subset U_i$.*

Proof. We may assume without loss of generality that the cover is finite. The functions $f_i(x) = d(x, M \setminus U_i)$ are continuous and $f_i(x) > 0$ if and only if $x \in U_i$. Therefore, the function $f(x) = \min\{f_i(x) : i \in I\}$ is strictly positive on M. Let $\xi > 0$ be the infimum value of f on M. A direct computation shows that ξ has the desired property. ∎

The product of compact spaces is compact in a wide topological context. Here we have the following which will be enough for some applications. Recall that the product topology is the coarser for which the projections are continuous.

Proposition 1.4.7. *The finite or countable product of compact metric spaces is also metrizable and compact with the product topology.*

Hint of proof. For the finite case it is enough to consider the product of two spaces (M_1, d_1) and (M_2, d_2). Endow $M_1 \times M_2$ with the metric

$$d((x_1, x_2), (y_1, y_2)) = d_1(x_1, y_1) + d_2(x_2, y_2).$$

With the topology associated to that metric the projections are continuous and a the diagonal method shows that any sequence has a convergent subsequence. A coarser topology coincides with that one. The infinite product case is done in similar way but with a metric defined by a series

$$d((x_n), (y_n)) = \sum_{n=1}^{\infty} 2^{-n} d_n(x_n, y_n)$$

where d_n is an equivalent metric on M_n bounded by 1. Again, compactness of the metric topology can be show by a diagonal method. ∎

1.5 Space of continuous functions

Take $C(K)$ the set of continuous real functions defined on K and define $\|f\|_\infty = \sup\{|f(x)| : x \in K\} < +\infty$ as the infimum is attained. Endow $C(K)$ with the metric $d(f, g) = \|f - g\|_\infty$ called the uniform metric.

Proposition 1.5.1. *The space $(C(K), \|\cdot\|_\infty)$ is complete.*

Proof. If (f_n) is a Cauchy sequence in $C(K)$, then $(f_n(x))$ is a convergent sequence in $\mathbb{R}$ for every $x \in K$ thus defining a real function by the formula $f(x) = \lim_n f_n(x)$. We claim that f is continuous. Indeed, given $\varepsilon > 0$ take $N \in \mathbb{N}$ such that if $n, m \geq N$ then $\|f_n - f_m\|_\infty < \varepsilon/3$. Then

$$|f_n(x) - f(x)| = \lim_m |f_n(x) - f_m(x)| \leq \varepsilon/3$$

for every $n \geq N$. Fix $U \ni x$ neighborhood such that $|f_N(y) - f_N(x)| < \varepsilon/3$ if $y \in U$. Triangle inequality gives that $|f(y) - f(x)| < \varepsilon$ for $y \in U$. Finally, the above inequality also gives that $\|f_n - f\|_\infty \leq \varepsilon/3$ for any $n \geq N$ which implies the convergence in the uniform distance to f of (f_n). ∎

The relation of pointwise and uniform convergence is delicate. The following is a quite useful result.

Theorem 1.5.2 (Dini)**.** *Let $(f_n) \subset C(K)$ a sequence of functions that converges to some $f \in C(K)$. If the sequence (f_n) is monotone (increasing or decreasing), then (f_n) converges uniformly to f.*

Hint of proof. Assume that (f_n) is decreasing, for instance. Then fix $\varepsilon > 0$ and prove that the sequence of sets

$$U_n = \{x : f_n(x) - f(x) > \varepsilon\}$$

is an open cover of K. ∎

Finally, we will prove this useful characterization of compactness for sets of continuous functions.

Theorem 1.5.3 (Arzelà-Ascoli)**.** *A subset $A \subset C(K)$ is compact if and only if its closed, bounded and equicontinuous.*

Proof. Obviously, compactness of A implies it is closed and bounded, to see that A is equicontinuous we will use its total boundedness. Given $\varepsilon > 0$ take $\{f_k\}_{k=1}^n \subset A$ a $\varepsilon/3$-net of A. Fixed $x \in K$, we may find $U \ni x$ open such that $\max_k |f_k(y) - f_k(x)| < \varepsilon/3$ whenever $y \in U$. Triangle inequality implies that $|f(y) - f(x)| < \varepsilon$ for any $f \in A$ and $x \in U$.
If A is equicontinuous, given $\varepsilon > 0$ then for every $x \in K$ there is $U_x \ni x$ open such that $|f(y) - f(x)| < \varepsilon/3$ whenever $f \in A$ and $y \in U_x$. Note that $\{U_x\}_{x \in K}$

is an open cover of K, so there are points $\{x_k\}_{k=1}^n$ such that $K = \bigcup_{k=1}^n U_{x_k}$. Since A is bounded, the set

$$\{(f(x_1), f(x_2), \ldots, f(x_n)) : f \in A\}$$

is bounded in $\mathbb{R}^n$ and therefore it has a $\varepsilon/3$-net (with the maximum distance) that we will denote $(\lambda_i)_{i=1}^N$ where $\lambda_i = (\lambda_i(1), \ldots, \lambda_i(n))$. Take $f_i \in A$ such that $f_i(x_k) = \lambda_i(k)$ for every $i \in \{1, \ldots, N\}$. We claim that $\{f_i\}_{i=1}^N$ is a ε-net for A. Indeed, given $f \in A$, there is i such that $|f(x_k) - f_i(x_k)| = |f(x_k) - \lambda_i(k)| < \varepsilon$ for every $k \in \{1, \ldots, n\}$. For any $x \in K$, there is some k such that $x \in U_{x_k}$. Therefore $|f(x) - f(x_k)| < \varepsilon/3$ and $|f_i(x) - f_i(x_k)| < \varepsilon/3$. Triangle inequality gives that $|f(x) - f_i(x)| < \varepsilon$ for arbitrary $x \in K$, that is, the ε-net property.■

1.6 Fractals

Let (M, d) be a complete metric space and denote by $\mathcal{K}(M)$ the set of nonempty compact subsets of M. For $A \in \mathcal{K}(M)$ and $r > 0$ define a closed "neighbourhood" as

$$D[A, r] = \{x \in M : d(A, x) \leq r\}.$$

And now a distance between elements from $\mathcal{K}(M)$ by

$$d(A, B) = \inf\{r > 0 : A \subset D[B, r], B \subset D[A, r]\}$$

where $A, B \in \mathcal{K}(M)$. There is no problem in using d for the distance on $\mathcal{K}(M)$ since it extends the metric of M considering the points as (compact) singletons. It is easy to verify that d is a metric in $\mathcal{K}(M)$. Indeed, clearly the less evident fact is the triangle inequality. If $A, B, C \in \mathcal{K}(M)$ and $\varepsilon > 0$ we may find $r < d(A, B) + \varepsilon$ and $s < d(B, C) + \varepsilon$ such that

$$A \subset D(B, r) \quad \text{and} \quad B \subset D(C, s).$$

That implies $A \subset D(C, r + s)$. The reverse containment is obtained likewise, thus

$$d(A, C) \leq r + s \leq d(A, B) + d(B, C) + 2\varepsilon$$

that proves the claim as ε was arbitrary. The metric d is known as the *Hausdorff metric*.

Our objective is the following,

Theorem 1.6.1. *If M is a complete metric space, then $(\mathcal{K}(M), d)$ is complete.*

Proof. Consider a Cauchy sequence $(A_n) \subset \mathcal{K}(M)$. We claim that for any choice $x_n \in A_n$, the sequence (x_n) has a cluster point. Indeed, fix $\varepsilon > 0$ and let n_ε such that if $n \geq n_\varepsilon$ then $A_n \subset D[A_{n_\varepsilon}, \varepsilon]$, and thus $(x_n) \subset D[A_{n_\varepsilon}, \varepsilon]$ except finitely many terms. It is clear that $D[A_{n_\varepsilon}, \varepsilon]$ can be covered by finitely many balls of radius $(3/2)\varepsilon$ and infinitely many x_n's are inside of one of those balls. Therefore $d(x_{n_k}, x_{n_j}) \leq 3\varepsilon$ for some subsequence (x_{n_k}). This selection process applied for $\varepsilon = 1/m$ and further diagonal argument will produce a Cauchy subsequence of (x_n)
Let $A \subset M$ be the set of all the cluster points of sequences obtained as before. Note that if $x \in A$, we can take $x_n \in A_n$ such that (x_n) converges to x. Note as well that A has to be closed since any cluster point of A can be reached by a suitable diagonal choice. Now, for $\varepsilon > 0$ note that $A \subset D(A_n, \varepsilon)$ and n large enough. That implies that A is totally bounded, and therefore $A \in \mathcal{K}(M)$. Also implies that A is "half limit" of (A_n). In order to complete, the proof we have to prove that $A_n \subset D(A, \varepsilon)$ for n large. If it is not the case, we can take $x_n \in A_n$ such that $d(A, x) \geq \varepsilon$ for infinitely many n's. That would produce a cluster point x such that $d(A, x) \geq \varepsilon$. On the other hand, by definition of A we have $x \in A$. The contradiction proves the theorem. ∎

Observe that if $f : M \to M$ is contractive, then the induced map $f : \mathcal{K}(M) \to \mathcal{K}(M)$ just taking $A \to f(A)$ is also contractive. Indeed, if $\lambda \in (0, 1)$ is the contraction ratio, then

$$f(D[A, r]) \subset D[f(A), \lambda r]$$

which implies $d(f(A), f(B)) \leq \lambda d(A, B)$, that is, f is contractive in $\mathcal{K}(M)$ too. Since a contractive map can have only a fixed point, that one has to be the singleton fixed by f in M. Nevertheless, on $\mathcal{K}(M)$ we can perform other operations with maps.

Proposition 1.6.2 (Iterated function system)**.** *Let $f_1, \ldots, f_n$ be contractive maps on M, and define $f : \mathcal{K}(M) \to \mathcal{K}(M)$ by*

$$f(A) = f_1(A) \cup \cdots \cup f_n(A).$$

Then f is contractive on $\mathcal{K}(M)$, and therefore f has a fixed point.

Proof. Indeed, if λ is the maximum of the contraction ratios we still have the containment

$$f(D[A, r]) = \bigcup_{k=1}^{n} f_k(D[A, r]) \subset \bigcup_{k=1}^{n} D[f_k(A), \lambda r]) = D[f(A), \lambda r]$$

which implies the contractivity of f. ■

Taking $f_1, \ldots, f_n$ affine and contractive on $\mathbb{R}^2$ we can obtain several *self-similar* typical fractals. Look up on Google for the *Barnsley fern algorithm*.

1.7 Rationale and remarks

Most likely the students already know some metric topology, so it is not necessary to stress on bizarre examples. The idea is to recall the role and importance of completeness and compactness. A curious application of Baire's theorem, the construction of fractals or the characterization of compactness on $C[a, b]$ can help the students to take topology more seriously.

1.8 Exercises

1. Prove that closed balls are closed sets, and open balls are open sets.

2. Prove that uniformly continuous maps between metric spaces preserve Cauchy sequences.

3. The distance $d(A, x)$ from a point x to a set $A \in M$ in a metric space is defined by $d(A, x) = \inf\{d(y, x) : y \in A\}$. Prove the following statements:

   (a) $|d(A, x) - d(A, y)| \leq d(x, y)$,
   (b) $\overline{A} = \{x \in X : d(A, x) = 0\}$,
   (c) $d(A, x) \leq d(B, x)$ if and only if $B \subset \overline{A}$.

4. Prove with the help of Dini's theorem that the uniform convergence on bounded subsets of the sequence

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

5. Define inductively a sequence of functions on $[0, 1]$ by $f_1(x) = 0$ and $f_{n+1}(x) = f_n(x) + \frac{1}{2}(x - f_n(x)^2)$. Prove that $(f_n(x))$ uniformly converges to $\sqrt{x}$. Deduce as a consequence that the function $|x|$ can be uniformly approached by polynomials on bounded intervals of $\mathbb{R}$.

6. Prove that separability of a metric space is hereditary to subsets.

7. A set in a metric space is said to be *perfect* if it has no isolated point. Given a set $A \subset M$, a point $x \in A$ is said of *condensation* if every of its neighbourhoods meets A at an uncountable set. Assume that the metric space M is separable. Show that for every $A \subset M$, the subset of the non-condensation points of A is countable, and the subset of condensation points is perfect. Deduce that M can be expressed as the union of a perfect set and a countable set.

8. Let M, N be complete metric spaces, $D \subset M$ a dense subset and $f : D \to N$ a uniformly continuous function. Prove that f can be extended to a unique uniformly continuous function $\tilde{f} : M \to N$. Show that if f is only continuous, then the extension can be done to a set $B \subset A$ such that $B \subset \overline{A}$ and $B = \bigcap_{n=1}^{\infty} U_n$ being (U_n) a sequence of open subsets of M.

9. Let $f \in C^{\infty}(\mathbb{R})$ be a function such that for every $x \in \mathbb{R}$ there is $n \in \mathbb{N}$ such that $f^{(n)}(x) = 0$. Prove that f is a polynomial.

10. Prove the following abstract version of the of Cantor diagonal method. Let (A_n) be a decreasing sequence of infinite subsets of $\mathbb{N}$. Then there is an infinite subset $A \subset \mathbb{N}$ such that $A \setminus A_n$ is finite for all $n \in \mathbb{N}$.

11. Prove that if M is a metric compact, then $\mathcal{K}(M)$ is compact too.

Chapter 2

Normed Spaces

2.1 Norms

A basic notion in Analysis is the notion of *normed space*, which is just a vector space together a *norm*. Let X be a vector space (either on $\mathbb{R}$ or $\mathbb{C}$, say $\mathbb{K}$). A function $\|\cdot\| : X \to [0, +\infty)$ is called a norm if:

1. $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$;
2. $\|\lambda x\| = |\lambda|\|x\|$ for all $x \in X$, $\lambda \in \mathbb{K}$;
3. $\|x\| = 0$ if and only if $x = 0$.

A norm induces a distance on X by means of $d(x, y) = \|x - y\|$, and that provides a topological structure, as a metric space. There are several weakenings of the notion of norm which are also interesting, see the section "Complements". Sometimes, we use $(X, \|\cdot\|)$ to denote a normed space, however that is not necessary when the norm we are dealing with is understood.

From now on we will focused on real normed spaces. The notation for open and closed balls will be the same that within metric spaces, however we will distinguish the unit ball

$$B_X := B[0, 1] = \{x \in X : \|x\| \leq 1\}.$$

All the closed balls in X can be obtained by translation and scaling of B_X. Note that the unit sphere

$$S_X := \{x \in X : \|x\| = 1\}$$

is the topological boundary of B_X and so the interior of B_X is exactly $B(0,1)$. That is not generally true in metric spaces.

Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on the same vector space are said *equivalent* is they generate the same topology. A nice consequence of the similarity of balls is the nice characterization of the equivalence of norms.

Proposition 2.1.1. *Let X be a vector space and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on X. Then the norms are equivalent if and only if there are constants $\alpha, \beta > 0$ such that*

$$\alpha\|x\|_1 \leq \|x\|_2 \leq \beta\|x\|_1$$

for all $x \in X$.

The notion of completeness has several nuances. Firstly, a consequence of the former Proposition.

Corollary 2.1.2. *The completeness (or its absence) of a normed space is invariant among the equivalent norms.*

Now we will consider series in normed spaces. As in the real case, a series $\sum_{n=1}^{\infty} x_n$ is just a symbolic expression. The series is said to be *convergent* if the partial sums $s_n = \sum_{k=1}^{n} x_k$ converge to some element in X called the sum of the series. We say that a series is *unconditionally convergent* if any rearrangement of its terms is convergent with the same sum. Finally, a series $\sum_{n=1}^{\infty} x_n$ is said to be *absolutely convergent* if $\sum_{n=1}^{\infty} \|x_n\| < +\infty$. Despite the name, absolute convergence does not always imply convergence.

Proposition 2.1.3. *A normed space $(X, \|\cdot\|)$ is complete if and only if every absolutely convergent series is convergent. In such a case, the series will also be unconditionally convergent.*

2.2 Finite-dimensional normed spaces

A vector space X of finite dimension n is algebraically isomorphic to $\mathbb{R}^n$ (or $\mathbb{C}^n$ in the complex case, that we will not consider here). The isomorphism is determined by fixing a basis $\{e_1, \ldots, e_n\}$ and after that we may consider as defined on X any of the functions on $\mathbb{R}^n$, in particular any of the standard norms. A key fact that we will use is that the subsets of $\mathbb{R}^n$ which are closed and bounded with respect the Euclidean norm are compact.

Theorem 2.2.1. *All the norms on a finite dimensional space X are equivalent.*

Proof. We will denote by $\|\cdot\|_2$ the Euclidean norm on X given by the isomorphism associated to a basis $\{e_1, \ldots, e_n\}$. Let $\|\cdot\|$ be an arbitrary norm on X. We will show that $\|\cdot\|$ is continuous as a function on $(X, \|\cdot\|_2)$. Indeed, note that

$$\|x\| = \|\lambda_1 e_1 + \cdots + \lambda_n e_n\| \leq |\lambda_1|\|e_1\| + \cdots + |\lambda_n|\|e_n\|$$
$$\leq (|\lambda_1|^2 + \cdots + |\lambda_n|^2)^{1/2}(\|e_1\|^2 + \cdots + \|e_n\|^2)^{1/2} = c\|x\|_2$$

by the Cauchy-Schwarz inequality and taking $c = (\|e_1\|^2 + \cdots + \|e_n\|^2)^{1/2}$. Now, we have

$$|\|x\| - \|y\|| \leq \|x - y\| \leq c\|x - y\|_2$$

that means that $\|\cdot\|$ is Lipschitz (with constant c) with respect to $\|\cdot\|_2$, and thus continuous as wanted. Let α and β the minimum and maximum respectively of $\|\cdot\|$ on the set $S = \{x \in X : \|x\|_2 = 1\}$. We have $\alpha > 0$ since $\|\cdot\|$ is a norm and S does not contain 0. If $x \in X \setminus \{0\}$ then $x/\|x\|_2 \in S$ and thus

$$\alpha \leq \left\| \frac{x}{\|x\|_2} \right\| \leq \beta$$

and so

$$\alpha\|x\|_2 \leq \|x\| \leq \beta\|x\|_2$$

which is the desired equivalence. ∎

Once we know that any norm on $\mathbb{R}^n$ is equivalent to the Euclidean norm $\|\cdot\|_2$, we have freedom to work with alternative norms that do not involve square roots, such as

$$\|(x_1, \ldots, x_n)\|_1 := |x_1| + \cdots + |x_n|,$$
$$\|(x_1, \ldots, x_n)\|_\infty := \max\{|x_1|, \ldots, |x_n|\}.$$

Corollary 2.2.2. *The finite dimensional subspace of a normed space are closed.*

Proof. The restriction of the norm to a finite dimensional subspace is equivalent to the Euclidean norm and so it is complete. As a complete subset it is closed in the overspace. ∎

Corollary 2.2.3. *Let X be a normed space and let $Y \subset X$ be a finite dimensional subspace. Then, for every $x \in X$ there is $y \in Y$ such that*

$$\|x - y\| = d(x, Y) := \inf\{\|x - z\| : z \in Y\}.$$

Proof. Note that $f(z) = \|z-x\|$ is a continuous function on Y whose infimum can be computed on a bounded subset. ∎

Now we will prove the existence of "almost orthogonal" elements in normed spaces.

Proposition 2.2.4. *Let X be a normed space, $Y \subset X$ a proper closed subspace and $\varepsilon \in (0,1)$. Then, there exists $x \in X \setminus Y$ with $\|x\| = 1$ and $d(x,Y) > 1-\varepsilon$. In case that Y is of finite dimension, then x can be taken such that $d(x,Y) = 1$.*

Proof. Take $x_0 \in X \setminus Y$. Then $d(Y, x_0) = d > 0$, and we may take $y_0 \in Y$ such that

$$d \leq \|x_0 - y_0\| \leq \frac{d}{1-\varepsilon},$$

and put

$$x := \frac{x_0 - y_0}{\|x_0 - y_0\|}.$$

Note that $x \in S_X$. Now we will estimate $d(Y, x)$. If $y \in Y$, note that

$$\|x - y\| = \left\| \frac{x_0 - y_0}{\|x_0 - y_0\|} - y \right\|$$

$$= \frac{1}{\|x_0 - y_0\|} \|(x_0 - y_0 - \|x_0 - y_0\| y)\| \geq \frac{d}{\|x_0 - y_0\|} \geq 1 - \varepsilon,$$

because $y_0 + \|x_0 - y_0\| y \in Y$. In case X is finite-dimensional we could have taken y_0 such that $\|x_0 - y_0\| = d$ which would have led to an equality. ∎

Corollary 2.2.5. *If X is an infinite-dimensional normed space, then there exists an infinite sequence $(x_n) \subset B_X$ such that $\|x_n - x_m\| \geq 1$ for $n, m \in \mathbb{N}$ with $n \neq m$.*

Proof. The construction is inductive: take any $x_1 \in S_X$. Assume $x_1, \ldots, x_n \in S_X$ already chosen. Let Y the finite dimensional subspace spanned by those elements. Clearly, $Y \neq X$, thus we can find $x_{n+1} \in S_X$ such that $d(Y, x_{n+1}) = 1$ by the proposition. In particular $\|x_k - x_{n+1}\| \geq 1$ for $1 \leq k \leq n$. ∎

This result is the culmination of the section.

Theorem 2.2.6. *A normed space X has finite dimension if and only if its unit ball B_X is compact.*

Proof. If X has finite dimension n, then it is isomorphic to $\mathbb{R}^n$, and therefore the unit ball, as closed bounded set, is compact. On the other hand, if X has infinite dimension, then B_X contains a sequence with no convergent subsequence. In such a case, B_X cannot be compact. ∎

Finite dimensional spaces are also characterized by the fact that unconditionally convergent series are absolutely convergent. On implication is clear, the other one is the celebrated *Dvoretsky-Rogers* theorem.

2.3 Linear operators

Linear continuous maps between normed spaces, also called *operators* are essential for the development of the theory. Firstly note the following.

Proposition 2.3.1. *Let X, Y be normed spaces (on the same field) and $T : X \to Y$ be linear. Then the following are equivalent:*

1. *T is continuous;*
2. *T is continuous at 0;*
3. *there is $c > 0$ such that $\|T(x)\| \leq c\|x\|$ for every $x \in X$;*
4. *$T(B_X)$ is a bounded set in Y.*

Proof. Note that continuity at one point for a linear function equals global continuity, *1⇔2*. Also *3⇒4*. The main trick to use the homogeneity to show that the continuity at 0 implies the boundedness of $T(B_X)$. If $c = \sup\{\|y\| : y \in T(B_X)\}$, again by homogeneity, we can deduce that $\|T(x)\| \leq c\,\|x\|$. ∎

A similar statement can be proved for bilinear or multilinear maps. The set of continuous operators from X to Y is denoted $\mathfrak{L}(X, Y)$. Note tha t$\mathfrak{L}(X, Y)$ becomes a normed space with the norm

$$\|T\| = \sup\{\|T(x)\| : x \in B_X\}$$

This norm inherits the completeness from Y, that is $\mathfrak{L}(X, Y)$ is a Banach space if and only if Y is.

The norm, by its very definition has the following remarkable property: If $T \in \mathfrak{L}(X,Y)$ and $S \in \mathfrak{L}(Y,Z)$, then $S \circ T \in \mathfrak{L}(X,Z)$ and $\|S \circ T\| \leq \|S\|\|T\|$.

Let us stress two particular interesting cases of spaces $\mathfrak{L}(X,Y)$. If $Y = \mathbb{K}$, that is the scalar field, then we set $X^* := \mathfrak{L}(X,\mathbb{K})$. The space X^*, which is always complete, is called the dual space of X. Duality theory studies how properties of a space induce properties of its duals, and viceversa. This is sometimes very useful in linear optimization problems. Note that for any $x \in X$ and $x^* \in X^*$ we always have $|x^*(x)| \leq \|x^*\|\|x\|$.

The other case is when $Y = X$, where we prefer the notation $\mathfrak{L}(X) := \mathfrak{L}(X,X)$, that enables with the structure of *algebra* (that will be important for spectral operator theory) since any two operators $T, S \in \mathfrak{L}(X)$ can always be composed, that is, $ST, TS \in \mathfrak{L}(X)$, and moreover $\|ST\| \leq \|S\|\|T\|$. Here we may consider the *invertible* operators: $T \in \mathfrak{L}(X)$ is said invertible if there exists $S \in \mathfrak{L}(X)$ such that $TS = I$ and $ST = I$, being I the identity operator on X. In such a case, S is unique and we denote $T^{-1} := S$. The invertible operators are also called *isomorphisms* when we want to express that they preserve the linear and topological structure of X. An operator T is called an *isometry* if it satisfies $\|T(x)\| = \|x\|$ for all $x \in X$. It is not difficult to see that the isometries are isomorphisms.

Finally, a very important observation in finite-dimensional spaces.

Proposition 2.3.2. *Every linear o multilinear map defined on a finite dimensional normed space is continuous.*

Proof. Using a base it is possible to write the map in terms of coordinates and so an explicit bound on the unit ball can be proved easily for the norms $\|\cdot\|_1$ or $\|\cdot\|_\infty$. Since all the norms are equivalent on the domain space, we deduce the statement. ∎

2.4 Spaces of functions

The uniform convergence of sequences and series of functions can be understood in the frame of normed spaces. Let us denote by $\ell_\infty(M)$ the set of all bounded real functions defined on a set M. For $f \in \ell_\infty(M)$ we denote

$$\|f\|_\infty = \sup\{|f(x) : x \in M|\}.$$

That norm makes $\ell_\infty(M)$ a complete normed space and the induced topology is usually referred as the topology of uniform convergence. When M has an additional structure as being a metric space we can study the properties that are preserved by uniform limits. We already know that it is the case with the continuity. Something more general can be said.

Proposition 2.4.1. *Assume that $(f_n) \subset \ell_\infty(M)$ and $(x_m) \subset M$ are such that:*

1. *the limit of (f_n) exists uniformly;*
2. *$\lim_m f_n(x_m)$ exists for every $n \in \mathbb{N}$.*

Then the following iterated limits exist and satisfies the equality

$$\lim_n (\lim_m f_n(x_m)) = \lim_m (\lim_n f_n(x_m)).$$

A great deal of Analysis is devoted to commutation of limits, understanding by "limit" some operations in Analysis that rely on the notion of limit as series, derivatives and integrals. The commutation of limits of sequences and series of functions with the integral will be treated in the corresponding chapter. As to the commutation of derivatives and limits of sequences, it is easy to put examples of the failing. However, the situation is more dramatic: the only linear spaces of C^1 functions where we can expect a good behaviour of limits are finite dimensional.

Theorem 2.4.2. *Let K be a compact metric space and $X \subset C(K)$ a (closed) subspace made up of Lipschitz functions. Then X is finite-dimensional.*

Hint of proof. Use Baire's theorem to show that the Lipschitz constant is bounded on B_X. Now, B_X is closed, bounded and equicontinuous, so by Arzèla-Ascoli B_X is compact. That implies that X has finite dimension. ∎

Despite the fact that a uniform limit of Lipschitz functions could not be Lipschitz, as for instance, the sequence $f_n(x) = n^{-1} \sin n^2 x$ on $[0, \pi]$, it is possible to endow the set of Lipschitz functions with a norm that makes it complete. Indeed, denote by $L(M)$ the set of Lipschitz functions defined on the metric space M and fix a point $x_0 \in M$. Then the number

$$\|f\| = |f(x_0)| + \sup \left\{ \frac{|f(x) - f(y)|}{d(x, y)} : x, y \in M, x \neq y \} \right\}$$

defines a norm on $L(M)$ that makes it complete. We left the proof to the reader. A variation for differentiable functions is asked among the exercises of the chapter.

2.5 Complements

In this section we include, without proof, several results that traditionally are reserved as topics for *Functional Analysis*, despite some proofs are accesible to this level.

A function $p : X \to [0, +\infty)$ is said to be sublinear if it satisfies the properties:

(a) $p(tx) = tp(x)$ for $t \geq 0$ and any $x \in X$ (positive homogeneity);

(b) $p(x + y) \leq p(x) + p(y)$ for all $x, y \in X$ (triangle inequality).

If p satisfies the stronger property

(A) $p(tx) = |t|p(x)$ for $t \in \mathbb{K}$ and any $x \in X$ (homogeneity)

then p is said to be a seminorm. Of course, p will be a norm provided that $p(x) = 0$ if and only if $x = 0$. Note that if

1. if p is sublinear then $x \to \max\{p(x), p(-x)\}$ is a seminorm,
2. if p is a seminorm, then $|p(x) - p(y)| \leq p(x - y)$.

A Hausdorff topology can be induced also by a family of seminorms $(p_i)_{i \in I}$ if for every $x \in X \setminus \{0\}$ there is some $i \in I$ such that $p_i(x) > 0$. In this case, a basis of neighbourhoods of a point x can be defined by

$$\{y \in X : p_{i_1}(y - x) < \varepsilon, \ldots, p_{i_n}(y - x) < \varepsilon\}$$

for $i_1, \ldots, i_n \in I$ and $\varepsilon > 0$. It is possible to prove that if a sequence of seminorns (p_n) induces a Hausdorff topology then it can be metrized by the *translation invariant* metric

$$d(x, y) = \sum_{n=1}^{\infty} 2^{-n} \min\{1, p_n(x - y)\}.$$

The topologies defined by a family of seminorms appear quite often and they will be considered in other chapters, for instance, uniform convergence on compact subsets of $\mathbb{R}^n$.

The Hahn-Banach theorem guarantees the existence of extensions of linear forms under very general conditions.

Theorem 2.5.1 (Hahn-Banach)**.** *Let X be a real vector space, $Y \subset X$ a subspace, p a sublinear homogeneous functional defined on X and f a linear form defined on Y such that $f(x) \leq p(x)$ for every $x \in Y$. Then there is a linear form $\tilde{f}$ defined on X such that $\tilde{f}|_Y = f$ and $\tilde{f}(x) \leq p(x)$ for all $x \in X$.*

The formulation with a sublinear functional is the key to prove separation results for convex sets, which is a matter we are not interested here. Nevertheless, if p is of the form $c\|\cdot\|$, we obtain the following consequence.

Theorem 2.5.2 (Hahn-Banach)**.** *Let X a normed space and $x \in X$. There exists $x^* \in X^*$ with $\|x^*\| = 1$ such that $x^*(x) = \|x\|$.*

Note that the result informally says that the dual X^* is, at least, as large as X, which was quite evident for a finite dimensional X. A straightforward application of that result says that X embeds isometrically into $X^{**} := (X^*)^*$. The closure of X as a subset of X^{**} provides a model for the completion of X.

Corollary 2.5.3. *Every normed space can be isometrically embedded into a complete normed space as a dense subset.*

A completion can be built "more directly" as a quotient of the space of Cauchy sequences.

In the finite dimensional case, the following result of Auerbach is very interesting, and far from being obvious.

Theorem 2.5.4 (Auerbach)**.** *Let X a normed space of dimension n. There exist bases $\{x_1, \ldots, x_n\}$ of X and $\{x_1^*, \ldots, x_n^*\}$ of X^* such that $\|x_k\| = \|x_k^*\| = 1$ for $1 \leq k \leq n$ and $x_j^*(x_k) = \delta_{jk}$ and $x \in X$. There exists $x^* \in X^*$ with $\|x^*\| = 1$ such that $x^*(x) = \|x\|$.*

For X a Banach space, the property enjoyed by $\mathfrak{L}(X)$ and $C(K)$, of being both Banach space and algebra (there is a *product* compatible with the sum) with the additional property $\|xy\| \leq \|x\|\|y\|$ leads to the notion of Banach algebras, whose theory is richer than the one of normed spaces, specially when the algebra is considered over the complex field.

2.6 Rationale and remarks

Normed spaces are the frame for the differential calculus in the next chapter. Once the metric topology is clear, it is necessary to point out that the normed

spaces have a richer theory. In particular, the properties of finite normed spaces (or subspaces) should be remarked.

Some aspects of the convergence of sequences and series of functions are better understood in the frame of normed spaces because the uniform convergence is a metric one. However, the most important cases, power and trigonometric series, have a particular treatment in other subjects along the degree studies.

2.7 Exercises

1. Prove that the notions of boundedness, Cauchy sequence and completeness are invariant by equivalence of norms. Show with an example that the same does not hold in general metric spaces.

2. For points $x, y \subset X$ in a normed space, the *segment* joining them is the set
$$[x, y] = \{\lambda\, x + (1 - \lambda)\, y : 0 \leq \lambda \leq 1\}.$$
A set $A \subset X$ is said *convex* if for every $x, y \in A$ then $[x, y] \subset A$. Prove that balls are convex sets and a closed set $A \subset M$ is convex if and only if $\frac{1}{2}(x + y) \in A$ for every $x, y \in A$.

3. Recall that a metric space M is said to be *connected* if the only subsets that are both closed and open are M and $\emptyset$. Prove that any to points in a connected open set of a normed space can be joined by a *polygonal*, that is, made up of segments, continuous line.

4. Find the optimal constants $a, b > 0$ for the equivalence of norms in $\mathbb{R}^n$
$$a\|\cdot\|_2 \leq \|\cdot\|_1 \leq b\|\cdot\|_2$$

5. Define on $\mathbb{R}^n$ the norm $\|\cdot\|_p$ for $p \geq 1$ by the formula
$$\|(x_1, \ldots, x_n)\|_p = \sqrt[p]{|x_1|^p + \cdots + |x_n|^p}.$$
Prove that for $x \in \mathbb{R}^n$ and $p \leq q$, then $\|x\|_p \geq \|x\|_q$ and
$$\lim_{p\to\infty} \|x\|_p = \|x\|_\infty.$$

6. Define on $C[a,b]$ the norm $\|\cdot\|_p$ for $p \geq 1$ by the formula

$$\|f\|_p = \sqrt[p]{\int_a^b |f(t)|^p \, dt}$$

Prove that $\|\cdot\|_p$ is actually a norm for $p = 1, 2$ (the other cases are mor difficult). Show that $\|\cdot\|_p$ is not equivalent to $\|\cdot\|_q$ if $p \neq q$, on $C[a,b]$. Prove also that for any $f \in C[a,b]$ then

$$\lim_{p\to\infty} \|f\|_p = \|f\|_\infty.$$

7. Prove that for every $n \in \mathbb{N}$ there is a constant C_n such that for all the $n \times n$ matrices with non-negative entries $(a_{i,j})$ the following inequality is verified

$$\sum_{i=1}^{n} (\sum_{j=1}^{n} a_{i,j})^2 \leq C_n \sum_{j=1}^{n} (\sum_{i=1}^{n} a_{i,j})^2.$$

8. Prove that the following formula

$$\|f\| = |f(0)| + \|f'\|_\infty$$

defines a norm on $C^1[0,1]$. Show also that $C^1[0,1]$ endowed with such a norm is complete.

9. Let X be a normed space and consider the unit sphere $S = \{x \in X : \|x\| = 1\}$. Show that

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

10. A function $f : X \to \mathbb{R}$ is said to be convex if it satisfies

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

for every $x, y \in X$ and $\lambda \in (0,1)$. Prove that:

    (a) a norm is a convex function;

    (b) $f(x) = \|x\|^2$ is convex too;

    (c) $d(A,x)$ is convex if and only if $\overline{A}$ is convex.

11. Prove that $(C[a,b], \|\cdot\|_\infty)$ is separable.

12. The set of real bounded sequences is denoted ℓ^∞, and the formula

$$\|(x_n)_{n=1}^\infty\|_\infty = \sup\{|x_n| : n \in \mathbb{N}\}$$

for $(x_n)_{n=1}^\infty \in \ell^\infty$ defines a norm. Show that $(\ell^\infty, \|\cdot\|_\infty)$ is complete and non separable.

13. We say that a function $Q : X \to \mathbb{R}$ defined on a vector space is a *quadratic form* if there exists a bilinear form $B : X \times E \to \mathbb{R}$ such that $Q(x) = B(x, x)$. Show that B is not determined, in general, by Q. However, if we ask the bilinear form to be symmetric, that is, $B(x, y) = B(y, x)$ for all $x, y \in X$ then

$$B(x, y) = \frac{1}{2}(Q(x + y) - Q(x) - Q(y)).$$

Find the generalization of that result for *cubic forms* and *symmetric trilinear forms*.

14. Show that a quadratic form $Q : X \to \mathbb{R}$ is continuous if and only if there is $k > 0$ such that $|Q(x)| \leq k\|x\|^2$ for all $x \in X$.

15. Let X be a finite dimensional normed space.

    (a) Show that all the quadratic forms on X are continuous.

    (b) If $Q : X \to \mathbb{R}$ is a quadratic form such that $Q(x, x) > 0$ for all $x \neq 0$, then there is $a > 0$ such that $Q(x) \geq a\|x\|^2$.

16. Let $h : \mathbb{R} \to \mathbb{R}$ be a differentiable function. Study the limit of the sequence

$$f_n(x) = n(h(x + n^{-1}) - h(x)).$$

Find a necessary and sufficient condition on h for the uniform convergence of (f_n) on bounded intervals.

17. Consider the sequence of functions $f_n : [0, 1] \to \mathbb{R}$ defined for $p > 0$ by

$$f_n(x) = n^p x(1 - x^2)^n$$

Study the convergence of the sequence (pointwise and uniform) depending on the parameter p. Is possible to say what happens with the limit of the derivatives?

18. Find the set where the series

$$\sum_{n=0}^{\infty} e^{-nx}$$

converges. Find also the sets where the convergence is uniform.

19. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. Consider $F : \mathbb{R} \to C[a, b]$ defined by $F(t)(x) = f(x + t)$. Prove that F is continuos as a function on $(C[a, b], \|\cdot\|_\infty)$.

20. Study the convergence of the sequence of functions defined on $\mathbb{R}$ by

$$f_n(x) = \left(1 - \frac{x^2}{n}\right)^n \quad \text{if} \quad |x| \leq \sqrt{n}$$

and $f_n(x) = 0$ si $|x| > \sqrt{n}$.

21. Prove *Hadamard's radius formula*: Let

$$R = \frac{1}{\limsup_n \sqrt[n]{|a_n|}}.$$

Then the power series $\sum_{n=1}^{\infty} a_n x^n$ converges if $|x| < R$ and diverges for $|x| > R$. What about the regions of uniform convergence?

22. Find a sequence of functions $f_n : [0, 1] \to [0, 1]$ such that $\sum_{n=1}^{\infty} f_n$ converges uniformly, but not absolutely.

Chapter 3

Functions of several real variables: a starter

3.1 Graphical representation

Actually, is more important to have a "mind representation", in other words, how to figure a function of several variables. In practise, those functions appears as a formula $f(x, y, \dots)$ that eventually may have a vector output. After, the introduction to normed spaces one could think that functions depends one vector variable instead of several numerical ones. That is almost right, but not quite. We should make the same distinction that in *Affine Geometry*: points and vectors are different objects. Eventually, both points and vectors are built from the same "ingredient" $\mathbb{R}^n$ but they are not the same. That distinction is compulsory when dealing with abstract manifolds: for every point there is a set of admisible directions, the tangent space, and two directions from tangent spaces at different points are even not comparable (a special "device" for that has to be introduced).

Also, in (Newtonian) Physics there is a notion of coordinate-free space. Coordinates only appear when you fix a coordinate system, made up of one point where three perpendicular axes meet. You, as *observer*, cannot even tell if the axes are moving with the time. However, two *frames of reference* can move with respect to each other. Then, the validity of the *Principle of Inertia* helps you to choose a good frame to develop the Mechanics. Therefore, we should think of the domain of a function on the physical space as composed of points so its representation as a formula is just a consequence of fixing a

coordinate system. In such a way is how we should think of the *intrinsicness* of vector operators (chapter on Vector Analysis).

Now, assume you have a function of two variables given by some formula $f(x, y)$. What is the simplest way to represent it? For pedagogical reasons, the best answer is the *graph*, that is, the set

$$\{(x, y, z) : z = f(x, y)\}.$$

However, in practise these functions occurs often and the graph is not advisable in some cases: think of (x,y) being a geographical position (longitude, latitude) and the function being the height (over the sea level) or the atmospheric pressure (at ground level). For that case, several curves (*level curves*) of the form

$$\{(x, y) : f(x, y) = c\}$$

for some values of the constant c provide a *contour map* that can be read as we, allegedly, can read a topographical map. Eventually, the curves, which necessarily are discrete, could be changed by the continuous variation of tone or colour.

Since thinking in four dimensions is difficult, unless you are A. Einstein or S. Dalí, for a function of three variables $f(x, y, z)$. The graph is not a good idea. Fortunately, level curves can be generalized to *level surfaces* (or more generally, to *level sets*) by taking

$$\{(x, y, z) : f(x, y, z) = c\}.$$

As the representation could be difficult to visualise (the surfaces cover one another like Russian dolls), it would be convenient to choose just one significative value of c, for instance, when it reach a maximum. That is exactly what you see in the pictures of *atomic orbitals* in Chemistry books.

An interesting situation is when the function has vector values. For functions of the form $f : \mathbb{R} \to \mathbb{R}^3$ the representation is a curve, and we could think of it as a *trajectory* regarding the variable as time. For functions as $f : \mathbb{R}^2 \to \mathbb{R}^2$ or $f : \mathbb{R}^3 \to \mathbb{R}^3$ we may think of them as deformations of the space, and we can visualise the functions by watching how they act on simple sets of points: curves, simple domains... In that way is usually done in *Complex Analysis* since a complex function is, in practise, a function from the plane to the plane. Another way to think of functions $f : \mathbb{R}^2 \to \mathbb{R}^2$ is to consider

the domain composed of points and, on the other hand, the images as vectors. That is a plane *vector field* that can be depicted by choosing a regularly ordered set of points and drawing an arrow on each of them (usually the arrow starts at the point) that represents the value of the function. In this way, the speed of wing is represented in weather forecast informations, for instance.

Be aware, that non-cartesian coordinate systems, as polar or spherical coordinates, can be used to convey a function defined on points of the plane or the space to a formula. On the other hand, some properties of functions are more or less evident depending on the way to represent them. For instance, the property of the exponential $e^{x+y} = e^x e^y$ is not evident in cartesian coordinates, but there is a sort of innuendo in polar coordinates $r = e^\theta$.

Example 3.1.1. *The importance of the choice of coordinates: the ellipse.*

The ellipse is defined as the curve made up of points from the plane such that the sum of the distances to two points (*foci*) is constant. The curve is symmetric with respect to the line passing through the foci, as well as the bisector line of the segment joining the foci. When referred the ellipse to those axes (X and Y, respectively) the well known cartesian equation is

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

where a is the long semi-axis and b the short one. The equation is particularly simple because of the good choice of the coordinate frame. Note that the sum of distances to the foci is $2a$ and the distance between them is $2c = 2\sqrt{a^2 - b^2}$. However, if we try to obtain the polar equation straight from the previous one, the result is

$$\frac{r^2 \cos^2\theta}{a^2} + \frac{r^2 \sin^2\theta}{b^2} = 1$$

and so

$$r = \frac{ab}{\sqrt{b^2 \cos^2\theta + a^2 \sin^2\theta}}$$

which is not specially nice. In orden to obtain a better expression, move the origin to one of the focus (the one on our right). The distance to the focus at the origin is r. The distance to the other focus is

$$\sqrt{(x+2c)^2 + y^2} = \sqrt{(r\cos\theta + 2c)^2 + r^2 \sin^2\theta} = \sqrt{r^2 + 4cr\cos\theta + 4c^2}.$$

Therefore,

$$\sqrt{r^2 + 4cr\cos\theta + 4c^2} = 2a - r.$$

Squaring we get

$$r^2 + 4cr\cos\theta + 4c^2 = 4a^2 - 4ar + r^2,$$

$$\text{thus,} \quad cr\cos\theta + ar = a^2 - c^2 = b^2,$$

and the polar expression now is

$$r = \frac{b^2}{a + c\cos\theta} = \frac{p}{1 + \epsilon\cos\theta}$$

being $p = b^2/a$ and $\epsilon = c/a$ the eccentricity. This equation, which is common for al the conics ($\epsilon = 0$ for the circle, $\epsilon \in (0, 1)$ ellipse, $\epsilon = 1$ parabola and $\epsilon > 1$ for the hyperbola) is useful for the description of the movement of planets.

3.2 Topology

The topology required to deal with functions of several variables is exactly the metric topology, where the metric is induced by any of the usual norms, which turn out to be equivalent. Moreover, guessing that geometry could be of any help when dealing with limits or continuity is a wrong idea. You may think that the limit of a function exits because it exits through all the lines going to that point (*radial limit*), and, however, the ordinary (topological) limit may not exits. That is the case of

$$f(x, y) = \frac{2xy^2}{x^2 + y^4}$$

having limit 0 at $(x, y) = (0, 0)$ through lines, however the limit is not null using suitable parabolas. In any case, radial and other variations of limits are useful as training exercises, and the more interesting fact that they relate limits in two o more variables to the *functional limit*. Indeed, the existence of radial limits at a point implies the existence of the ordinary limit if the radial limit is *uniform* with respect to the angle.

Despite the example, a radially continuous function is not so bad. For instance, if we assume that a function is *separately continuous*, that is, for every fixed value of all the variables except one the restricted function is continuous

with respect the remaining variable. It is possible to prove with the help of Baire's theorem that a separately continuous function has a dense set of points of actual continuity.

In finite dimension there is a great availability of compactness, therefore a map which is continuous and injective behaves locally as a homeomorphism. It is quite easy to show, that the segment $[0,1]$ is not homeomorphic to the square $[0,1]^1$: the map on the border of the square cannot be injective. However, when we drop injectivity, strange things happen. For instance, there is a continuous map from $[0,1]$ *onto* $[0,1]^2$. That is the *Peano map*, that can be built as the limit of a uniformly convergent sequence of maps. Of course, by Heine, that map is uniformly continuous also. However, it cannot be *Lipschitz* (recall that Lipschitz for a map between metric spaces $f : (M_1, d_1) \to (M_2, d_2)$ means the existence of a constant $\lambda > 0$ such that $d_2(f(x), f(y)) \leq \lambda\, d_1(x, y)$ for any $x, y \in M_1$). The reason can be derived from elementary facts about the Lebesgue measure. That shows the existence of a huge gap between continuous maps and Lipschitz maps (in particular C^1). The only simple fact abut continuous maps is its definition, almost nothing more. Intuition can be dangerous.

3.3 Genuine functions on $\mathbb{R}^n$?

The reader could be disappointed if after stressing the idea that a function of several variables actually depends on a point-set of $\mathbb{R}^n$, all the examples are actually a *superposition* of functions of one variable applied to the several scalar variables available. Well, Kolmogorov proved that any continuous function of several variables can be expressed as a certain superposition (two compositions and a linear combination). The general result is quite complex to write, so we will restrict ourselves to only two variables. Given a continuous function $f(x, y)$ there exist *six* continuous functions $g_0, g_1, \ldots, g_5$ and a number λ such that

$$f(x, y) = \sum_{k=1}^{5} g_0(g_k(x) + \lambda g_k(y)).$$

The complicated formulation is due to the fact that the functions $g_1, \ldots, g_5$ given by Kolmogorov's theorem are universal, that is, they do not depend on f. Nevertheless, the idea is clear: there is not a "genuine" continuous real function of two or more variables.

3.4 Rationale and remarks

Despite the very general frame with metric and normed spaces, sometimes is necessary to point out that the matter is "Functions of several real variables", not a few, but not too many.

This chapter is a reflection about the notion of function of several real variables. The idea is to blend some "philosophical" comments with simple examples and pictures. There is a lot of information to awake the curiosity: *Mechanics* as an axiomatic theory; separately continuous functions (Baire classes and so); Peano curves (they can satisfy a Hölder condition but not a Lipschitz one); and the surprising Kolmogorov theorem, of course. All good for a TFG.

3.5 Exercises

1. Use polar coordinates to express the set limited by the triangle with vertices at $(0,0)$, $(0,1)$ y $(1,0)$.

2. Use polar coordinates to find the equation of a circle that passes by the origin.

3. Express the set $\{(x,y,z) : 0 \leq 2z \leq 1 - x^2 - y^2\}$ into spherical coordinates.

4. Find a two variable function whose level curves is the family of circles that are tangent to the Y axis at the origin.

5. Parameterize the curve resulting form the intersection of these surfaces

   $$S_1 = \{(x,y,z) : (x-1)^2 + y^2 + z^2 = 1\}; \quad S_2 = \{(x,y,z) : z^2 = x^2 + y^2\}$$

6. Cancel the parameter in the parametric equation of the curve

   $$(r\cos at, r\sin at, bt)$$

   with $t \in \mathbb{R}$, that is, find two functions $f(x,y,z), g(x,y,z)$ such that the curve is the set

   $$\{(x,y,z) : f(x,y,z) = 0,\ g(x,y,z) = 0\}.$$

7. Prove that $f(x,y) = \sqrt[4]{x^2+y^2}$ does not satisfy Lipschitz condition and yet it is uniformly continuous on $\mathbb{R}^2$.

8. Prove the existence and compute the limit

$$\lim_{(x,y)\to(0,0)} \frac{x^3+y^3}{x^2+y^2}.$$

9. Let $h : \mathbb{R} \to \mathbb{R}$ be a differentiable function. Consider the two variable function

$$f(x,y) = \frac{f(x)-f(y)}{x-y}$$

for $(x,y) \in \mathbb{R}^2$ with $x \neq y$. Compute then iterated, the radial and the double limits at the points of the form (x,x). Find conditions on h for f to be continuously extended to $\mathbb{R}^2$.

10. Show that a separately continuous function of $[a,b] \times [c,d]$ is a pointwise limit of a sequence of continuous functions.

Chapter 4

Differentiable mappings

4.1 The basics

Differentiable maps are those whose increments behave almost linearly on a small scale (it is hard to think that things could be different in the real world).

Definition 4.1.1. *A map $f : D \subset E \to F$ between normed spaces is differentiable at $x_0 \in D$ if there exists $A \in \mathfrak{L}(E, F)$ such that*

$$f(x) - f(x_0) = A(x - x_0) + o(\|x - x_0\|). \tag{4.1}$$

It is immediate from the definition that

1. the map f is continuous at x_0;
2. the element $A \in \mathfrak{L}(E, F)$ satisfying (4.1) is unique, thus we will write
$$df(x_0) := A;$$
3. the assignment $f \to df(x_0)$ is linear among the maps that are differentiable at x_0;
4. if f is linear, then $df(x) = f$ at any $x \in E$.

For real valued functions, the geometrical idea behind the notion of differentiability can be understood as that the "graph of the function f is well approximated by the tangent plane". However, to be rigorous the definition of tangent plane should depend on the notion of differentiability.

Definition 4.1.2. *We call the tangent plane to the graph of $f : D \subset \mathbb{R}^n \to \mathbb{R}$ at $(x_0, f(x_0))$ to the set*

$$\mathcal{T}_{x_0}(f) = \{(x, y) \in \mathbb{R}^{n+1} : y = f(x_0) + df(x_0)(x - x_0)\}$$

provided that f is differentiable at x_0.

Our aim now is to develop the rules to compute differentials likewise it is done with derivatives for functions of one variable. This is Leibniz differentiation of a product rule. It could be stated for any finite number of "factors".

Proposition 4.1.3. *Let $B : E_1 \times E_2 \to F$ a continuous bilinear map. Then B is differentiable at any point and*

$$dB(x_0, y_0)(x, y) = B(x, y_0) + B(x_0, y)$$

Proof. Indeed

$$B(x, y) - B(x_0, y_0) = B(x - x_0, y_0) + B(x_0, y - y_0) + B(x, x_0, y - y_0)$$

and $B(x, x_0, y - y_0) = O(\|x - x_0\| \cdot \|y - y_0\|) = o(\sqrt{\|x - x_0\|^2 + \|y - y_0\|^2})$. ■

Now we will prove the chain rule.

Theorem 4.1.4. *Let $f : E \to F$ and $g : F \to G$ and $x_0 \in dom(f)$ and $y_0 = f(x_0) \in dom(g)$. Then $g \circ f$ is differentiable at x_0 and*

$$d(g \circ f)(x_0) = dg(y_0) \circ df(x_0).$$

Proof. By hypotheses we have

$$f(x) - f(x_0) = df(x_0)(x - x_0) + o(\|x - x_0\|),$$

$$g(y) - g(y_0) = dg(y_0)(y - y_0) + o(\|y - y_0\|).$$

Putting $y = f(x)$ into the second equation we get

$$(g \circ f)(x) - (g \circ f)(x_0) = dg(y_0)(f(x) - f(x_0)) + o(\|x - x_0\|)$$

because the differentiability of f implies $\|f(x) - f(x_0)\| = O(\|x - x_0\|)$. A further replacement leads to

$$(g \circ f)(x) - (g \circ f)(x_0)$$

$$= dg(y_0)(df(x_0)(x - x_0) + o(\|x - x_0\|)) + o(\|x - x_0\|)$$
$$= (dg(y_0) \circ df(x_0))(x, x_0) + o(\|x - x_0\|)$$

which implies the differentiability of the composed map and $d(g \circ f)(x_0) = dg(y_0) \circ df(x_0)$ as wished. ∎

The action of the differential on a given $h \in E$ can be computed as a directional derivative

$$df(x_0)(h) = \left(\frac{df(x_0 + th)}{dt} \right)_{t=0} = \lim_{t \to 0} \frac{f(x_0 + th) - f(x_0)}{t}.$$

The differentiability implies that the limit is uniform on bounded sets with respect to h. For vector valued functions of one real variable, there is essentially a unique direction so we can keep the standard notation

$$f'(x_0) = df(x_0) \in \mathfrak{L}(\mathbb{R}, F) \sim F$$

and we say that f is *derivable* at x_0 instead of differentiable.

Simple examples such as $f(t) = (\cos t, \sin t, t)$ for $t \in [0, 2\pi]$ show that the mean value theorem with an *equality* is not longer true for vector valued functions

$$f(2\pi) - f(0) \neq f'(t)(2\pi - 0)$$

for all $t \in [0, 2\pi]$. However, we have the following that it is enough for most applications.

Theorem 4.1.5. *Let E be a normed space and let $f : [a, b] \to E$ and $g : [a, b] \to \mathbb{R}$ be continuous functions such that they are derivable on (a, b) and satisfy $\|f'(t)\| \leq g'(t)$ for all $t \in (a, b)$. Then*

$$\|f(b) - f(a)\| \leq g(b) - g(a).$$

Proof. Take $\varepsilon > 0$ and consider the set

$$I = \{t \in [a, b] : \forall\ a \leq s \leq t, \|f(s) - f(a)\| \leq g(s) - g(a) + \varepsilon(s - a) + \varepsilon\}.$$

By construction and continuity of the functions it is clear that $I = [a, s]$ for some $a < s \leq b$. If $s = b$ for all $\varepsilon > 0$ we are done, so we may assume $s < b$ in order to get a contradiction. Assume that there exists a decreasing sequence $(s_n) \subset (s, b)$ with limit s such that

$$\|f(s_n) - f(a)\| > g(s_n) - g(a) + \varepsilon(s_n - a) + \varepsilon.$$

Therefore

$$g(s_n) - g(a) + \varepsilon(s_n - a) < \|f(s_n) - f(s)\| + \|f(s) - f(a)\|$$

$$\leq \|f(s_n) - f(s)\| + g(s) - g(a) + \varepsilon(s - a)$$

and thus

$$g(s_n) - g(s) + \varepsilon(s_n - s) \leq \|f(s_n) - f(s)\|.$$

Dividing by $s_n - s$ we have

$$\frac{g(s_n) - g(s)}{s_n - s} + \varepsilon \leq \left\| \frac{f(s_n) - f(s)}{s_n - s} \right\|$$

and taking limits we get $g'(s) + \varepsilon \leq \|f'(s)\|$ which is a contradiction. ∎

Corollary 4.1.6. *Let $f : D \subset E \to F$ be a differentiable map, $x, y \in D$ two points that the segment $[x, y]$ that join them by is contained in D. Then*

$$\|f(y) - f(x)\| \leq \sup\{\|df(z)\| : z \in [x, y]\} \cdot \|y - x\|.$$

4.2 Partial derivatives

In practice, functions defined on $\mathbb{R}^n$ are given by means of a formula involving the coordinates of the point $x = (x_1, \ldots, x_n)$ that we can represent simply as $f(x_1, \ldots, x_n)$, understanding an identification of the function and its formula. For low dimensions we may use $f(x, y)$, $f(x, y, z)\ldots$ with the same meaning. In this setting we may study the behaviour of a function with respect to one variable x_i leaving the others constant, and in particular to compute the derivative with respect to x_i if possible. This is called the *partial derivative* with respect to x_i and it is denoted as

$$\frac{\partial f}{\partial x_i}(x_1, \ldots, x_n).$$

From the point of view of the previous section, a partial derivative is just a directional derivative for the direction given by a vector of the canonical basis. Therefore, the differentiability implies the existence of the partial derivatives. However, the partial derivatives are easier to compute, thus for function that is differentiable at $x^0 = (x_1^0, \ldots, x_n^0)$ we have

$$df(x^0)(\lambda_1 e_1 + \cdots + \lambda_n e_n) = \lambda_1 df(x^0)(e_1) + \cdots + \lambda_n df(x^0)(e_n)$$

$$= \lambda_1 \frac{\partial f}{\partial x_1}(x^0) + \cdots + \lambda_n \frac{\partial f}{\partial x_n}(x^0).$$

If we denote by dx_i the linear map $\lambda_1 e_1 + \cdots + \lambda_n e_n \to \lambda_i$, then for any $x \in \mathbb{R}^n$ we may write the previous identity as

$$df(x^0)(x) = dx_1(x)\frac{\partial f}{\partial x_1}(x^0) + \cdots + dx_n(x)\frac{\partial f}{\partial x_n}(x^0)$$

that can be rewritten in a more aesthetically way

$$df(x^0) = \frac{\partial f}{\partial x_1}(x^0)dx_1 + \cdots + \frac{\partial f}{\partial x_n}(x^0)dx_n$$

or simply

$$df = \frac{\partial f}{\partial x_1}dx_1 + \cdots + \frac{\partial f}{\partial x_n}dx_n$$

despite the fact that the partial derivatives $\frac{\partial f}{\partial x_i}$ may have vector values. For real valued functions compare with the formula of the tangent plane in terms of the partial derivatives

$$y - f(x^0) = \frac{\partial f}{\partial x_1}(x^0)(x_1 - x_1^0) + \cdots + \frac{\partial f}{\partial x_n}(x^0)(x_n - x_n^0).$$

In the old times before the arrival of rigor in Calculus, was usual to think that $dx_1, \ldots, dx_n$ where infinitesimal increments of the variables. That spirit still last in reasonings that can be found in some Physics and Engineering books.

The chain rule for the differential implies a chain rule for partial derivatives. Indeed, assume that the variables $x_1, \ldots, x_n$ are replaced by derivable functions $X_1(t), \ldots, X_n(t)$. Take $X(t) = (X_1(t), \ldots, X_n(t))$. Then

$$\frac{d}{dt}(f(X(t))) = \frac{\partial f}{\partial x_1}(X(t))\frac{dX_1}{dt}(t) + \cdots + \frac{\partial f}{\partial x_n}(X(t))\frac{dX_n}{dt}(t).$$

Typical abuse of language and removal of variables that are obvious leads to this neat expression

$$\frac{df}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \cdots + \frac{\partial f}{\partial x_n}\frac{dx_n}{dt}$$

that reminds of the expression of the differential above divided by "dt". If t were one of several other variables, then the expression of the chain rule would be with partial derivatives

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t}.$$

Let us stress once more that the chain rule is valid provided that the (second) function is differentiable. So far we have not provided a differentiability criterion based on the partial derivatives. The following will fill the gap.

Theorem 4.2.1. *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a function such that its first partial derivatives are defined on a neighbourhood of $x^0 \in D$ and they are also continuous at x^0. Then f is differentiable at x^0.*

Proof. The idea of the proof is the same in n dimensions than 2. In order not to complicate much the notation we will assume something in the middle, say 3 dimensions. The point from the hypothesis will be denoted $p = (x_0, y_0, z_0)$. Fix $\varepsilon > 9$ and let $\delta > 0$ be such that the partial derivatives exists $B(p, \delta)$ (we shall consider the Euclidean norm) and its values on points of $B(p, \delta)$ differs less than ε from the value at (x_0, y_0, z_0). Assume that $\|(x, y, z) - p\| < \delta$, then the four points $(x_0, y_0, z_0), (x, y_0, z_0), (x, y, z_0)$ and (x, y, z) are in the ball and so the segments joining them. We have

$$f(x, y, z) - f(x_0, y_0, z_0) =$$

$$f(x, y_0, z_0) - f(x_0, y_0, z_0) + f(x, y, z_0) - f(x, y_0, z_0) + f(x, y, z) - f(x, y, z_0)$$
$$= \frac{\partial f}{\partial x}(\overline{x}, y_0, z_0)(x - x_0) + \frac{\partial f}{\partial y}(x, \overline{y}, z_0)(y - y_0) + \frac{\partial f}{\partial z}(x, y, \overline{z})(z - z_0),$$

where $\overline{x} \in [x_0, x]$, $\overline{y} \in [y_0, y]$ and $\overline{z} \in [z_0, z]$ are given by the finite increments theorem. Now

$$\left| f(x, y, z) - f(p) - \frac{\partial f}{\partial x}(p)(x - x_0) - \frac{\partial f}{\partial y}(p)(y - y_0) - \frac{\partial f}{\partial z}(p)(z - z_0) \right|$$

$$\leq \varepsilon |x - x_0| + \varepsilon |y - y_0| + \varepsilon |z - z_0| \leq \sqrt{3}\varepsilon \|(x, y, z) - p\|.$$

That means f is differentiable at p as wished. ∎

Corollary 4.2.2. *Let f be a function whose first derivatives are null on a connected domain. Then f is constant.*

Proof. In that case f is differentiable by the the previous theorem and its differential is null everywhere. Two arbitrary points can be joined by a C^1 path γ. As $f \circ \gamma$ has null derivative, it is constant and thus the function has the same value at the butts. ∎

We could skip the use of Theorem 4.2.1 in the Corollary by showing that two points in a connected (open) domain can be joined by a path made of finitely many segments which are parallel to the axes.

4.3 Second order differentiability and more

Assume that a map $f : D \subset E \to F$ is differentiable at any point of D. In such a case we may consider the differential map $df : D \subset E \to \mathfrak{L}(E, F)$. We may consider the continuity of df with respect to the norm on $\mathfrak{L}(E, F)$ and, moreover, we may consider its further differentiability at some point $x_0 \in D$. In such a case, note that $d(df)(x_0) \in \mathfrak{L}(E, \mathfrak{L}(E, F))$. For simplicity, we have the identification

$$\mathfrak{L}(E, \mathfrak{L}(E, F)) = \mathfrak{B}(E \times E, F)$$

which means bilinear maps on E valued in F. Therefore, $d^2 f(x_0) = d(df)(x_0)$ can be interpreted as a bilinear form. The relation of that bilinear form to the increment of the function is depicted in the following result.

Theorem 4.3.1. *Let $f : D \subset E \to F$ be twice differentiable at $x_0 \in D$. Then*

$$f(x_0 + h) = f(x_0) + df(x_0)(h) + \frac{1}{2} d^2 f(x_0)(h, h) + o(\|h\|^2).$$

Proof. By the very definition, given $\varepsilon > 0$ there is $\delta > 0$ such that if $\|h\| < \delta$ then

$$\|df(x_0 + h) - df_0(x_0) - d^2 f(x_0)(h)\| < \varepsilon \|h\|.$$

The definition of the norm for linear operators implies

$$|df(x_0 + h)(v) - df(x_0) - d^2 f(x_0)(h)(v)| < \varepsilon \|h\| \|v\|$$

for every $v \in E$. Fix $h \in E$ with $\|h\| < \delta$ and consider the functions $h(t) = t^2$ and

$$g(t) = f(x_0 + th) - f(x_0) - t df(x_0)(h) - \frac{t^2}{2} d^2 f(x_0)(h, h).$$

The finite increment theorem for two functions says that

$$\frac{g(1) - g(0)}{h(1) - h(0)} = \frac{g'(\tau)}{h'(\tau)}$$

for some $\tau \in (0, 1)$. Therefore we have

$$|f(x_0 + h) - f(x_0) - df(x_0)(h) - \frac{1}{2} d^2 f(x_0)(h, h)| = |g(1)|$$

$$= (2\tau)^{-1} |df(x_0 + \tau h)(h) - df(x_0)(h) - \tau d^2 f(x_0)(h, h)|$$

$$= (2\tau)^{-1}|df(x_0+\tau h)(h) - df(x_0)(h) - d^2f(x_0)(\tau h, h)|$$
$$< \frac{\varepsilon\|\tau h\|\|h\|}{2\tau} = \frac{\varepsilon\|h\|^2}{2}$$

which proves the theorem. ■

Second order differentiability is related to second order derivation. The previous result implies that

$$d^2f(x_0)(h,h) = \left(\frac{d^2f(x_0+th)}{dt^2}\right)_{t=0}.$$

In order to discuss what happens in finite dimension, $E = \mathbb{R}^n$, we need to introduce the second order derivatives

$$\frac{\partial^2 f}{\partial x_j \partial x_k} := \frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_j}\right).$$

Do not mind much that convention about the derivation order since the derivatives commute in very general conditions. With the new notation and putting $h = (h_1, \ldots, h_n)$ we have

$$\frac{df(x_0+th)}{dt} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_0+th)h_j,$$

and the second derivative at $t = 0$ is

$$\left(\frac{d^2f(x_0+th)}{dt^2}\right)_{t=0} = \sum_{k=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_k}(x_0)h_j h_k.$$

The matrix of coefficients of the quadratic form $d^2f(x_0)$ given by

$$\left(\frac{\partial^2 f}{\partial x_j \partial x_k}(x_0)\right)_{j,k}$$

is called the *Hessian matrix*. The Hessian is symmetric under very general conditions, actually if f is twice differentiable at x_0, however we will prove a fairly general result with a simpler proof. We will use a more compact notation $f_x = \frac{\partial f}{\partial x}$, $f_y = \frac{\partial f}{\partial y}$, $f_{xy} = \frac{\partial^2 f}{\partial x \partial y}$ and $f_{yx} = \frac{\partial^2 f}{\partial y \partial x}$ for the statement and the proof of the following result.

In all what follows we will consider only real valued functions. Some results can be extended in an obvious way to functions taking values in finite dimensional space.

Theorem 4.3.2. *Let f be a real function defined on a neighbourhood of (x_0, y_0) such that f_x, f_y, f_{xy} and f_{yx} are also defined and continuous. Then*

$$f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0).$$

Proof. Take h, k small enough and for a function $g(x, y)$ we introduce the notation

$$\Delta_x g(x, y) = g(x + h, y) - g(x, y),$$
$$\Delta_y g(x, y) = g(x, y + k) - g(x, y).$$

Note that $\Delta_x(\Delta_y f) = \Delta_y(\Delta_x f)$. Now we will work with one of the terms, being the other one similar. The finite increments theorem implies

$$\Delta_x(\Delta_y f)(x_0, y_0) = (\Delta_y f)_x(x_0 + \theta_1 h, y_0)h$$

for some $0 < \theta_1 < 1$. Note that $(\Delta_y f)_x = \Delta_y f_x$, therefore

$$\Delta_x(\Delta_y f)(x_0, y_0) = \Delta_y f_x(x_0 + \theta_1 h, y_0)h.$$

A second application of the finite increments theorem gives

$$\Delta_x(\Delta_y f)(x_0, y_0) = f_{xy}(x_0 + \theta_1 h, y_0 + \theta_2 k)hk.$$

That implies

$$\lim_{h,k} \frac{\Delta_x(\Delta_y f)(x_0, y_0)}{hk} = f_{xy}(x_0, y_0).$$

The commutation of the increments claimed at the beginning implies the commutation of the derivatives. ∎

The Taylor formula for several variables. The commutativity of the derivations can be extended to orders higher that 2 if the hypothesis is satisfied. In particular, for a C^k function all the derivatives commute till the order k, meaning by order of a derivative the sum of of the orders with respect to each variable. In order to consider formulae involving complicated derivatives is convenient to introduce a *multi-index notation*: let $\alpha = (k_1, \ldots, k_n)$ be an n-uple of positive integers (including 0) and let

$$k = k_1 + \cdots + k_n.$$

We will denote

$$\frac{\partial^k f}{\partial x^\alpha} = \frac{\partial^k f}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}}.$$

The fact that the derivations are ordered with the variables implicitly means that the commutation is assumed. For the next result we will also need factorials, multi-powers and related functions. Put

$$\alpha! = k_1! \dots k_n!$$

and

$$\binom{k}{\alpha} = \frac{k!}{\alpha!}.$$

If $x = (x_1, \dots, x_n)$, then put $x^\alpha = x_1^{k_1} \dots x_n^{k_n}$. With this notation, we can prove Newton's multinomial formula

$$(x_1 + x_2 + \cdots + x_n)^m = \sum_{|\alpha|=m} \binom{m}{\alpha} x^\alpha.$$

Analogously, the *Taylor polynomial* up to grade m of a function f at the point $p = (x_1^0, \dots, x_n^0)$ is the following

$$T_m(p, x) = \sum_{|\alpha| \leq m} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(p)(x - p)^\alpha.$$

The reasons for that choice will be clear along the proof of the following result.

Theorem 4.3.3. *Let f be a $C^{m+1}(\mathbb{R}^n)$ function and let $c > 0$ be bound for the absolute value of the derivatives of order $(m + 1)$ on $B(p, r)$. Then for $x \in B(p, r)$, where the ball is taken with respect to the $\| \cdot \|_1$ norm, we have*

$$|f(x) - T_m(p, x)| \leq \frac{c\, r^{m+1}}{(m + 1)!}.$$

Proof. Without loss of generality we may assume $p = (0)$. Consider that following auxiliary function $h(t) = f(tx_1, \dots, tx_n)$. The derivatives of h are

$$h'(t) = \sum_i \frac{\partial f}{\partial x_i}(tx_1, \dots, tx_n)\, x_i$$

$$h''(t) = \sum_i \sum_j \frac{\partial^2 f}{\partial x_i \partial x_j}(tx_1, \dots, tx_n)\, x_i x_j$$

and so on, following the schema of the powers of a multinomial. Therefore, we can gather the terms in this way

$$h^{(k)}(t) = \sum_{|\alpha|=k} \binom{k}{\alpha} \frac{\partial^k f}{\partial x^\alpha}(tx_1, \dots, tx_n)\, x^\alpha.$$

Now we have

$$h(1) = \sum_{k=0}^{m} \frac{h^{(k)}(0)}{k!} + \frac{h^{(m+1)}(\theta x_1, \dots, \theta x_n)}{(m+1)!}$$

where $\theta \in (0,1)$ by the one variable Taylor formula with Lagrange remainder. Clearly

$$\sum_{k=0}^{m} \frac{h^{(k)}(0)}{k!} = \sum_{|\alpha|\le m} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(0)\, x^\alpha$$

and

$$\left| \frac{h^{(m+1)}(\theta x_1, \dots, \theta x_n)}{(m+1)!} \right| = \left| \sum_{|\alpha|=m+1} \frac{1}{\alpha!} \frac{\partial^{m+1} f}{\partial x^\alpha}(\theta x_1, \dots, \theta x_n)\, x^\alpha \right|$$

$$\le \frac{c}{(m+1)!} \sum_{|\alpha|=m+1} \binom{m+1}{\alpha} |x|^\alpha = \frac{c}{(m+1)!} (|x_1| + \cdots + |x_n|)^{m+1}$$

$$\le \frac{c\, \|x\|_1^{m+1}}{(m+1)!}$$

as wished. ∎

4.4 Applications to extrema

We will use derivatives in order to investigate the (relative) extreme values of a function f on a domain D. A point $x_0 \in D$ is said critical if $df(x_0) = 0$. Points where df does not exist are considered critical too in the literature, however we will not consider them here. Let us start by the following observation.

Proposition 4.4.1. *If a differentiable function f a relative extremum at an interior point of its domain, then necessarily the point is critical.*

In this case the domain is not assumed open. Typically, problems about extrema are posed on a compact domain. The existence of extrema is assured by Weierstrass theorem, however finding them is a different matter.

The strategy to compute the relative extremum on a "regular" compact domain is the following:

- the extremum is attained at an interior point, so we could find it among the critical points;
- otherwise, the extremum is on the border: then parametrise the border, which reduces the dimension by one, and start again.

There is method that spare us from parameterise the border, the so called *Lagrange multipliers* (see Section 6.3.1), whose explanation relies on the implicit functions theorem. In any case, the iteration of the previous algorithm will produce a list of points among where the maximum and the minimum are attained. We only have to check the values of f.

In case we are interested in *relative* extrema, the method above cannot distinguish them. If the function is C^2 we may consider the second order approximation Theorem 4.3.1. Assuming that $df(x_0) = 0$, the local behaviour of f at x_0 is the same that $d^2f(x_0)(x, x)$ at 0 in these cases:

- if $d^2f(x_0)(x, x) \geq c\|x\|^2$ for some $c > 0$, f has a local minimum at x_0;
- if $d^2f(x_0)(x, x) \leq -c\|x\|^2$ for some $c > 0$, f has a local maximum at x_0;
- if $d^2f(x_0)(x, x)$ takes both positive and negative values, then no relative extremum is attained at x_0 (*saddle point*).

The cases where merely $d^2f(x_0)(x, x) \geq 0$ or $d^2f(x_0)(x, x) \leq 0$ cannot be decided with this method. The proof for the two first cases follows straight from Theorem 4.3.1. The third case can be reduced to one variable: there are directions such that the restriction of the function to the line has a minimum at x_0, and other directions such that the restriction has a maximum.

The discussion in the finite dimensional case is done with the help of the Hessian matrix

$$A = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

In Linear Algebra is shown a method called *Sylvester criterion* to know if the associated quadratic form is positive or negative by checking n determinants. The quadratic form d^2f, and so the Hessian, can also be used to check the *convexity* of a C^2 function.

We will show how differential calculus works in an infinite dimensional context. The *Calculus of Variations* appeared as a collection of techniques to find (or characterize) the extrema of certain kind of functionals, that usually involve the function together its derivatives. The following is the most basic, but it is enough to solve the *brachistochrone* problem, by reducing the variational problem to an ordinary differential equation. Let $\phi(y, \dot{y}, x)$ a function defined on $\mathbb{R}^2 \times [a, b]$ and consider the associated functional

$$\Phi(f) = \int_a^b \phi(f(x), f'(x), x)\, dx$$

defined for C^1 functions f on $[a, b]$.

Theorem 4.4.2 (Euler - Lagrange)**.** *Let $A, B \in \mathbb{R}$. The relative extremes of the functional Φ among the C^1 functions f such that $f(a) = A$ and $f(b) = B$ satisfies the differential equation*

$$\frac{d}{dx}\left(\frac{\partial \phi}{\partial \dot{y}}\right) = \frac{\partial \phi}{\partial y}.$$

Proof. In order to show the extremality of f will consider only C^2 perturbations of the form $h(x)$ such that it and $\dot{h}(x)$ vanish at a, b. By hypothesis, the directional derivative

$$\frac{d}{ds}\int_a^b \phi(f(x) + sh(x), f'(x) + sh'(x), x)\, dx$$

must be 0 at $s = 0$. The derivation with respect s can be performed under the integral sign this way (we omit the variable x in some of the functions for the sake of readability)

$$\int_a^b \Big(\frac{\partial \phi}{\partial y}(f + sh, f' + sh', x)h' + \frac{\partial L}{\partial \dot{y}}(f + sh, f' + sh', x)h''\Big)\, dx.$$

Now, the formula of integration by parts gives that

$$\int_a^b \frac{\partial \phi}{\partial \dot{y}}(f + sh, f' + sh', t)h''\, dx =$$

$$\frac{\partial\phi}{\partial\dot{y}}(f+sh,f'+sh',t)h'\Big|_a^b - \int_a^b \frac{d}{dx}\Big(\frac{\partial\phi}{\partial\dot{y}}(f+sh,f'+sh',x)\Big)h'\,dx$$
$$= -\int_a^b \frac{d}{dx}\Big(\frac{\partial\phi}{\partial\dot{y}}(f+sh,f'+sh',x)\Big)h'\,dt.$$

Going back to the derivative of the functional, we have

$$\frac{d}{ds}\int_a^b \phi(f+sh,f'+sh',x)\,dx =$$

$$\int_a^b \Big(\frac{\partial\phi}{\partial y}(f+sh,f'+sh',x) - \frac{d}{dx}\Big(\frac{\partial\phi}{\partial\dot{y}}(f+sh,f'+sh',t)\Big)h'\,dx.$$

For $s=0$ we get that

$$\int_a^b \Big(\frac{\partial\phi}{\partial y}(f,f',x) - \frac{d}{dx}\Big(\frac{\partial\phi}{\partial\dot{y}}(f,f',x)\Big)h'\,dx = 0$$

for all the perturbations h satisfying the required assumptions. As it is possible to take h with arbitrarily small support contained into (a,b), we deduce that

$$\frac{\partial\phi}{\partial y}(f,f',x) - \frac{d}{dx}\Big(\frac{\partial\phi}{\partial\dot{y}}(f,f',x)\Big) = 0$$

for $x\in(a,b)$ as we wanted. ∎

4.5 Two applications to Algebra

Here we shall present to interesting applications of the calculus of several variables to proof to results of Algebra, that turns out to be useful in Analysis.

4.5.1 The Fundamental Theorem of Algebra

Theorem 4.5.1. *Every non-constant polynomial with complex coefficients has a complex root.*

Proof. Let $p(z)=a_Nz^N+\cdots+a_0$ with $a_N\neq 0$. Note that

$$|p(z)|\geq |a_N||z|^N - |a_{N-1}z^{N-1}+\cdots+a_0|.$$

Therefore, the real function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x, y) = |p(x+iy)|$ satisfies

$$\lim_{(x,y)\to\infty} f(x, y) = +\infty.$$

Together the continuity, that implies the existence of (x_0, y_0) such that

$$m := f(x_0, y_0) = \inf\{f(x, y) : (x, y) \in \mathbb{R}^2\}.$$

Our aim is to prove that $m = 0$. Let $z_0 = x_0 + iy_0$. We may assume that $z_0 = 0$ just by replacing p by $p(z - z_0)$. Using conjugates, we have

$$|p(z)|^2 = (a_0 + a_n z^n + \cdots + a_N z^N)(\overline{a_0} + \overline{a_n}\overline{z}^n + \cdots + \overline{a_N}\overline{z}^N)$$

where n is the index of the first non-null coefficient after a_0. We have

$$|p(z)|^2 = a_0\overline{a_0} + \overline{a_0}a_n z^n + a_0\overline{a_n}\overline{z}^n + \dots$$

where the following non-written terms are of degree greater than n. Put $a_0\overline{a_0} = A(\cos\alpha + i\sin\alpha)$ and $z = r(\cos\theta + i\sin\theta)$ and note that $A \neq 0$. Now we can use the fact that $|a_0| = m$ and De Moivre 's to get

$$\begin{aligned} |p(z)|^2 &= m^2 + Ar^n(\cos(n\theta+\alpha) + i\sin(n\theta+\alpha)) \\ &+Ar^n(\cos(n\theta+\alpha) - i\sin(n\theta+\alpha)) + O(r^{n+1}) \\ &= m^2 + 2A\cos(n\theta+\alpha) + O(r^{n+1}). \end{aligned}$$

We have

$$0 \leq R(r, \theta) := |p(z)|^2 - m^2 = R(r, \theta) = 2A\cos(n\theta+\alpha) + O(r^{n+1}).$$

Fix $\theta \in [0, 2\pi]$. We deduce thus that

$$0 \leq \lim_{r\to 0} r^{-n} R(r, \theta) = 2A\cos(n\theta+\alpha),$$

which is impossible as we can choose θ such that $\cos(n\theta+\alpha) = -1$. ∎

4.5.2 Diagonalization of symmetric matrices

Recall that the linear homeomorphism of $\mathbb{R}^n$ that preserve the Euclidean metric are characterised by *orthogonal* matrices, for which the inverse coincide with the transpose. We will prove the following well known result from Linear Algebra.

Theorem 4.5.2. *Let A be a symmetric matrix. Then there exists an orthogonal matrix such that Q^tAQ is diagonal.*

Proof. For an $n \times n$ matrix $B = (b_{ij})$ define $\sigma(B) = \sum_{i \neq j} b_{ij}^2$. The set of orthogonal $n \times n$ matrices, name it Ω can be identified with a closed and bounded subset of $\mathbb{R}^{n^2}$. Therefore Ω is compact. Assume that the symmetric matrix A is fixed and consider a function $f : \Omega \to \mathbb{R}$ defined by

$$f(Q) = \sigma(Q^tAQ)$$

That function attains its minimum value at some matrix, say B. We claim that B is diagonal. The idea is to prove that if B is not diagonal, then we could obtain a smaller value for σ by applying an orthogonal transformation to B. Thus, assume $b_{sr} = b_{rs} \neq 0$ for some $r \neq s$. Consider a rotation

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

on the 2-dimensional space corresponding to the coordinates r, s and enlarge it to an orthogonal Q matrix by taking the identity on the $(n-2)$-dimensional orthogonal complement. Consider now the matrix $C = Q^tAQ$ and note that for any pair of indices i, j such that $\{i, j\} \cap \{r, s\} = \emptyset$ we have $b_{ij} = c_{ij}$. In case that the sets have only one element in common, the coefficients change, but not enough to modify the value of σ. For instance, if $i \notin \{r, s\}$, then it can be computed that

$$b_{ir}^2 + b_{is}^2 = c_{ir}^2 + c_{is}^2.$$

And for the same pair r, s of entries we have

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} b_{rr} & b_{rs} \\ b_{sr} & b_{ss} \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} c_{rr} & c_{rs} \\ c_{sr} & c_{ss} \end{pmatrix}$$

A tedious computation and the help of some trigonometry gives

$$c_{rs}^2 + c_{sr}^2 = \frac{1}{2}((b_{ss} - b_{rr})\sin 2\theta + 2b_{rs}\cos 2\theta)^2.$$

For $\theta = 0$ we recover the value $2b_{rs}^2 = b_{rs}^2 + b_{sr}^2$. If $b_{ss} = b_{rr}$ any small perturbation of θ reduces the value of the function σ. If $b_{ss} \neq b_{rr}$ is posible to choose the sign of the perturbation, and so the sing of $\sin 2\theta$ in such a way that we reduce the value of σ. ∎

4.6 Rationale and remarks

Differentiability is a fundamental notion in Analysis. The students should be get used to all the different versions as well as the right use of the chain rule. In practise, people do not derive functions, they derive one variable with respect another. From the formal point of view, that always entails abuse of language. Nevertheless, future mathematicians should get used to that use in order to communicate with scientists and engineers.

The Taylor polynomial of second degree with vector values and infinitesimal remainder is presented mostly to show the complexity of the second differential from the point of view of the involved operator spaces.

The use of the Euler-Lagrange equation to solve a particular problem depends on the knowledge of differential equations. Diagonalization of symmetric matrices is more important in Physics than in Algebra. I like to recall the notion of *tensor of inertia* when I have the opportunity.

4.7 Exercises

1. Compute at any point the directional derivative of

   $$\ln(e^x + e^y)$$

   along directions parallel to the line $x = y$. Provide a simple explanation for the result.

2. Calculate

   $$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} + z\frac{\partial u}{\partial z}$$

   for

   $$u = \arctan\left(\frac{y^2 + z^2}{x^2}\right).$$

3. Let the function

   $$u = xy\,\frac{x^2 - y^2}{x^2 + y^2}.$$

   Compute at $(0, 0)$ the following functions

   $$\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial y}\right) \quad \text{y} \quad \frac{\partial}{\partial y}\left(\frac{\partial u}{\partial x}\right)$$

4. Consider the function on $\mathbb{R}^2 \setminus \{(0,0)\}$ defined by

$$f(x,y) = \frac{x^3y^3}{x^4+y^4}.$$

Show that is possible to continuously extend the function to $\mathbb{R}^2$. Then, study its differentiability.

5. Consider a function $F : \mathbb{R}^2 \to \mathbb{R}$ of the form $F(x,y) = yf(x) + xg(y)$ where $f, g : \mathbb{R} \to \mathbb{R}$ are continuous at 0. Prove that F is differentiable at $(0,0)$. Find a reasonable hypothesis to guarantee that F is twice differentiable at $(0,0)$.

6. Let $w = f(x,y,z)$ and $z = g(x,y)$. Then

$$\frac{\partial w}{\partial x} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial x} = \frac{\partial w}{\partial x} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial x}$$

since $\frac{\partial x}{\partial x} = 1$ y $\frac{\partial y}{\partial x} = 0$. Therefore $\frac{\partial w}{\partial z}\frac{\partial z}{\partial x} = 0$. Now, assume that $w = x + y + z$ and $z = x + y$. Then we get $\frac{\partial w}{\partial z} = \frac{\partial z}{\partial x} = 1$ an so $1 = 0$. Please, find the mistake.

7. Find and classify the critical points of $f(x,y) = (x^2+y^2)\, e^{x^2-y^2}$.

8. Find the maximum volume of the straight parallelepiped contained into the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

9. Find the minimum volume ellipsoid

$$E(a,b,c) = \{(x,y,z) : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1\}$$

passing at the point $(1,2,3)$.

10. Prove that the maximum value of the function

$$f(x_1, x_2, \ldots, x_n) = x_1^2 x_2^2 \ldots x_n^2$$

on the sphere $S = \{(x_1, x_2, \ldots, x_n) : x_1^2 + x_2^2 + \cdots + x_n^2 = 1\}$ is n^{-n}. Find, as an application, the arithmetic-geometric mean inequality: for every $n \in \mathbb{N}$ and $a_k \geq 0$ with $1 \leq k \leq n$ we have

$$\sqrt[n]{a_1 a_2 \ldots a_n} \leq \frac{a_1 + a_2 + \cdots + a_n}{n}.$$

11. Let the function $f : \mathbb{R}^3 \to \mathbb{R}$ be defined by $f(x,y,z) = \frac{x^2z^2}{x^2+y^2}$ for $(x,y,z) \neq (0,0,z)$ and $f(0,0,z) = 0$. Find $D_v f(1,1,\sqrt{2})$ being v an unitary vector which tangent at $(1,1,\sqrt{2})$ to the curve

$$x^2 + y^2 = 2x;$$

$$x^2 + y^2 + z^2 = 4.$$

12. A function $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ is said to be *homogeneous of degree* m if $f(tx) = t^m f(x)$ for every $x \in \mathbb{R}^n \setminus \{0\}$ and $t > 0$. Prove that f is homogeneous of degree m if and only if

$$x_1 \frac{\partial f}{\partial x_1} + \cdots + x_n \frac{\partial f}{\partial x_n} = mf$$

for all $x \in \mathbb{R}^n \setminus \{0\}$.

13. Prove that the function

$$f(x,y) = \frac{xy}{(x+y)(1+x)(1+y)}$$

defined on $\Omega = \{(x,y) : x > 0, y > 0\}$ is uniformly continuous and find its maximum value.

14. Consider the function $\langle x|\, a\rangle\, e^{-\|x\|_2}$ defined on $\mathbb{R}^n$, and find its maximum and minimum values.

15. Determine the values of the parameters $a, b \in \mathbb{R}$ for which the surface

$$z = e^{ax+y^2} + b\cos(x^2+y^2)$$

is below or above its tangent plane at $(0,0)$.

16. Prove that $x^3 + y^3 + 6xy$ is convex on $A = \{(x,y) : xy > 1, x > 0\}$.

17. Find a ball centred at $(0,0,0)$ where is convex the function

$$\log(1 + x^2 + y^2 + z^2).$$

18. We say that a function $f : \Omega \to \mathbb{R}$ is *analytic* on an open domain $\Omega \subset \mathbb{R}^n$ if it admits a power expansion centred at any point of Ω, that is, for every $(x_1^0, \ldots, x_n^0) \in \Omega$ there are coefficients $(a_{k_1,\ldots,k_n})$ such that

$$f(x_1, \ldots, x_n) = \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} a_{k_1,\ldots,k_n} (x_1 - x_1^0)^{k_1} \ldots (x_n - x_n^0)^{k_n}$$

for $(x_1, \ldots, x_n)$ in a a neighbourhood of $(x_1^0, \ldots, x_n^0)$. Prove that an analytic function in Ω is $C^\infty(\Omega)$.

19. Prove that the function defined by

$$f(x, y) = \frac{\sin y - \sin x}{y - x}$$

if $x \neq y$ and $f(x, x) = \cos x$ is analytic on $\mathbb{R}^2$.

20. Find the minimal distance to the origin from the line where meet the planes $x + 2y + z = 4$ and $3x + y + 2z = 3$.

21. Find the closest and the farthest point from the ellipsoid $x^2+2y^2+3z^2 = 1$ to the plane $x + y + z = 10$.

22. Find the closest and the farthest point form the ellipse $x^2/9 + y^2/4 = 1$ to $(1, 0)$.

23. Find the maximum and minimum values of $f(x, y) = x^2 - y^2$ on the set

$$x^2/16 + y^2/9 \leq 1.$$

Chapter 5

Theorems of the inverse mapping and implicit functions

5.1 Theorem of the inverse mapping

A continuous real function defined on an interval has (continuous) inverse if and only if it is strictly monotone (increasing or decreasing). For two or more variables, the existence of global inverse is more complicated so we will relax our exigences to the existence of local inverse for regular enough maps. In order to set the suitable hypotheses, let us analyse the one variable case first. Monotonicity on the neighbourhood of a point x_0 is related to the condition $f'(x_0) \neq 0$, but this condition is not necessary if we do not ask regularity to the inverse: the function $f(x) = x^3$ fails the condition at 0 and it has local, and global, continuous inverse, however $f^{-1}(y)$ fails to be derivable at 0. On the other hand, the condition $f'(x_0) \neq 0$ is neither sufficient for the existence of the local inverse if we do not ask regularity to the derivative: consider the function $f(x) = x + 2x^2 \sin(1/x)$ for $x \neq 0$ and $f(0) = 0$, which is derivable at every point and $f'(0) = 1$, but fails to be monotone at every neighbourhood of 0 and so it fails to have local inverse too.

The first aim in this chapter is to prove the inverse mapping theorem for maps defined on subsets of $\mathbb{R}^d$. In this context, the one variable condition $f'(x_0) \neq 0$ is replaced by the non degeneracy of $df(x_0)$, that is, it is invertible as map on $\mathbb{R}^d$. For that, we need a couple of lemmata to understand the local behaviour of a C^1 map on $\mathbb{R}^d$. That information will play a crucial role in the proof of theorem for change of variable in multiple integrals.

We will prove the first auxiliary results in the frame of Banach spaces because we do not need any special property of $\mathbb{R}^d$. Before stating the first lemma, let us state this "mean value theorem" for vector valued functions, which is just a corollary of Theorem 4.1.5:

$$\|f(x) - f(y)\| \leq \sup\{\|df(tx + (1-t)y)\| : t \in [0,1]\} \cdot \|x - y\|$$

where we suppose that f is differentiable at every point the segment joining x and y which, of course, is contained in the domain of f.

Lemma 5.1.1. *Let E be a Banach space and let $f : D \subset E \to E$ be a differentiable map such that $B[0,r] \subset D$ for some $r > 0$, $f(0) = 0$, $df(0) = \mathbb{I}$ and there is $0 < \eta < 1$ such that*

$$\|df(x) - \mathbb{I}\| \leq \eta$$

for every $x \in B[0,r]$. Then, we have

(a) $(1-\eta)\|x-y\| \leq \|f(x) - f(y)\| \leq (1+\eta)\|x-y\|$ *for any $x, y \in B[0,r]$;*

(b) $B[0,(1-\eta)r] \subset f(B[0,r]) \subset B[0,(1+\eta)r]$;

(c) there are neighbourhood U and V of 0 such that f is a homeomorphic bijection from U onto V;

(d) Moreover, f^{-1} (defined on V) is differentiable at 0 with $d(f^{-1})(0) = \mathbb{I}$.

Proof. Applying the mean value theorem to $g(x) = f(x) - x$ and noticing that $\|dg(x)\| = \|df(x) - \mathbb{I}\| \leq \eta$ if $x \in B[0,r]$ we have

$$\|f(x) - f(y) - (x-y)\| \leq \eta\|x-y\|.$$

In particular, for $y = 0$, we have $\|f(x) - x\| \leq \eta\|x\|$ that we will need later. The triangle inequality implies that

$$|\|f(x) - f(y)\| - \|x-y\|| \leq \eta\|x-y\|$$

and so we obtain the desired inequality of statement (a)

$$\|x-y\| - \eta\|x-y\| \leq \|f(x) - f(y)\| \leq \|x-y\| + \eta\|x-y\|.$$

Now observe that if $x \in B[0,r]$ then $\|f(x)\| = \|f(x) - f(0)\| \leq (1+\eta)\|x\|$ and therefore $f(x) \in B[0,(1+\eta)r]$, which is the right hand-side set inclusion

of statement (b).
The other set inclusion is more delicate and requieres Banach's fixed point theorem. Assume that $y \in B[0,(1-\eta)r]$. We want to find $x \in B[0,r]$ such that $y = f(x)$. Observe that such a point x is a fixed point of the map

$$\phi(x) := x - f(x) + y$$

We claim that ϕ is a contractive map from $B[0,r]$ into itself. Indeed, if $x \in B[0,r]$ then

$$\|\phi(x)\| \leq \|x - f(x)\| + \|y\| \leq \eta\|x\| + \|y\| \leq \eta\, r + (1-\eta)r = r$$

so $\phi(x) \in B[0,r]$. Assume now that $x, z \in B[0,r]$. Then

$$\|\phi(x)-\phi(z)\| = \|x-f(x)+y-z+f(z)-y\| = \|f(z)-f(x)-(z-x)\| \leq \eta\|x-z\|.$$

Since $\eta < 1$, the map ϕ is contractive as desired.
In order to prove (c), note that the inequality in (a) implies that f is one-to-one on $B[0,r]$, Lipschitz and the inverse f^{-1} defined on $f(B[0,r])$ is also Lipschitz. Therefore, $f|_{B[0,r]}$ is an homeomorphism of $B[0,r]$ onto $f(B[0,r])$. Note that $f(\partial B[0,r])$ is closed in E because f^{-1} preserves Cauchy sequences and completeness equals closedness in E. Since $0 \notin f(\partial B[0,r])$, we may take $\delta > 0$ such that $B(0,\delta) \cap f(\partial B[0,r]) = \emptyset$. Then we set $V = B(0,\delta)$ and $U = B(0,r) \cap f^{-1}(V)$. Clearly the choice of U and V implies that f is a bijection between them.
Finally statement (d). In order to show that f^{-1} is differentiable at 0, we have to prove that $\|f^{-1}(y) - y\| = o(\|y\|)$ where $y \in V$. Fix $\varepsilon > 0$. As f is differentiable at 0, there is $\delta > 0$ such that

$$\|f(x) - x\| = \|f(x) - f(0) - \mathbb{I}(x-0)\| < (1-\eta)\varepsilon\|x\|$$

for $\|x\| < \delta$. Put $x = f^{-1}(y)$ and observe that if $y \in f^{-1}(B(0,\delta))$ we have

$$\|f^{-1}(y) - y\| = \|x - f(x)\| \leq (1-\eta)\varepsilon\|x\| \leq \varepsilon\|y\|.$$

where we are using one of the inequalities from (a). Note that the last inequality holds for every $y \in f(B(0,\delta))$ which is a neighbourhood of 0. As $\varepsilon > 0$ was arbitrary, that implies the differentiability of f^{-1} at 0 with $d(f^{-1})(0) = \mathbb{I}$. ∎

Lemma 5.1.2. *Let $f : D \subset E \to E$ be a differentiable map and let $x_0 \in D$ such that df is continuous at x_0 and $df(x_0)$ has a continuous inverse. Then*

for every $0 < \eta < 1$ there exists $\delta > 0$ such that $f|_{B[x_0,\delta]}$ is one-to-one, f^{-1} is differentiable at $f(x_0)$ and

$$f(x_0) + df(x_0)(B[0,(1-\eta)r]) \subset f(B[x_0,r]) \subset f(x_0) + df(x_0)(B[0,(1+\eta)r])$$

for every $0 \leq r \leq \delta$ In particular, the image through f of a neighbourhood of x_0 is a neighbourhood of $f(x_0)$. Moreover, $f(U)$ is open whenever $U \subset D$ is open and $df(x)$ has a continuous inverse at every point $x \in U$.

For the proof we will use the shift map $\tau_h(x) = x + h$.

Proof. Put $y_0 = f(x_0)$ and consider the map

$$g := df(x_0)^{-1} \circ \tau_{y_0}^{-1} \circ f \circ \tau_{x_0}$$

Observe that $g(0) = 0$ and $dg(0) = \mathbb{I}$ by the chain rule. Moreover, dg is continuous at 0 and thus there is $\delta > 0$ such that $\|dg(x) - \mathbb{I}\| \leq \eta$ for every $x \in B[0,\delta]$. The application of the previous lemma for $0 < r \leq \delta$ gives us

$$B[0,(1-\eta)r] \subset g(B[0,r]) \subset B[0,(1+\eta)r]$$

and applying $df(x_0)$ to all the members we have

$$df(x_0)(B[0,(1-\eta)r]) \subset f(B[x_0,r]) - y_0 \subset df(x_0)(B[0,(1+\eta)r])$$

which is equivalent to the set inclusion of the statement. The differentiability of f^{-1} at y_0 is consequence of the expression

$$f^{-1} = \tau_{x_0} \circ g^{-1} \circ df(x_0)^{-1} \circ \tau_{y_0}^{-1}.$$

For the last part, just observe that $x + df(x)(B(0,\xi))$ is a neighbourhood of x for every $\xi > 0$ whenever $df(x)$ is nonsingular. ∎

Remark 5.1.3. *Note that the $\delta > 0$ given by the previous lemma depends on the modulus of continuity of $df(x)$ and an upper bound for $\|df^{-1}(x)\|$. Therefore, if f is C^1 then it is easy to modify the statement to get the same δ for all the points on a certain neighbourhood of x_0.*

We can state now the inverse mapping theorem.

Theorem 5.1.4. *Let $f : D \subset \mathbb{R}^d \to \mathbb{R}^d$ be a C^k map with $k \geq 1$ and let $x_0 \in D$ such that $df(x_0)$ is nonsingular. Then there exist neighbourhoods U of x_0 and V of $y_0 = f(x_0)$ such that f is a bijection of U onto V and the inverse map f^{-1} defined on V is C^k.*

Proof. Being f of class C^1 we may restrict our attention to a neighbourhood of x_0 where df is nonsingular. By the previous lemma we may fix neighbourhoods U of x_0 and V of y_0 such that f is a bijection of U onto V. For any $x \in U$ the application of the previous lemma gives us that f^{-1} is differentiable at $f(x)$, thus $d(f^{-1})(y)$ is defined for every $y \in V$. Before proving that $f^{-1}(y)$ is C^k, note that the map sending the nonsingular linear maps A on $\mathbb{R}^d$ to their inverses A^{-1} is C^∞. Indeed, use the matrix expression for A and observe that the coefficients of the matrix of A^{-1} are polynomials on the coefficients of A divided by the determinant, which is a non vanishing polynomial of the coefficients of A. Now we will proceed by induction: if $k = 1$ then $d(f^{-1})(y) = (df(f^{-1}(y)))^{-1}$ is continuous as a composition of continuous maps and so f^{-1} is C^1. Assume that the theorem is proven for $k-1$ and f is C^k. In such a case we know that $df(x)$ is C^{k-1} and, by the induction hypothesis $f^{-1}(y)$ is C^{k-1}. Therefore, $d(f^{-1})(y) = (df(f^{-1}(y)))^{-1}$ is C^{k-1} as composition of C^{k-1} maps, which means that f^{-1} is C^k. ∎

5.2 The implicit function theorem and smooth manifolds

Now we are interested in the following problem: consider an equation $F(x, y) = 0$, find conditions to solve it in the form $y = f(x)$ for some range of values of x. Of course, not for every x there is y such that $F(x, y) = 0$, and when such a y exists it is not unique necessarily. It seems natural to find conditions ensuring that the function f is as regular as possible and not a random choice of solutions. For that, a reference value for f should be fixed in form of a particular solution (x_0, y_0) of the equation. If we think of the set $F(x, y) = 0$ like a curve on the plane, we need that it looks like a graph around (x_0, y_0), so we should skip having a vertical tangent there.

Theorem 5.2.1. *Let $F : D \subset \mathbb{R}^{n+m} \to \mathbb{R}^m$ be a C^k map, $k \geq 1$. Assume that $(x_0, y_0) \in D$ is such that $F(x_0, y_0) = 0$ and $dF(x_0, y_0)$ is nonsingular when restricted to the subspace $0 \times \mathbb{R}^m$. Then there are neighbourhoods U of x_0 and V of y_0 with $U \times V \subset D$ and such that for every $x \in U$ there exists a unique $y \in V$ such that $F(x, y) = 0$ and the map $f : U \to V$ defined by $f(x) = y$ in that way is C^k.*

Proof. Consider the map $G : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ defined by $G(x, y) = (x, F(x, y))$ and observe that $dG(x_0, y_0)$ is nonsingular because of the box decomposition of its matrix. Therefore, G has a C^k inverse defined on a neighbourhood of

$G(x_0, y_0) = (x_0, 0)$, that we may assume of the form $U \times B(0, \delta)$, with values in a neighbourhood of (x_0, y_0), that we may assume of the form $U \times V$. The condition $U \times V \subset D$ can be achieved by shrinking U and $\delta > 0$. Note that if $x \in U$, then $G^{-1}(x, 0) = (x, y)$ with $y \in V$, and thus $F(x, y) = 0$. That point y is unique because G is injective on $U \times V$, therefore we may define $f(x) = y$. Now, the map f can be written as the composition of G^{-1} with a couple of linear maps, and thus it is C^k. ∎

Once we know the existence of the implicit function $f = (f_1, \ldots, f_m)$ around $(\bar{x}_0, \bar{y}_0)$ we may be interested in computing its the partial derivatives. For that aim it is enough to derivate with respect to x_j the equalities

$$F_i(x_1, \ldots, x_n, f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)) = 0$$

for $1 \leq i \leq m$. Then assign the value $\bar{x}_0$ to $(x_1, \ldots, x_n)$ and the derivatives $\frac{\partial f_i}{\partial x_j}$ with $1 \leq i \leq m$ will show up as the solution of a linear system whose matrix is composed of the last m rows from the matrix of $dF(x_0, y_0)$, which is nonsingular by hypothesis (note that the possibility of obtaining the derivative of as unique solution of the system implies nonsingularity). The derivatives of superior order can be obtained by further derivation of the formules.

Smooth manifolds. A d-dimensional smooth manifold in $\mathbb{R}^n$, for $1 \leq d < n$, defined implicitly is a set of the form

$$M = \{x \in \mathbb{R}^n : F(x) = 0\}$$

where $F : \mathbb{R}^n \to \mathbb{R}^{n-d}$ is at least C^1 and $dF(x)$ has rank $n - d$ at every $x \in M$. Smooth manifolds defined that way appear very often in Analysis, for instance in optimization problems. We will prove that every point of M has a neighbourhood that can be parameterized by d variables, that is, like a graph of a map from $\mathbb{R}^d$ to $\mathbb{R}^n$. That allows to reduced constrained optimization problems to nonconstrained ones, for instance.

Proposition 5.2.2. *Let $M \subset \mathbb{R}^n$ a d-dimensional smooth manifold and $x_0 \in M$. Then there is a subset $A \subset \{1, \ldots, n\}$ of cardinality d, a neighbourhood U of x_0 and $f : \pi_A(U) \to \mathbb{R}^{n-d}$ such that $M \cap U$ is the graph of f.*

Note that the tangent space to an implicitly defined smooth manifold can be expressed also implicitly. Indeed, at a point $x_0 \in M$, the tangent space is the d-dimensional subspace

$$\mathcal{T}_{x_0} M = \{x \in \mathbb{R}^n : dF(x_0)(x) = 0\}.$$

Do not confuse with the *tangent manifold* at x_0, that is the *affine* space $x_0 + \mathcal{T}_{x_0}M$. It s not difficult to prove that for a $(n-1)$-dimensional manifold the tangent manifold that can be expresed as the graph of a real function coincides with tangent plane introduced in Section 4.1.

The previous proposition only gives local information on the set, that is, being a manifold. Additional properties have to be obtained by different techniques.

Example 5.2.3. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ a convex C^1 function that attains its minimum m exactly at $(0,0)$. Show that for any $\lambda > m$, the set*

$$M_\lambda = \{(x,y) \in \mathbb{R}^2 : f(x,y) = \lambda\}$$

is a C^1 manifold that is homeomorphic to the circle $\mathbb{T}$.

Firstly note the unique critical points that a convex function can have are those where the minimum is attained. In this case, the unique point were both partial derivatives vanishes at once is $(0,0)$. That implies that

$$\left(\frac{\partial f}{\partial x}(x,y), \frac{\partial f}{\partial y}(x,y)\right) \neq (0,0)$$

for $(x,y) \neq (0,0)$, or equivalently, for $f(x,y) > m$. Therefore, M_λ is a C^1 manifold.
The global result requires a more detailed analysis. Let $r > m$ and take

$$s = \inf\{f(x,y) = x^2 + y^2 = r^2\}.$$

The convexity implies that $f(x,y) \geq (s/r)\sqrt{x^2+y^2}$ for $x^2 + y^2 \geq r^2$. We easily deduce that M_λ is bounded, and thus compact (it is evidently closed). We also deduce that f is strictly increasing on any line starting at $(0,0)$. The mapping $\phi : M_\lambda \to \mathbb{T}$ defined by

$$\phi(x,y) = \frac{(x,y)}{\sqrt{x^2+y^2}}$$

is continuous and one-to-one. That implies ϕ is an homeomorphism and therefore M_λ is homeomorphic to $\mathbb{T}$ as wished.

Elimination of variables. When we have a system of equations like

$$\begin{cases} f(x,y,z)=0 \\ g(x,y,z)=0 \end{cases}$$

we may consider the possibility of eliminating one of the variables, say z, in order to obtain a simpler relation satisfied by the two remaining variables, namely $h(x,y)=0$. Geometrically that is equivalent to find the equation of satisfied by the orthogonal projection of the one-dimensional manifold (curve) in $\mathbb{R}^3$ defined by the previous system. Typically, this projection is an one-dimensional manifold in $\mathbb{R}^2$, however there could be points of singularity, for instance when the tangent vector has null projection on the XY plane. Indeed, assume that the curve in $\mathbb{R}^3$ is parameterized by t and take the derivatives of the composed functions

$$\frac{\partial f}{\partial x}x' + \frac{\partial f}{\partial y}y' + \frac{\partial f}{\partial z}z' = 0,$$

$$\frac{\partial g}{\partial x}x' + \frac{\partial g}{\partial y}y' + \frac{\partial g}{\partial z}z' = 0.$$

If t is such that $(x'(t),y'(t)) = (0,0)$ and $z'(t) \neq 0$. That enforces that $(\frac{\partial f}{\partial z}, \frac{\partial g}{\partial z}) = (0,0)$ at that point. For that reason, the result about elimination of variables have to be local. Given a point (x_0,y_0,z_0) we can assure that z can be eliminated locally if $\frac{\partial f}{\partial z}(x_0,y_0,z_0)\cdot\frac{\partial g}{\partial z}(x_0,y_0,z_0) \neq 0$. Indeed, in that case Theorem 5.2.1 will give functions ϕ,γ defined on a neighbourhood of (x_0,y_0) such that

$$f(x,y,\phi(x,y)) = 0,$$

$$g(x,y,\gamma(x,y)) = 0$$

with $\phi(x_0,y_0) = z_0$ and $\gamma(x_0,y_0) = z_0$). In that case $h(x,y) = \phi(x,y) - \gamma(x,y)$ realises the elimination of z on a neighbourhood of (x_0,y_0,z_0).

5.3 Some applications

Our exposition along the section will be carried out with the less necessary number of variables, that means two, for the sake of simplicity.

5.3.1 Lagrange multipliers

Let $f : M \to \mathbb{R}$ be a C^1 function defined on a 1-dimensional C^1 manifold $M = \{(x, y) : g(x, y) = 0\}$ (a curve in $\mathbb{R}^2$). We look for the relative extrema (maximum or minimum) of f on M. Assume $(x_0, y_0) \in M$ is one of such points. We can represent M around $(x_0, y_0) \in M$ by a C^1 parameterization $(x(t), y(t))$ with $x_0 = x(0), y_0 = y(0)$. Necessarily we have

$$\left(\frac{d}{dt} f(x(t), y(t))\right)_{t=0} = 0.$$

Applying the chain rule that is equivalent to

$$\frac{\partial f}{\partial x}(x_0)x'(0) + \frac{\partial f}{\partial y}(y_0)y'(0) = 0.$$

The chain rule applied to $g(x(t), y(t)) = 0$ at $t = 0$ gives

$$\frac{\partial g}{\partial x}(x_0)x'(0) + \frac{\partial g}{\partial y}(y_0)y'(0) = 0.$$

As $(x'(0), y'(0)) \neq (0, 0)$ the vectors

$$\nabla f(x_0, y_0) = \left(\frac{\partial f}{\partial x}(x_0), \frac{\partial f}{\partial y}(y_0)\right),$$

$$\nabla g(x_0, y_0) = \left(\frac{\partial g}{\partial x}(x_0), \frac{\partial g}{\partial y}(y_0)\right)$$

are linearly dependent. From a geometrical point of view, that means that the level curves of f and g are tangent at (x_0, y_0). The hypothesis on g (M is a manifold) implies the existence of $\lambda \in \mathbb{R}$ such that

$$\nabla f(x_0, y_0) + \lambda \nabla g(x_0, y_0) = 0.$$

Note that the argument works the other way around, so the existence of such a λ implies that (x_0, y_0) is a critical point of f on M. Consequently, we deduce that the extrema of f on M are contained among the solutions of the equations

$$\begin{aligned} \nabla f(x_0, y_0) + \lambda \nabla g(x_0, y_0) &= 0, \\ g(x, y) &= 0. \end{aligned} \tag{5.1}$$

Curiously, the solutions of the system (5.1) is equivalent to the search of critical points of the function

$$F(x, y, \lambda) = f(x, y) + \lambda g(x, y).$$

The new variable λ is called *Lagrange multiplier* and its introduction reduces the constrained problem of extrema to an unconstrained problem. That can be done in similar terms with more variables and constraints, adding one multiplier by each constraint. For instance, looking for the extrema of $f(x, y, z)$ on the 1-dimensional manifold $\{(x, y, z) : g(x, y, z) = h(x, y, z) = 0\}$ is equivalent to investigate the critical points of

$$F(x, y, z, \lambda, \nu) = f(x, y, z) + \lambda g(x, y, z) + \nu h(x, y, z).$$

Example 5.3.1. *The production function of Cobb-Douglas (with* 3 *variables) is a function that modelizes the profits after a investment in different stages of the manufacturing of a product: materials, machinery... and maybe tech and marketing too. The function has the form*

$$f(x, y, z) = c\, x^\alpha y^\beta z^\gamma,$$

where $c, \alpha, \beta, \gamma > 0$ *and by homogeneity we should have* $\alpha + \beta + \gamma = 1$.

We wish to maximize the production f with a limited budget $x + y + z \leq m$. Obviously, we can restrict ourselves to a budget equal to m. As Lagrange auxiliary function we can take

$$F(x, y, z, \lambda) = x^\alpha y^\beta z^\gamma + \lambda(x + y + z).$$

The partial derivatives should be zero

$$\alpha x^{\alpha-1} y^\beta z^\gamma + \lambda = 0$$

$$\beta x^\alpha y^{\beta-1} z^\gamma + \lambda = 0$$

$$\gamma x^\alpha y^\beta z^{\gamma-1} + \lambda = 0$$

Multiplying by x, y, z respectively and adding we get

$$(\alpha + \beta + \gamma) x^\alpha y^\beta z^\gamma + \lambda(x + y + z) = x^\alpha y^\beta z^\gamma + \lambda m = 0.$$

Using that information in the first equation we get

$$\alpha x^{\alpha-1} y^\beta z^\gamma = m^{-1} x^\alpha y^\beta z^\gamma,$$

therefore $x = \alpha m$. Analogously, $y = \beta m$ and $z = \gamma m$.

5.3.2 Functional dependence

Now we will discuss functional dependence. It is an easy task to check that the functions $\cos x$ and $\sin x$ are linearly independent. However they are *algebraically dependent* since $\cos^2 x + \sin^2 x = 1$. More generaly, we say that the functions $f_k : D \subset \mathbb{R}^n \to \mathbb{R}$ for $k = 1, \dots, m$ are *functionally dependent* if there is a nontrivial $F : \Omega \subset \mathbb{R}^m \to \mathbb{R}$ such that

$$F(f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)) = 0.$$

Here "nontrivial" means dF of maximal rank. Note that any pair of C^1 one variable functions f and g are functionally dependent on some interval. Indeed, we may assume that $(f'(x_0), g'(x_0)) \neq (0, 0)$, otherwise both functions are constant and so dependent. Therefore, one of the functions is locally monotone. Let us assume it is f, and thus f^{-1} is defined on some neighbourhood. Now note that $F(f(x), g(x)) = 0$ where $F(u, v) = g(f^{-1}(u)) - v$ is non trivial (rank 1). For a couple of functions f and g defined on an open subset of $\mathbb{R}^2$ its functional dependence is locally equivalent to another one of the form $g(x, y) = G(f(x, y))$ thanks to the implicit function theorem. Observe that $\nabla g = G' \nabla f$ at every point and thus

$$\frac{\partial(f, g)}{\partial(x, y)} = 0,$$

where we are using the standard notation for the Jacobian determinant.

In general, if $m > n$ in the definition above the functions are necessarily functionally dependent. That is a consequence of this important result.

Theorem 5.3.2. *Let $f_k : D \subset \mathbb{R}^n \to \mathbb{R}$ for $k = 1, \dots, n$ functions, at least C^1. If they are functionally dependent, then*

$$\frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)} = 0$$

on D. Conversely, if the Jacobian vanishes on D then the functions (f_k) are functionally dependent locally at any point of D.

Proof. Suppose that the last function can be expressed by means of the others by means of some nontrivial C^1 function $G(y_1, \dots, y_{n-1})$ as

$$f_n(x_1, \dots, x_n) = G(f_1(x_1, \dots, x_n), \dots, f_{n-1}(x_1, \dots, x_n)).$$

Then the chain rule implies

$$\nabla f_n = \sum_{k=1}^{n-1} \frac{\partial G}{\partial y_k} \nabla f_k$$

and so the jacobian vanishes. The converse is a little more technical, thus we will prove the particular case $n = 2$ which enough to show the ideas behind. Suppose we are given functions $f(x, y)$ and $g(x, y)$ such that

$$\frac{\partial(f, g)}{\partial(x, y)} = 0$$

on $D \subset \mathbb{R}^2$. Consider the system of equations

$$u - f(x, y) = 0$$

$$v - g(x, y) = 0$$

Assume that the coefficients of $\frac{\partial(f,g)}{\partial(x,y)}$ do not vanishes at once on any open subset, otherwise all the functions are constant there and so they are functionally dependent. Without loss of generality we may assume that $\frac{\partial f}{\partial y} \neq 0$ on some open subset. In that case, we may use the first equation to solve y as a function of (x, u), that is, $y = \phi(x, u)$. Later we will need the derivative $\frac{\partial \phi}{\partial x}$ expressed in terms of f. That can be done by implicit derivation

$$0 = \frac{\partial}{\partial x}(u - f(x, \phi(x, u)) = -\frac{\partial f}{\partial x} - \frac{\partial f}{\partial y}\frac{\partial \phi}{\partial x}$$

therefore $\frac{\partial \phi}{\partial x} = -(\frac{\partial f}{\partial y})^{-1}\frac{\partial f}{\partial x}$. Consider the composition

$$G(x, u) = g(x, \phi(x, u))$$

Now we compute the partial derivative with respect to x

$$\frac{\partial G}{\partial x} = \frac{\partial g}{\partial x} + \frac{\partial g}{\partial y}\frac{\partial \phi}{\partial x} = \frac{\partial g}{\partial x} - \frac{\partial g}{\partial y}\left(\frac{\partial f}{\partial y}\right)^{-1}\frac{\partial f}{\partial x} = -\left(\frac{\partial f}{\partial y}\right)^{-1}\frac{\partial(f, g)}{\partial(x, y)} = 0.$$

That means that actually G depends only on u, the substitution $u = f(x, y)$ and $\phi(x, u) = y$ in order to remove u gives

$$G(f(x, y)) = g(x, y)$$

which is a functional dependence valid on some open subset. ∎

5.3.3 Envelope of a family of curves.

Consider a family of curves in $\mathbb{R}^2$ depending of a parameter. The more general way to express such a family is the implicit form

$$f(x, y, t) = 0$$

where t is the parameter. Sometimes it happen that there exists a curve $\phi(x, y) = 0$ that meets all the curves of the family exactly at one point for every value of t and the curves $\phi(x, y) = 0$ and $f(x, y, t) = 0$ are tangent. For instance, the family of straight lines

$$x \cos t + y \sin t = 1$$

are tangent to the circle $x^2 + y^2 = 1$. We say that such a curve $\phi(x, y) = 0$ is the (or a) envelope of the family. We are going to show how to obtain ϕ from f. First of all, note that the curves $f(x, y, t_0) = 0$ and $f(x, y, t) = 0$ for $t \neq t_0$ meet typically at some point $(x(t), y(t))$. If $t \sim t_0$ the point $(x(t), y(t))$ is close to the envelope. We could formalize this by saying that

$$(x_0, y_0) = \lim_{t \to t_0} (x(t), y(t)) \in \{(x, y) : \phi(x, y) = 0\}$$

if such a limit exists. Moreover, in such a case we have

$$0 = \lim_{t \to t_0} \frac{f(x(t), y(t), t) - f(x(t), y(t), t_0)}{t - t_0} = \frac{\partial f}{\partial t}(x_0, y_0, t_0)$$

provided that f is C^1. Therefore, if $f(x, y, t_0)$ and $\phi(x, y)$ meet at the point (x_0, y_0) then

$$\begin{aligned} f(x_0, y_0, t_0) &= 0, \\ \frac{\partial f}{\partial t}(x_0, y_0, t_0) &= 0. \end{aligned} \tag{5.2}$$

We can find $\phi(x, y)$ by removing t from the system of equations

$$\begin{cases} f(x, y, t) = 0, \\ \dfrac{\partial f}{\partial t}(x, y, t) = 0. \end{cases} \tag{5.3}$$

That can be done likewise for envelopes of families of surfaces depending on one parameter, or families of spatial curves depending on two parameters.

Example 5.3.3. *Find the envelope of all the trajectories of an object which is thrown from the same point, at the same speed and only affected by the (uniform) gravitational force.*

Without loss of generality, the objects departs from the origin. We will consider only the trajectories contained in a vertical plane XY (the spatial case will follow by symmetry). Let v be the speed, θ the angle of depart and g denote the gravitational force per unit of mass. Elementary Newtonian Physics gives the trajectory as a function of the time t ($t = 0$ at the depart moment)

$$x = (v\cos\theta)\,t,$$

$$y = (v\sin\theta)\,t - (g/2)\,t^2.$$

The parameter time can be eliminated (put $t = x(v\cos\theta)^{-1}$ in the second equation)

$$y = \frac{xv\sin\theta}{v\cos\theta} - \frac{g}{2}\left(\frac{x}{v\cos\theta}\right)^2 = (\tan\theta)x - \frac{g}{2v^2}(\tan^2\theta + 1)x^2.$$

Therefore, the family of trajectories in terms of the angle θ is given by

$$y - (\tan\theta)x + \frac{g}{2v^2}(\tan^2\theta + 1)x^2 = 0.$$

Derivation with respect to θ gives

$$-(\tan^2\theta + 1)x + \frac{g}{2v^2}(2\tan\theta)(\tan^2\theta + 1)x^2 = 0.$$

The factor $\tan^2\theta + 1$ can be eliminated, so $x\,\tan\theta = v^2/g$. The substitution above produces

$$0 = y - \frac{v^2}{g} + \frac{g}{2v^2}\left(\frac{v^2}{g}\right)^2 + \frac{g}{2v^2}x^2 = y - \frac{v^2}{g} + \frac{g}{2v^2}x^2,$$

that shows that the envelope is a larger parabola.

5.4 Rationale and remarks

Note that Lemma5.1.1 and Lema 5.1.2 were stated in the Banach space setting, however Theorem 5.1.4 and Theorem 5.2.1 are restricted to $\mathbb{R}^d$. The inverse

mapping theorem and the implicit functions theorem are still valid in the Banach frame, but the proof requires to know that the inversion of operators is a C^∞ mapping, analytic in fact.

The notion of smooth manifold is studied at large in Differential Geometry. We merely need the existence of local parameterizations for our purposes. I have included a discussion on the elimination of variables because is a very usual operation, and yet is seldom treated from a theoretical point of view in modern texts. Functional dependence is also a forgotten classic topic.

There are interesting directions to suggest for a TFG. For instance, Saint-Raymond proved the inverse mapping theorem on $\mathbb{R}^n$ under weaker hypotheses, or the study of global invertibility following the ideas of Hadamard.

5.5 Exercises

1. Study the local and global invertibility of the mapping $f : D \subset \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$f(x,y,z) = (\frac{x}{1-x-y-z}, \frac{y}{1-x-y-z}, \frac{z}{1-x-y-z}),$$

   where $D = \{(x,y,z) \in \mathbb{R}^3 : x+y+z \neq 1\}$.

2. Study the local and global invertibility of the mapping $f : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f(x,y) = (x^2 - y^2, 2xy)$.

3. Consider the mapping $J : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}^2$ defined by means of polar coordinates on the domain and Cartesian for the image by

$$(r,\theta) \to ((r+1/r)\cos\theta, (r-1/r)\sin\theta).$$

   Prove that every point of $\mathbb{R}^2 \setminus \{(-2,0),(2,0)\}$has exactly two preimages. Find the maximal regions in $\mathbb{R}^2$ where J is a diffeomorphism.

4. Show that the equation $x^2 + xy + y^3 - 11 = 0$ defines y as a function of x around $x = 1$, taking the value $y = 2$. Compute the first and second derivatives of that function at $x = 1$.

5. The mapping $f(x,y,z) = (y^3+z^5, x+z^5, x+y^3)$ is globally invertible on $\mathbb{R}^3$ ¿Does it satisfies the hypotheses of the inverse mapping theorem at $(0,0,0)$? ¿What is the relation with the possible differentiability of f^{-1}?

6. Consider the mapping $f : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f(x, y) = (u, v)$ where $u = x$, $v = y - x^2$ if $x^2 \leq y$, $v = (y^2 - x^2 y)/x^2$ if $0 \leq y < x^2$ and $v(x, y) = -v(x, -y)$ in case that $y < 0$. Prove that f is differentiable at $(0, 0)$, compute its differential and show that it is one-to-one ¿It verified the inverse mapping theorem around $(0, 0)$?

7. Assume that the equation $f(x, y, z) = 0$ defines every variable as a function of the remaining two ones. Show that

$$\frac{\partial x}{\partial y}\frac{\partial y}{\partial z}\frac{\partial z}{\partial x} = -1.$$

8. Find the extreme values of the implicit functions defined by the equation

$$y^3 - x^2 y + x^3 - 3 = 0.$$

9. Prove that the equation

$$\sin x + \cos y = 1,$$

defines y as a function of x around $(\pi/2, \pi/2)$. Find the first and second derivatives of the implicit function at that point.

10. Prove that the equation $x^y - y^x = 0$ defines $y = f(x)$ around $(2, 4)$ and compute $f'(2)$. Find the largest open interval where $f(x)$ is defined. Is there any implicit function defined around (e, e)? Find the largest open interval where $f(x)$ is defined.

11. The polynomial $x^3 - \lambda x^2 + \lambda^2 x - 1$ has a unique real root $\rho(\lambda)$ when the parameter λ runs on a neighbourhood of 0. Suppose that $\lambda > 0$ is very near to 0 ¿Is it possible that $\rho(\lambda) < 1$?

12. Prove that the set of roots of an algebraic polynomial

$$x^n + a_{n-1}x^{n-1} + \cdots + a_0$$

regarded as a compact subset in $\mathbb{C}$ with the Hausdorff metric on $\mathbb{R}^2$, is a continuous function of the n coefficients of the polynomial.

13. Prove that the equation

$$\cos(x^2 + y) + \sin(x + y) + e^{x^2 y} = 2$$

defines y as a function g of class C^∞ on a a neighbourhood of $x = 0$, and $g(0) = \pi/2$. Show that g has a local minimum at $x = 0$.

14. Prove that the equation
$$\pi \cos\theta = t\,\theta$$
has a unique solution $\theta(t)$ for t in a neighbourhood of $3/2$ and find $\theta(3/2)$. Prove also that $\theta'(t)$ exists on a neighbourhood and compute $\theta'(3/2)$.

15. Prove that the equations
$$4x^2 - 3y^2 - z = 0$$
$$x^2 + y^2 + z^2 = 24$$
define a C^∞ curve on a neighbourhood of $(2, -2, 4)$. Find the tangent line at that point. Show that, actually, the equations define a closed C^∞ curve.

16. Let $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ be a C^1 mapping and $x_0 \in D$. Prove that:
    (a) if $df(x_0)$ is one-to-one, then there is a neighbourhood V of x_0 such that $f|_V$ is one-to-one;
    (b) if $df(x_0)$ is onto, then there is a neighbourhood V of x_0 such that $f(V)$ is a neighbourhood of $f(x_0)$.

17. Check that these functions are functionally dependent and find their relation
$$f(x, y) = \frac{x}{y}; \quad g(x, y) = \frac{x - y}{x + y}.$$

18. Check that these functions are functionally dependent and find their relation
$$f(x, y) = 2xy + 2x + 1; \quad g(x, y) = x^2y^2 + 2x^2y + x^2 - 1.$$

19. Check that these functions are functionally dependent and find their relation
$$f(x, y, z) = x^2 + y^2 + z^2;\ g(x, y, z) = x + y + z;\ h(x, y, z) = xy + yz + zx.$$

20. Let $M \subset \mathbb{R}^3$ be C^1 a compact manifold of dimension 2, that is, a C^1 surface in $\mathbb{R}^3$. Assume that M is oriented and there is continuous normal field $\vec{N}$. Prove that $\vec{N}$ takes all the values in $\mathbb{R}^3$. Show that the statement is not true if we drop compactness or the manifold is piecewise C^1.

21. Consider the set $P \subset \mathbb{R}^3$ defined by the equations

$$x^2 + 4y^2 = 16$$

$$9x^2 + 16z^2 = 144$$

Show that y, z are defined as C^∞ functions of x around $(0, 2, 3)$. Find the first and second derivatives at that point and the tangent line. Prove that P is not a manifold.

22. Assume that the function f implicitly defined in a neighbourhood of x_0 by $F(x, y) = 0$ and $f(x_0) = y_0$ has a critical point at x_0. Prove that if F is C^2 and

$$\frac{\partial^2 F}{\partial x^2}(x_0, y_0) \cdot \frac{\partial F}{\partial y}(x_0, y_0) > 0,$$

then f has a local maximum at x_0.

23. Let $F : B_{\mathbb{R}^n} \to B_{\mathbb{R}^n}$ be a C^1 contractive mapping. Prove that for every $t \in (0, 1]$ the mapping $F_t(x) := tF(x)$, which is also defined on $B_{\mathbb{R}^n}$, has a unique fixed point $x(t)$ that continuously depends on the parameter t. Compute $\lim_{t \to 0^+} x(t)$.

24. Let $f : \mathbb{R} \to \mathbb{R}$ be a C^2 function and let $x_0 \in \mathbb{R}$ be such that $f''(x_0) > 0$. Show that the equality

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

define $y = g(x)$ implicitly as a function of x on an interval (x_0, x_1). Prove the existence and find the value of the limit

$$\lim_{x \to x_0^+} \frac{g(x) - x_0}{x - x_0}.$$

25. An object is thrown from the origin of $\mathbb{R}^2$ with the same speed and variable direction, and its movement is affected only by its weight so it follows parabolic trajectories. Find the envelope of the family of all possible trajectories.

26. Let $f(x_1, \ldots, x_n) = a_1 x_1 + \cdots + a_n x_n$. Find the maximum of f on the set

$$B_p^n = \{(x_1, \ldots, x_n) : |x_1|^p + \cdots + |x_n|^p \leq 1\}.$$

27. Let X be a Banach space and let $\mathfrak{L}(X)$ denote the linear continuous operators acting on X with the operator norm.

    (a) Let $A \in \mathfrak{L}(X)$ such that $\|\mathbb{I} - A\| < 1$, where $\mathbb{I}$ lis the identity map on X. Show that A is invertible.

    (b) Assume that $A \in \mathfrak{L}(X)$ is an invertible. Prove the existence of $\delta > 0$ such that if $B \in \mathfrak{L}(X)$ with $\|B - A\| < \delta$, then B is invertible.

    (c) Deduce that the assignation $A \to A^{-1}$ within the invertible operators of $\mathfrak{L}(X)$ is continuous and, moreover, it is C^∞.

Chapter 6

Riemann Integral

6.1 Rectangles and partitions

In all that follows we will assume that the dimension of the space is a fixed number $d \in \mathbb{N}$. The case $d = 1$ is the one dimensional Riemann integral that has been studied previously in first year, but the characterizations of Riemann integrability in terms of continuity point are not likely studied in that setting. A rectangle in $\mathbb{R}^d$ will always be a *compact rectangle* $R = [a_1, b_1] \times \cdots \times [a_d, b_d]$ unless we specify other kind of rectangle (open). The d-dimensional volume of the rectangle is the nonnegative number

$$\mathbf{m}(R) = (b_1 - a_1)(b_2 - a_2) \dots (b_d - a_d)$$

The rectangle is non degenerate if $\mathbf{m}(R) > 0$. Clearly, the topological interior of R is the set

$$(a_1, b_1) \times \cdots \times (a_d, b_d)$$

Two rectangles R_1 and R_2 are said not overlapping if they meet only on their borders.

Any non degenerate rectangle R can be tiled with smaller non degenerate rectangles $\{R_i\}_{i=1}^n$ which are pairwise not overlapping. To see that, just consider the rectangles of the form $I_1 \times \cdots \times I_d$ where each I_k is an interval coming from a finite partition of $[a_k, b_k]$. Then arrange all these rectangles into a sequence $\{R_i\}_{i=1}^n$. The tiling $\{R_i\}_{i=1}^n$ of R obtained in this way is called a *grill* of R. It is not difficult to see that $\mathbf{m}(R) = \sum_{i=1}^n \mathbf{m}(R_i)$ in this case, but something more general is true. Given a rectangle R, a collection $\pi = \{R_i\}_{i=1}^n$ is said a partition of R if they are not overlapping and $\bigcup_{i=1}^n R_i = R$.

Proposition 6.1.1. *If $\{R_i\}_{i=1}^n$ is a partition of a rectangle R, then*

$$\mathbf{m}(R) = \sum_{i=1}^{n} \mathbf{m}(R_i).$$

Proof. Assume that $R = [a_1, b_1] \times \cdots \times [a_d, b_d]$ is non degenerate, since in other case the result is trivial. As well, we may assume that $\{R_i\}_{i=1}^n$ contain no degenerate rectangle, since after removing them we still have $\bigcup_{i=1}^n R_i = R$ (the union of the interiors are dense in R). Fix a coordinate $1 \leq k \leq d$. The k-projection of R_i is a subinterval of $[a_k, b_k]$. Consider the one-dimensional partition of $[a_k, b_k]$ generated for all the endpoints of such intervals for $1 \leq i \leq n$, and then consider the grill $\{R'_j\}_{j=1}^m$ obtained form those intervals by cartesian products. For each $1 \leq j \leq m$ there is exactly one $1 \leq i \leq n$ such that $R'_j \subset R_i$ since both have nonempty interior. Consider the sets $A_i = \{j : R'_j \subset R_i\}$ for $1 \leq i \leq n$ which are disjoint and $\bigcup_{i=1}^n A_i = \{1, \ldots, m\}$. Observe that $\{R'_j\}_{j \in A_i}$ is a partition of R_i. Now

$$\sum_{i=1}^{n} \mathbf{m}(R_i) = \sum_{i=1}^{n} \sum_{j \in A_i} \mathbf{m}(R'_j) = \sum_{j=1}^{m} \mathbf{m}(R'_j) = \mathbf{m}(R)$$

■

With similar arguments it is possible to prove the following

Proposition 6.1.2. *If $\{R_i\}_{i=1}^n$ is collection of non overlapping rectangles and $\{R'_j\}_{i=1}^m$ is another collection of rectangles such that $\bigcup_{i=1}^n R_i \subset \bigcup_{j=1}^m R'_j$, then $\sum_{i=1}^n \mathbf{m}(R_i) \leq \sum_{j=1}^m \mathbf{m}(R'_j)$.*

A partition $\pi' = \{R'_j\}_{j=1}^m$ is finer than $\pi = \{R_i\}_{i=1}^n$ if for every $j : 1 \ldots m$ there is $i : 1 \ldots n$ such that $R'_j \subset R_i$. Observe that in this case we have

$$R_i = \bigcup \{R'_j : R'_j \subset R_i\}.$$

Given two partitions $\pi = \{R_i\}_{i=1}^n$ and $\pi' = \{R'_j\}_{j=1}^m$ is always possible to find a third partition which is finer. Just take the rectangles $R_i \cap R'_j$ having nonempty interior.

6.2 Integrals on compact rectangles

Given a bounded function $f : R \to \mathbb{R}$ defined on a rectangle and partition $\pi = \{R_i\}_{i=1}^n$ of R, we consider the numbers

$$L(f,\pi) = \sum_{i=1}^{n} \inf\{f, R_i\}\mathbf{m}(R_i)$$

$$U(f,\pi) = \sum_{i=1}^{n} \sup\{f, R_i\}\mathbf{m}(R_i)$$

named lower and upper sums respectively. Observe that for $\pi_1 \leq \pi_2$ partitions of R we always have

$$L(f,\pi_1) \leq L(f,\pi_2) \leq U(f,\pi_2) \leq U(f,\pi_1)$$

The Darboux lower and upper integrals of f (on R) are defined this way

$$\underline{\int} f = \sup\{L(f,\pi) : \pi \text{ partition of } R\}$$

$$\overline{\int} f = \inf\{U(f,\pi) : \pi \text{ partition of } R\}.$$

Definition 6.2.1. *A bounded function $f : R \to \mathbb{R}$ is said Riemann integrable (on R) if $\underline{\int} f = \overline{\int} f$. In that case, its integral (in Riemann sense) is that common value $\int f = \int_R f := \underline{\int} f = \overline{\int} f$.*

Recall that the oscillation of a function $f : R \to \mathbb{R}$ on a set $A \subset R$ is the number

$$\operatorname{osc}(f, A) = \sup\{|f(x) - f(y)| : x, y \in A\}$$

In order to establish the properties of integrable functions the following criterion will be very useful.

Proposition 6.2.2. *A bounded function $f : R \to \mathbb{R}$ is Riemann integrable if and only if for every $\varepsilon > 0$ there is a partition $\pi = \{R_i\}_{i=1}^n$ of R such that*

$$\sum_{i=1}^{n} \operatorname{osc}(f, R_i)\mathbf{m}(R_i) < \varepsilon$$

Hint of Proof. Just notice that $\operatorname{osc}(f, R_i) = \sup\{f, R_i\} - \inf\{f, R_i\}$. ∎

The first application provides us with an important class of integrable functions.

Corollary 6.2.3. *If $f : R \to \mathbb{R}$ is continuous, then it is Riemann integrable.*

Proof. Since R is compact, then f is uniformly continuous. Given $\varepsilon > 0$, take a partition $\pi = \{R_i\}_{i=1}^n$ made of rectangles small enough to guarantee that $\operatorname{osc}(f, R_i) < \varepsilon/\mathbf{m}(R)$. ∎

The reader that is acquainted with the properties of the Riemann integral for one variable functions will not see anything new in the following result.

Proposition 6.2.4. *Let $\mathfrak{R}(R)$ denote the set of functions which are Riemann integrable on R. Then*

1. *$\mathfrak{R}(R)$ is a vector space and $\int_R(\alpha f + \beta g) = \alpha \int_R f + \beta \int_R g$ whenever $f, g \in \mathfrak{R}(R)$ and $\alpha, \beta \in \mathbb{R}$.*
2. *$\mathfrak{R}(R)$ is stable by products (so it is an algebra).*
3. *If $f, g \in \mathfrak{R}(R)$ and $f \leq g$, then $\int_R f \leq \int_R g$.*
4. *If $f \in \mathfrak{R}(R)$, then $f^+, f^-, |f| \in \mathfrak{R}(R)$ and $|\int_R f| \leq \int_R |f|$.*
5. *If $f \in \mathfrak{R}(R)$ and $S \subset R$ a rectangle, then $f|_S \in \mathfrak{R}(S)$.*
6. *If $f \in \mathfrak{R}(R)$ and $\{R_i\}_{i=1}^n$ is a partition of R, then $\int_R f = \sum_{i=1}^n \int_{R_i} f$.*

Hint of Proof. Observe that

$$\underline{\int_R} f + \underline{\int_R} g \leq \underline{\int_R}(f+g) \leq \overline{\int_R}(f+g) \leq \overline{\int_R} f + \overline{\int_R} g$$

and

$$\overline{\int_R} \alpha f = \alpha \overline{\int_R} f, \quad \underline{\int_R} \alpha f = \alpha \underline{\int_R} f$$

for $\alpha > 0$, while if $\alpha < 0$ then

$$\overline{\int_R} \alpha f = \alpha \underline{\int_R} f, \quad \underline{\int_R} \alpha f = \alpha \overline{\int_R} f.$$

Integrability of products can be reduced to integrability of squares of positive functions. In such a case, we have

$$\operatorname{osc}(f^2, A) \leq 2 \sup\{f, A\} \operatorname{osc}(f, A)$$

which is suitable for that purpose. ∎

6.3 Integrability and continuity points

The goal of this section is to give a characterization of Riemann integrability by means of the set of continuity points of the function. Let us begin with a simple but useful observation.

Proposition 6.3.1. *If $f \in \mathfrak{R}(R)$, $f \geq 0$ and $\int_R f = 0$, then $f(x) = 0$ whenever $x \in R$ is a point of continuity of f.*

A bounded set $A \subset \mathbb{R}^d$ is said of null content if for every $\varepsilon > 0$ there is a family of rectangles $\{R_i\}_{i=1}^n$ such that $A \subset \bigcup_{i=1}^n R_i$ and $\sum_{i=1}^n \mathbf{m}(R_i) < \varepsilon$. Notice that being of content null is stable by subsets, finite unions and closures.

A set $A \subset \mathbb{R}^d$ is said of null measure if for every $\varepsilon > 0$ there is a family of rectangles $\{R_i\}_{i=1}^\infty$ such that $A \subset \bigcup_{i=1}^\infty R_i$ and $\sum_{i=1}^\infty \mathbf{m}(R_i) < \varepsilon$. Measure null sets are stable by subsets and countable unions. Of course, content null sets are measure null, but the converse is not true: just consider $A = [0,1] \cap \mathbb{Q}$. As the countable union of its singletons it is of null measure. On the other hand, $\overline{A} = [0,1]$ so this set cannot be of null content. Notice that a compact set of null measure is of null content since there is no restriction in considering the cover made of open rectangles (slightly larger ones).

Given a function $f : R \to \mathbb{R}$, we may define its oscillation at some $x \in R$ as

$$\operatorname{osc}(f, x) = \inf\{\operatorname{osc}(f, U) : U \text{ neighborhood of } x\}.$$

Observe that f is continuous at x if and only if $\operatorname{osc}(f, x) = 0$. Moreover, for every $\delta > 0$ the set $\{x \in R : \operatorname{osc}(f, x) < \delta\}$ is open (relatively to R). The following is the celebrated Riemann-Lebesgue characterization of the Riemann integrability.

Theorem 6.3.2. *Let $f : R \to \mathbb{R}$ be a bounded function defined on a non degenerate compact rectangle $R \subset \mathbb{R}^d$. The following statements are equivalent:*

i) f is Riemann integrable on R;

ii) $\{x \in R : \operatorname{osc}(f,x) \geq \delta\}$ is of null content for every $\delta > 0$;

iii) the set of discontinuity points of f is of null measure.

Proof. Note that the equivalence between *ii)* and *iii)* is consequence of this set equality

$$\{x \in R : \operatorname{osc}(f,x) > 0\} = \bigcup_{n=1}^{\infty} \{x \in R : \operatorname{osc}(f,x) \geq 1/n\}$$

bearing in mind that the first are the discontinuity points of f and the second is represented as a union of compact subsets of R.
Suppose that f is Riemann integrable. For $\varepsilon, \delta > 0$, take a partition $\{R_i\}_{i=1}^n$ of R into rectangles such that

$$\sum_{i=1}^{n} \operatorname{osc}(f,R_i)\mathbf{m}(R_i) < \delta\varepsilon$$

Consider the open set $O = \bigcup_{i=1}^n R_i^\circ$. If $y \in O \cap \{x \in R : \operatorname{osc}(f,x) > \delta\}$, then $\operatorname{osc}(f,R_i) > \delta$ if $y \in R_i$. Take $N = \{i : 1 \leq i \leq n, \operatorname{osc}(f,R_i) > \delta\}$ and observe that

$$\delta \sum_{i \in N} \mathbf{m}(R_i) < \sum_{i \in N} \operatorname{osc}(f,R_i)\mathbf{m}(R_i) < \delta\varepsilon$$

following that $O \cap \{x \in R : \operatorname{osc}(f,x) > \delta\}$ is covered by $\{R_i\}_{i \in N}$. Since $R \setminus O = \bigcup_{i=1}^n \partial R_i$ is of null content and $\varepsilon > 0$ arbitrary, we deduce that $\{x \in R : \operatorname{osc}(f,x) > \delta\}$ is of null content.
Supose now that statement *ii)* holds. Given $\varepsilon > 0$, set $M = \operatorname{osc}(f,R)$ and take a cover $\{S_j\}_{j=1}^m$ by open rectangles of the set $\{x : \operatorname{osc}(f,x) \geq \varepsilon/\mathbf{m}(R)\}$ such that $\sum_{j=1}^m \mathbf{m}(S_j) < \varepsilon/M$. If $O = \bigcup_{j=1}^m S_j$, then $R \setminus O$ is compact. Every $x \in R \setminus O$ has an open neighborhood U_x such that $\operatorname{osc}(f,U_x) < \varepsilon/\mathbf{m}(R)$. Let $\xi > 0$ be the Lebesgue number of the covering $\{U_x\}_{x \in R \setminus O}$. Note that $R \setminus O$ is a finite union of non overlaping rectangles, that can be decomposed into smaller nonoverlaping rectangles of diameter less than ξ. That family of rectangles can be extended to a partition $\{R_i\}_{i=1}^n$ of R adding rectangles filling $R \cap O$. With all these ingredients we have

$$\sum_{i=1}^{n} \operatorname{osc}(f,R_i)\mathbf{m}(R_i) = \sum_{R_i^\circ \subset O} \operatorname{osc}(f,R_i)\mathbf{m}(R_i) + \sum_{R_i \subset R \setminus O} \operatorname{osc}(f,R_i)\mathbf{m}(R_i)$$

$$\leq M \sum_{R_i^\circ \subset O} \mathbf{m}(R_i) + \frac{\varepsilon}{\mathbf{m}(R)} \sum_{R_i \subset R \setminus O} \mathbf{m}(R_i) \leq M \frac{\varepsilon}{M} + \frac{\varepsilon}{\mathbf{m}(R)} \mathbf{m}(R) = 2\varepsilon.$$

That proves the Riemann integrability of f. ∎

Corollary 6.3.3. *If $f \in \mathfrak{R}(R)$, $f \geq 0$ and $\int_R f = 0$, then $\{x \in R : f(x) \neq 0\}$ is of null measure.*

6.4 Integration on general domains

Let $D \subset \mathbb{R}^d$ a bounded subset and $f : D \to \mathbb{R}$ a bounded function. We say that f is Riemann integrable on D if given a compact rectangle $R \supset D$, the function $\tilde{f} : R \to \mathbb{R}$ defined as $\tilde{f}(x) = f(x)$ if $x \in D$ and $f(x) = 0$ if $x \in R \setminus D$ is Riemann integrable on R. In such a case, we take

$$\int_D f := \int_R \tilde{f}.$$

It is not difficult to check that the definition is independent of the chosen rectangle R, and taking $\mathfrak{R}(D)$. Properties of function integrables on rectangles extend naturally to $\mathfrak{R}(D)$. In a similar fashion, for $f : \mathbb{R}^d \to \mathbb{R}$ with compact support, that is, if the set $\{x \in \mathbb{R}^d : f(x) \neq 0\}$ is bounded, we may define $\int f$ in terms of integration on rectangles.

The integrability of $f : D \to \mathbb{R}$ depends on the continuity points of the extended function $\tilde{f}$ which in turn depends both on the values of f and the "distribution" of D into $\mathbb{R}^d$. It seems to be a good idea to investigate the sets of $\mathbb{R}^d$ where the continuous functions, at least, are integrable.

Definition 6.4.1. *A bounded subset $A \subset \mathbb{R}^d$ is said Jordan measurable (or Jordan domain) if its indicator function χ_A is Riemann integrable. In such a case, the number $\mathbf{c}(A) = \int \chi_A$ is called the Jordan content of A.*

Observe that null content sets are those Jordan measurable sets having content zero. For a bounded set $A \subset \mathbb{R}^d$ we may define the inner content $\mathbf{c}_*(A) = \underline{\int} \chi_A$ and the outer content as $\mathbf{c}^*(A) = \overline{\int} \chi_A$. We have that a bounded set$A$ is measurable Jordan if and only if $\mathbf{c}_*(A) = \mathbf{c}^*(A)$, whose interpretation is related to the Greek's exhaustion method for areas and volumes.

Proposition 6.4.2. *A bounded subset $A \subset \mathbb{R}^d$ is Jordan measurable if and only if its boundary ∂A is of null content.*

Proof. The discontinuities of χ_A happen exactly at the points of ∂A. ∎

We have defined the Jordan content from the Riemann integral. The other way around is posible as shows the following result. The details of the proof are left to the reader.

Proposition 6.4.3. *Let $R \subset \mathbb{R}^d$ be a rectangle.*

1. *If $f : R \to [0, +\infty)$ a bounded function and consider $F = \{(x, t) : x \in R, 0 \leq t \leq f(x)\}$. Then*

$$\underline{\int_R} f = \mathbf{c}_*(F), \qquad \overline{\int_R} f = \mathbf{c}^*(F)$$

   where the Jordan content is taken in $\mathbb{R}^{d+1}$. In particular, f is Riemann integrable if and only if F is Jordan measurable, and then $\int_R f = \mathbf{c}(F)$.

2. *Bounded sets defined by subgraphs and epigraphs of Riemann integrable functions are Jordan measurable.*

3. *A bounded function $f : R \to \mathbb{R}$ is Riemann integrable if and only if its graph $\{(x, f(x)) : x \in R\}$ is of null content in $\mathbb{R}^{d+1}$.*

We have the mean value property of the integral.

Proposition 6.4.4. *If D is a Jordan set and $f \in \mathfrak{R}(D)$ then*

$$\inf\{f, D\} \leq \frac{1}{\mathbf{c}(D)} \int_D f \leq \sup\{f, D\}.$$

Proof. Just compare f with $\lambda\chi_D$ with $\lambda \in \{\inf\{f, D\}, \sup\{f, D\}\}$ and integrate. ∎

The characterization Theorem 6.3.2 is extended with no trouble.

Proposition 6.4.5. *A bounded function $f : D \to \mathbb{R}$ is Riemann integrable on a Jordan domain D if and only if the set of its points of discontinuity is of null measure (equivalently, the set of points where the oscillation is bigger than δ is of null content for every $\delta > 0$).*

Note that Jordan sets are stable by finite unions, finite intersections and differences. We say that two Jordan sets A and B do not overlap if $A \cap B \subset \partial A \cup \partial B$. The problem of measuring sets in $\mathbb{R}^d$ is solved in the frame of Jordan sets.

Proposition 6.4.6. *If $A_{i\,i=1}^{n} \subset \mathbb{R}^d$ is a non overlapping finite family of Jordan sets, then its union is Jordan as well and $\mathbf{c}(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n \mathbf{c}(A_i)$.*

We can easily deduce the following uniqueness result.

Corollary 6.4.7. *Let k be a positive, monotone and additive function defined on a class of subsets of $\mathbb{R}^d$ that includes the rectangles and is invariant by translations. Then there exists some $\lambda > 0$ such that $k = \lambda \mathbf{c}$.*

A Jordan partition of a Jordan set D is a non overlapping finite family $\{D_i\}_{i=1}^n$ of Jordan sets such that $D = \bigcup_{i=1}^n D_i$. Jordan partitions provide a good frame for Riemann sums, which provide a more explicit way for the computation of integrals.

Theorem 6.4.8. *Let $f \in \mathfrak{R}(D)$ where D is a Jordan domain. For every $\varepsilon > 0$ there is $\delta > 0$ such that if $\{D_i\}_{i=1}^n$ is a Jordan partition of D into sets of diameter less than δ, then*

$$|\sum_{i=1}^{n} f(t_i)\mathbf{c}(D_i) - \int_D f| < \varepsilon$$

for any choice of points $t_i \in D_i$.

Proof. Without loss of generality we may assume that D is compact. Indeed, take $\overline{D}$ and extend f to $\overline{D} \setminus D$ as zero. Fix $\varepsilon > 0$. Let $M = \operatorname{osc}(f, D)$. The set $x \in D : \operatorname{osc}(f, x) \geq \varepsilon/\mathbf{c}(D)$ is covered by finitely many open rectangles such that its union is an open set O with $\mathbf{c}(O) < \varepsilon/2M$. For any point $x \in D \setminus O$, take $U_x \ni x$ such that $\operatorname{osc}(f, U_x) < \varepsilon/2\mathbf{c}(D)$ and consider the Lebesgue number ξ of the open cover $\{O\} \cup \{U_x\}_{x \in D \setminus O}$. If $\{D_i\}_{i=1}^n$ is a Jordan partition such any set D_i has diameter less than ξ, take N to be the set of such indices i for which $D_i \subset O$. We have

$$|\sum_{i=1}^{n} f(t_i)\mathbf{c}(D_i) - \int_D f| \leq \sum_{i=1}^{n} \int_{D_i} |f(t_i) - f|$$

$$= \sum_{i \in N} \int_{D_i} |f(t_i) - f| + \sum_{i \notin N} \int_{D_i} |f(t_i) - f|$$

$$\leq \sum_{i \in N} M\mathbf{c}(D_i) + \sum_{i \notin N} \frac{\varepsilon}{2\mathbf{c}(D)}\mathbf{c}(D_i) \leq M\mathbf{c}(O) + \frac{\varepsilon}{2\mathbf{c}(D)}\mathbf{c}(D) \leq \varepsilon$$

whenever the points $t_i \in D_i$ are chosen. ∎

In fact, the thesis in the previous statement implies the Riemann integrability suitably reformulated. Indeed, if the Riemann sums

$$\sum_{i=1}^{n} f(t_i)\mathbf{c}(D_i)$$

have a common limit when the Jordan partition $\{D_i\}_{i=1}^n$ is either refined or the maximum diameter of its sets goes to zero, then the function f must be integrable on D.

The convergence of Riemann sum can be applied to prove the change of variables formula in a very important particular case.

Theorem 6.4.9. *Let $E \subset [0,+\infty) \times [0,2\pi]$ a Jordan domain mapped on the Jordan domain $D \subset \mathbb{R}^2$ by the map $(\theta, r) \to (r\cos\theta, r\sin\theta)$. Then for any $f \in \mathfrak{R}(D)$ we have*

$$\iint_D f(x,y)\,dxdy = \iint_E f(r\cos\theta, r\sin\theta)\,r\,drd\theta$$

Proof. Without loss of generality we may assume that E is a rectangle, since the extension of f to be zero on the complement do not change the value of the integrals. Set $\tilde{f}(r,\theta) = f(r\cos\theta, r\sin\theta)$. Take a partition on E with nodes $\{(r_i,\theta_j)\}_{i=1,j=1}^{n,m}$. The rectangles are mapped on sectors $D_{i,j}$ having area

$$\mathbf{c}(D_{i,j}) = \frac{r_{i-1}+r_i}{2}(r_i - r_{i-1})(\theta_j - \theta_{j-1})$$

The associate Riemann sum over D with the evaluation on central points is

$$\sum_{i=1}^{n}\sum_{j=1}^{m} f\Big(\frac{r_{i-1}+r_i}{2}\cos(\frac{\theta_{i-1}+\theta_i}{2}), \frac{r_{i-1}+r_i}{2}\cos(\frac{\theta_{i-1}+\theta_i}{2})\Big)\mathbf{c}(D_{i,j})$$

which approaches $\iint_D f(x,y)\,dxdy$. On the other hand, the sum coincides with

$$\sum_{i=1}^{n}\sum_{j=1}^{m} \tilde{f}\Big(\frac{r_{i-1}+r_i}{2}, \frac{\theta_{j-1}+\theta_j}{2}\Big)\frac{r_{i-1}+r_i}{2}(r_i - r_{i-1})(\theta_j - \theta_{j-1})$$

which is a Riemann sum associate to $\iint_E f(r\cos\theta, r\sin\theta)\,r\,dr\,d\theta$. The refining of the partition in the sense of Theorem 6.4.8 gives the equality of the two integrals of the thesis. ∎

6.5 Iterated integrals

Until this moment we have not said how Riemann integrals in $\mathbb{R}^d$ are computed. The idea is to reduce to iterated integral in spaces of lesser dimension, which in practice means that all can be reduced to one dimensional integrals where the calculus of primitive functions is the main device for its computation.

Next result is known as Fubini theorem for Riemann integral.

Theorem 6.5.1. *Let $R \subset \mathbb{R}^{d_1}$ and $S \subset \mathbb{R}^{d_2}$ rectangles and $f \in \mathfrak{R}(R \times S)$. For $x \in R$ take $f_x(y) = f(x, y)$ defined on S and consider its Darboux integrals*

$$L(x) = \underline{\int_S} f_x, \qquad U(x) = \overline{\int_S} f_x.$$

Then $L, U \in \mathfrak{R}(R)$ and

$$\int_{R\times S} f = \int_R L = \int_R U$$

Moreover, $f_x \in \mathfrak{R}(S)$ for $x \in R$ except a null measure set.

Proof. Consider partitions into rectangles $\{R_i\}_{i=1}^n$ and $\{S_j\}_{j=1}^m$ of R and S respectively. Observe that $\mathbf{m}(R_i \times S_j) = \mathbf{m}(R_i)\mathbf{m}(S_j)$, where each volume is understood according to the dimension of the space. If $x \in R_i$ then

$$\sup\{f_x, S_j\} \leq \sup\{f, R_i \times S_j\}$$

that implies

$$U(x) = \overline{\int_S} f_x \leq \sum_{j=1}^m \sup\{f_x, S_j\}\mathbf{m}(S_j) \leq \sum_{j=1}^m \sup\{f, R_i \times S_j\}\mathbf{m}(S_j)$$

Taking supremum on $x \in R_i$ we get to

$$\sup\{U, R_i\} \leq \sum_{j=1}^m \sup\{f, R_i \times S_j\}\mathbf{m}(S_j)$$

that implies

$$\overline{\int_R} U \leq \sum_{i=1}^n \sup\{U, R_i\}\mathbf{m}(R_i) \leq \sum_{i=1}^n \sum_{j=1}^m \sup\{f, R_i \times S_j\}\mathbf{m}(R_i \times S_j)$$

Taking infimum on the left hand side we get that

$$\overline{\int_R} U \leq \overline{\int_{R\times S}} f = \int_{R\times S} f$$

A similar argument will show that

$$\underline{\int_R} L \geq \underline{\int_{R\times S}} f = \int_{R\times S} f$$

On the other hand, we have these obvious inequalities

$$\underline{\int_R} L \leq \underline{\int_R} U \leq \overline{\int_R} U$$

All together implies that $\underline{\int_R} U = \overline{\int_R} U$, so U is Riemann integrable on R. Similarly, we have $\underline{\int_R} L = \overline{\int_R} L$ and so the Riemann integrability of L, as well as the equality with $\int_{R\times S} f$. Now observe that $\int_R (U - L) = 0$ and the function $U - L$ is positive, so $U = L$ except a null measure set. ■

Not only multiple integrals are reduced to iterated integrals. A "typical exercise" is to transform an impossible iterated integral into a feasible one.

Example 6.5.2.

$$\int_0^1 \left(\int_x^1 e^{y^2} dy \right) dx$$

Firstly, note that

$$D = \{(x, y) : 0 \leq x \leq 1, x \leq y \leq 1\} = \{(x, y) : 0 \leq y \leq 1, 0 \leq x \leq y\}.$$

Therefore,

$$\int_0^1 \left(\int_x^1 e^{y^2} dy \right) dx = \iint_D e^{y^2} dx\, dy = \int_0^1 \left(\int_0^y e^{y^2} dx \right) dy$$

$$= \int_0^1 \left(x e^{y^2} \Big|_{x=0}^{y} \right) dy = \int_0^1 y e^{y^2} dy = \frac{1}{2} e^{y^2} \Big|_{y=0}^{1} = \frac{e-1}{2}.$$

6.6 Improper integrals

Not always the interesting integrals satisfy the Riemann requirements: boundedness of the function or boundedness of the domain. In such a case we have an *improper Riemann integral*. The approach is simple. Assume that $\int_D f$ is improper. If we can take an increasing sequence of bounded (Jordan) domains (D_n) such that $\bigcup_{n=1}^{\infty} D_n = D$ and $f|_{D_n}$ is bounded (and integrable Riemann, of course), we may define

$$\int_D f = \lim_n \int_{D_n} f$$

if the limit exists. That is not totally arbitrary: the way to produce the sequence (D_n) is standard: on $\mathbb{R}$ we take intervals and $\mathbb{R}^2$ rectangles or circles, depending on the geometry of the domain. In the problem is just a singularity of the function, the domains consists in removing a neighbourhood of the singularity, usually an Euclidean ball centred at the singularity.

For positive functions the existence of the limit is guarantied by Lebesgue theory independently of the geometry of the domains. If the function takes positive and negative values around a singularity, the limit of the integrals could depend on the choice of the domains. We say that the convergence happens in *principal value* if there is convergence when the singularity is skipped symmetrically. For instance $\int_{-\infty}^{+\infty} f = \lim_{S\to+\infty} \int_{-S}^{S} f$ or $\int_a^b f = \lim_{\varepsilon\to 0^+} (\int_a^{c-\varepsilon} f + \int_{c+\varepsilon}^b)$ if $c \in [a, b]$ is the singularity of f.

Now, we will combine the iterated integration technique, the change to polar variables and the notion of improper integral to obtain a very important example: the famous Gaussian integral.

Example 6.6.1.

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$$

Once we are convinced about the existence (finiteness) of

$$I = \int_{-\infty}^{+\infty} e^{-x^2} dx = \lim_{S\to+\infty} \int_{-S}^{S} e^{-x^2} dx$$

we will consider the following improper integral on the plane

$$J = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-x^2-y^2} dxdy.$$

Using integration on squares $[-S,S]^2$ we deduce that $J = I^2$. However, the plane integral can be calculated through circles with the same result. Indeed, the function is positive and every circle is contained into a square and viceversa. With the help of polar coordinates we have

$$I^2 = \lim_{R\to+\infty} \int_0^{2\pi} \int_0^R e^{-r^2} r\, dr\, d\theta = \lim_{R\to+\infty} 2\pi \frac{-1}{2} e^{-r^2}\Big|_{r=0}^{R} = \pi$$

and therefore we get that $I = \sqrt{\pi}$.

6.7 Rationale and remarks

The Riemann integral is more constructive, so more pedagogical, than Lebesgue integral (I do not share Dieudonne's views [14, p. 146]). The chapter is develop so we could eventually skip Lebesgue theory, using the Jordan measurability and content, which is enough for a practical use of integration.

With the same spirit, we include an independent proof of the transformation of the integral to polar coordinates if there is no time to fully develop the change of variables theorem. With the same idea is possible to justify the change to spherical coordinates and some other ones simple enough.

Note that the content is denoted **c**, although it coincides with Lebesgue measure denoted **m** later.

6.8 Exercises

1. Prove that the null content sets in $\mathbb{R}^d$ are stable by finte unions and closures.

2. Let $R \subset \mathbb{R}^d$ be a rectangle, $D \subset R$ a null content subset and $f : R \to \mathbb{R}$ a bounded function such that $f(x) = 0$ for every $x \in R \setminus D$. Prove that f is Riemann integrable on R and compute its integral.

3. For $x \in [0,1]$ consider the set

$$N(x) = \{n \in \mathbb{N} : \lfloor 3^n x \rfloor - 1 \text{ is multiple of } 3\}$$

where $\lfloor \cdot \rfloor$ is the integer part of a real number. Consider also the set

$$D = \{x \in [0,1] : N(x) \neq \emptyset\}.$$

Finally, for $x \in D$ take $n(x) = \min(N(x))$.

(a) Prove that $[0,1] \setminus D$ is a null measure set.

(b) Prove that the function $f : [0,1] \to \mathbb{R}$ defined by $f(x) = 3^{-n(x)}$ if $x \in D$ and $f(x) = 0$ otherwise, is Riemann integrable.

(c) Calculate $\int_0^1 f(x)\,dx$.

4. A *cylindrical cradle* is a body limited by the horizontal plane, a vertical cylinder and a tilted plane that meets the horizontal plane at a diameter of the cylinder base. Find the volume of cylindrical cradle of radius r and height h.

5. Let $D = [0,1]^2$. Find the values $\alpha > 0$ for which is finite the improper integral

$$\iint_D \frac{dx\,dy}{|x-y|^\alpha}$$

and calculate its value .

6. Compute

$$\int_0^{\sqrt{\pi}} \left(\int_y^{\sqrt{\pi}} \sin(x^2)\,dx \right) dy.$$

7. Prove the convergence and compute for $D = \{(x,y) : x, y \geq 0, x^2 + y^2 \leq 1\}$ the integral

$$\iint_D \frac{dxdy}{x+y}.$$

8. Let $f : R \to \mathbb{R}$ a bounded function defined on a rectangle $R \subset \mathbb{R}^d$. Prove that $f$is Riemann integrable on R if an only if its graph

$$\text{graf}(f) = \{(x, f(x)) : x \in R)\}$$

has null content. Prove, as a consequence, that the compact smooth manifolds in $\mathbb{R}^d$ of dimension $d-1$ have null content.

9. A function $f : [a, b] \to \mathbb{R}$ is said to be *step* if there exists a partition of $[a, b]$ such that f is constant on the interior of each of the intervals defined by the partition. A function is said *ruled* if it is a uniform limit of step functions. Prove the following statements:

   (a) Ruled functions has countably many discontinuities, at most.

   (b) Ruled functions are Riemann integrable on their domains.

   (c) A function is ruled if and only if at each $c \in [a, b)$ exists $\lim_{x\to c^+} f(x)$ and at each $c \in (a, b]$ exists $\lim_{x\to c^-} f(x)$.

10. Let $f : [0, a] \to [0, b]$ be a decreasing continuous bijection. Prove with the help of a plane integral that

$$\int_0^a f(x)\,dx = \int_0^b f^{-1}(x)\,dx.$$

11. Generalize the Riemann-Lebesgue theorem for functions defined on Jordan domains.

12. Let $D_i \subset \mathbb{R}^d$ $i = 1, \ldots, n$ be pairwise disjoint Jordan measurable sets and let $f_i : D_i \to \mathbb{R}$ be Riemann integrable functions. Prove the Riemann integrability of the "glued" function $f : D \to \mathbb{R}$ where $D = \bigcup_{i=1}^n D_i$ and $f(x) = f_i(x)$ if $x \in D_i$. Show also that

$$\int_D f = \sum_{i=1}^n \int_{D_i} f_i.$$

Chapter 7

Change of Variables in Integration

7.1 Linear volume transformations

The objetive of this section is to prove that for any linear map $T : \mathbb{R}^n \to \mathbb{R}^n$ one has that

$$\mathbf{m}(T(A)) = |\det(T)|\, \mathbf{m}(A)$$

where $\mathbf{m}$ means volume in the sense of the Jordan content or the Lebesgue measure (Chapter 9). The statement implicitly includes the measurability of $T(A)$. Firstly note the measurability is clear if $\det(T) = 0$ because in such a case the image is include into a subspace of dimension $n-1$ so it has measure 0 or content 0 if it is moreover bounded. If $\det(T) \neq 0$, then T is an homeomorphism, so it preserves topology and Borel measurability. We will discuss Lebesgue or Jordan measurability. Lebesgue measurability is characterized by the fact that a measurable set differs from a Borel set in a set of null measure. Jordan measurable sets are characterized for having Lebesgue null boundary. Thus in both cases it is enough to study how T transforms null sets. The main tool for this purpose is an estimation of the measure of images of Lipschitz maps.

Proposition 7.1.1. *Assume that $\mathbb{R}^n$ is endowed with the supremum norm. Let $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$ be a Lipschitz map with constant λ. Then*

$$\mathbf{m}^*(f(A)) \leq \lambda^n\, \mathbf{m}^*(A)$$

where m^* denotes the outer Lebesgue measure.

Proof. The result is consequence of three easy observations. Firstly, the Lebesgue outer measure can be approximated by coverings of balls (actually cubes) if the norm we are using is $\|\cdot\|_\infty$ (actually, thanks to a result of Vitali we may use any norm for the same purpose). Indeed, the outer measure is defined by coverings of generalized rectangles and those rectangles can be arbitrarily approached by non-overlapping unions of cubes. The second observation is that for any ball we have

$$f(B[x,r]) \subset B[f(x),\lambda].$$

Finally we have $\mathbf{m}(B[x,\lambda r]) = \lambda^n \mathbf{m}(B[x,r])$. ■

As the notion of null measure does not depend on the norm we have.

Corollary 7.1.2. *A locally Lipschitz map $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$ carries Lebesgue null sets to Lebesgue nul sets.*

Proof. The domain D can be decomposed in countably many domains where the restriction of f is Lipschitz, so the previous theorem is applicable. ■

Now that the measurability of $T(A)$ is not a problem, remember that the Lebesgue measure is the unique non trivial translation invariant Borel measure on $\mathbb{R}^n$, but a multiplicative positive constant. Therefore, for every T linear there is $k(T) \geq 0$ such that

$$\mathbf{m}(T(A)) = k(T)\,\mathbf{m}(A).$$

Obviously $k(T)$ is the volume of the image of the unitary cube through T. Note also that the constant is multiplicative

$$k(TS) = k(T)\,k(S).$$

Now we will prove the following result.

Theorem 7.1.3. *Let k be a nonnegative nontrivial multiplicative function defined on the square matrices of size n. Assume that k is continuous at the identity matrix I. Then there is $\lambda \in \mathbb{R}$ such that*

$$k(T) = |\det(T)|^\lambda.$$

Proof. Clearly $k(I) = k(II) = k(I)^2$. Since k is not trivial we have $k(I) = 1$. If T is invertible we have $k(T)k(T^{-1}) = k(I) = 1$. Therefore $k(T) \neq 0$ and $k(T^{-1}) = k(T)^{-1}$. We deduce that k takes the same value for similar matrices

$$k(S^{-1}TS) = k(S^{-1})\,k(T)\,k(S) = k(T).$$

Consider now the diagonal matrices D_x having 1's on the diagonal except the first entry which takes the value $x \in \mathbb{R}$. Observe that

$$k(D_x)^2 = k(D_xD_x) = k(D_{-x}D_{-x}) = k(D_{-x})^2$$

and so $k(D_{-x}) = k(D_x)$. If we define $f(x) = k(D_x)$ then $f(-x) = f(x)$ and $f(xy) = f(x)f(y)$. Since the assignment $x \to D_x$ is linear (among other properties) and it is continuous at $x = 1$. The function $g(t) = \log(f(e^t))$ defined on $\mathbb{R}$ satisfies the equation $g(t+s) = g(t)+g(s)$ and it is continuous at 0. It is well know that there is $\lambda \in \mathbb{R}$ such that $g(t) = \lambda t$ and thus $f(x) = |x|^\lambda$. Now, a diagonal matrix can be written as a product of matrices which are similar to D_x's matrices (a permutation of the basis is a similarity operation) being the x's the eigenvalues. We deduce that $k(D) = |\det(D)|^\lambda$ for a diagonal matrix. The result now extends to symmetric matrices which are known to be similar to diagonal matrices. Finally, an arbitrary matrix T is similar to its transpose implying that $k(T) = k(T^t)$. As we have that $T\,T^t$ is symmetric we deduce

$$k(T) = \sqrt{k(T\,T^t)} = |\det(T\,T^t))|^{\lambda/2} = |\det(T)|^\lambda$$

which concludes the proof. ∎

Now we can achieve the objetive stated at the beginning of the section.

Theorem 7.1.4. *Given a linear map $T : \mathbb{R}^n \to \mathbb{R}^n$ we have*

$$\mathbf{m}(T(A)) = |det(T)|\,\mathbf{m}(A)$$

for any measurable, either in sense of Jordan or Lebesgue, set $A \subset \mathbb{R}^n$.

Proof. Assume $\mathbb{R}^n$ is equipped with the supremum norm, and take $C = B[0, 1/2]$. The set C has n-dimensional volume 1. If we denote by $k(T) = \mathbf{m}(C)$, we already know that $\mathbf{m}(T(A)) = k(T)\,\mathbf{m}(A)$ for any measurable set and $k(TS) = k(T)\,k(S)$. Therefore, in order to prove the result we only have to reduce it to the previous theorem checking that k is continuous at I and the constant λ must be equal to 1. Let $0 < \varepsilon < 1$. Since the operation of taking

inverse is continuous, there is $0 < \delta < \varepsilon/2$ such that $\|T - I\| < \delta$ implies $\|T^{-1} - I\| < \varepsilon/2$. We have $\|T\|, \|T^{-1}\| \leq 2$ and

$$T(C) \subset C + B[0, \varepsilon/2] = (1 + \varepsilon)C;$$

$$T^{-1}(C) \subset C + B[0, \varepsilon/2] = (1 + \varepsilon)C.$$

Applying T to the last we get $C \subset (1 + \varepsilon)T(C)$, following that

$$(1 + \varepsilon)^{-1}C \subset T(C) \subset (1 + \varepsilon)C$$

which implies the continuity of k at I because $(1 + \varepsilon)^{-d} \leq k(T) \leq (1 + \varepsilon)^n$. Now we have that $k(T) = |\det(T)|^\lambda$ for some λ. If we set $T = 2I$ we have

$$2^n = \mathbf{m}(T(C)) = k(T) = |\det T|^\lambda = 2^{\lambda d}$$

following that $\lambda = 1$ as wanted. ■

The proof of the formula for the transformation of volumes through linear maps can be obtained also by geometrical considerations which are specially clear for $\mathbb{R}^2$: showing that a parallelogram is equivalent to a rectangle by decomposing it into 2 pieces.

7.2 The change of variables theorem

Now we will turn our attention to nonlinear transformations. Let start with this fact which is just Lemma 5.1.2 stated in $\mathbb{R}^n$.

Lemma 7.2.1. *Let $T : D \subset \mathbb{R}^n \to \mathbb{R}^n$ be a C^1 map and let $x_0 \in D$ such that $dT(x_0)$ is nonsingular. Then for every $0 < \eta < 1$ there exists $\delta > 0$ such that $T|_{B[x_0,r]}$ is one-to-one, T^{-1} is differentiable at $f(x_0)$ and*

$$T(x_0) + dT(x_0)(B[0, (1 - \eta)\delta]) \subset T(B[x_0, \delta]) \subset T(x_0) + dT(x_0)(B[0, (1 + \eta)\delta]).$$

In particular, the image through T of a neighbourhood of x_0 is a neighbourhood of $T(x_0)$. Moreover, $f(U)$ is open whenever $U \subset D$ is open and $dT(x)$ is nonsingular at every point $x \in U$.

After the proof of the lemma a remark was done: the number $\delta > 0$ only depends on the modulus of continuity of $dT(x)$ and an upper bound for $\|dT^{-1}\|$

Theorem 7.2.2. *Let $R \subset \mathbb{R}^n$ be a compact rectangle, let $T : R \to \mathbb{R}^n$ be an one-to-one C^1 map with dT non singular on R. Then $T(R)$ is Jordan measurable and for every Riemann integrable function $f : T(R) \to \mathbb{R}$ then $f \circ T$ is Riemann integrable on R and*

$$\int_{T(R)} f = \int_R f \circ T \, |\det(dT)|$$

where the determinant is computed for the matrix of dT with respect to the canonical bases.

Proof. First of all, we may assume that R has nonempty interior. Otherwise R would be measure 0 and so its image $T(R)$ being the result true trivially. We may assume that $f \geq 0$ as well. By Theorem 5.1.4 we know that the interior of R is transformed into an open set by T, therefore the boundary of $T(R)$ is contained in $T(\partial R)$ which has null measure. That implies $T(R)$ is Jordan measurable.
Observe that $\|(dT)^{-1}\|$ is bounded on $T(R)$ which implies that T^{-1} is Lipschitz. If D is the null measure set of discontinuities of f then $T^{-1}(D)$ is also null. Since the set of discontinuities of $f \circ T$ is exactly $T^{-1}(D)$ we get that f is Riemann integrable.
We may set the norm of $\mathbb{R}^n$ to have the unit ball a translation of R. Take $0 < \eta < 1$ and note that now R can be decomposed into N^n non overlapping balls of radius $1/N$. By the continuity of dT on a larger open containing R we may take N large enough to guarantee that the set containment of the Lemma can be applied with such η to all the balls of radius $1/N$. Let x_k with $1 \leq k \leq 2^N$ the centres of the balls covering R and $B_k = B[x_k, \frac{1}{N}]$. We have now

$$T(x_k) + dT(x_k)\Big(B\big[0, \frac{1-\eta}{N}\big]\Big) \subset T(B_k) \subset T(x_k) + dT(x_k)\Big(B\big[0, \frac{1+\eta}{N}\big]\Big).$$

Having in mind that $\mathbf{m}(L(S)) = |\det(L)|\mathbf{m}(S)$ for any linear map L and any compact rectangle S, we get

$$(1-\eta)^n \mathbf{m}(B_k)|\det(dT(x_k))| \leq \mathbf{m}(T(B_k)) \leq (1+\eta)^n \mathbf{m}(B_k)|\det(dT(x_k))|.$$

Multiplying by $f(T(x_k))$ and adding we get

$$(1-\eta)^n \sum_{k=1}^{2^N} f(T(x_k))\, \mathbf{m}(B_k)\, |\det(dT(x_k))|$$

$$\leq \sum_{k=1}^{2^N} f(T(x_k))\,\mathbf{m}(B_k) \leq (1+\eta)^n \sum_{k=1}^{2^N} f(T(x_k))\,\mathbf{m}(B_k)\,|\det(dT(x_k))|$$

The sums are of Riemann type, standard ones at the ends and associated to a Jordan partition of $T(R)$ in the middle. so letting n going to infinity we will get

$$(1-\eta)^n \int_R f\circ T\,|\det(dT)| \leq \int_{T(R)} f \leq (1+\eta)^n \int_R f\circ T\,|\det(dT)|$$

As η can be taken arbitrarily close to 1 we get the desired result. ∎

The result can be extended to general Jordan domains.

Theorem 7.2.3. *Let $D \subset \mathbb{R}^n$ an open Jordan domain, let $T : \overline{D} \to \mathbb{R}^n$ be a C^1 map such that T is one-to-one and dT is non singular on D. Then $T(D)$ is Jordan measurable and for every Riemann integrable function $f : T(D) \to \mathbb{R}$ then $f \circ T$ is Riemann integrable on D and*

$$\int_{T(D)} f = \int_D f\circ T\,|\det(dT)|.$$

Proof. The arguments employed above for the Jordan measurability of $T(D)$ and the Riemann integrability of $f \circ T$ can be adapted here with some small changes. As before $T(D)$ is open and the boundary of $T(D)$ is included into $T(\partial D)$. However, T^{-1} is locally Lipschitz which implies that the set of discontinuities of $f \circ T$ is null.
To prove the formula, cover ∂D with a finite union of compact rectangles whose volumes sums less than ε. Then $D \setminus S$ can be decomposed into a finite union of non-overlapping rectangles. The previous theorem applied on each rectangle and having in mind that the images by T of the rectangles are non-overlapping give us

$$\int_{T(D\setminus S)} f = \int_{D\setminus S} f\circ T\,|\det(dT)|.$$

If M is an upper bound to f we have

$$\left|\int_{T(D)} f - \int_{T(D\setminus S)} f\right| \leq \int_{T(D\cap S)} |f| \leq M\mathbf{m}(S)$$

and

$$\left|\int_D T\circ f - \int_{D\setminus S} T\circ f\right| \leq \int_{D\cap S} |T\circ f| \leq M\lambda^n\,\mathbf{m}(S)$$

can be done arbitrarily small which leads to the desired equality. ∎

Corollary 7.2.4. *If D is an open Jordan domain and $T : \overline{D} \to \mathbb{R}^n$ be a C^1 map such that T is one-to-one and dT is non singular on D then $T(D)$ is a Jordan domain*

$$\mathbf{m}(T(D)) = \int_D |\det(dT)|.$$

7.3 The Morse-Sard theorem

So far we have been asking the map T to have non singular differential dT on the interior of the domain. We will see that the singular points of dT are actually negligible when it comes to integration. That is the spirit of the following result known as the Morse-Sard theorem, that we will prove only the version for mapping between spaces of the same dimension.

Theorem 7.3.1. *Let $T : D \subset \mathbb{R}^n \to \mathbb{R}^n$ be a C^1 map and consider the set of singular points*

$$S = \{x \in D : dT(x) \text{ is singular}\}.$$

Then $T(S)$ has null measure.

Proof. First of all note that S is closed. The proof will be by induction on the dimension n.
Suppose $d = 1$. It is enough to show that $S \cap (a, b)$ has null measure. Note that in this case $dT(x)$ is singular if and only if $T'(x) = 0$. Given $\varepsilon > 0$, consider the set

$$U = \{x \in (a, b) : |T'(x)| < \varepsilon\}$$

Then $S \cap (a, b) \subset U$ and T is ε-Lipschitz on every interval composing U thanks to the mean value theorem. Now apply Proposition 7.1.1 to obtain that

$$\mathbf{m}^*(T(S \cap (a, b))) \leq \varepsilon\, \mathbf{m}^*(S \cap (a, b)) \leq \varepsilon(b - a).$$

That implies $\mathbf{m}^*(T(S \cap (a, b))) = 0$ as ε is arbitrary.
Assume now that the statement is proven for $n - 1$. We will write

$$T(x) = (f_1(x_1, x_2, \ldots, x_n), f_2(x_1, x_2, \ldots, x_n), \ldots, f_n(x_1, x_2, \ldots, x_n)).$$

Consider the set

$$Z = \{x \in D : dT(x) = 0\} = \{x \in D : \frac{\partial f_i}{\partial x_j}(x) = 0, 1 \leq i, j \leq d\}.$$

Obviously $Z \subset S$. If B is an arbitrary closed ball and $\varepsilon > 0$, then its is possible to cover $Z \cap B$ with finitely many non-overlapping convex sets such that T is ε-Lipschitz on each of them thanks to the mean value theorem (in several variables). Reasoning as in the 1-dimensional case that gives $\mathbf{m}^*(T(Z \cap B)) \leq \varepsilon \mathbf{m}^*(B)$ which implies $\mathbf{m}^*(T(Z)) = 0$ on account of B and ε.
The objetive now is to show that $T(S \setminus Z)$ has null measure. Note that it is enough to show that every $x_0 \in S \setminus Z$ has a neighbourhood U such that $T((S \setminus Z) \cap U)$ is null. As $x_0 \in S \setminus Z$ there are i, j such that $\frac{\partial f_i}{\partial x_j}(x) \neq 0$. Reordering the variables and the coordinate functions we may assume that $\frac{\partial f_1}{\partial x_1}(x) \neq 0$. Consider the map $G(x) = (f_1(x), x_2, \ldots, x_n)$ and note that $dG(x_0)$ is not singular. By the inverse mapping theorem there is a neighbourhood U of x_0 and V of $G(x_0)$ such that G is a bijection form U onto V. The composition $H = T \circ G^{-1}$ defined on V is of the form

$$H(y) = (y_1, h_2(y), \ldots, h_n(y))$$

and $dH(y)$ is singular if and only if $y \in A = G((S \setminus Z) \cap U)$ because they come from the singular points of dT. On the other hand, $dH(y)$ is singular if and only if the $n-1$ dimensional Jacobian

$$\begin{pmatrix} \frac{\partial h_2}{\partial y_2} & \cdots & \frac{\partial h_2}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_n}{\partial y_2} & \cdots & \frac{\partial h_n}{\partial y_n} \end{pmatrix}$$

is singular at y, thanks to particular form of H. Evidently $T((S \setminus Z) \cap U) = H(A)$ so it is enough to prove that $H(A)$ is null. For $t \in \mathbb{R}$ we will consider the affine hyper-plane

$$P_t = \{y = (y_1, \ldots, y_n) \in \mathbb{R} : y_1 = t\}.$$

and the map $H_t(y_2, \ldots, y_n) = H(t, y_2, \ldots, y_n)$ defined on $V \cap P_t$ whose Jacobian is just above. The set of singular points of dH_t is exactly $A \cap P_t$. By the induction hypothesis we get that $H_t(A \cap P_t)$ is null and therefore

$$\{t\} \times H_t(A \cap P_t) = P_t \cap H(A)$$

is a $(n-1)$-dimensional null set, and this is true for every $t \in \mathbb{R}$. A well known consequence of Fubini's theorem says that $H(A)$ is a n-dimensional null set. ■

7.4 Brouwer fixed point theorem

A spectacular application of the change of variables formula is a simple proof of the topological theorem about fixed points due to Brouwer.

Theorem 7.4.1. *A continuous map from $B_{\mathbb{R}^n}$ into itself has a fixed point.*

Along the section $n \in \mathbb{N}$ is fixed and we will write $B = B_{\mathbb{R}^n}$ and $S = \partial B$. By standard techniques it is easy to prove the equivalence of the *fixed point property* (FPP) for B with the nonexistence of a retraction of B onto S, that is, a continuous map from B onto S that fixes the points of S. That can be done not only in the category of continuous maps but also C^1, which will important for the proof.

Lemma 7.4.2. *The FPP of B for C^1 maps implies the FPP of B for continuous maps.*

Proof of the Theorem 7.4.1. Let P be a C^1 retraction of B onto S. We will arrive to a contradiction after a witty construction. For every $t \in [0,1]$ take

$$P_t(x) = (1-t)x + tP(x)$$

and note that P_t is a C^1 map from B onto itself that fixes S. We claim that for t small enough P_t is an homeomorphism onto its image. Indeed, let L be the Lipschitz constant of P. Then

$$\begin{gathered}\|P_t(x) - P_t(y)\| = \|(1-t)(x-y) + t(P(x) - P(y))\| \\ \geq (1-t)\|x-y\| - t\|P(x) - P(y)\| \geq (1-t)\|x-y\| - Lt\|x-y\| \\ \geq (1-(L+1)t)\|x-y\|.\end{gathered}$$

Therefore, taking $t < (L+1)^{-1}$ the inverse of P_t is defined and Lipschitz. Moreover, for t small enough the map and its inverse are open. Indeed, that is consequence on the Inverse Map Theorem since $\det(dP_t)$ is nearly 1 for t close to 0. That implies P_t carries one-to-one S onto the $\partial P_t(B)$. As P_t fixes S we deduce that $P_t(B) = B$ for t small enough.
On the other hand, note that $\det(dP_t)$ is a polynomial in t, so it is the function

$$h(t) = \int_B \det(dP_t(x))\, dx$$

defined for $t \in [0,1]$. As for t small enough P_t is a diffeomorphism, the change of variables formula Theorem 7.2.3 says that $h(t) = \mathbf{m}(B) > 0$. As h is a polynomial, being constant in an interval implies to be constant everywhere. However, $h(1) = 0$ because $P_1 = P$ collapses on S. That is a contradiction. ■

Corollary 7.4.3. *Any compact set that is homeomorphic to, or a retract of, an Euclidean ball has the FPP.*

7.5 Assorted changes of variables

In practise we do not need introduce much new letters to do a change of variables: if we put

$$T(u_1, \ldots, u_n) = (x_1(u_1, \ldots, u_n), \ldots, x_n(u_1, \ldots, u_n))$$

then the Jacobian of T is usually denoted

$$\frac{\partial(x_1, \ldots, x_n)}{\partial(u_1, \ldots, u_n)} = \begin{pmatrix} \frac{\partial x_1}{\partial u_1} & \cdots & \frac{\partial x_1}{\partial u_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_2} & \cdots & \frac{\partial x_n}{\partial u_n} \end{pmatrix}$$

so the change of variables in the integral becomes

$$\int \cdots \int_{T(D)} f(x_1, \ldots, x_n)\, dx_1 \ldots dx_n$$

$$= \int \cdots \int_{D} (f \circ T)(u_1, \ldots, u_n) \left| \frac{\partial(x_1, \ldots, x_n)}{\partial(u_1, \ldots, u_n)} \right| du_1 \ldots du_n$$

that it is easy to remember.

7.5.1 Sum of the inverse of the squared integers

Consider the integral

$$\int_0^1 \int_0^1 \frac{dxdy}{1 - xy}$$

Since $(1 - xy)^{-1} = \sum_{n=0}^{\infty} x^n y^n$, the integration term by term gives that

$$\int_0^1 \int_0^1 \frac{dxdy}{1 - xy} = \sum_{n=0}^{\infty} \int_0^1 \int_0^1 x^n y^n = \sum_{n=1}^{\infty} \frac{1}{n^2}$$

where the equality can be justified by the monotone convergence theorem (application of Riemann theory needs a more detailed analysis).

Let T be the triangle with vertices $(0,0)$, (1.0) and $(1,1)$. By symmetry we have

$$\int_0^1 \int_0^1 \frac{dxdy}{1-xy} = 2 \iint_T \frac{dxdy}{1-xy}$$

Consider now the change of variables given by $x = v + u$, $y = v - u$ where (u, v) runs over the triangle D with vertices $(0,0)$, $(1/2, 1/2)$ and $(0,1)$. As the jacobian is 2 we have

$$\iint_T \frac{dxdy}{1-xy} = \iint_D \frac{dudv}{1-v^2+u^2}$$

In order to compute the last integral we do the decomposition

$$\iint_D \frac{dudv}{1-v^2+u^2} = \int_0^{1/2} \int_0^v \frac{dudv}{1-v^2+u^2} + \int_{1/2}^1 \int_0^{1-v} \frac{dudv}{1-v^2+u^2}$$

The first integral is calculated as follows

$$\int_0^{1/2} \frac{1}{\sqrt{1-v^2}} \arctan\left(\frac{v}{\sqrt{1-v^2}}\right) dv = \int_0^{1/2} \arcsin(v) \frac{dv}{\sqrt{1-v^2}}$$

$$= \frac{1}{2} \arcsin(v)^2 |_0^{1/2} = \frac{\pi^2}{72}$$

As to the second integral, we have after a first integration that

$$\int_{1/2}^1 \frac{1}{\sqrt{1-v^2}} \arctan\left(\frac{1-v}{\sqrt{1-v^2}}\right)$$

Putting $v = \cos t$, then $1 - v = 2\sin^2(t/2)$ and

$$\sqrt{1-v^2} = \sin t = 2\sin(t/2)\cos(t/2).$$

Therefore

$$\frac{1-v}{\sqrt{1-v^2}} = \frac{\sin(t/2)}{\cos(t/2)} = \tan(t/2)$$

$$\arctan\left(\frac{1-v}{\sqrt{1-v^2}}\right) = \frac{t}{2} = \frac{1}{2}\arccos(v)$$

and with this

$$\int_{1/2}^1 \frac{1}{\sqrt{1-v^2}} \arctan\left(\frac{1-v}{\sqrt{1-v^2}}\right) dv = \frac{1}{2}\frac{-1}{2} \arccos(v)^2 |_{1/2}^1 = \frac{\pi^2}{36}$$

Finally,

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 4\left(\frac{\pi^2}{72} + \frac{\pi^2}{36}\right) = \frac{\pi^2}{6}$$

as desired. ■

7.5.2 Integrals of Euler

They are defined as parametric one variable integrals, however the relation between them is explained with the help of a two variable integral. Consider for $p, q > 0$ the functions

$$\Gamma(p) = \int_0^{+\infty} t^{p-1} e^{-t}\, dt$$

$$B(p,q) = \int_0^1 t^{p-1}(1-t)^{q-1} dt$$

Observe that

$$\Gamma(p)\Gamma(q) = \left(\int_0^{+\infty} t^{p-1} e^{-t}\, dt\right)\left(\int_0^{+\infty} s^{q-1} e^{-s}\, ds\right)$$

$$= \iint_Q t^{p-1} s^{q-1} e^{-t-s} dt ds = \int_0^{+\infty} \left(\int_0^1 (rw)^{p-1}((1-w)r)^{q-1} e^{-r} r\, dw\right) dr$$

$$= \left(\int_0^{+\infty} r^{p+q-1} e^{-r} dr\right)\left(\int_0^1 w^{p-1}(1-w^{q-1}) dw\right) = \Gamma(p+q)B(p,q)$$

where Q stands for the first quadrant and the change of variables is

$$t = rw, s = r(1-w),$$

whose Jacobian (absolute value) is r. Therefore

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}.$$

7.5.3 Integrals of Dirichlet

The transformation of the pyramid

$$\{(x_1,\dots,x_n) : x_1, x_1 \geq 0, \dots, x_n \geq 0, x_1 + \cdots + x_n \leq 1\}$$

into a cube $[0,1]^n$ can be perform with this change of variables

$$\begin{aligned} x_1 + x_2 + \cdots + x_n &= u_1 \\ x_2 + \cdots + x_n &= u_1 u_2 \\ \dots \quad \vdots \quad &\dots \\ x_n &= u_1 u_2 \dots u_n \end{aligned}$$

that leads to $x_1 = u_1(1-u_2)$, $x_2 = u_1 u_2 (1-u_3)$ and so on. The transformation is a diffeomorphism for the interior of the domains, however the boundary collapses (that will not be a problem). The Jacobian can be computed with the usual tricks to reduce the complexity of the determinant

$$\frac{\partial(x_1,\dots,x_n)}{\partial(u_1,\dots,u_n)} = u_1^{n-1} u_2^{n-2} \dots u_{n-1}$$

With the help of this transformation one can generalise the formula of Euler above, for instance. Let $p_0, p_1, \dots, p_n > 0$ and consider the pyramid

$$D = \{(x_1,\dots,x_n) : x_1, x_1 \geq 0, \dots, x_n \geq 0, x_1 + \cdots + x_n \leq 1\}.$$

Then

$$\int \cdots \int_D x_1^{p_1-1} x_2^{p_2-1} \dots x_n^{p_n-1} (1 - x_1 - \cdots - x_n)^{p_0-1} dx_1 dx_2 \dots dx_n$$

$$= \frac{\Gamma(p_0)\Gamma(p_1)\dots\Gamma(p_n)}{\Gamma(p_0 + p_1 + \cdots + p_n)}.$$

7.6 Rationale and remarks

The changes of variables theorem is stated only for Riemann integral, nevertheless it can be obviously adapted to Lebesgue integral. We will not do that explicitly, but is clear that in Lebesgue theory some troubles of Riemann integral disappear.

The theorem of Brouwer is usually proved in Topology books with discretisation and combinatorics. The proof using the change of variables is due to Milnor and Rogers.

The change of variables of Dirichlet deserves a discussion in the classroom because is not evident how a triangle or a pyramid can be transformed into a square or a cube, respectively.

7.7 Exercises

1. Find the volume of the body limited by the sphere $x^2+y^2+z^2=1$ and the cylinder $x^2+y^2=2x$.
2. Let $D=\{(x,y): x^2+y^2\leq 1\}$, and calculate
$$\iint_D \sqrt{1+x^2+y^2}dxdy.$$
3. Let $B=\{(x,y,z): x^2+y^2+z^2\leq 1\}$, and calculate
$$\iiint_B \frac{dxdydz}{x^2+y^2+(z-2)^2}.$$
4. Let $E=\{(x,y,z): 4x^2+9y^2+36z^2\leq 36\}$, and calculate
$$\iiint_E (2x+3y+6z)^2 dxdydz.$$
5. Prove the convergence and find the value of the integral
$$\iint_{\mathbb{R}^2} \frac{e^{-x^2-y^2}}{\sqrt{x^2+y^2}}dxdy.$$
6. Prove that $\Gamma(n)=(n-1)!$ for $n\in\mathbb{N}$.
7. Prove that $\Gamma(1/2)=\sqrt{\pi}$.
8. Prove that, for $p>0$, the n-dimensional volume of the set
$$B_p^n=\{(x_1,\ldots,x_n): |x_1|^p+\cdots+|x_n|^p\leq 1\}.$$
is
$$\mathrm{vol}(B_p^n)=\frac{2^n\,\Gamma(\frac{1}{p})^n}{p^n\,\Gamma(\frac{n}{p}+1)}.$$

Chapter 8

Measure Theory and Lebesgue Integral

8.1 Motivation

A moment's reflection on how the notion of area for polygons is treated in elementary texts shows that the existence of the area and its additive property are mostly assumed a priori, so the actual task is to compute the areas of progressively more complicate polygons. In a second stage the area can be extended to some nonpolygons, as the circle, assuming monotonicity. Well, it is possible to provided a sound basis to the elemental method: define the area for triangles, show that it is independent of the position of the triangle, prove that it is additive within the triangles, and finally extend it to polygons by decompositions into triangles... However, knowing in advance the scope of this method, we could opt for an easier approach which will lead to the same results.

An elemental measure theory on $\mathbb{R}^n$ can be developed as follows. We will consider in the first step rectangles $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$ whose measure is the number $\mathbf{m}(R) = (b_1 - a_1) \dots (b_n - a_n)$. We may also consider products of open intervals with the same measure (the border of the rectangles is negligible) or infinite intervals with the convention $+\infty \cdot 0 = 0 \cdot (+\infty) = 0$. When a rectangle can be decomposed into a finite number of disjpoint rectangles, or merely non overlapping, namely $R = \bigcup_{k=1}^{n} R_k$ then $\mathbf{m}(R) = \sum_{k=1}^{n} \mathbf{m}(R_k)$. Indeed, any decomposition can be refined to a grid decomposition, for whom the additive property is just a consequence of the distributivity of the sum with respect to the product.

We will say that a set is elemental if it is union of finitely many rectangles. The measure can be extended to elemental sets and the measure is extended additively using non overlapping decompositions into rectangles. The reason why an elemental set can be reduced to a finite union of non overlapping rectangles lies in the fact the difference of two rectangles is a finite union of rectangles. Checking that the definition of m does not depend on how the decomposition is chosen and the additivity of m with respect to finite disjoint (or not overlapping) unions offers no challenge.

The last step of the construction is the extension of the measure to a more general class of sets. Since the measure must be *monotone*, given an arbitrary set $A \subset \mathbb{R}^n$ its the measure $\mathbf{m}(A)$, in case it is possible to define it, must satisfy $\mathbf{m}(E_1) \leq \mathbf{m}(A) \leq \mathbf{m}(E_2)$ whenever $E_1, E_2 \subset \mathbb{R}^n$ are elemental and $E_1 \subset A \subset E_2$. Let us say that a set $A \subset \mathbb{R}^n$ is measurable (in the sense of Jordan) if for every $\varepsilon > 0$ there are elemental sets $E_1, E_2 \subset \mathbb{R}^n$ with $E_1 \subset A \subset E_2$ and $\mathbf{m}(E_2 \setminus E_1) < \varepsilon$. In that case, we can assign a measure to A by

$$\mathbf{m}(A) = \sup\{\mathbf{m}(E) : E \subset A \text{ elemental}\} = \inf\{\mathbf{m}(E) : E \supset A \text{ elemental}\}.$$

The numbers appearing are called *inner content* and *outer content* (in Jordan's sense) of the set A respectively, so *measurability* appears as the agreement of inner and outer measures. Finite unions and finite intersections of measurable sets are measurable and the measure $\mathbf{m}$ is finitely additive for finite disjoint unions of measurable sets.

It is not difficult to prove that a bounded set A is measurable if and only if its border ∂A has measure 0, which in this context means that it can be covered by finitely many rectangles such that the sum of their measures can be made arbitrarily small. That implies easily that all the simple geometrical objects (polygons, circles, etc.) are measurable and their standard measurement is back up by a rigorous construction.

The method just sketched above, namely Jordan theory of measure (see Chapter 7), is somehow related to Riemann integral. It serves well at elementary level but it has many limitations. For instance, sets as simple as the rational numbers between 0 and 1 are not measurable. Moreover, the approximation of the circle area from within using polygons can be understood as a limit process which implies a decomposition of the circle into countably many

rectangles. The measure $\mathbf{m}$ should be countably additive as the geometric interpretation deserves, however it cannot. Otherwise, the set of rationals would have measure 0, since it is a countable union of points.

It would be desirable to define a countably additive measure for (some) sets of $\mathbb{R}^n$ that generalizes the Jordan construction. This is actually possible and the measure is called the *Lebesgue measure*. The next sections of the chapter will be devoted to measures and their construction and Lebesgue measure and its properties will appear as a by product of more general results.

8.2 Measures

We need a family of sets where we can perform all the required operations with a countably additive function. Motived by the previous section we will introduce algebras and σ-algebras. An algebra of subsets of a set Ω is family $\mathcal{A} \subset \mathcal{P}(\Omega)$ which satisfies

1. $\emptyset, \Omega \in \mathcal{A}$;
2. $A \in \mathcal{A}$ implies $A^c \in \mathcal{A}$;
3. $\bigcup_{k=1}^{n} A_k \in \mathcal{A}$ whenever $A_1, \ldots, A_n \in \mathcal{A}$.

We say that a family $\Sigma \subset \mathcal{P}(\Omega)$ is a σ-algebra if it is an algebra and satisfies moreover

3'. $\bigcup_{n=1}^{\infty} A_k \in \Sigma$ whenever $(A_n)_{n=1}^{\infty} \subset \Sigma$.

As we will see algebras and measures on them appears naturally, however the theory works nicer with σ-additivity on σ-algebras. On the other hand to build nontrivial σ-additive measures is a delicate task that we will face in later sections. Here we will study some properties of systems of sets and measures provided they are given.

Note that all the σ-algebras that one can define on a nonempty set Ω lie between the smallest one $\{\emptyset, \Omega\}$ and the biggest one $\mathcal{P}(\Omega)$. Since the intersection of σ-algebras is again a σ-algebra, given $\mathcal{F} \subset \mathcal{P}(\Omega)$ there is a smaller σ-algebra containing $\mathcal{F}$ called the σ-algebra generated by $\mathcal{F}$ and denoted $\sigma(\mathcal{F})$. This σ-algebra can be actually built explicitly from $\mathcal{F}$ using transfinite induction. Among the σ-algebras generated by families of sets in a topological space we

will consider the *Borel σ-algebra* which is generated by the open (eq. closed) sets and the *Baire σ-algebra* which is the smaller making measurable the continuous functions. Borel and Baire sets coincide for a metrizable space but they are different in general.

Sometimes it is necessary to check that a given family of sets is a σ-algebra however the stability by complements is far from being obvious. The following notion can be helpful in those cases. A *monotone class* on Ω is a family of sets $\mathcal{M} \subset \mathcal{P}(\Omega)$ which is stable by countable monotone unions and intersections, namely if $A_1 \subset A_2 \subset \cdots \in \mathcal{M}$ then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}$ and if $B_1 \supset B_2 \supset \cdots \in \mathcal{M}$ then $\bigcap_{n=1}^{\infty} B_n \in \mathcal{M}$.

Theorem 8.2.1. *Let $\mathcal{A}$ be an algebra of subsets of Ω. Then $\sigma(\mathcal{A})$ is the smallest monotone class that contains $\mathcal{A}$.*

Proof. The existence of a smallest monotone class $\mathcal{M}$ containing $(\mathcal{A})$ is clear, and obviously it is contained into $\sigma(\mathcal{A})$ as any σ-algebra is a monotone class, therefore we have to show the reverse implication. Take a set $A \subset \Omega$ and consider the class

$$\mathcal{M}(A) = \{B \subset \Omega : A \setminus B, B \setminus A, A \cup B \in \mathcal{M}\}.$$

Observe that $\mathcal{M}(A)$ is a monotone class for any $A \subset \Omega$. Suppose that $A \in \mathcal{A}$. The definition of the set above implies clearly that $\mathcal{A} \subset \mathcal{M}(A)$ and therefore $\mathcal{M} \subset \mathcal{M}(A)$ by minimality. Now suppose that $B \in \mathcal{M}$. Since $B \in \mathcal{M}(A)$, the definition of $\mathcal{M}(A)$ implies that $A \in \mathcal{M}(B)$ and this is true for any $A \in \mathcal{A}$. Therefore $\mathcal{M} \subset \mathcal{M}(B)$ for any $B \in \mathcal{M}$. Appealing again to the definition of $\mathcal{M}(B)$ we deduce that if $A, B \in \mathcal{M}$ then $A \setminus B, A \cup B \in \mathcal{M}$. The first containment implies the stability by complements, the second one the stability by countable unions because now they can be reduced to monotone unions.■

A nonempty set Ω endowed with a σ-algebra Σ is called a *measurable space*. A (countably additive or σ-additive, if we want to stress the notion) *measure* μ defined on a measurable space (Ω, Σ) is a function $\mu : \Sigma \to [0, +\infty]$ which satisfies $\mu(\emptyset) = 0$ and

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$$

whenever $(A_n)_{n=1}^{\infty} \subset \Sigma$ are mutually disjoint. Sometimes we will have to consider the weaker notion of finitely additive measure. In that case the qualification is always necessary to avoid confusions. Finitely additive (an so σ-additive) measures has these properties whose verification is left to the reader:

1. If $A, B \in \Sigma$, $A \subset B$ and $\mu(A) < +\infty$ then $\mu(B \setminus A) = \mu(B) - \mu(A)$.

2. If $A, B \in \Sigma$ then $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$.

3. If $(A_i)_{i=}^n \subset \Sigma$ have finite measure then

$$\mu\left(\bigcup_{i=1}^n A_i\right) =$$

$$\sum_i \mu(A_i) - \sum_{i \neq j} \mu(A_i \cap A_j) + \sum_{\#\{i,j,k\}=3} \mu(A_i \cap A_j \cap A_k) - \cdots \pm \mu\left(\bigcap_{i=1}^n A_i\right).$$

Now some properties that take advantage of the σ-additivity.

Proposition 8.2.2. *Let (Ω, Σ, μ) be a measure space and $(A_n) \subset \Sigma$.*

1. *If $A_1 \subset A_2 \subset A_3 \subset \ldots$ then $\mu(\bigcup_{n=1}^\infty A_n) = \lim_n \mu(A_n)$.*

2. *If $A_1 \supset A_2 \supset A_3 \supset \ldots$ and $\mu(A_1) < +\infty$ then $\mu(\bigcap_{n=1}^\infty A_n) = \lim_n \mu(A_n)$.*

Proof. In the first case define sets $B_n = A_n \setminus \bigcup_{k<n} A_k$ and note $A_n = \bigcup_{k=1}^n B_k$ being the last union disjoint. Therefore

$$\lim_n \mu(A_n) = \lim_n \sum_{k=1}^n \mu(B_k) = \sum_{k=1}^\infty \mu(B_k) = \mu(\bigcup_{k=1}^\infty B_k) = \mu(\bigcup_{k=1}^\infty A_k).$$

The second case is consequence of the first when applied to the sets $C_n = A_1 \setminus A_n$. Indeed, we have

$$\mu(\bigcap_{n=1}^\infty A_n) = \mu(A_1 \setminus \bigcup_{n=1}^\infty C_n) = \mu(A_1) - \mu(\bigcup_{n=1}^\infty C_n)$$

$$= \mu(A_1) - \lim_n \mu(C_n) = \lim_n \mu(A_1 \setminus C_n) = \lim_n \mu(A_n). \qquad \blacksquare$$

Now we will give an example in order to discuss a classification of measure spaces. Consider a set Γ with the σ-algebra of all its subsets $\mathcal{P}(\Gamma)$. Take numbers $(a_\gamma) \subset [0, +\infty]$ for $\gamma \in \Gamma$ and define $\mu(A) = \sum_{\gamma \in A} a_\gamma$. it is not difficult to check that the measure is σ-additive despite the series could be uncountable. The particular case with $a_\gamma = 1$ for all $\gamma \in \Gamma$ is called the *cardinal measure*. Note that the singletons "$\{\gamma\}$" of positive measure cannot

be decomposed into sets of smaller positive measure. Given a measure space (Ω, Σ, μ) a set $A \in \Sigma$ is called an *atom* if $0 \in \{\mu(B), \mu(A \setminus B)\}$ whenever $B \in \Sigma$ with $B \subset A$. We say that two atoms $A, B \in \Sigma$ are equivalent if $\mu((A \setminus B) \cup (B \setminus A)) = 0$. Sometimes we may requiere to work with finite measure sets in order to have a result and then, in second step, to extend the result to countable unions of those sets. We say that a measure space (Ω, Σ, μ) is σ-finite if there exists $(A_n) \subset \Sigma$ with $\mu(A_n) < +\infty$ such that $\Omega = \bigcup_{n=1}^{\infty} A_n$. Note that our example $(\Gamma, \mathcal{P}(\Gamma), \mu)$ is σ-finite if and only if $\{\gamma : a_\gamma = +\infty\} = \emptyset$ and $\{\gamma : a_\gamma \neq 0\}$ is countable. In particular, the cardinal measure on Γ is σ-finite if and only if Γ is countable. With the previous definitions we can prove the following results.

Proposition 8.2.3. *A σ-finite measure space (Ω, Σ, μ) has countably many non equivalent atoms, at most, whose union is called atomic part, and its complement is called atom-free part is case it has positive measure.*

Proof. In this case the atoms must have finite measure. A maximal set of nonequivalent atoms is necessarily countable at most again by the σ-finiteness. Indeed, if Ω is decomposed into disjoint parts with finite measure (Ω_n) and A is an atom then there is only one n such that $\mu(A \setminus \Omega_n) = 0$, that is $A \subset \Omega_n$ except a measure null set. That enforces that given $m \in \mathbb{N}$ only finitely many atoms essentially contained into A_n have measure greater than $1/m$. ∎

Theorem 8.2.4. *Let (Ω, Σ, μ) be an atom-free measure space. Then*

$$\{\mu(A) : A \in \Sigma\} = [0, \mu(\Omega)].$$

Proof. The application of Zorn's lemma allows us to find a maximal family $\mathcal{A}$ of subsets from Σ which is totally ordered and $\mu|_{\mathcal{A}}$ is injective. For every $0 < t < \mu(\Omega)$ the sets

$$A_t = \bigcup \{A \in \mathcal{A} : \mu(A) \leq t\} \quad \text{and} \quad A^t = \bigcap \{A \in \mathcal{A} : \mu(A) \geq t\}$$

belong to Σ because they equal a countable union and a countable intersection respectively. Evidently $\mu(A_t) \leq t \leq \mu(A^t)$. We claim that $\mu(A_t) = \mu(A^t)$. Otherwise $\mu(A^t \setminus A_t) > 0$ and there is $E \subset A^t \setminus A_t)$ such that $0 < \mu(E) < \mu(A^t \setminus A_t)$. The set $A_t \cap E$ can added to $\mathcal{A}$ preserving the total ordering and the injectivity of μ, and therefore violating the maximal property. Now we have $\mu(A_t) = t$ (actually $A_t = A^t$). ∎

As a consequence of these results we have that any σ-finite space can be decomposed into two parts: the atomic one, that behaves like a measure on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$; and the atom-free part, which reminds the Lebesgue measure on $\mathbb{R}$.

8.3 Construction of measures

The notion of *outer measure* plays an essential role here. Let Ω be a nonempty set. A function $\mu^* : \mathcal{P}(\Omega) \to [0, +\infty]$ is called an outer measure if satisfies:

1. $\mu^*(\emptyset) = 0$;
2. $\mu^*(A) \leq \mu^*(B)$ if $A \subset B$;
3. $\mu^*(\bigcup_{n=0}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$.

Obviously an outer measure is nor a measure in the sense of the previous section. The idea is that outer measures are easier to define and we will show that an outer measure behaves as a measure on a "rich" σ-algebra.

It is not very difficult to check that the following function for sets of $\mathbb{R}^n$ is an outer measure

$$\mathbf{m}^*(A) = \inf\{\sum_{n=1}^{\infty} \mathbf{m}(R_n) : (R_n)_{n=1}^{\infty} \text{ rectangles}, A \subset \bigcup_{n=1}^{\infty} R_n\},$$

namely *Lebesgue's outer measure.*

Given an outer measure μ^* on a set Ω, we say that $A \subset \Omega$ is measurable (in the sense of Caratheodory, with respecto to μ^*) if

$$\mu^*(B) = \mu^*(B \cap A) + \mu^*(B \setminus A)$$

for any $B \subset \Omega$. Note that one inequality is guaranteed, so checking measurability reduces to prove that "$\geq$" holds above. The definition of measurabilty could be explained using an ephemeral notion of *inner measure* in $\mathbb{R}^n$ with respect to a bounded rectangle R. For simplicity assume $A \subset R$ and define

$$\mathbf{m}_*(A) = \mathbf{m}(R) - \mathbf{m}^*(R \setminus A).$$

In that case, the equality $\mathbf{m}_*(A) = \mathbf{m}^*(A)$ expressing the measurability of A witnessed by R is analogous to Jordan's notion of measurability.

Theorem 8.3.1. *Let be a nonempty set Ω and an outer measure μ^* defined on its subsets.Then the family of measurable sets is a σ-algebra Σ and $\mu^*|_\Sigma$ is a (σ-additive) measure.*

Proof. Denote by Σ the family of measurable sets. Clearly $\emptyset, \Omega \in \Sigma$ and $A \in \Sigma$ if and only of $A^c \in \Sigma$. As to the union of sets, we will begin by showing that the union of two sets: assume $A_1, A_2 \in \Sigma$ and $B \subset \Omega$ is arbitrary. The measurability of A_1 witnessed by $B \cap (A_1 \cup A_2)$ gives

$$\begin{aligned}\mu^*(B \cap (A_1 \cup A_2)) &= \mu^*(B \cap (A_1 \cup A_2) \cap A_1) + \mu^*(B \cap (A_1 \cup A_2) \cap A_1^c) \\ &= \mu^*(B \cap A_1) + \mu^*(B \cap A_2 \cap A_1^c).\end{aligned}$$

And the measurability of A_2 witnessed by $B \cap A_1^c$ gives

$$\mu^*(B \cap A_1^c \cap A_2) + \mu^*(B \cap A_1^c \cap A_2^c) = \mu^*(B \cap A_1^c).$$

Now we have

$$\begin{gathered}\mu^*(B \cap (A_1 \cup A_2)) + \mu^*(B \cap (A_1 \cup A_2)^c) = \\ \mu^*(B \cap A_1) + \mu^*(B \cap A_2 \cap A_1^c) + \mu^*(B \cap A_1^c \cap A_2^c) \\ = \mu^*(B \cap A_1) + \mu^*(B \cap A_1^c) = \mu^*(B)\end{gathered}$$

which implies the measurability of $A_1 \cup A_2$.
Clearly, that implies that Σ is closed for finite union of sets, and so it is closed for finite intersections and differences via complements. Therefore, in order to show that Σ is closed for countable unions it is enough to consider sequences of disjoint sets $(A_n) \subset \Sigma$. Firstly we will prove by induction the following formula

$$\mu^*(B) = \sum_{k=1}^{n} \mu^*(B \cap A_k) + \mu^*(B \cap \bigcap_{k=1}^{n} A_k^c).$$

Indeed, for $n = 1$ is just the measurability of A_1. Now, the measurability of A_n implies

$$\begin{aligned}\mu^*(B \cap \bigcap_{k=1}^{n-1} A_k^c) &= \mu^*(B \cap \bigcap_{k=1}^{n-1} A_k^c \cap A_n) + \mu^*(B \cap \bigcap_{k=1}^{n} A_k^c) \\ &= \mu^*(B \cap A_n) + \mu^*(B \cap \bigcap_{k=1}^{n} A_k^c).\end{aligned}$$

If we assume the formula is true for $n-1$, the last equality added will imply the formula is true for n.
The formula easily implies

$$\mu^*(B) \geq \sum_{k=1}^{\infty} \mu^*(B \cap A_k) + \mu^*(B \cap \bigcap_{k=1}^{\infty} A_k^c) \geq$$

$$\mu^*(B\cap\bigcup_{k=1}^{\infty}A_k)+\mu^*(B\cap(\bigcup_{k=1}^{\infty}A_k)^c)\geq\mu^*(B)$$

that gives both the measurability of $\bigcup_{k=1}^{\infty}A_k$ and the σ-additivity of $\mu^*|_\Sigma$. ∎

Note that the class of Caratheodory measurable sets with respect to an exterior measure μ^* has the following property: if $\mu^*(A)=0$ then A is measurable. In particular, the σ-algebra Σ given by the previous theorem satisfies $A\in\Sigma$ whenever $A\subset B$ and $\mu(B)=0$. Any measure space with such a property is called *complete*.

The previous theorem is devoid of content if we do not ensure that there are many other measurable sets besides $\emptyset$ and Ω.

Theorem 8.3.2. *Let $\mathcal{A}$ be an algebra of subsets of Ω and $\mu:\mathcal{A}\to[0,+\infty]$ which is σ-additive (on $\mathcal{A}$) and consider the outer measure generated by μ as*

$$\mu^*(A)=\inf\{\sum_{n=1}^{\infty}\mu(A_n):A\subset\bigcup_{n=1}^{\infty}A_n,(A_n)\subset\mathcal{A}\}$$

for any $A\subset\Omega$ and let Σ be the σ-algebra of μ^*-measurable sets. Then we have $\mathcal{A}\subset\Sigma$ and $\mu^*|_{\mathcal{A}}=\mu$.

Proof. Firstly we will show the measurability of any $A\in\mathcal{A}$. Let $B\subset\Omega$ be arbitrary. For any cover $(A_n)\subset\mathcal{A}$ of B we have $B\cap A\subset\bigcup_{n=1}^{\infty}(A_n\cap A)$ and $B\setminus A\subset\bigcup_{n=1}^{\infty}(A_n\setminus A)$ being both covers made of sets from $\mathcal{A}$. We have

$$\mu^*(B\cap A)+\mu^*(B\setminus A)\leq\sum_{n=1}^{\infty}\mu(A_n\cap A)+\sum_{n=1}^{\infty}\mu(A_n\setminus A)=\sum_{n=1}^{\infty}\mu(A_n)$$

which implies $\mu^*(B\cap A)+\mu^*(B\setminus A)\leq\mu^*(B)$. As to recover μ from μ^* assume $A\in\mathcal{A}$ and take a cover $(A_n)\subset\mathcal{A}$. The cover can be improved in the sense that $\sum_{n=1}^{\infty}\mu(A_n)$ does not increase changing A_n by $A_n\cap A$ and by making the sequence disjoint inductively $A_1,A_2\setminus A_1,A_3\setminus(A_1\cup A_2),\ldots$ so we may assume that (A_n) is a countable decomposition of A. By hypothesis $\sum_{n=1}^{\infty}\mu(A_n)=\mu(A)$, therefore $\mu^*(A)=\mu(A)$. ∎

Perhaps a delicate point in order to apply the previous results is to prove that the "pre-measure" μ is σ-additive. This verification can be reduced to a simpler class of subsets $\mathcal{R}\subset\mathcal{A}$ provided that a few properties are satisfied.

Proposition 8.3.3. *Let $\mathcal{R} \subset \mathcal{P}(\Omega)$ be a class of sets and let Σ be the σ-algebra that generates. Suppose that:*

1. *There is a function $\mu : \mathcal{R} \to [0, +\infty]$ which is σ-additive (on $\mathcal{R}$);*
2. *$\mathcal{R}$ is a "rectangular class", namely if $R, S \in \mathcal{R}$ implies $R \cap S \in \mathcal{R}$ and there are mutually disjoint sets $R_1, \ldots, R_n \in \mathcal{R}$ such that $R \backslash S = \bigcup_{i=1}^{n} R_i$.*

Then μ can be extended to Σ as a σ-additive measure.

Sketch of proof. Consider $\mathcal{A}$ the algebra generated by $\mathcal{R}$. Note that

$$\mathcal{A} = \{\bigcup_{i=1}^{n} R_i : R_1, \ldots, R_n \in \mathcal{R}\}.$$

It is easy to check the following facts:

1. Every $A \in \mathcal{A}$ can be expressed as a disjoint union $A = \bigcup_{i=1}^{n} R_i$ with $R_1, \ldots, R_n \in \mathcal{R}$
2. The formula $\mu(A) = \sum_{i=1}^{n} \mu(R_i)$ using the disjoint decomposition above extends unambiguously the measure μ to all $\mathcal{A}$.
3. μ is σ-additive on $\mathcal{A}$.

Now we can apply Theorem 8.3.2 in order to finish the proof. ∎

The family $\mathcal{R}$ plays the role of the rectangles in the construction of measures on $\mathbb{R}^n$ and that is the reason for the choice of the name rectangular, obviously. At this point we can resume the construction of the Lebesgue measure. Recall that we have a finitely additive measure $\mathbf{m}$ defined on the algebra generated by the rectangles and exterior measure $\mathbf{m}^*$ built from countable covers with rectangles. We can use Theorem 8.3.2 to show that we recover $\mathbf{m}$ from $\mathbf{m}^*$, and according to the reduction to rectangular families, we only have to show that $\mathbf{m}$ is σ-additive within the rectangles.

Fact 8.3.4. *If a rectangle $R \subset \mathbb{R}^d$ is a countable union $R = \bigcup_{n=1}^{\infty} R_n$ of disjoint (or merely non overlapping) rectangles then $\mathbf{m}(R) = \sum_{n=1}^{\infty} \mathbf{m}(R_n)$.*

Proof. If one of the dimensions of R collapse to 0 then all the measures are 0 and so the equality holds, so we may assume that all the sides of R has length greater than 0. The case R has an infinite edge so $\mathbf{m}(R) = +\infty$ can be reduced

to the bounded case by intersecting $\|\cdot\|_\infty$-balls centred at the origin. Assume then R is bounded and closed (the faces have d-dimensional measure 0). Fix $\varepsilon > 0$ and for every $n \in \mathbb{N}$ take B_n an open $\|\cdot\|_\infty$-ball centred at the origin such that

$$\mathbf{m}(R_n + B_n) < \mathbf{m}(R_n) + 2^{-n}\varepsilon$$

which is possible by the continuous dependence of the measure on the lengths of the edges. Since the enlarged rectangles $R_n + B_n$ are open an cover R there are finitely many such that $R \subset \bigcup_{k=1}^{n}(R_n + B_n)$. We deduce that

$$\mathbf{m}(R) \leq \sum_{k=1}^{n} \mathbf{m}(R_k) + \varepsilon \leq \sum_{k=1}^{\infty} \mathbf{m}(R_k) + \varepsilon$$

and thus $\mathbf{m}(R) \leq \sum_{n=1}^{\infty} \mathbf{m}(R_n)$ since $\varepsilon > 0$ is arbitrary. The reverse inequality follows from the finite additivity of $\mathbf{m}$. ∎

That proves the existence of a σ-additive measure on $\mathbb{R}^d$ that extends the measure of the rectangles, and so the Jordan measure, which is called the *Lebesgue measure*. Some of the additional properties and features of the Lebesgue measure will be developed along the next sections as particular cases.

8.4 Measurable functions

The underlying idea behind Riemann integration theory is that the integral can be defined for those functions that can be approximated suitably by functions constant on intervals. Uniform approximation by functions constant on intervals is enough for the integration of continuous functions. The class of Riemann integrable functions is slightly greater, but their functions are bound to be continuous almost everywhere. Lebesgue integration theory carries the definition of integral to functions that can be approximated suitably by functions constant on measurable sets. Although Lebesgue measure and the integration of functions defined on $\mathbb{R}^n$ is always a main motivation, integration theory will be developed for a general measure space.

Let (Ω, Σ) a measurable space. We call a function $s : \Omega \to \mathbb{R}$ *simple* it is a linear combination of characteristic functions of sets from Σ, namely $s = \sum_{k=1}^{n} a_k \chi_{A_k}$ with $a_k \in \mathbb{R}$ and $A_k \in \Sigma$ for $k = 1, \ldots, n$. If μ is a finite

measure (even finitely additive) defined on (Ω, Σ) the number

$$\int s\, d\mu = \sum_{k=1}^{n} a_n \mu(A_k)$$

noes not depend on the particular expression of s. This is a tedious but elementary verification based on the algebra structure of Σ and the additivity of μ. Note that the integral $\int\, d\mu$ defines a linear operator on the space of simple functions $\mathcal{S}$ that can be naturally extended to any closure of $\mathcal{S}$ whith respect to a topology which makes $\int\, d\mu$ continuous. For instance, the topology of uniform convergence in case of $\mu(\Omega) < +\infty$ would do the work. However this is not the way, and the theory is more powerful if the extension of the integral is done by monotonicity and the set of integrable functions can be described in an easier way.

Let (Ω_1, Σ_1) and (Ω_2, Σ_2) be measurable spaces. We say that a map $f : \Omega_1 \to \Omega_2$ is measurable if $f^{-1}(A) \in \Sigma_1$ whenever $A \in \Sigma_2$. Note that this definition reminds the continuity. Clearly, the composition of measurable maps is again measurable. The statement can be easily proved by the reader using the minimality of the generated σ-algebra.

Proposition 8.4.1. *If $\Sigma_2 = \sigma(\mathcal{F})$ then $f : \Omega_1 \to \Omega_2$ is measurable if and only if $f^{-1}(A) \in \Sigma_1$ whenever $A \in \mathcal{F}$.*

As we will consider mainly real functions defined on a measurable space (Ω, Σ), the measurability of $f : \Omega \to \mathbb{R}$ will be understood with respect to the Borel σ-algebra on $\mathbb{R}$. The measurability criterion just said implies that $f : \Omega \to \mathbb{R}$ is measurable if one of the sets $f^{-1}((-\infty, a))$, $f^{-1}((-\infty, a])$, $f^{-1}((a, +\infty))$ or $f^{-1}([a, +\infty))$ lies in Σ for all $a \in \mathbb{R}$. The properties of the measurable functions are summarized in the following result.

Proposition 8.4.2. *Let (Ω, Σ) be a measurable space and let $\mathcal{M}$ denote the set of measurable real functions defined on it and let $\mathcal{M}_\infty$ denote the set of measurable functions valued into $\overline{R} = [-\infty, +\infty]$ (also with the Borel σ-algebra). Then:*

1. *If $f_1, \ldots, f_n \in \mathcal{M}$ (or $\mathcal{M}_\infty$) then $(f_1, \ldots, f_n) : \Omega \to \mathbb{R}^n$ (or $\overline{R}^n$) is measurable for the Borel σ-algebra on $\mathbb{R}^n$ (resp. $\overline{R}^n$).*

2. *$\mathcal{M}$ is an algebra, $\mathcal{M}_\infty$ is stable by inverses, $\mathcal{M}$ and $\mathcal{M}_\infty$ are lattices.*

3. *$\mathcal{M}_\infty$ is stable by supremums and infimums of countable sets.*

4. *$\mathcal{M}_\infty$ is stable by* $\liminf$ *and* $\limsup$ *of sequences, and thus it is also stable by limits of pointwise convergent sequences.*

5. *If $f \in \mathcal{M}$ is bounded then it can be uniformly approximated by simple functions.*

6. *If $f \in \mathcal{M}_\infty$ and $f \geq 0$ there is an increasing sequence of simple functions*
$$0 \leq s_1 \leq s_2 \leq \cdots \leq f$$
which converges pointwise to f.

Proof. (1) In both cases the topologies are generated by rectangles. Note that
$$(f_1, \ldots, f_n)^{-1}([a_1, b_1] \times \cdots \times [a_n, b_n]) = \bigcap_{k=1}^{n} f_k^{-1}([a_k, b_k]) \in \Sigma.$$
(2) Binary algebra operations can be expressed by a composition with a continuous function $* : \mathbb{R}^2 \to \mathbb{R}$. That apply also to lattice operations on the extended reals. On the other hand, note that $t \to t^{-1}$ is continuous on $[-\infty, +\infty]$.
(3) If $f = \sup\{f_n : n \in \mathbb{N}\}$ then
$$f^{-1}([-\infty, a]) = \bigcap_{n=1}^{\infty} f_n^{-1}([-\infty, a]) \in \Sigma$$
and the infimum is done likewise.
(4) Note that
$$\limsup_n f_n = \inf\{\sup\{f_k : k \geq n\} : n \in \mathbb{N}\};$$
$$\liminf_n f_n = \sup\{\inf\{f_k : k \geq n\} : n \in \mathbb{N}\}$$
so both are measurable.
(5) Assume $f(\Omega) \subset [a, b)$ and fix $\varepsilon > 0$. Take $(t_k)_{k=1}^n$ a partition of $[a, b]$, that is, $a = t_1 < t_2 < \cdots < t_n = b$ such that $t_{k+1} - t_k < \varepsilon$ and the sets $A_k = f^{-1}([t_k, t_{k+1}))$. The simple function
$$s = \sum_{k=1}^{n-1} a_k \chi_{A_k}$$
satisfies $s \leq f$ and $\|f - s\|_\infty < \varepsilon$.
(6) The previous construction applied to $f_n = \min\{f, n\}$ can produce a simple function $s_n' \leq f_n$ such that $\|f_n - s_n'\|_\infty < 1/n$. Now take $s_n = \max\{s_1', \ldots, s_n'\}$ which is simple and the sequence (s_n) converges pointwise to f. ∎

8.5 Integration

Now we are ready to define the integral for positive functions. Let (Ω, Σ, μ) be a measure space. Recall that the integral was already defined for simple functions in case that $\mu(\Omega) < +\infty$. If we limit ourselves to positive simple functions $s = \sum_{k=1}^{n} a_n \chi_{A_k}$ with $a_k \geq 0$ we may remove the finiteness hypothesis and the formula

$$\int s \, d\mu = \sum_{k=1}^{n} a_n \mu(A_k)$$

will make sense in $[0, +\infty]$. It is easy to see that also in this case the value does not depend on the particular (positive) representation of s.

We define the integral of $f : \Omega \to [0, +\infty]$ with respect to μ as the value in $[0, +\infty]$ given by

$$\int f \, d\mu = \sup\{ \int s \, d\mu : 0 \leq s \leq f, s \in \mathcal{S}\}$$

and $\int_A f \, d\mu = \int \chi_A f \, d\mu$ for $A \in \Sigma$. Note that the computation of $\int s \, d\mu$ could need operations involving $+\infty$, however the limitation to positive values avoids us possible troubles. The very definition implies these almost obvious properties that we will need later.

Proposition 8.5.1. *Under the notation and assumptions above we have:*

1. *if $0 \leq f \leq g$ are measurable then $\int f \, d\mu \leq \int g \, d\mu$;*
2. *if $A, B \in \Sigma$, $A \subset B$ and f is measurable then $\int_A f \, d\mu \leq \int_B f \, d\mu$;*
3. *if $f \geq 0$ is measurable and $\lambda \geq 0$ a real number then $\int \lambda f \, d\mu = \lambda \int f \, d\mu$.*

Theorem 8.5.2 (Monotone convergence theorem)**.** *Let (Ω, Σ, μ) be a measure space and let $0 \leq f_1 \leq f_2 \leq \cdots \leq f$ a sequence of measurable functions defined on Ω with values in $[0, +\infty]$ that pointwise converges to f. Then*

$$\lim_n \int f_n \, d\mu = \int f \, d\mu.$$

Proof. The limit of the lefthand side exists in $[0, +\infty]$ by monotony and it is obvious the inequality

$$\lim_n \int f_n \, d\mu \leq \int f \, d\mu.$$

For the converse, fix a simple function $s \leq f$ and a number $\lambda \in (0,1)$. Note that the sequence of measurable sets

$$A_n = \{x \in \Omega : f_n(x) \geq \lambda s(x)\}$$

is increasing and $\bigcup_{n=1}^{\infty} A_n = \Omega$. We have

$$\int f_n \, d\mu \geq \int_{A_n} f_n \, d\mu \geq \lambda \int_{A_n} s \, d\mu.$$

Note that $\nu(A) = \int_A s \, d\mu$ defines a positive measure on Σ, so taking limits we have

$$\lim_n \int f_n \, d\mu \geq \lambda \lim_n \nu(A_n) = \lambda \, \nu(\Omega) = \lambda \int s \, d\mu.$$

Being $\lambda < 1$ arbitrary and taking into account the definition of the integral we get

$$\lim_n \int f_n \, d\mu \geq \int f \, d\mu$$

as wished. ∎

Corollary 8.5.3. *If $f, g : \Omega \to [0, +\infty]$ are measurable then*

$$\int (f+g) \, d\mu = \int f \, d\mu + \int g \, d\mu.$$

Proof. Take increasing sequences of simple functions pointwise converging $s_n \to f$ and $z_n \to g$. Then

$$\int (f+g) \, d\mu = \lim_n \int (s_n + z_n) \, d\mu$$

$$= \lim_n \int s_n \, d\mu + \lim_n \int z_n \, d\mu = \int f \, d\mu + \int g \, d\mu$$

applying the monotone convergence theorem and having in mind that the additivity of the integral was established for simple functions. ∎

Corollary 8.5.4. *Let (Ω, Σ, μ) be a measure space and let (f_n) be a sequence of measurable functions valued in $[0, +\infty]$. Then*

$$\int \left(\sum_{n=1}^{\infty} f_n \right) d\mu = \sum_{n=1}^{\infty} \int f_n \, d\mu.$$

Proof. Just apply the monotone convergence theorem to the increasing sequence of functions $g_n = \sum_{k=1}^{n} f_k$ whose limit is $\sum_{k=1}^{\infty} f_k$. ∎

Proposition 8.5.5 (Fatou's lemma)**.** *Let (f_n) be a sequence of non negative measurable functions. Then*

$$\int \liminf_n f_n \, d\mu \leq \liminf_n \int f_n \, d\mu.$$

Proof. Consider the increasing sequence $g_n = \inf\{f_k : k \geq n\}$ and apply the monotone convergence theorem

$$\int \liminf_n f_n \, d\mu = \int \lim_n g_n \, d\mu = \lim_n \int g_n \, d\mu \leq \liminf_n \int f_n \, d\mu$$

since $\int g_n \, d\mu \leq \int f_n \, d\mu$ for all $n \in \mathbb{N}$. ∎

Now we are ready to extend the notion of integral to non positive functions. We say that $f : \Omega \to [-\infty, +\infty]$ is *integrable* if it is measurable and $\int |f| \, d\mu < +\infty$. In such a case we define the integral of f as the real number

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu.$$

We will also consider the integrals over sets $\int_A f \, d\mu := \int \chi_A f \, d\mu$. The following properties are not a surprise.

Proposition 8.5.6. *Let $\mathcal{L}^1(\mu)$ denote the set of integrable functions defined on the measure space (Ω, Σ, μ). Then*

1. *$\mathcal{L}^1(\mu)$ is vector lattice;*
2. *the integral is a linear functional on $\mathcal{L}^1(\mu)$;*
3. *$|\int f \, d\mu| \leq \int |f| \, d\mu$*

The following result is the key of the versatility of Lebesgue integral.

Theorem 8.5.7 (Dominated convergence theorem)**.** *Let $(f_n) \subset \mathcal{L}^1(\mu)$ a sequence which converges pointwise to f. Assume that there is $g \in \mathcal{L}^1(\mu)$ such that $|f_n| \leq g$ for all $n \in \mathbb{N}$. Then $f \in \mathcal{L}^1(\mu)$ and*

$$\lim_n \int f_n \, d\mu = \int f \, d\mu \quad \textit{and} \quad \lim_n \int |f_n - f| \, d\mu = 0.$$

Proof. The integrability of f is clear from the inequality $|f| \leq g$. We may apply Fatou's lemma to the positive sequence $(2g - |f_n - f|)$ we get

$$\int 2g\, d\mu = \int \lim_n (2g - |f_n - f|)\, d\mu \leq \liminf_n \int (2g - |f_n - f|)\, d\mu$$

$$= \int 2g\, d\mu - \limsup_n \int |f_n - f|\, d\mu.$$

We deduce $\limsup_n \int |f_n - f|\, d\mu = 0$ and thus $\lim_n \int |f_n - f|\, d\mu = 0$, which easily implies the other part of the statement. ∎

So far we have Lebesgue measure on $\mathbb{R}^d$ and the basics of abstract integration theory. It is time to put them together and to compare the result to Riemann integration theory.

Proposition 8.5.8. *Let $f : D \to \mathbb{R}$ be a Riemann integrable function defined on a Jordan measurable set $D \subset \mathbb{R}^d$. Then f is measurable and both integrals, Lebesgue and Riemann, coincide for f.*

Proof. Jordan measurable sets are Lebesgue measurable as they differs from an open set in a null measure set. Lower and upper sums associated to Riemann integral can be understood as the integrals of simple functions encompassing f. Choosing a sequence of partitions, which are ordered and approaching the integral for Riemann sums, the lower and upper sequences of simple functions converge to limits $\underline{f}$ and $\overline{f}$ which are measurable, satisfy $\underline{f} \leq f \leq \overline{f}$ and by construction

$$\int (\overline{f} - \underline{f})\, d\mathbf{m} = 0.$$

That implies f coincides with $\underline{f}$ and $\overline{f}$ almost everywhere. That implies the Labesgue measurability of f and the coincidence of Riemann and Lebesgue integrals. ∎

Unfortunately there are some important integrals which are not covered by Lebesgue theory. For instance, the following one exists in improper Riemann sense but not in Lebesgue

$$\int_0^{+\infty} \frac{\sin x}{x}\, dx.$$

The convergence theorems cast some light on the following question: when can we commute derivation and integration? That is, whether is true the

following formula

$$\frac{\partial}{\partial y}\int f(x,y)\,d\mu(x)=\int\frac{\partial f}{\partial y}(x,y)\,d\mu(x).$$

If we express the derivation by its very definition at y_0

$$\frac{\partial}{\partial y}\int f(x,y)\,d\mu(x)\bigg|_{y=y_0}=\lim_{h\to 0}\int h^{-1}(f(x,y_0+h)-f(x,y_0))\,d\mu(x)$$

and this last limit can be written sequentially, taking $h=h_n$ with $\lim_n h_n=0$, for instance. Thus the question is reduced to know the limit

$$\lim_n\int h_n^{-1}(f(x,y_0+h_n)-f(x,y_0))\,d\mu(x)=\lim_n\int\frac{\partial f}{\partial y}(x,y_0+\theta(x,n))\,d\mu(x)$$

where $|\theta(x,n)|<|h_n|$ is given by the finite increments theorem. If the family of functions $\{\frac{\partial f}{\partial y}(x,y):y\}$ were dominated by a positive integrable function for y in a neighbourhood of y_0 we could apply the dominated convergence theorem. The analysis for interesting integrals is sometimes more tricky. Lets go back to the improper Riemann non-Lebesgue integral above.

We will consider the auxiliary function defined for $y>0$ by an integral

$$F(y)=\int_0^{+\infty}e^{-xy}\,\frac{\sin x}{x}\,dx$$

We claim that

$$F'(y)=\int_0^{+\infty}\frac{\partial}{\partial y}\left(e^{-xy}\,\frac{\sin x}{x}\right)\,dx=-\int_0^{+\infty}e^{-xy}\sin x\,dx$$

In order to show the domination by a positive integrable function observe that

$$0\le xe^{-xy}\le xe^{-xy_1}$$

if $y\ge y_1>0$ and $x\in[0,+\infty)$. Now, using elementary calculus of primitives

$$-\int_0^{+\infty}e^{-xy}\sin x\,dx=\frac{-1}{1+y^2}$$

Therefore $F(y)$ and $-\arctan y$ should differ in a constant. Taking limits as y goes to $+\infty$, it is clear that $F(y)$ goes to 0, thus we have

$$F(y)=\frac{\pi}{2}-\arctan y$$

We claim that

$$\int_0^{+\infty} \frac{\sin x}{x}\, dx = \lim_{y\to 0^+} F(y) = \frac{\pi}{2}.$$

Indeed, take

$$a_n(y) = \int_{(n-1)\pi}^{n\pi} e^{-xy}\, \frac{\sin x}{x}\, dx$$

and notice that $\sum_{n=1}^{\infty} a_n(y)$ is a Leibniz series, that is, signs are alternating and $|a_n(y)|$ goes monotonically to 0. For $n \in \mathbb{N}$ odd, we have

$$0 \leq \sum_{k=n}^{\infty} a_k(y) \leq a_n(y) \leq \frac{e^{-\pi n y}}{\pi n} \int_{(n-1)\pi}^{n\pi} \sin x\, dx \leq \frac{2}{\pi n}$$

Decompose the integral as follows

$$\int_0^{+\infty} e^{-xy}\, \frac{\sin x}{x}\, dx - \int_0^{(n-1)\pi} e^{-xy}\, \frac{\sin x}{x}\, dx = \int_{(n-1)\pi}^{+\infty} e^{-xy}\, \frac{\sin x}{x}\, dx = \sum_{k=n}^{\infty} a_k(y)$$

Taking limits for $y \to 0^+$, the first term converges to

$$\frac{\pi}{2} - \int_0^{(n-1)\pi} \frac{\sin x}{x}\, dx$$

while the last one remains bounded by $2/(\pi n)$ for n odd. Taking limits in n, we get that

$$\int_0^{+\infty} \frac{\sin x}{x}\, dx = \frac{\pi}{2}.$$

8.6 Approximation and topology

We already know that the approximation of measurable functions by simple ones is possible uniformly in the bounded case. On the space $\mathcal{L}^1(\mu)$ we may define a seminorm by the formula

$$\|f\|_1 = \int |f|\, d\mu.$$

Proposition 8.6.1. *The set of simple functions on finite measure sets is dense in* $(\mathcal{L}^1(\mu), \|\cdot\|_1)$.

Proof. Given $f \in \mathcal{L}^1(\mu)$ the sequence $f_n = \min\{n, \max\{f, -n\}\}$ is dominated by $|f|$ and converges to f in $\|\cdot\|_1$, so we may assume f is bounded. Now consider the sequence (f_n) where $f_n(x) = f(x)$ if $|f(x)| \geq 1/n$ and $f_n(x) = 0$ otherwise. This sequence is also dominated by $|f|$ and converges to f, and so for the seminorm $\|f\|_1$. The functions f_n have supports of finite measure, so they can be uniformly approached by simple functions also with supports of finite measure. ∎

The previous result makes clear that the approximation of integrable functions by others reduces to the approximation of simple functions, and thus the approximation of characteristic functions. Define a pseudometric on Σ by $d_\mu(A, B) = \mu(A \Delta B)$ where $A \Delta B = (A \setminus B) \cup (B \setminus A)$ is the symmetric difference. Note that d is actually the restriction of the seminorm $\|\cdot\|_1$ through characteristic functions $d(A, B) = \|\chi_A - \chi_B\|_1$.

Proposition 8.6.2. *Let (Ω, Σ, μ) a finite measure space and assume that Σ is generated by an algebra $\mathcal{A}$. Then $\mathcal{A}$ is dense in (Σ, d_μ)*

Proof. Consider the set

$$\mathcal{M} = \{A \in \Sigma : \text{for all } \varepsilon > 0 \text{ there is } B \in \mathcal{A} \text{ with } d_\mu(A, B) < \varepsilon\}.$$

It is obvious that $\emptyset, \Omega \in \mathcal{M}$ and $A \in \mathcal{M}$ implies $A^c \in \mathcal{M}$ as $A \Delta B = A^c \Delta B^c$. For the union of two sets note that

$$(A_1 \cup A_2) \Delta (B_1 \cup B_2) \subset (A_1 \Delta B_1) \cup (A_2 \Delta B_2).$$

Now, as we have stability by complements and unions, we have also stability by intersections and differences. In particular, in order to see that $\mathcal{M}$ is stable by countable unions it is enough to consider unions of disjoint sequences $(A_n) \subset \mathcal{M}$. Given $\varepsilon > 0$, as $\mu(\Omega) < +\infty$ we may take $n \in \mathbb{N}$ such that $\mu(\bigcup_{k>n} A_k) < \varepsilon/2$. If $(B_k)_{k=1}^n \subset \mathcal{A}$ are such that $\sum_{k=1}^n d_\mu(A_k, B_k) < \varepsilon/2$. We have

$$\left(\bigcup_{k=1}^{\infty} A_k\right) \Delta \left(\bigcup_{k=1}^{n} B_k\right) \subset \bigcup_{k=1}^{n} (A_k \Delta B_k) \cup \bigcup_{k>n} A_k$$

which implies $d_\mu(\bigcup_{k=1}^\infty A_k, \bigcup_{k=1}^n B_k) < \varepsilon$. Since $\mathcal{M} \subset \Sigma$ is a σ-algebra that contains $\mathcal{A}$ they must be the same. ∎

With similar ideas we can deal with the *completion* of a measure space.

Proposition 8.6.3. *Let (Ω, Σ, μ) a measure space. There exists a complete measure space over the same set $(\Omega, \overline{\Sigma}, \overline{\mu})$ which is the smaller possible and has the following property: for every $A \in \overline{\Sigma}$ there is $B \in \Sigma$ such that $\overline{\mu}(A \Delta B) = 0$, that is, Σ is dense in $\overline{\Sigma}$ with respect to $d_{\overline{\mu}}$.*

Proof. Evidently, a completion of (Ω, Σ, μ) must contain the family of sets

$$\mathcal{N} = \{M \subset \Omega : \exists N \in \Sigma, \mu(N) = 0, M \subset N\}.$$

Using the same ideas than in the previous proposition it is possible to prove that

$$\overline{\Sigma} = \{A \subset \Omega : \exists B \in \Sigma, A \Delta B \in \mathcal{N}\}$$

is a σ-algebra and $\overline{\mu}(A) = \mu(B)$ if $A \Delta B \in \mathcal{N}$ is well defined. ∎

Corollary 8.6.4. *Let $(\Omega, \overline{\Sigma}, \overline{\mu})$ be the completion of (Ω, Σ, μ). If f is $\overline{\Sigma}$ measurable, then there is a Σ-measurable function g such that $f = g$ almost everywhere with respect to $\overline{\mu}$.*

Proof. For every $t \in \mathbb{Q}$ take a set $N_t \in \Sigma$ with $\mu(N_t) = 0$ such that there is $A_t \in \Sigma$ such that $A_t \subset \{f \leq t\}$ and $\{f \leq t\} \setminus A_t \subset N_t$. The set $N = \bigcup_{t \in \mathbb{Q}} N_t$ is null. Define $g(x) = f(x)$ if $x \notin N$ and $g(x) = 0$ otherwise. By construction g fulfils the requirements. ∎

Under the hypotheses of the dominated convergence theorem a sequence $(f_n) \subset \mathcal{L}^1(\mu)$ converges to its limit with respect to the seminorm $\| \cdot \|_1$. It is interesting that the converse is true through passing to a suitable subsequence.

Theorem 8.6.5. *If $\lim_n \|f_n - f\|_1 = 0$ then there is a subsequence (f_{n_k}) which converges to f almost everywhere.*

Proof. For every $\varepsilon > 0$ we have

$$\mu(\{|f_n - f| > \varepsilon\}) \leq \frac{1}{\varepsilon} \int |f_n - f| \, d\mu \ \to \ 0.$$

Therefore it is possible to find n_1 such that

$$\mu(\{|f_{n_1} - f| > 1\}) \leq 1/2.$$

Inductive it is possible to build an increasing sequence $n_1 < n_2 < \dots$ such that the sets

$$A_k = \{|f_{n_k} - f| > 1/k\}$$

satisfy $\mu(A_k) \leq 2^{-k}$. Take $A = \bigcap_{k=1}^{\infty} \bigcup_{j \geq k} A_j$. And note that $\mu(A) = 0$. By construction we have for any $x \in A^c$ that $|f_{n_k}(x) - f(x)| \leq 1/k$ from a certain k on, and so the theorem is proven. ■

Now we will consider a topological space X (see Appendix A) endowed with its Borel σ-algebra $\mathfrak{B}$ and a measure μ. We say that $A \subset X$ is *regular* if for every $\varepsilon > 0$ there is a closed set $F \subset A$ and an open set $U \supset A$ such that $\mu(U \setminus F) < \varepsilon$. We say that μ is regular if all the sets in $\mathfrak{B}$ are regular. We may also use this weaker form of regularity which is more convenient for infinite measures. We say that $A \in \mathfrak{B}$ is *inner regular* if

$$\mu(A) = \sup\{\mu(F) : F \subset A \text{ closed}\}$$

and *outer regular* if

$$\mu(A) = \inf\{\mu(u) : U \supset A \text{ closed}\}.$$

Proposition 8.6.6. *Let $(X, \mathfrak{B}, \mu)$ be a topological space with a finite Borel measure. If all the open (or closed) sets are regular then μ is regular.*

Proof. Defining the family of sets

$$\mathcal{M} = \{A \in \mathfrak{B} : \forall\ \varepsilon > 0\ \exists F \subset A \subset U, \mu(U \setminus F) < \varepsilon\}$$

we may proceed by applying the same ideas of the proof of Proposition 8.6.2.■

Since in a metric space the open sets are a countable union of closed sets we have the following.

Corollary 8.6.7. *Every finite Borel measure in a metrizable space is regular.*

The possibility of changing closed sets by compact sets in the inner approximation.

Theorem 8.6.8. *Assume that X is separable and completely metrizable and μ a finite Borel measure on it then*

$$\mu(A) = \sup\{\mu(K) : K \subset A \textit{ compact}\}$$

for every $A \in \mathfrak{B}$.

Proof. After the previous proposition it is enough to show the result is true for A closed. Fix $\varepsilon > 0$. For every $n \in \mathbb{N}$ take a countable cover $(B_{n,m})_{m=1}^{\infty}$ of A by balls of radius less than $1/n$. Now fix m_n such that

$$\mu(A \setminus \bigcup_{m=1}^{m_n} B_{n,m}) < 2^{-n}\varepsilon.$$

Now we have

$$B = \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{m_n} B_{n,m}$$

is totally bounded and $\mu(A \setminus B) < \varepsilon$ by construction. The set $K = \overline{A \cap B} \subset A$ is compact and satisfies $\mu(K) > \mu(A) - \varepsilon$. ∎

The results discussed so far could be adapted for σ-finite measures with some additional hypotheses. For instance, it is easy that Theorem 8.6.8 is still true if the space X can be covered by countably many closed sets of finite measure. Let us mention that the result is still true even in the σ-finite case since any Borel subset of the completely metrizable space X can be completely metrized for the relative topology. In any case, for our most important case we have the following.

Theorem 8.6.9. *A Lebesgue measurable set of $\mathbb{R}^d$ differs from a Borel set in a null measure set and it is regular for the Lebesgue measure.*

Proof. Let $A \subset \mathbb{R}^d$ a Lebesgue measurable set. Since the Lebesgue outer measure can be computed by open covers we may find a G_δ-set A (countable intersection of covers) $E \supset A$ such that $\mathbf{m}(E) = \mathbf{m}(A)$. That implies $E \setminus A$ has null measure. In order to prove the regularity it is enough to work with Borel sets. If A were bounded the result could be deduced from Proposition 8.6.7. Otherwise, fix $\varepsilon > 0$ and take the sets $C_n = B(0, n+1) \setminus B(0, n)$. Find a compact $K_n \subset C_n \cap A$ and an open $U_n \supset C_n \cap A$ such that $\mathbf{m}(U_n \setminus K_n) < 2^{-n}\varepsilon$. Obviously $U = \bigcup_{n=1}^{\infty} U_n$ is open, and $F = \bigcup_{n=1}^{\infty} K_n$ is closed. Indeed, converging sequences stays in only one K_n. Clearly $\mathbf{m}(U \setminus F) < \varepsilon$. ∎

Now we will discuss the approximation of measurable functions by more regular ones.

Proposition 8.6.10. *Let $(X, \mathfrak{B}, \mu)$ be a normal topological space endowed with a regular Borel measure. Then the continuous functions with support of finite measure are dense in $(\mathcal{L}^1(\mu), \|\cdot\|_1)$.*

Proof. After Proposition 8.6.1 it is enough to prove the statement for characteristic functions χ_A with $A \in \mathfrak{B}$ and $\mu(A) < +\infty$. Fix $\varepsilon > 0$ and take closed an open sets $F \subset A \subset U$ such that $\mu(\overline{U}) < \infty$ and $\mu(U \setminus F) < \varepsilon$. By Urysohn's lemma there is $f : X \to [0,1]$ continuous such that $f|_F = 1$ and $f|_{U^c} = 0$. Note that the support of f has finite measure and $\|\chi_A - f\|_1 < \varepsilon$. ■

8.7 Product measures

Consider two measure spaces $(\Omega_1, \Sigma_1, \mu_1)$ and $(\Omega_2, \Sigma_2, \mu_2)$. Denote by

$$\Sigma_1 \otimes \Sigma_2 = \sigma(\{A \times B : A \in \Sigma_1, B \in \Sigma_2\}).$$

Note that the sets of the form $A \times B$ with $A \in \Sigma_1$ and $B \in \Sigma_2$ is a rectangular class and the function $\mu(A \times B) = \mu_1(A)\mu_2(B)$ is finitely additive on the algebra generated by them, just as the are for rectangles on $\mathbb{R}^2$. If we prove that μ is σ-additive within the class of "rectangles" then μ has an extension to a σ-algebra that includes $\Sigma_1 \otimes \Sigma_2$

Proposition 8.7.1. *Given measure spaces $(\Omega_1, \Sigma_1, \mu_1)$ and $(\Omega_2, \Sigma_2, \mu_2)$ there exist a measure $\mu_1 \otimes \mu_2$ on $\Sigma_1 \otimes \Sigma_2$ such that*

$$(\mu_1 \otimes \mu_2)(A \times B) = \mu_1(A)\mu_2(B).$$

If $(\Omega_1, \Sigma_1, \mu_1)$ and $(\Omega_2, \Sigma_2, \mu_2)$ are σ-finite then the measure on $\Sigma_1 \otimes \Sigma_2$ with such a property is unique.

Proof. After the preliminary discussion showing that $\Sigma_1 \otimes \Sigma_2$ is a rectagular class we only have to check the σ-additivity. Assume $R \times S = \bigcup_{n=1}^{\infty} R_n \times S_n$. Now define measurable functions on R by

$$f_n = \mu_2(S_n)\chi_{R_n}.$$

For every $x \in R$ the sets $\{S_n : x \in R_n\}$ are disjoint and their union is S. Therefore

$$\mu_2(S) = \sum_{x \in R_n} \mu_2(S_n) = \sum_{n=1}^{\infty} f_n(x)$$

and thus $\sum_{n=1}^{\infty} f_n = \mu_2(S)\chi_R$. The monotone convergence for series gives

$$\sum_{n=1}^{\infty} \mu_1(R_n)\mu_2(S_n) = \sum_{n=1}^{\infty} \int f_n \, d\mu_1 = \int \mu_2(S)\chi_R \, d\mu_1 = \mu_1(R)\mu_2(S)$$

as wished. If $(\Omega_1, \Sigma_1, \mu_1)$ and $(\Omega_2, \Sigma_2, \mu_2)$ are σ-finite then $(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2, \mu_1 \otimes \mu_2)$ also is. Assume first that $\mu_1 \otimes \mu_2$ is finite. Proposition 8.6.2 implies that the measure is determined by the value on the algebra generated by the rectangles. This can be extended to the σ-finite case in an obvious way.■

Once that the uniqueness of the product measure is guarantied under σ-finiteness we will assume that hypothesis for the rest of the discussion. In that way, the computation of product measure can be reduced to integration with respect to the factor measures (cross section integral).

Theorem 8.7.2. *Suppose that $(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2, \mu_1 \otimes \mu_2)$ is σ-finite. Take a set $A \in \Sigma_1 \otimes \Sigma_2$ and for $x \in \Omega_1$ and $y \in \Omega_2$ denote*

$$A_x = \{z \in \Omega_2 : (x, z) \in A\};$$

$$A^y = \{z \in \Omega_1 : (z, y) \in A\}.$$

Then the functions $f(x) = \mu_2(A_x)$ and $g(y) = \mu_1(A^y)$ are measurable with respect the corresponding σ-algebras and

$$\mu(A) = \int f \, d\mu_1 = \int g \, d\mu_2.$$

Proof. Consider the class $\mathcal{M} \subset \mathcal{P}(\Omega_1 \times \Omega_2)$ for which the statement of the theorem is true. Clearly, $\Sigma_1 \times \Sigma_2 \subset \mathcal{M}$ and the sets of the algebra generated by $\Sigma_1 \times \Sigma_2$ because of the reduction to disjoint unions. In order to prove that $\mathcal{M}$ actually contains $\Sigma_1 \otimes \Sigma_2$ we will use Theorem 8.2.1. Indeed, if $(A_n) \subset \mathcal{M}$ is an increasing sequence, then $f_n(x) = \mu_2((A_n)_x)$ and $g_n(y) = \mu_1((A_n)^y)$ are also increasing, so the monotone convergence applies to get that

$$\mu(\bigcup_{n=1}^{\infty} A) = \int \lim_n f_n \, d\mu_1 = \int \lim_n g_n \, d\mu_2.$$

Note that $\lim_n f(x) = \mu_2((\bigcup_{n=1}^{\infty} A)_x)$ and $\lim_n g(y) = \mu_2((\bigcup_{n=1}^{\infty} A)^y)$ and so $\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}$. The proof for decreasing sequences is similar but using dominated convergence instead if we assume that the measure is finite. Now, the σ-finite case follows straight: the intersection of $\mathcal{M}$ with every finite measure set lies on $\Sigma_1 \otimes \Sigma_2$. ■

After the result for sets we will prove the corresponding for functions. In order the result be more powerful, we will consider measurability with respect the completion of the product measure. In this way, cross section technique for integration on $\mathbb{R}^d$ will be covered by the result.

Theorem 8.7.3 (Fubini, Tonelli)**.** *Suppose that* $(\Omega_1, \Sigma_1, \mu_1)$ *and* $(\Omega_2, \Sigma_2, \mu_2)$ *are complete and* σ*-finite. Let* $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$ *measurable with respect of the completion of* $\Sigma_1 \otimes \Sigma_2$ *and assume either* f *is positive or integrable and put* $f_x(\ ) = f(x,\ )$ $f^y(\ ) = f(\ , y)$ *for* $x \in \Omega_1$ *and* $y \in \Omega_2$*. Then* $\int f_x \, d\mu_2$ *and* $\int f^y \, d\mu_1$ *exists for almost* x *and* y *(with respect to* μ_1 *and* μ_2*), they are measurable on their respective spaces and*

$$\int f \, d(\mu_1 \otimes \mu_2) = \int \left(\int f_x \, d\mu_2 \right) d\mu_1 = \int \left(\int f^y \, d\mu_1 \right) d\mu_2.$$

In order to avoid possible confusions we will write the integration variables sometimes in this way

$$\int \left(\int f(x,y) \, d\mu_2(y) \right) d\mu_1(x) \quad \text{and} \quad \int \left(\int f(x,y) \, d\mu_1(x) \right) d\mu_2(y).$$

Proof. Firstly note that the result is true for simple functions built on subsets from $\Sigma_1 \otimes \Sigma_2$ and the result extends to simple functions because the expression is linear. If f where positive and measurable with respect $\Sigma_1 \otimes \Sigma_2$ the result would be consequence of the observation and the monotone convergence theorem. Obviously the result extends to f measurable with respect $\Sigma_1 \otimes \Sigma_2$ and integrable. Now, if f is measurable with respect of the completion of $\Sigma_1 \otimes \Sigma_2$, then there is g which is $\Sigma_1 \otimes \Sigma_2$ measurable and coincides with f almost everywhere. The support of $|f - g|$ is contained in a set $N \in \Sigma_1 \otimes \Sigma_2$ of null measure. Theorem 8.7.2 implies that the set N_x has null measure for almost all $x \in \Omega_1$. Then $f(x,\ )$ is measurable for those x and coincides with $g(x,\ )$ almost everywhere. A similar reasoning works for $f(\ , y)$. ∎

We can prove a result on the derivation of parametric integrals with the help of Fubini's theorem.

Proposition 8.7.4. *Assume* (Ω, Σ, μ) *is* σ*-finite and let* $f : \Omega \times (a,b) \to \mathbb{R}$ *be measurable with respect to the product measure. Suppose moreover that* $f(x,\lambda)$ *is derivable with respect to* $\lambda \in (a,b)$ *for almost all* $x \in \Omega$ *and*

$$\int_{\Omega \times (a,b)} \left| \frac{\partial f}{\partial \lambda}(x,\lambda) \right| d\mu(x) \, d\lambda < +\infty.$$

Then, the integral is derivable with respect to λ *and the following equality holds*

$$\frac{\partial}{\partial \lambda} \int f(x,\lambda) \, d\mu = \int \frac{\partial f}{\partial \lambda}(x,\lambda) \, d\mu$$

at the points $\lambda \in (a,b)$ *where the second term is continuous.*

Proof. Put $F(\lambda) = \int f(x,\lambda)\,d\mu$. For $\lambda_1, \lambda_2 \in (a,b)$ we have

$$\int_{\lambda_1}^{\lambda_2} \int \frac{\partial f}{\partial \lambda}(x,\lambda)\,d\mu\,d\lambda = \int \left(\int_{\lambda_1}^{\lambda_2} \frac{\partial f}{\partial \lambda}(x,\lambda)\,d\lambda \right) d\mu$$

$$= \int (f(x,\lambda_2) - f(x,\lambda_1))d\mu = F(\lambda_2) - F(\lambda_1).$$

Therefore,

$$\frac{F(\lambda_2) - F(\lambda_1)}{\lambda_2 - \lambda_1} = \frac{1}{\lambda_2 - \lambda_1} \int_{\lambda_1}^{\lambda_2} \int \frac{\partial f}{\partial \lambda}(x,\lambda)\,d\mu\,d\lambda$$

and the result follows from the mean value theorem. ∎

8.8 Signed measures

Let (Ω, Σ) be a measurable space. We may consider "measures" eventually taking negative values. We say that $\nu : \Sigma \to \mathbb{R}$ is a signed measure if

$$\nu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \nu(A_n)$$

whenever $(A_n)_{n=1}^{\infty} \subset \Sigma$ are mutually disjoint. Note that any permutation of the sets in the union on the lefthand-side leaves the value unchanged so the series on the righthand-side have to be unconditionally convergent, which is the same that absolutely convergent for real numbers, namely

$$\sum_{n=1}^{\infty} |\nu(A_n)| < +\infty$$

whenever $(A_n)_{n=1}^{\infty} \subset \Sigma$ are mutually disjoint.

The first task to do with a signed measure is finding sets where the measure behaves monotonically. Let us say that $A \in \Sigma$ is *positive* if $\nu(B) \geq 0$ for any $B \in \Sigma$ with $B \subset A$. Analogously *negative* sets can be defined.

Lemma 8.8.1. *Let ν be a signed measure. Then any set $A \in \sigma$ with $\nu(A) > 0$ contains a positive set $P \in \Sigma$ with $\nu(P) \geq \nu(A)$.*

Proof. Consider

$$d_1 = \inf\{\nu(B) : B \subset A, B \in \Sigma\} \in [-\infty, +\infty)$$

If $d_1 \geq 0$ the set A is already positive and there is nothing to do. In other case $d_1 < 0$. Take a set B_1 such that $\nu(B_1) < \max\{d_1/2, -1\}$. Assume the sets $B_1, \ldots, B_{n-1}$ are already built and take

$$d_n = \inf\{\nu(B) : B \in \Sigma, B \subset A \setminus \bigcup_{k=1}^{n-1} B_k\}.$$

If this number is $d_n \geq 0$ the process stop because we have already a positive set whose complement in A has negative measure, otherwise take $B_n \in \Sigma$ such that $\nu(B_n) < \max\{d_n/2, -1\}$. Assume we have the disjoint sequence of sets (B_n) already built. Note that the sequence d_n either satisfies $d_n < -1$ or there is some n_0 such that $d_n \geq -1$ for all $n \geq n_0$. The first possibility is not possible because it it would imply

$$\nu(\bigcup_{n=1}^{\infty} B_n) = -\infty$$

that is impossible since ν takes values in $\mathbb{R}$ only. Therefore, from some n on we have $0 > d_n > 2\nu(B_n)$ which implies the convergence of the series $\sum_{n=1}^{\infty} d_n$. In particular, we have $\lim_n d_n = 0$. We claim that $P = A \setminus \bigcup_{n=1}^{\infty} B_n$ is positive. Indeed, if for $\nu(B) < 0$ some $B \subset P$, there is n such that $d_n > \nu(B)$ and this contradicts the definition of d_n. ∎

Let us say that $A \in \Sigma$ is ν-null if for every $B \in \Sigma$ with $B \subset A$, then $\nu(B) = 0$.

Theorem 8.8.2 (Hahn decomposition)**.** *Given a signed measure of bounded variation ν on a measurable space (Ω, Σ) there exists sets $P, N \in \Sigma$ such that $P \cap N = \emptyset$, $P \cup N = \Omega$, P is positive and N negative. This decomposition is unique up to ν-null sets.*

Proof. Consider

$$s = \sup\{\nu(B) : B \in \Sigma \text{ positive}\} < +\infty$$

and take $P_n \in \Sigma$ positive such that $\lim_n \nu(P_n) = s$. It is obvious that $P = \bigcup_{n=1}^{\infty} P_n$ is positive and $\nu(P) = s$. On the other hand, $N = P^c$ is negative. Otherwise, if $A \subset N$ is such that $\nu(A) > 0$, we may assume that A is positive

after the lemma. Then $P \cup A$ would be positive and $\nu(P \cup A) > s$ which violates the definition of s. Now, it is clear that if $A \subset \Sigma$ is positive then $\nu(A \setminus P) = 0$ and if A is negative then $\nu(A \setminus N) = 0$, which implies the uniqueness of the decomposition up to null measure sets. ∎

Corollary 8.8.3. *A signed measure is the difference of two positive finite measures.*

Proof. Let (P, N) be the Hahn decomposition of ν. Take $\nu^+(A) = \nu(P \cap A)$ and $\nu^-(A) = -\nu(N \cap A)$. Obviously, we have $\nu = \nu^+ - \nu^-$. ∎

Define the *variation* of ν as the finite positive measure

$$|\nu|(A) = \nu^+(A) + \nu^-(A).$$

It is not difficult to see that the variation can be recovered by this formula

$$|\nu|(A) = \sup\{\sum_{n=1}^{\infty} |\nu(A_n)| : (A_n)_{n=1}^{\infty} \text{ disjoint}, \bigcup_{n=1}^{\infty} A_n = A\}.$$

The fact that $|\nu|(\Omega) < +\infty$ is expressed usually by saying that ν has *finite variation*. The formula above is more interesting for vector *valued measures* (we will skip the definition, but the reader can easily guess it) because it allows to define a positive measure $|\nu|$ that accurately controls ν. However, in the infinitely dimensional case $|\nu|$ could be not finite.

A signed measure ν can be expressed also in form of an integral with respect a positive measure. For that, take $g = \chi_P - \chi_N$ and note that $|g| = 1$ and

$$\nu(A) = \int_A g \, d|\nu|.$$

Now we will study more general representations of measures as *indefinite integrals* with respect to a positive measure. We say that a signed measure ν is *absolutely continuous* with respect to a positive measure μ, both defined on (Ω, Σ) if $\nu(A) = 0$ whenever $A \in \Sigma$ with $\mu(A) = 0$. Note that for the same set A we have $|\nu|(A) = 0$ which implies that also $|\nu|$ is absolutely continuous with respect to μ. The name of the property refers to the following characterization quite similar to continuity of real functions.

Proposition 8.8.4. *A signed measure ν is absolutely continuous with respect to μ if and only if for every $\varepsilon > 0$ there is $\delta > 0$ such that $\mu(A) < \delta$ for $A \in \Sigma$ implies that $\nu(A) < \varepsilon$.*

Proof. Without loss of generality we may assume ν positive. One of the implications is clear. For the converse just assume the ε-δ property is false. Namely, there is some $\varepsilon > 0$ such that for all $n \in \mathbb{N}$ there is $A_n \in \Sigma$ such that $\mu(A_n) < 2^{-n}$ and $\nu(A_n) > \varepsilon$. Note now that the set

$$A = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

satisfies that $\mu(A) = 0$ and $\nu(A) > \varepsilon$ contradicting absolute continuity. ■

Note that a signed measure defined by a function $f \in \mathcal{L}^1(\mu)$ by the formula

$$\nu(A) = \int_A f \, d\mu$$

is absolutely continuous with respect to μ. The less trivial converse is true under very general hypotheses.

Theorem 8.8.5 (Radon-Nikodym)**.** *Let (Ω, Σ, μ) be a σ-finite measure space and let ν be a positive measure defined on Σ and absolutely continuous with respect to μ. Then there exists $f : \Omega \to [0, +\infty)$ measurable such that*

$$\nu(A) = \int_A f \, d\mu$$

for every $A \in \Sigma$. The function f is uniquely determined μ almost everywhere.

Proof. It is enough to prove the result for a finite measure space as the general result can be deduced by gluing the functions defined on each finite measure part. Note that if $\mu(\Omega) < +\infty$ then $\nu(\Omega) < +\infty$. Consider the set

$$\mathcal{F} = \{f \geq 0 \text{ measurable} : \int_A f \, d\mu \leq \nu(A) \text{ for all } A \in \Sigma\}.$$

It is easy to see that if $f_1, f_2 \in \mathcal{F}$ then $\max\{f_1, f_2\} \in \mathcal{F}$. Therefore it is posible to build an increasing sequence $(f_n) \subset \mathcal{F}$ such that

$$\lim_n \int f_n \, d\mu = \sup\{\int f \, d\mu : f \in \mathcal{F}\} \leq \nu(\Omega) < +\infty.$$

Take $f = \lim_n f_n$ which is integrable and so finite μ almost everywhere. Without loss of generality we may assume f takes only finite values and also $f \in \mathcal{F}$.

We claim f is the function we are looking for. Note that $\nu_0(A) = \nu(A) - \int_A f\,d\mu$ defines a positive measure which is also absolutely continuous with respect to μ. Suppose that $\nu_0(\Omega) > 0$ in order to get a contradiction. Fix $\varepsilon > 0$ such that $\nu_0(\Omega) > \varepsilon\mu(\Omega)$. Now applying the Hahn decomposition to $\nu_0 - \varepsilon\mu$ we get a positive part P with respect to such a measure. Since $(\nu_0 - \varepsilon\mu)(\Omega) > 0$ we get $(\nu_0 - \varepsilon\mu)(P) > 0$ and also

$$\nu_0(A \cap P) \geq \varepsilon\mu(A \cap P)$$

for every $A \in \Sigma$. Now we have

$$\nu(A) = \int_A f\,d\mu + \nu_0(A) \geq \int_A f\,d\mu + \nu_0(A \cap P) \geq$$

$$\int_A f\,d\mu + \varepsilon\mu(A \cap P) = \int_A (f + \varepsilon\chi_P)\,d\mu$$

for every $A \in \Sigma$, which implies $f + \varepsilon\chi_P \in \mathcal{F}$. However

$$\int f\,d\mu \geq \int (f + \varepsilon\chi_P)\,d\mu = \int f\,d\mu + \varepsilon\mu(P)$$

implies that $\mu(P) = 0$ and so $\nu_0(P) = 0$ too by absolute continuity. This contradicts the assumptions above so the theorem is proven. ∎

Corollary 8.8.6 (Signed Radon-Nikodym)**.** *Let (Ω, Σ, μ) be a σ-finite measure space and let ν be a signed measure defined on Σ and absolutely continuous with respect to μ. Then there exists $f \in \mathcal{L}^1(\mu)$ such that*

$$\nu(A) = \int_A f\,d\mu$$

for every $A \in \Sigma$. The function f is uniquely determined μ almost everywhere.

Proof. It is enough to apply Theorem 8.8.5 to the positive measures given by Hahn decomposition of ν, Theorem 8.8.2. ∎

8.9 Differentiation

We will develop the Lebesgue theory of differentiation on $\mathbb{R}^d$ endowed with the d-dimensional Lebesgue measure $\mathbf{m}$. The chosen norm on $\mathbb{R}^d$ will not play

an essential role, however it must fixed from the beginning. Instead of the difference quotients we will use integral averages

$$\mathfrak{A}_r(f)(x) = \frac{1}{\mathbf{m}(B(x,r))} \int_{B(x,r)} f \, d\mathbf{m}$$

for every $r > 0$ and $f \in \mathcal{L}^1(\mu)$. We have to prepare the tools for the main result. The first one is expresses a general property of the convolution, actually.

Proposition 8.9.1. *The average $\mathfrak{A}_r(f)$ is norm 1 operator on $f \in \mathcal{L}^1(\mu)$.*

Proof. Consider the (in)equalities

$$\begin{aligned}
\|\mathfrak{A}_r(f)\|_1 &= \int \left| \frac{1}{\mathbf{m}(B(y,r))} \int \chi_{B(y,r)}(x) f(x) \, d\mathbf{m}(x) \right| \, d\mathbf{m}(y) \\
&\leq \frac{1}{\mathbf{m}(B(0,r))} \int \int \chi_{B(y,r)}(x) |f(x)| \, d\mathbf{m}(x) \, d\mathbf{m}(y) \\
&= \frac{1}{\mathbf{m}(B(0,r))} \int \int \chi_{B(y,r)}(x) |f(x)| \, d\mathbf{m}(y) \, d\mathbf{m}(x) \\
&= \frac{1}{\mathbf{m}(B(0,r))} \int \int \chi_{B(x,r)}(y) |f(x)| \, d\mathbf{m}(y) \, d\mathbf{m}(x) \\
&= \int |f(x)| \, d\mathbf{m}(x) = \|f\|_1
\end{aligned}$$

since $\chi_{B(x,r)}(y) = \chi_{B(y,r)}(x)$ and $\int \chi_{B(x,r)}(y) d\mathbf{m} = \mathbf{m}(B(0,r))$. ∎

The previous result (and its proof) can be interpreted in terms of *convolution* with a family of kernels.

Now we will prove a version of Vitali's covering lemma.

Proposition 8.9.2. *Let $A \subset \mathbb{R}^d$ be a measurable set which is covered by a family $\mathcal{F}$ of balls with radii uniformly bounded and whose union covers A. Then there is a disjoint sequence $(B_n) \subset \mathcal{F}$ (maybe finite) such that*

$$\sum_n \mathbf{m}(B_n) \geq 5^{-d} \, \mathbf{m}(A).$$

Proof. Write $\mathcal{R}$ for the radius of a ball or the supremum of the radii of a family of balls. The choice of balls will be by induction. Take $B_1 \in \mathcal{F}$ with a $\mathcal{R}(B_1) \geq 2^{-1}\mathcal{R}(\mathcal{F})$. Suppoose now that B_k are already chosen for $k = 1, \ldots, n-1$. Find $B_n \in \mathcal{F}$ such that

$$\mathcal{R}(B_n) \geq \frac{1}{2}\mathcal{R}(\{B \in \mathcal{F} : B \cap B_k = \emptyset, \quad k = 1, \ldots, n-1\})$$

if that choice is possible, otherwise the construction stops. In order to show that the sequence satisfies the statement, we may assume $\sum_n \mathbf{m}(B_n) < \infty$. Let B_n' be a ball with the same center that B_n and radius 5 times bigger. We claim that (B_n) meets every set in $\mathcal{F}$. Indeed, take $B \in \mathcal{F}$ and assume $B_n \cap B = \emptyset$. That would imply $\mathcal{R}(B_{n+1}) > \mathcal{R}(B)/2$ and therefore the sequence is infinite and $\sum_n \mathbf{m}(B_n) = \infty$ against the previous assumption. Now, since B meets some B_n, assume n is minimum. Then we have $\mathcal{R}(B_n) > \mathcal{R}(B)/2$ and so $B \subset B_n'$ (draw a picture). In consequence, $A \subset \bigcup_n B_n'$, and thus

$$\mathbf{m}(A) \leq \sum_n \mathbf{m}(B_n') = 5^d \sum_n \mathbf{m}(B_n)$$

as desired. ∎

Given a measurable function f its *maximal function* is defined as

$$M(f)(x) = \sup_{r>0} \frac{1}{\mathbf{m}(B(x,r))} \int_{B(x,r)} |f| \, d\mathbf{m}$$

Proposition 8.9.3. *Assume $f \in \mathcal{L}^1(\mathbb{R}^d)$, then for every $\varepsilon > 0$*

$$\mathbf{m}(\{M(f) > \varepsilon\}) \leq \frac{5^d}{\varepsilon} \|f\|_1$$

Proof. Fix ε and put $A = \{M(f) > \varepsilon\}$. By the definition of the maximal function, for every $x \in A$ there is a ball B_x centred at x such that

$$\int_{B_x} |f| \, d\mathbf{m} > \varepsilon \, \mathbf{m}(B_x)$$

Clearly the radii of the balls are uniformly bounded. Using Vitali's lemma there is disjoint sequence (maybe finite) in $\{B_x : x \in A\}$ that we denote (B_n). We have

$$\int_{\bigcup_n B_n} |f| \, d\mathbf{m} > \varepsilon \sum_n \mathbf{m}(B_n) \geq \varepsilon \, 5^{-d} m(A)$$

which implies the statement. ∎

Theorem 8.9.4. *Let $f \in \mathcal{L}^1(\mathbb{R}^d)$ then for almost every point $x \in \mathbb{R}^d$ there exists the limit*

$$\lim_{r\to 0^+} \frac{1}{\mathbf{m}(B(x,r))} \int_{B(x,r)} f \, d\mathbf{m} = f(x).$$

Proof. Firstly we will show that the convergence of the averages happens with respect to $\|\cdot\|_1$, namely

$$\lim_{r\to 0^+} \|\mathfrak{A}_r(f) - f\|_1 = 0.$$

Indeed, if f where continuos with compact support then $\mathfrak{A}_r(f)$ would converge uniformly to f. Since $\mathfrak{A}_r(f)$ has support contained into $\sup(f) + B(0,r)$, that the convergence is also with respect to $\|\cdot\|_1$. Now take $\varepsilon > 0$ and $f \in \mathcal{L}^1(\mathbb{R}^d)$. Take g continuous with compact support such that $\|f - g\|_1 < \varepsilon/3$. That also implies $\|\mathfrak{A}_r(f) - \mathfrak{A}_r(g)\|_1 < \varepsilon/3$. Now, if $r > 0$ is small enough to have $\|\mathfrak{A}_r(g) - g\|_1 < \varepsilon/3$ then

$$\|\mathfrak{A}_r(f) - f\|_1 \leq \|\mathfrak{A}_r(f) - \mathfrak{A}_r(g)\|_1 + \|\mathfrak{A}_r(g) - g\|_1 + \|g - f\|_1 < \varepsilon.$$

Define the mean oscillation

$$\operatorname{osc}(f,x) = \limsup_{r\to 0^+} \frac{1}{\mathbf{m}(B(x,r))} \int_{B(x,r)} f \, d\mathbf{m} - \liminf_{r\to 0^+} \frac{1}{\mathbf{m}(B(x,r))} \int_{B(x,r)} f \, d\mathbf{m}.$$

Note that $\operatorname{osc}(f,x) \leq 2\, M(f)(x)$ and the desired convergence at x is equivalent to $\operatorname{osc}(f,x) = 0$. If f were continuous the result would be trivial, as we have already seen before. Take $\varepsilon > 0$. In general, for any $f \in \mathcal{L}^1(\mathbb{R}^d)$ we can find continuous (compactly supported) functions g such that $\|f - g\|_1$ is arbitrarily small. Note that $\operatorname{osc}(f - g, x) = \operatorname{osc}(f, x)$ and so

$$\mathbf{m}(\{\operatorname{osc}(f,x) > \varepsilon/2\}) \leq \frac{5^d}{\varepsilon} \|f - g\|_1$$

that implies $\mathbf{m}(\{\operatorname{osc}(f,x) > \varepsilon/2\}) = 0$. Therefore the integral averages of f converge almost everywhere and the limit must coincide with f almost everywhere by the first part of the proof and Theorem 8.6.5. ∎

Is clear that Theorem 8.9.4 can be extended to functions that are integrable on a bounded open subset of $\mathbb{R}^d$ (locally integrable). In this way the result naturally applies to characteristic functions.

Corollary 8.9.5. *Given a measurable set $A \subset \mathbb{R}^d$ the limit*

$$\lim_{r\to 0^+} \frac{\mathbf{m}(B(x,r)\cap A)}{\mathbf{m}(B(x,r))}$$

exists for almost all $x \in \mathbb{R}^d$ with values 0 or 1.

Now we will discuss the derivation of real functions of one variable from the point of view given by measure theory. Firstly note that we should change the previous averages by these more convenient ones

$$\lim_{r\to 0^+} \frac{1}{r}\int_{x-r}^{x} f(t)\,dt \qquad \text{and} \qquad \lim_{r\to 0^+} \frac{1}{r}\int_{x}^{x+r} f(t)\,dt.$$

It is not difficult to check that all the previous theory can can be adapted to these averages despite x is not the central point. As a consequence we have the following.

Corollary 8.9.6. *The indefinite integral $F(x) = \int_a^x f(t)\,dt$ of a locally integrable function on $\mathbb{R}$ is differentiable almost everywhere and the equality $F'(x) = f(x)$ holds almost everywhere on its domain.*

A related important question is to recognise those functions which are indefinite integrals of $\mathcal{L}^1(\mathbb{R})$ functions. The key idea is the fact that the measure $\mu(A) = \int_A f\,dm$ is absolutely continuous with respect the Lebesgue measure on $\mathbb{R}$. We say that a function F defined on an interval (a,b) of $\overline{\mathbb{R}}$ is *absolutely continuous* if for every $\varepsilon > 0$ there is $\delta > 0$ such that for any choice of points

$$a < a_1 < b_1 < a_2 < b_2 < \cdots < a_n < b_n < b$$

with $\sum_{k=1}^n (b_k - a_k) < \delta$ then $\sum_{k=1}^n |F(b_k) - F(a_k)| < \varepsilon$. Evidently, an absolutely continuous function is continuous and also it is of *bounded variation* on bounded intervals.

Theorem 8.9.7. *Let $F : (a,b) \to \mathbb{R}$ be an absolutely continuos function. Then F is differentiable almost everywhere and its derivative $F'(x) = f(x)$ is locally integrable and satisfies*

$$\int_c^d f(x)\,dx = F(d) - F(c)$$

for every interval $[c,d] \subset (a,b)$.

Proof. We may assume that $[a, b]$ is finite and the butts belong to the domain. The members of the algebra $\mathcal{A}$ generated by the intervals can be represented as a disjoint finite union of intervals. We define a function on $\mathcal{A}$ by

$$\nu(A) = \sum_{k=1}^{n} (F(b_k) - F(a_k))$$

where the intervals $|a_k, b_k|$ made up the decomposition of A. As a function it is uniformly continuous with respect to the pseudometric $d(A, B) = \mathbf{m}(A \Delta B)$, Proposition 8.6.2. Since $\mathcal{A}$ is dense with respect to d in the measurable subsets of (a, b), ν extends as a uniformly continuous function to all those sets. It is not difficult to check that ν is a signed σ-additive measure of bounded variation and absolutely continuous with respect to $\mathbf{m}$. By the Radon-Nikodym Theorem 8.8.5 there is $f \in \mathcal{L}^1(a, b)$ such that $\nu(A) = \int_A f \, dm$. Therefore

$$F(x) = F(a) + \int_a^x f \, d\mathbf{m}.$$

By the previous Corollary F is differentiable almost everywhere and the equality $F'(x) = f(x)$ holds almost everywhere. ∎

8.10 Rationale and remarks

This chapter is noticeably longer than others for the sake of its autonomy and it could be used in an advanced course together Appendix B. Nevertheless, the choice of topics is free and may depend on the time employed for Riemann integral. In case one decides to withdraw Riemann integral from the course, the chapter on Change of Variables should be adapted for the Lebesgue integral. My personal choice is to keep Riemann integral (actually, the integration of continuous functions with bounded support) in order to regard Lebesgue integral as an extended linear operator.

In the classroom, I like to begin the topic with the "proof" of Pythagoras Theorem based in different decompositions of a square. In this way, I discuss in what extent we are using an intuitive notion of area and how we could calm the need for rigor. We can devise a "naïve" theory of area for polygons based in *equidecomposability* and then I would mention the theorems of Bolyai-Gerwien and Hadwiger-Glur. However, the same program cannot be developed in three dimensions because of the solution of Dehn to Hilbert's third problem. That

shows that integral methods are needed even if we restrict ourselves to volumes of polyhedral bodies.

We arrive just to the differentiation results, for measures and integrals. Going on will requiere special properties of $\mathcal{L}^1(\mu)$, namely *completeness* with respect to (semi)norm, which is traditionally considered Functional Analysis. That material can be found in the appendix. We also managed to skip, explicitly, the *convolution*. Some topics can be properly treated in Advanced Analysis topics.

As to the methods to carry out the topic, I want to point out the struggle to make clear the need for countable additive measures and the meaning of Caratheodoy's measurability definition. Our proof of the Radon-Nikodym theorem is constructive, instead of von Neumann's idea using Hilbert properties of L^2. In our vision, some specific properties of function spaces belong to Functional Analysis, so they are relegated to the auxiliary chapter.

8.11 Exercises

1. Use the formula for the measure of a union of sets to deduce the area of a spherical triangle in terms of its angles.

2. Let $\mathcal{S}_n$ be the permutation group action on a set of n elements. Let F_n be the subset of $\mathcal{S}_n$ that fixes at least one element. Show the existence and find the value of the limit

$$\lim_n \frac{\#(F_n)}{n!}.$$

3. Find the values of α for which

$$\lim_n \int_0^1 n^\alpha (1-x)x^n \cos(\pi x/n)\, dx = 0.$$

4. Find the domain of the function

$$f(x) = \int_0^{\pi/2} t^x \sin t\, dt.$$

   Show that $f(x)$ is convex and it attains an absolute minimum between 0 and 1.

5. Let $f(x) = \int_0^1 \frac{\log(1+xt)}{1+t^2} dt$ defined for $x > 0$. Show that $f(1) = \pi \log(2)/8$ and
$$f'(x) = \frac{\log 4 + \pi x - 4\log(1+x)}{4(1+x^2)}.$$

6. Prove that the function
$$f(x) = \int_0^{+\infty} \left(x\sqrt{t} + \frac{\cos(xt)}{\sqrt{t}} \right) e^{-t} dt$$
is defined on $\mathbb{R}$, it continuous and monotone.

7. Let $f : \mathbb{R} \to \mathbb{R}$ be integrable, and for $n \in \mathbb{N}$ take
$$f_n(x) = \frac{n}{\pi} \int_{-\infty}^{+\infty} \frac{f(t)\, dt}{1 + n^2(t-x)^2}.$$
Prove that the integral is finite and if f is continuous at some $x \in \mathbb{R}$ then $\lim_n f_n(x) = f(x)$.

8. Let $f : \mathbb{R}^d \to \mathbb{R}$ be an integrable function.
   (a) Show that $\lim_{t\to+\infty} tm(\{x \in \mathbb{R}^d : |f(x)| > t\}) = 0$.
   (b) If $K \subset \mathbb{R}^d$ is compact, show that
$$\lim_{\|x\|\to+\infty} \int_{x+K} |f(t)|\, dt = 0.$$

9. Prove that a closed interval cannot be written as a countable disjoint union of smaller closed intervals.

10. Let $\mathcal{F} \subset \mathcal{P}(\Omega)$ be a family of sets. Prove that for every $A \in \sigma(\mathcal{F})$ there is $\mathcal{F}_A \subset \mathcal{F}$ countable such that $A \in \sigma(\mathcal{F}_A)$.

11. Prove that the cardinality of the Borel sets of $\mathbb{R}$ is $\mathfrak{c} = 2^{\mathbb{N}}$ and the cardinality of Lebesgue measurable sets of $\mathbb{R}$ es $2^{\mathfrak{c}}$.

12. Find a non measurable Lebesgue set (use an equivalent version of the Axiom of Choice).

13. Prove that there is no probability μ on $\mathcal{P}(\mathbb{N})$ such that $\mu(n\mathbb{N}) = 1/n$ for all $n \in \mathbb{N}$.

14. For any set $A \subset \mathbb{R}$, let $\mathcal{D}(A) = \{x-y : x, y \in A\}$. Prove that if $m(A) > 0$ then $\mathcal{D}(A)$ is a neighbourhood of 0. Show that the reciproque does not hold by computing $\mathcal{D}(T)$ where T is the ternary Cantor set.

15. Let $f : \mathbb{R} \to \mathbb{R}$ a *first Baire class function*, that is, a pointwise limit of continuous functions. Show that $f^{-1}(U)$ is a countable union of closed sets for every open $U \subset \mathbb{R}$. Deduce that the indicator fuction of the rationals is not first Baire class, but it is *second Baire class* (pointwise limit of first Baire class functions).

16. Let $\langle x \rangle$ denote the not-integer part of a number $x \in \mathbb{R}$. Prove that for all $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ and every $f \in C[0,1]$ then

$$\lim_n \frac{1}{n} \sum_{k=1}^{n} f(\langle k\alpha \rangle) = \int_0^1 f(x)\, dx.$$

17. Let $(f_n) \subset C[a,b]$ a bounded sequence of continuous function pointwise converging to 0. Show that $\lim_n \int_a^b f_n = 0$ without invoking the dominated convergence theorem.

18. Let $g : \mathbb{R} \to \mathbb{R}$ be an increasing left-continuous function. Prove that there is a unique Borel measure μ such that $\mu([a,b)) = g(b) - g(a)$ for every $a < b$. The integration with respect to μ is called *Lebesgue–Stieltjes*, because it extends the older Riemann–Stieltjes integral, and it is denoted

$$\int_a^b f\, dg.$$

19. Let $f, g : [0, +\infty) \to \mathbb{R}$ be functions such that f is decreasing with $\lim_{x\to+\infty} f(x) = 0$ and there is $M > 0$ such that $|\int_x^y g(t)dt \leq M|$ for every $x, y \in [0, +\infty)$. Show that if $x < y \in [0, +\infty)$, then

$$\left| \int_x^y f(t)g(t)\, dt \right| \leq M\, f(x).$$

20. Show that the product of two absolutely continuous functions (defined on the same compact interval) is absolutely continuous too.

Chapter 9

Integration on curves and surfaces

9.1 Functions of bounded variation

We will discuss the notion of length of a curve in a normed space, but for that aim is better to start by a related notion in the setting of real valued functions.

Given a real function $f : [a, b] \to \mathbb{R}$ we may consider the following number, eventually $+\infty$,

$$V_a^b(f) = \sup\{\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| : (x_i)_{i=0}^n \text{ partition of } [a, b]\}$$

called the *variation of* f on $[a, b]$. If $V_a^b(f) < +\infty$ we say that f is of *bounded variation*. Note that monotone functions are trivially of bounded variation. A less trivial example: the existence and boundedness of the derivative implies bounded variation since

$$\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| = \sum_{i=1}^{n} |f'(\xi_i)|(x_i - x_{i-1}) \leq M\,(b - a)$$

for some $\xi_i \in (x_{i-1}, x_i)$ and $M > 0$ being a bound for $f'(x)$ on (a, b).

Elementary properties of variation are summarised here.

Proposition 9.1.1. *Let $f : [a, b] \to \mathbb{R}$ be a function, then:*

(a) $|f(b) - f(a)| \leq V_a^b(f)$;

(b) If $[c,d] \subset [a,b]$ *then* $V_c^d(f) \leq V_a^b(f)$;

(c) if $c \in [a,b]$ *then* $V_a^c(f) + V_c^b(f) = V_a^b(f)$.

Proof. (a) Note that $|f(b)-f(a)|$ is the sum associated to the trivial partition of $[a,b]$. Statements (b) and (c) follow just taking suitable partitions of $[a,b]$ including points a, b. ∎

The following summarises some properties of the variation when it is finite.

Proposition 9.1.2. *Let* $f, g : [a,b] \to \mathbb{R}$ *be functions of bounded variation, then:*

(a) $V_a^b(f+g) \leq V_a^b(f) + V_a^b(g)$, $V_a^b(\lambda f) = |\lambda| V_a^b(f)$ *for* $\lambda \in \mathbb{R}$;

(b) $V_a^x(f)$ *is increasing for* $x \in [a,b]$;

(c) $V_a^x(f) - f(x)$ *is increasing for* $x \in [a,b]$.

Proof. (a) It is just the triangle property and homogeneity of the absolute value. (b) It is consequence of (c) of previous proposition. In order to prove (c), for $x \leq y$ we have

$$f(y) - f(x) \leq V_x^y(f) = V_a^y(f) - V_a^x(f)$$

and thus

$$V_a^x(f) - f(x) \leq V_a^y(f) - f(y)$$

which finishes the proof. ∎

We are ready for the main characterization.

Theorem 9.1.3. *A function* $f : [a,b] \to \mathbb{R}$ *is of bounded variation if and only if it is the difference of two monotone functions.*

Proof. If f is of bounded variation, by the previous proposition we have

$$f(x) = V_a^x(f) - (V_a^x(f) - f(x))$$

which a representation of f as a difference of two increasing functions. On the other hand, monotone functions are of bounded variation and this property is stable by sums. ∎

Corollary 9.1.4. *A function of bounded variation has at most countably many discontinuities of jump type.*

Now we are interested in knowing if continuity of the function is inherited by its variation.

Theorem 9.1.5. *If $f : [a, b] \to \mathbb{R}$ is continuous and of bounded variation then $V_a^x(f)$ is continuous for $x \in [a, b]$.*

Proof. Suppose that $V_a^x(f) \not\to V_a^c(f)$ for some $c \in [a, b]$. We may assume that $c > a$ and $x < c$, otherwise we could made a similar argument. Therefore there is some $\eta > 0$ such that $V_x^c(f) > \eta$ for every $x < c$. Take $a_1 = a$ and find a partition $(x_i)_{i=1}^n$ of $[a_1, c]$ such that

$$\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| > \eta.$$

By the continuity of f we may assume that $|f(x_{n-1}) - f(c)| < \eta/2$. Now take $a_2 = x_{n-1} < c$ and observe that $V_{a_1}^{a_2} > \eta/2$. Proceed likewise to find a_3 and so an increasing sequence (a_n) such that $V_{a_n}^{a_{n+1}}(f) > \eta/2$. Thus

$$V_{a_1}^{a_n}(f) = V_{a_1}^{a_2} + \cdots + V_{a_{n-1}}^{a_n}(f) \geq (n-1)\eta/2.$$

That clearly violates the boundedness of the variation of f. ∎

Corollary 9.1.6. *A continuous function of bounded variation is the difference of two continuous monotone functions.*

9.2 Curves in normed spaces

Consider a parameterized (non necessarily continuous) curve $\gamma : [a, b] \to X$ where X is a normed space. We say that the curve is *rectifiable* if

$$L_a^b(\gamma) = \sup\{\sum_{i=1}^{n} \|\gamma(t_i) - \gamma(t_{i-1})\| : (t_i)_{i=0}^n \text{ partition of } [a, b]\} < +\infty.$$

It is easy to check that a rectifiable curve is rectifiable for any equivalent norm on X unless the length is obviously not invariant.

Note the similarities of the length with the variation. Some of the arguments in the preceding section can be adapted to prove the following properties.

Proposition 9.2.1. *Let $\gamma : [a,b] \to X$ be a rectifiable continuous curve. Then*

(a) $\|\gamma(b) - \gamma(a)\| \leq L_a^b(f)$;

(b) if $c \in [a,b]$ *then* $L_a^c(f) + L_c^b(f) = L_a^b(f)$;

(c) $L_a^t(\gamma)$ *is increasing for* $t \in [a,b]$;

(d) $L_a^t(\gamma)$ *is continuous for* $t \in [a,b]$.

In the following, we may assume that parameterized curves are always continuous. We will begin with the characterization in finite dimensional spaces.

Theorem 9.2.2. *A curve $\gamma : [a,b] \to \mathbb{R}^d$ is rectifiable if and only if its coordinatewise functions are of bounded variation.*

Proof. Being rectifiable is independent of the norm on $\mathbb{R}^d$ since all the norms are equivalent. We will use the $\|\cdot\|_1$ norm to prove the equivalence. Write $\gamma(t) = (x_1(t), \ldots, x_d(t))$ and observe that

$$\sum_{i=1}^{n} |x_k(t_i) - x_k(t_{i-1})| \leq \sum_{i=1}^{n} \|\gamma(t_i) - \gamma(t_{i-1})\|_1 = \sum_{j=1}^{d} \sum_{i=1}^{n} |x_j(t_i) - x_j(t_{i-1})|$$

where $(t_i)_{i=0}^n$ is a partition of $[a,b]$ and $k = 1, \ldots, d$. The first inequality implies that x_k is of bounded variation when γ is rectifiable. The equality on the right hand side implies that γ is rectifiable if all the functions x_j for $j = 1, \ldots, d$ are of bounded variation. ∎

Using a deep result saying that monotone functions have derivative almost everywhere we can deduce the following corollary.

Corollary 9.2.3. *A rectifiable curve $\gamma : [a,b] \to \mathbb{R}^d$ has a tangent line at $\gamma(t)$ for almost every $t \in [a,b]$.*

However, this derivative is of little use as it doesn't show the global behaviour of of the curve unless we assume extra regularity. Indeed, remind that there exist non trivial monotone functions with null derivative at almost every point (Cantor's staircase e.g.).

We will fix the standard of regularity in order to get profit of the derivative of the curve. We will say that a curve $\gamma : [a,b] \to X$ is C^1 (please, remark: on $[a,b]$) if it has derivative at every point of $[a,b]$ including the endpoints with

side derivatives and the derivative is continuous on $[a,b]$. It is no difficult to prove that this is equivalent to say that there exists a C^1 extension of γ to an open interval containing $[a,b]$. We will say that a curve $\gamma : [a,b] \to X$ is *piecewise* C^1 if it continuous and there exists a finite partition of $[a,b]$ such that γ is C^1 on every subinterval of the partition.

Theorem 9.2.4. *Let $\gamma : [a,b] \to X$ be a piecewise C^1 curve. Then γ is rectifiable and*

$$L_a^b(\gamma) = \int_a^b \|\gamma'(t)\|\, dt.$$

Proof. Firstly we will assume that γ is C^1 on $[a,b]$, which implies the uniform continuity of $\gamma'(t)$ on $[a,b]$. Given $\varepsilon > 0$ find $\delta > 0$ such that $|t-\xi| < \delta$ implies $\|\gamma'(t) - \gamma'(\xi)\| < \varepsilon$. Take a partition $(t_i)_{i=0}^n$ of $[a,b]$ such that $|t_i - t_{i-1}| < \delta$. Using the mean value theorem on the auxiliary function

$$f(t) = \gamma(t) - \gamma(t_{i-1}) - \gamma'(\xi_i)(t - t_{i-1})$$

with $\xi_i \in [t_{i-1}, t_i]$ we get that

$$\|\gamma(t_i) - \gamma(t_{i-1}) - \gamma'(\xi_i)(t_i - t_{i-1})\| = \|f(t_i) - f(t_{i-1})\|$$

$$\leq \sup\{\|f'(t)\| : t \in [t_i, t_{i-1}]\}(t_i - t_{i-1}) \leq \varepsilon(t_i - t_{i-1}).$$

Therefore

$$|\|\gamma(t_i) - \gamma(t_{i-1})\| - \|\gamma'(\xi_i)\|(t_i - t_{i-1})| \leq \varepsilon(t_i - t_{i-1}).$$

Using that on any interval of the partition we have

$$|\sum_{i=0}^n \|\gamma(t_i) - \gamma(t_{i-1})\| - \sum_{i=0}^n \|\gamma'(\xi_i)\|(t_i - t_{i-1})| \leq \sum_{i=0}^n \varepsilon(t_i - t_{i-1}) = \varepsilon(b-a)$$

Since we could take partitions such that the first term approaches the length and the second one the Riemann integral, taking limits we get

$$|L_a^b(\gamma) - \int_a^b \|\gamma'(t)\|\, dt| \leq \varepsilon(b-a)$$

following that $L_a^b(\gamma) < +\infty$. Now, as $\varepsilon > 0$ was arbitrary we get the equality between both numbers. Finally, the general case with γ being piecewise C^1 reduces to the last equality by the additivity of the length and the integral with respect to intervals. ∎

9.3 Some formulas

Despite the generality of the results of the previous section, the Euclidean norm still plays a fundamental role. If the curve γ is parameterized as $(x(t), y(t), z(t))$ for $t \in [a, b]$, the length is given by

$$L_a^b(\gamma) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\, dt.$$

The important case of the graph $y = f(x)$ of a function we have for the length for $x \in [a, b]$

$$L_a^b = \int_a^b \sqrt{1 + f'(x)^2}\, dx.$$

However, a plane curve could be given in polar form $r = \phi(\theta)$ with $\theta \in [\alpha, \beta]$. We have (locally)

$$x(\theta) = \phi(\theta)\cos\theta,$$
$$y(\theta) = \phi(\theta)\sin\theta.$$

The derivation gives

$$x'(\theta) = -\phi(\theta)\sin\theta + \phi'(\theta)\cos\theta,$$
$$y'(\theta) = \quad \phi(\theta)\cos\theta + \phi'(\theta)\sin\theta.$$

A simple calculus shows that

$$x'(\theta)^2 + y'(\theta)^2 = \phi(\theta)^2 + \phi'(\theta)^2.$$

Therefore, the length in this case is

$$L_\alpha^\beta = \int_\alpha^\beta \sqrt{\phi(\theta)^2 + \phi'(\theta)^2}\, d\theta.$$

9.4 Integration with respect to the arc length

Suppose that $\gamma : [a, b] \to X$ is rectifiable and f is a function defined on $\gamma([a, b])$. Sometimes is necessary to consider the convergence of Riemann-like sums of the form

$$\sum_{i=0}^{n} f(\gamma(\xi_i)) L_{t_{i-1}}^{t_i}(\gamma)$$

for a partition $(t_i)_{i=0}^n$ of $[a,b]$ and $\xi_i \in [t_{i-1}, t_i]$. For instance, the mass of a curve in terms of its linear density can be obtained this way. The reader acquainted with the Riemann-Stieltjes integral can see that the integral could be expressed as

$$\int_a^b f(\gamma(t))\, dL_a^t(\gamma)$$

That implies, in particular, that the existence is guarantied for f continuous. A direct proof can be obtained by just mimicking the proof of Riemann integrability of continuous functions. We will follow the following notation

$$\int_\gamma f\, d\ell = \lim \sum_{i=0}^n f(\gamma(\xi_i)) L_{t_{i-1}}^{t_i}(\gamma)$$

and $d\ell$ is called "arc element". It worth noticing that the same limit, when existing, can be obtained by the Riemann-like sums

$$\sum_{i=0}^n f(\Xi_i)\|\gamma(t_i) - \gamma(t_{i-1})\|$$

where the point Ξ_i can be taken from $\gamma([t_{t-1}, t_i])$ or from $[\gamma(t_{i-1}), \gamma(t_i)]$ in case of f is defined on a neighbourhood of $\gamma([a,b])$ where it is uniformly continuous.

We will prove a formula to compute that integral in case of regularity of γ.

Proposition 9.4.1. *Let $\gamma : [a,b] \to X$ be a piecewise C^1 curve and $f : \gamma([a,b]) \to \mathbb{R}$ then*

$$\int_\gamma f\, d\ell = \int_a^b f(\gamma(t))\, \|\gamma'(t)\|\, dt$$

Proof. Take a partition $(t_i)_{i=0}^n$ of $[a,b]$ and find $\xi_i \in [t_{i-1}, t_i]$ such that

$$L_{t_{i-1}}^{t_i}(\gamma) = \|\gamma'(\xi_i)\|(t_{i-1} - t_i)$$

and so

$$\sum_{i=0}^n f(\gamma(\xi_i)) L_{t_{i-1}}^{t_i}(\gamma) = \sum_{i=0}^n f(\gamma(\xi_i))\|\gamma'(\xi_i)\|(t_{i-1} - t_i)$$

Taking limits with respect to the partitions we will get the desired identity. ■

For the following we will restrict ourselves to the Euclidean norm on $\mathbb{R}^d$ since the scalar product is involved. A very alike notion appears when we wish formalise the path integral used to compute the work done by a force. Suppose that $f : \gamma([a,b]) \to \mathbb{R}^d$ is continuous and consider the convergence of Riemann-like sums of the form

$$\sum_{i=0}^{n} f(\gamma(\xi_i)) \cdot (\gamma(t_{i-1}) - \gamma(t_i))$$

where "$\cdot$" is the scalar product, $(t_i)_{i=0}^n$ a partition of $[a,b]$ and $\xi_i \in [t_{i-1}, t_i]$. Again, the existence of this limit called the *line integral* can be proved by standard methods and its value is denoted by

$$\int_\gamma \overrightarrow{f} \cdot d\overrightarrow{\ell}$$

(the notation with $d\overrightarrow{s}$ is also popular but we will try to avoid when it could lead to confusion). Note that we could work in a Hilbert space instead of $\mathbb{R}^d$ because of the properties of the scalar product, or more generally, we could assume that f takes values in X^*, so we may consider sums of terms $f(\gamma(\xi_i))(\gamma(t_i) - \gamma(t_{i-1}))$ which is actually what appear in the theory of integration of differential forms.

In case of regularity we have an explicit formula for the line integral. We will state it only in the case of $\mathbb{R}^d$ with the scalar product.

Proposition 9.4.2. *Let $\gamma : [a,b] \to \mathbb{R}^d$ be a piecewise C^1 curve and $f : \gamma([a,b]) \to \mathbb{R}^n$ then*

$$\int_\gamma \overrightarrow{f} \cdot d\overrightarrow{\ell} := \int_a^b f(\gamma(t)) \cdot \gamma'(t)\, dt.$$

Proof. Given $\varepsilon > 0$ we have establish in the proof of Theorem 9.2.4 that

$$\|\gamma(t_{i-1}) - \gamma(t_i) - \gamma'(\xi_i)(t_i - t_{i-1})\| < \varepsilon(t_i - t_{i-1})$$

for a fine enough partition, and so

$$|f(\gamma(\xi_i)) \cdot (\gamma(t_{i-1}) - \gamma(t_i)) - f(\gamma(\xi_i)) \cdot \gamma'(\xi_i)(t_i - t_{i-1})| < \varepsilon M(t_i - t_{i-1})$$

where $M > 0$ is an upper bound for f. Summing all the terms we have

$$\Big| \sum_{i=1}^{n} f(\gamma(\xi_i)) \cdot (\gamma(t_{i-1}) - \gamma(t_i)) - \sum_{i=0}^{n} f(\gamma(\xi_i)) \cdot \gamma'(\xi_i)(t_i - t_{i-1}) \Big| < \varepsilon M.$$

Clearly, that implies the desired result as $\varepsilon > 0$ is arbitrary. ∎

9.5 Alternative parameterizations

A change of parameterization of a curve $\gamma : [a, b] \to X$ can be easily done just taking an increasing onto function $h : [c, d] \to [a, b]$ and considering $\gamma \circ h$. The regularity of the new parameterization depends on que quality of both γ and h. We will try to solve the inverse problem: assume that we have two parameterizations giving the same curve (image and orientation), are these two parameterizations linked by a regular change of variables?

We will show that the arc length provides a natural parameterization of curves. Note that $L_a^t(\gamma)$ is an increasing continuous function defined on $[a, b]$, and take

$$\tau(s) = \inf\{t : L_a^t(\gamma) = s\}$$

which is actually it is a minimum and therefore $L_a^{\tau(s)}(\gamma) = s$ for every $s \in [0, L_a^b(\gamma)]$. Note as well that τ is strictly increasing and discontinuous, however, the composition $\tilde{\gamma}(s) = \gamma(\tau(s))$ is continuous with respect to $s \in [0, L_a^b(\gamma)]$. Indeed, it is 1-Lipschitz

$$\|\tilde{\gamma}(s) - \tilde{\gamma}(s_0)\| \leq L_{\tau(s)}^{\tau(s_0)}(\gamma) = |L_a^{\tau(s)}(\gamma) - L_a^{\tau(s_0)}(\gamma)| = |s - s_0|.$$

This is the parameterization of γ with respect to the arc length, which is usually denoted by the choice of the letter s as a variable.

The regularity of this parameterization depends on the regularity of γ, and so from a previous parameterization. If γ is C^1 then $dL_a^t(\gamma)/dt = \|\gamma'(t)\|$ by the length formula, and thus a continuous function. Note that $\tau(s)$ is differentiable at s whenever $\gamma'(\tau(s)) \neq 0$ by the rule of the derivative of the inverse function. But if we want to guarantee that τ is C^1 we have to keep $\gamma'(t)$ away from zero.

Proposition 9.5.1. *Let $\gamma : [a, b] \to X$ be a C^1 curve such that $\gamma'(t) \neq 0$ for all $t \in [a, b]$. Then its re-parameterization $\tilde{\gamma}(s)$ with respect to the arc length is C^1 too and $\|\tilde{\gamma}'(s)\| = 1$ for all $s \in [0, L_a^b(\gamma)]$.*

Proof. Under the hypotheses $L_a^t(\gamma)$ is strictly increasing and thus τ is an actual inverse function whose derivative $\tau'(s) = 1/\|\gamma'(\tau(s))\|$ which is continuous too. Moreover

$$\tilde{\gamma}'(s) = \gamma'(\tau(s))\tau'(s)$$

which has norm one by the previous formula. ∎

The answer to the question of the beginning of the section turns out as a corollary.

Corollary 9.5.2. *If $\gamma_1 : [a, b] \to X$ and $\gamma_2 : [c, d] \to X$ are two C^1 parameterizations of the same curve (image and orientation) whose derivatives do not vanish, then there exists a C^1 increasing bijection $h : [a, b] \to [c, d]$ such that $\gamma_1 = \gamma_2 \circ h$.*

Proof. We have two re-parameterizations by the arc length $\tilde{\gamma}_1 = \gamma_1 \circ \tau_1$ and $\tilde{\gamma}_2 = \gamma_2 \circ \tau_2$. Note that we must have $\tilde{\gamma}_1 = \tilde{\gamma}_2$ and therefore $\gamma_1 = \gamma_2 \circ \tau_2 \circ \tau_1^{-1}$. ■

9.6 Another way to compute the length

In order to discuss the notion of area of a surface we will assume that the "ambient space" $\mathbb{R}^3$ is equipped with the Euclidean norm. Even in that intuitive case, the area of a surface cannot be defined in the same way that we defined the length of a curve. In the case of curves, the length is the limit of the lengths of the polygonal lines obtained from the nodes placed following partitions. In this case, the segments the polygonal is made of approach tangent lines of the curve. However, the faces of the polyhedric surfaces built likewise may not approach the tangent plane, even for surfaces as simple as the cylinder. For that reason we will propose an alternative method to compute a curve's length that can be adapted to define a notion of area for surfaces. This will be done in $\mathbb{R}^2$ in order to keep a greater analogy for the next section.

Let $\gamma(s)$ with $t \in [0, L]$ a C^2 curve in $\mathbb{R}^2$ already parameterized by the arc length (so its length is L). We claim that

$$L = \lim_{\varepsilon \to 0^+} \frac{\text{Area}(\gamma([0, L]) + B[0, \varepsilon])}{2\varepsilon}$$

This formula means that we can recover the length of γ from its image with the help of 2-dimensional Lebesgue measure (please, make a picture). The idea of the proof we are going to sketch is to compute the measure of the set $\gamma([0, L]) + B[0, \varepsilon])$ with the help of a parameterization. As $\gamma(s) = (x(s), y(s))$ is parameterized by the length the vector $\gamma'(s)$ has norm one and so the normal vector $n(s) = (y'(s), -x'(s))$. Define a 2 variables function by

$$F(s, t) = \gamma(s) + tn(s) = (x(s) + ty'(s), y(s) - tx'(s)).$$

Now, if $\varepsilon > 0$ is small enough then F is injective when defined on $[0, L] \times [-\varepsilon, \varepsilon]$ and its image covers $\gamma([0, L]) + B[0, \varepsilon])$ except the butts which are covered by two semicircles of radius ε so

$$\text{Area}(\gamma([0, L]) + B[0, \varepsilon]) = \text{Area}(F([0, L] \times [-\varepsilon, \varepsilon])) + \pi\varepsilon^2$$

The area of $F([0, L] \times [-\varepsilon, \varepsilon])$ can be computed by integrating the absolute value of the Jacobian of F over $[0, L] \times [-\varepsilon, \varepsilon]$. Firstly, the computation leads to

$$\frac{\partial F}{\partial(s, t)} = \begin{vmatrix} x'(s) + ty''(s) & y'(s) \\ y'(s) - tx''(s) & -x'(s) \end{vmatrix}$$

$$= -x'(s)^2 - y'(s)^2 - ty''(s)x'(s) + tx''(s)y'(s) = -1 + O(t)$$

Thus

$$\text{Area}(F([0, L] \times [-\varepsilon, \varepsilon])) = \iint_{[0,L]\times[-\varepsilon,\varepsilon]} 1 \, ds \, dt + O(t) = 2\varepsilon L + O(\varepsilon^2)$$

(actually, the term $O(\varepsilon^2)$ is null). Having in mind the butts of the curve, we still have $\text{Area}(\gamma([0, L]) + B[0, \varepsilon]) = 2\varepsilon L + O(\varepsilon^2)$. Dividing by 2ε and taking limits we will recover L, proving so the claim.

9.7 Area of a C^1 surface with boundary

Some simple examples show that the area of a surface cannot be defined as simply as the length of a curve, that means, as the supremum of areas of polyhedral surfaces built upon the points of a triangulation. Based on the ideas of the previous section we will propose a method for the definition of the area of a surface. For the computations, we assume already known the definition and the properties of the *vector product* (see Section 11.1 for some historical information).

Let $\Gamma(u, v)$ be a parameterized C^1 surface with boundary, which means that:

1. $(u, v) \in \overline{D} \subset \mathbb{R}^2$, $\Gamma(u, v) \in \mathbb{R}^3$, Γ is injective;
2. there is $\overline{D} \subset C \subset \mathbb{R}^2$ where $\Gamma(u, v)$ extends as C^1 function;
3. $\frac{\partial \Gamma}{\partial u}(u, v) \times \frac{\partial \Gamma}{\partial v}(u, v) \neq 0$ for $(u, v) \in \overline{D}$.

Any implicitly defined surface can always be parameterized in that way locally. Note that

$$\frac{\partial \Gamma}{\partial u}(u,v) \quad \text{and} \quad \frac{\partial \Gamma}{\partial v}(u,v)$$

are tangent vectors at the point $\Gamma(u,v)$. The condition in terms of the vector product "$\times$" means that they generate the tangent plane at that point. Moreover, $\frac{\partial \Gamma}{\partial u}(u,v) \times \frac{\partial \Gamma}{\partial v}(u,v)$ is a normal vector to that plane.

Definition 9.7.1. *The area (2-dimensional measure) of a compact subset $S \in \mathbb{R}^3$ is defined by the limit, whether it exists, as*

$$Area(S) = \lim_{\varepsilon \to 0^+} \frac{vol(S + B[0,\varepsilon])}{2\varepsilon}$$

Any C^1 surface with boundary has an area that can be calculated with the formula given in the following result, however the result is proved for C^2 surfaces in order to simplify the proof.

Theorem 9.7.2. *Assume that $\Gamma(u,v)$ is a parameterized C^2 surface with boundary. Then its area is given by the formula*

$$\iint_D \left\| \frac{\partial \Gamma}{\partial u} \times \frac{\partial \Gamma}{\partial v} \right\| du\, dv.$$

Proof. Set the unitary normal vector as

$$N = \left\| \frac{\partial \Gamma}{\partial u} \times \frac{\partial \Gamma}{\partial v} \right\|^{-1} \frac{\partial \Gamma}{\partial u} \times \frac{\partial \Gamma}{\partial v}$$

The points that are at distance less or equal than ε from $\Gamma(\overline{D})$ with exception of the points that attains the minimum distance to $\Gamma(\partial D)$ are covered with the image of the map

$$F(u,v,t) = \Gamma(u,v) + tN(u,v)$$

with $(u,v,t) \in \Omega_t = \overline{D} \times [-\varepsilon, \varepsilon]$. The computation of the jacobian, done more explicitly later, will show that it does not vanish for a small choice of $\varepsilon > 0$. That implies that F is locally injective on $\overline{D} \times [-\varepsilon, \varepsilon]$ for such a small ε. Now we will show that a more careful choice of $\varepsilon > 0$ will make F actually injective. Indeed, by the Inverse Mapping theorem and the Lebesgue covering theorem there is $\delta > 0$ such that F is injective on every subset of $\overline{D} \times [-\varepsilon, \varepsilon]$

whose diameter does not exceed δ. Now we may assume that $\varepsilon > 0$ is small enough in order to satisfy $\varepsilon < \delta/3$ and $\|\Gamma(u_1, v_1) - \Gamma(u_2, v_2)\| < 2\varepsilon$ implies $\|(u_1, v_1) - (u_2, v_2)\| < \delta/3$ which is possible by the uniform continuity of Γ^{-1} defined on $\Gamma(\overline{D})$. In order to prove global injectivity assume that $F(u_1, v_1, t_1) = F(u_2, v_2, t_2)$. Since $|t_1|, |t_2| < \varepsilon$ we have $\|\Gamma(u_1, v_1) - \Gamma(u_2, v_2)\| < 2\varepsilon$ and so $\|(u_1, v_1, t_1) - (u_2, v_2, t_2)\| < \delta$ which contradicts the local injectivity.
The volume of $\Gamma(\overline{D}) + B[0, \varepsilon]$ differs from the volume of $F(\Omega_t)$ in $O(t^2)$, obtained by estimation of the volume of those points in $\Gamma(\overline{D}) + B[0, \varepsilon]$ whose distance, less than ε, is attained at some point from ∂D which is composed of finitely many C^1 curves.
In order to compute this volume, as F is injective and C^1, we may use the change of variable formula

$$\iiint_{\Omega_t} \left| \frac{\partial F}{\partial(u, v, t)} \right| du\, dv\, dt$$

Note that the partial derivatives that we need for the computation of the Jacobian can be expressed in vector notation as

$$\frac{\partial \Gamma}{\partial u} + t\frac{\partial N}{\partial u}; \quad \frac{\partial \Gamma}{\partial v} + t\frac{\partial N}{\partial v};$$

and N. Therefore, the Jacobian can be computed as the mixed product of the three vectors

$$\left(\frac{\partial \Gamma}{\partial u} + t\frac{\partial N}{\partial u}\right) \times \left(\frac{\partial \Gamma}{\partial v} + t\frac{\partial N}{\partial v}\right) \cdot N$$

$$= \left(\frac{\partial \Gamma}{\partial u} \times \frac{\partial \Gamma}{\partial v} + t\frac{\partial \Gamma}{\partial u} \times \frac{\partial N}{\partial v} - t\frac{\partial \Gamma}{\partial v} \times \frac{\partial N}{\partial u} + t^2\frac{\partial N}{\partial u} \times \frac{\partial N}{\partial v}\right) \cdot N$$

$$= \left\| \frac{\partial \Gamma}{\partial u} \times \frac{\partial \Gamma}{\partial v} \right\| + tf + t^2 g$$

where f, g are continuous functions on $\overline{D}$. Thus

$$\mathrm{vol}(F(\Omega_t)) = 2\varepsilon \iint_D \left\| \frac{\partial \Gamma}{\partial u} \times \frac{\partial \Gamma}{\partial v} \right\| du\, dv + O(\varepsilon^2)$$

Therefore

$$\lim_{\varepsilon \to 0^+} \frac{\mathrm{vol}(\Gamma(\overline{D}) + B[0, \varepsilon])}{2\varepsilon} = \iint_D \left\| \frac{\partial \Gamma}{\partial u} \times \frac{\partial \Gamma}{\partial v} \right\| du\, dv$$

which completes the proof. ∎

Actually the formula in the previous Theorem is true for C^1 parameterized surfaces with boundary but with a more complicate proof. The real purpose of the Theorem is to show that the standard formula for the area does not actually depend on the parameterization but on the set of points. There are different alternative methods to introduce the area of a surface as a measure which depends only on the point sets. One of them is the notion of 2-dimensional Hausdorff measure.

Example 9.7.3. *Compute the area of the torus defined by the parametric equation*

$$\Gamma(\theta,\phi) = ((R + r\cos\phi)\cos\theta, (R + r\cos\phi)\sin\theta, r\sin\phi),$$

for $\theta, \phi \in [0, 2\pi]$ *and with* $0 < r < R$.

The condition on $0 < r < R$ ensures the regularity of the surface. We have

$$\frac{\partial\Gamma}{\partial\theta}(\theta,\phi) = (-(R + r\cos\phi)\sin\theta, (R + r\cos\phi)\cos\theta, 0),$$

$$\frac{\partial\Gamma}{\partial\phi}(\theta,\phi) = (-r\sin\phi\cos\theta, -r\sin\phi\sin\theta, r\cos\phi).$$

These vectors are orthogonal, so the modulus of their vector product is just the product of their modules

$$\left\|\frac{\partial\Gamma}{\partial u} \times \frac{\partial\Gamma}{\partial v}\right\| = (R + r\cos\phi)r.$$

Then we can compute the area

$$\text{Area} = \int_0^{2\pi}\int_0^{2\pi} (R + r\cos\phi)r\, d\theta\, d\phi = 4\pi^2 Rr.$$

9.8 Alternative expressions for the area

We will introduce some classic notation

$$E = \left\|\frac{\partial\Gamma}{\partial u}\right\|^2; \quad F = \frac{\partial\Gamma}{\partial u}\cdot\frac{\partial\Gamma}{\partial v}; \quad G = \left\|\frac{\partial\Gamma}{\partial v}\right\|^2$$

which are the coefficients of the so called *first fundamental form* in differential geometry of surfaces. With this notation we have

$$\left\| \frac{\partial \Gamma}{\partial u} \times \frac{\partial \Gamma}{\partial v} \right\|^2 = EG - F^2$$

and therefore

$$\text{Area} = \iint_D \left\| \frac{\partial \Gamma}{\partial u} \times \frac{\partial \Gamma}{\partial v} \right\| du\, dv = \iint_D \sqrt{EG - F^2}\, du\, dv$$

which is easier to compute, specially in the case of an *orthogonal parameterization* because $F = 0$ in that case. An important feature of the formula for the area based in the first fundamental form coefficients is that it is independent of the dimension of the ambient space: the formulas are valid in $\mathbb{R}^d$ for any $d \geq 3$.

Now we will discuss two important particular cases. The first one is the form adopted by the integral for a surface given as the graph of a function $z = f(x, y)$ with $(x, y) \in D$. In this case the parameters are the variables x, y and $\Gamma(x, y) = (x, y, f(x, y))$ and so

$$\frac{\partial \Gamma}{\partial x} = (1, 0, \frac{\partial f}{\partial x}); \qquad \frac{\partial \Gamma}{\partial x} = (0, 1, \frac{\partial f}{\partial y}).$$

We have then

$$E = 1 + \left(\frac{\partial f}{\partial x} \right)^2; \quad F = \frac{\partial f}{\partial x} \frac{\partial f}{\partial y}; \quad G = 1 + \left(\frac{\partial f}{\partial y} \right)^2.$$

After an easy computation we get that

$$EG - F^2 = 1 + \left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2$$

and

$$\text{Area} = \iint_D \sqrt{1 + \left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2}\, dx\, dy.$$

The second case of surfaces admitting a particular formula for their area we are going to discuss is the case of the *surfaces of revolution*. Assume that we rotate around the the X-axis the graph of an one variable function $y = f(x) \geq 0$ with

$x \in [a,b]$. In such a case, in addition to the variable $x \in [a,b]$, we have to consider a rotation parameter $\theta \in [0, 2\pi]$, so the surface can be expressed as

$$\Gamma(x,\theta) = (x, f(x)\cos\theta, f(x)\sin\theta)$$

and for the partial derivatives

$$\frac{\partial\Gamma}{\partial x} = (1, f'(x)\cos\theta, f'(x)\sin\theta); \quad \frac{\partial\Gamma}{\partial\theta} = (0, -f(x)\sin\theta, f(x)\cos\theta)$$

which points out the (evident) fact that the parameterization is orthogonal, so $F = 0$ and

$$E = \sqrt{1 + f'(x)^2}; \qquad G = f(x).$$

Therefore the area can be write as

$$\text{Area} = \int_0^{2\pi}\int_a^b f(x)\sqrt{1+f'(x)^2}\,dx\,d\theta = 2\pi\int_a^b f(x)\sqrt{1+f'(x)^2}\,dx$$

which admits an intuitive interpretation based on the fact that $\int_a^b \sqrt{1+f'(x)^2}\,dx$ is the length of the graph of f. There is another *physical* interpretation for this integral. If we consider that the curve has a constant linear density, the vertical coordinate of the center of mass of the segment of graph $\{(x, f(x)) : x \in [a,b]\}$ is

$$\mathfrak{CM}_Y = \frac{\int_a^b f(x)\sqrt{1+f'(x)^2}\,dx}{\int_a^b \sqrt{1+f'(x)^2}\,dx}.$$

With this notation, the area generated by the rotation of the curve is

$$\text{Area} = 2\pi\,\mathfrak{CM}_Y\,\text{Length}$$

which is the classical *first Pappus–Guldin* theorem. The reader can check that the result of Example 9.7.3 can be easily obtained in this way.

For the sake of completeness, let us comment also the *second Pappus–Guldin theorem*. Consider now the vertical coordinate of the center of mass of the trapezoid

$$\{(x,y) : x \in [a,b], y \in [0, f(x)]\}$$

that is

$$\mathfrak{CM}_Y = \frac{\int_a^b f(x)^2\,dx}{2\int_a^b f(x)\,dx}.$$

Then the volume generated by the rotation can be calculated as

$$\text{Volume} = 2\pi\, \mathfrak{CM}_Y \text{ Area}$$

being "Area" the area of the trapezoid. The formula is evident by integrating through cross sections. Note that the center of mass involves quadratic degree. That was used by Archimedes in his *Method* to reduce one degree in "integration", reducing in that way, for instance, the computation of the area limited by a parabola to a property of the triangle.

9.9 Area measure and integration on surfaces

Once we know how compute the area of a parameterized surface with boundary we can extend the notion to other kinds of surfaces. If the surface is simply the finite union of parameterized surfaces with boundary (think of a polyhedron) we may take the union. In case we have a compact 2-dimensional C^1 manifold it is possible to decompose it into finitely many parameterized surfaces with boundary with the help of the implicit function theorem. This is intuitive clear but the details are quite tricky. What is more important is the fact that the value of the area does not depend on how the decomposition was done. We may cast some light using notions from measure theory.

On a parameterized C^1 surface with boundary Γ (the parameter domain is D) embedded into $\mathbb{R}^3$ we can define a positive measure S by the formula

$$S(A) = \iint_{\Gamma^{-1}(A)} \left\| \frac{\partial \Gamma}{\partial u} \times \frac{\partial \Gamma}{\partial v} \right\| du\, dv$$

where A is a relative Borel subset of S, that is, the intersection of $\Gamma(D)$ with a Borel subset of $\mathbb{R}^3$, and the integral is taken in Lebesgue sense with respect to the Lebesgue measure on $\mathbb{R}^2$. Note that there is no trouble considering the measure also on $\partial\Gamma(D) = \Gamma(\overline{D}) \setminus \Gamma(D)$ and its measure is 0. We know that this measure does not depends on the particular C^2 parameterization but a proof is required for C^1 surfaces.

Theorem 9.9.1. *Given a piezewise C^1 surface Σ in $\mathbb{R}^n$ (2-dimensional manifold) there exists a Borel measure on $\mathbb{R}^n$ concentrated on Σ such that for any Borel set $A \subset \Sigma$ which is covered by a parameterized C^1 piece Γ, then*

$$S(A) = \iint_{\Gamma^{-1}(A)} \left\| \frac{\partial \Gamma}{\partial u} \times \frac{\partial \Gamma}{\partial v} \right\| du\, dv.$$

Proof. The surface can be represented as finite or countable union of non overlapping parameterized C^1 surfaces with boundaries. The formula above defines a measure on each piece and the sum (series) of all those measures will be the wanted measure. We have to check that the measure does not depends on the decomposition into parameterized C^1 surfaces neither the parameterizations.
Suppose that Σ has two different decompositions. The intersection of both decompositions induce a finer decomposition, at most countable. It is not difficult to see that the problem of uniqueness for S reduces to check if it is the same on each of such a pieces.
Suppose that Γ_1 and Γ_2 with domains D_1 and D_2. Firstly, note that $h = \Gamma_2^{-1} \circ \Gamma_1$ is an injective C^1 map form D_1 onto D_2. Therefore the theorem of change of variables is applicable

$$\iint_{\Gamma_2^{-1}(A)} \left\| \frac{\partial \Gamma_2}{\partial u} \times \frac{\partial \Gamma_2}{\partial v} \right\| du\, dv = \iint_{(h^{-1} \circ \Gamma_2^{-1})(A)} \left\| \frac{\partial \Gamma_2}{\partial u} \times \frac{\partial \Gamma_2}{\partial v} \right\| \left| \frac{\partial (u,v)}{\partial (s,t)} \right| ds\, dt$$

where $h(s,t) = (u(s,t), v(s,t))$. Now we have to express the tangent vectors in terms of Γ_1 and the variables s, t. We have $\Gamma_1 = h \circ \Gamma_2$, thus

$$\frac{\partial \Gamma_1}{\partial s} = \frac{\partial \Gamma_2}{\partial u}\frac{\partial u}{\partial s} + \frac{\partial \Gamma_2}{\partial v}\frac{\partial v}{\partial s}; \qquad \frac{\partial \Gamma_1}{\partial t} = \frac{\partial \Gamma_2}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial \Gamma_2}{\partial v}\frac{\partial v}{\partial t}.$$

The vector product gives

$$\frac{\partial \Gamma_1}{\partial s} \times \frac{\partial \Gamma_1}{\partial t} = \left(\frac{\partial u}{\partial s}\frac{\partial v}{\partial t} - \frac{\partial v}{\partial s}\frac{\partial u}{\partial t} \right) \frac{\partial \Gamma_2}{\partial u} \times \frac{\partial \Gamma_2}{\partial v} = \frac{\partial (u,v)}{\partial (s,t)} \frac{\partial \Gamma_2}{\partial u} \times \frac{\partial \Gamma_2}{\partial v}$$

and thus

$$\iint_{\Gamma_2^{-1}(A)} \left\| \frac{\partial \Gamma_2}{\partial u} \times \frac{\partial \Gamma_2}{\partial v} \right\| du\, dv = \iint_{\Gamma_1^{-1}(A)} \left\| \frac{\partial \Gamma_1}{\partial s} \times \frac{\partial \Gamma_1}{\partial t} \right\| ds\, dt$$

since $(h^{-1} \circ \Gamma_2^{-1})(A) = (h \circ \Gamma_2)^{-1}(A) = \Gamma_1^{-1}(A)$. That proves that the measure S is independent on the parameterization and so on the decomposition. ∎

Remark 9.9.2. *There are more general approach to the problem of define the area that do not require parameterizations, like the* 2*-dimensional Hausdorff measure. However, dealing with Hausdorff measures is not a quite easy task.*

Integration on surfaces with respect to the area measure S plays a very important role in applications, like integration with respect to the arc length

for curves. The construction of the measure S by means of an integral together standard methods of measure theory (approximation by simple functions) implies for function f defined on a surface Σ which is integrable with respect to S we have

$$\int_\Sigma f\,dS = \sum_n \iint_{D_n} f\circ\Gamma_n \left\| \frac{\partial \Gamma_n}{\partial u} \times \frac{\partial \Gamma_n}{\partial v} \right\| du\,dv.$$

where (Γ_n, D_n) is a finite or countable decomposition of Σ.

We will discuss some cases which may happen in physical applications in $\mathbb{R}^3$. In this setting is preferred to write the surface integral with double integration sign $\iint_\Sigma f\,dS$ and vectors are often distinguished with little arrows above.

Firstly, in many occasions it will be necessary to integrate vector fields $\overrightarrow{F} = (f_1, f_2, f_3)$. In that case we have

$$\iint_\Sigma \overrightarrow{F}\,dS = (\iint_\Sigma f_1\,dS, \iint_\Sigma f_2\,dS, \iint_\Sigma f_3\,dS)$$

However y many occasions, the vector field will be normal to the surface, so it can be write as $\overrightarrow{F} = f\overrightarrow{N}$, being $\overrightarrow{N}$ a normal unitary vector field and f a scalar function. Let us remark that at any point there are two unitary vectors which are normal to the surface. To set a continuous normal field $\overrightarrow{N}$ is to choose the orientation of Γ between the two possible ones for a parameterized surface with boundary (general surfaces cannot always be oriented as the *Moebius strip*. In that case there is a specific notation

$$\iint_\Sigma f\overrightarrow{N}\,dS = \iint_\Sigma f\,d\overrightarrow{S}$$

The notation using the *oriented element of area* $d\overrightarrow{S}$ is very suitable because in the case of a parameterized surface Γ we have

$$\iint_\Gamma f\,d\overrightarrow{S} = \iint_D f\circ\Gamma\,\frac{\partial\Gamma}{\partial u}\times\frac{\partial\Gamma}{\partial v}\,du\,dv.$$

in case the unitary normal $\overrightarrow{N}$ points in the same direction that $\frac{\partial\Gamma}{\partial u}\times\frac{\partial\Gamma}{\partial v}$. Indeed, in such a case we have

$$\frac{\partial\Gamma}{\partial u}\times\frac{\partial\Gamma}{\partial v} = \left\|\frac{\partial\Gamma}{\partial u}\times\frac{\partial\Gamma}{\partial v}\right\|\overrightarrow{N}.$$

Finally we will consider the so called *flux of a vector field throughout a surface*. If the vector field represents the speeds at a given moment of the particles a moving fluid is composed of, the flux integral computes the volume of fluid crossing the surface by time unit at that moment. Let Γ be a parameterized C^1 surface with boundary and assume that there is an orientation given to Γ which agrees with $\frac{\partial \Gamma}{\partial u} \times \frac{\partial \Gamma}{\partial v}$. Then the flux of a field $\overrightarrow{F}$ throughout Γ is defined as

$$\iint_{\Gamma} \overrightarrow{F} \cdot d\overrightarrow{S} = \iint_{\Gamma} \overrightarrow{F} \cdot \overrightarrow{N}\, dS$$

and, obviously, it can be computed by

$$\iint_{\Gamma} \overrightarrow{F} \cdot d\overrightarrow{S} = \iint_{D} \overrightarrow{F} \cdot \frac{\partial \Gamma}{\partial u} \times \frac{\partial \Gamma}{\partial v}\, du\, dv$$

and the scalar-vector product inside can be computed straightly by means of a determinant of the coefficients of the three vectors. The flux integrals can be interpreted as integrals of *differential forms* of second degree.

9.10 Rationale and remarks

In this chapter we define the length of parameterized curves in normed spaces and the area of parameterized surfaces in the Euclidean $\mathbb{R}^3$. The difference of generality between both approaches is due to these facts:

1. Rectangles provide the most simple example of computation of area. Of course, "rectangles" requiere a notion of orthogonality that is not obvious without Euclidean structure.

2. The restriction to dimension 3 is a consequence of the chosen method of definition. The proof also involves the use of the vector product, which is a peculiarity of $\mathbb{R}^3$.

We choose the parameterized form of the manifolds because it provides explicit formulas. If the manifold cannot fully parameterized, the measure is extended piecewise. In the case of curves, we consider simply piecewise C^1. For surfaces the construction is a bit more complicated using that any surface admits locally parameterizations.

In order to characterize rectifiable curves in finite dimensional spaces we start with the theory of functions of bounded variation. As a consequence, the

fact of being rectifiable implies that the curve necessarily is continuous but countably many points and continuous rectifiable curves have tangents almost everywhere.

The definition of area of a surface is motivated by an alternative approach to the computation of the length of a curve in $\mathbb{R}^2$. The purpose of our approach is to have a "coordinate free" definition for the area, however the method of proof gives the existence of the area (and the formula) only for C^2 surfaces with boundary. Later, the use of the area formula for piecewise C^1 surfaces is discussed. Other coordinate free approaches, as the Hausdorff 2-dimensional measure are not simpler.

Integration on curves and surfaces is mainly intended for scalar functions, since the integration of differential forms is done in other lesson. However, line and flux integrals are addressed briefly here because the theory of differential forms, despite its elegance, does not provide a constructive approach.

9.11 Exercises

1. A *logarithmic spiral* is a curve that can be represented in polar coordinates by $r = ae^{b\theta}$. Calculate the length of the arc for $\theta \in [0, 2\pi]$.

2. Calculate the length of the *cardioid*, a classic curve whose polar formula is
$$r = 1 + \cos\theta$$
where $\theta \in [-\pi, \pi]$.

3. Calculate the length of the closed curve
$$x = a\sin^3 t; \quad y = a\cos^3 t.$$

4. Find the length of the Viviani curve
$$x^2 + y^2 + z^2 = 1; \quad x^2 + y^2 - x = 0.$$

5. Prove that the length of a continuous rectifiable curve is a continuous function of the parameter.

6. Find the explicit formula for the length of a curve contained in a C^1 surface $T(u,v) = (x(u,v), y(u,v), z(u,v))$ in terms of the first fundamental form and the this expression for the curve $\gamma = T \circ \eta$ donde $\eta : [a,b] \to \mathbb{R}^2$ also C^1.

7. Calculate the area of the portion of sphere $x^2 + y^2 + z^2 = 2x$ inside the cone $x^2 + y^2 = z^2$.

8. Calculate the area of the cone $x^2 + y^2 = z^2$ with $z \geq 0$ inside the sphere $x^2 + y^2 + z^2 = 2ax$.

9. Consider the surface $z = Axy$ con $x^2 + y^2 \leq R^2$ where $A, R \geq 0$. Estimate the value of the parameter A for the area of the surface be twice the area of its orthogonal projection on the plane XY.

10. Find the area of the portion of cone $z = \sqrt{2x^2 + 2y^2}$ below the plane $z = x + 1$.

11. Calculate the area of the surface $z = \sqrt{2xy}$ with $(x,y) \in [0,a] \times [0,b]$.

12. Calculate the area of the piece of paraboloid $y^2 + z^2 = 2px$ limited by the plane $x = a$.

13. Calculate the area of the piece of paraboloid $x^2/a + y^2/b = 2z$ inside the cylinder $x^2/a^2 + y^2/b^2 = 1$.

14. Calculate the area of the cone $z^2 = 2xy$ limited by the planes $x = 0$, $x = a$, $y = 0$, $y = b$.

15. Find the are of the surface

$$z = \arcsin(\sinh x \sinh y)$$

limited between the planes $x = a$, $x = b$ with $a, b > 0$.

16. Calculate the area of the piece of paraboloid $y^2 + z^2 = 4ax$ intercepted by the cylinder $y^2 = ax$ and the plane $x = 3a$.

17. Calculate $\iint_S z \, dS$ where S is the half-sphere $x^2 + y^2 + z^2 = a^2$, $z > 0$.

18. Calculate the area and volume of a *torus*.

19. Find with the help of the Pappus-Guldin theorem the position of the center of mass of a solid homogeneous semicircle.

20. The *catenary* is the curve with the shape of a hanging chain and it is modelled by the hyperbolic cosine. The *catenoid* is the surface generated by a catenary that rotates around its symmetry axis.

    (a) Find the length of an arc of catenary.

    (b) Find the area of a catenoid between its vertex and a circular section.

21. Consider the arc of *cycloid* $x(t) = t - \sin t$, $y(t) = 1 - \cos t$ for $t \in [0.2\pi]$.

    (a) Find the length of the arc.

    (b) Find the area of the surface generated by the rotation of the arc around the X axis.

Chapter 10

Differential forms of low degree

10.1 Forms of degree 1

The aim of the following device is to understand the line integral with not appeal to the scalar product. Remember that a very common interpretation of the path integral is the work done by a force along a trajectory and scalar product is a keystone in such a formulation.

Let X denote a normed space. The set of continuous linear functionals defined on X is also a normed space denoted usually by X^* (the *dual* space). The elements of X^* act on the elements of X giving a real number: $x^*(x) \in \mathbb{R}$ for $x \in X$ and $x^* \in X^*$. Note that in the case that $X = \mathbb{R}^d$ we may identify X^* with $\mathbb{R}^d$ too, however they have different nature and the norm may differ.

Definition 10.1.1. *A differential form ω of degree 1 (also called 1-form) is a function defined on a open subset of X with values on X^*.*

The 1-form is said continuous, differentiable... if it is so considered as a map between normed spaces. Usually we will consider 1-forms which are continuous at least, but later we will requiere more regularity for some applications.

The most typical example of 1-form is the differential of a differentiable real function f defined on some open subset $D \subset X$. Indeed, df is defined on D and $df(x)$ is a continuous linear map from X to $\mathbb{R}$, that is, an element of X^* for every $x \in D$. By the way, we may think of scalar functions as differential forms of degree 0, so the differentiation increases by one the degree of the form. Later we will increase to degree 2 by a further derivation.

Now we will address our attention to the finite dimensional case $X = \mathbb{R}^n$. If we denote a generic point as $x = (x_1, x_2, \ldots, x_n)$ we may consider the linear forms given by the assignation to the k-th coordinate $x \to x_k$ an to denote by dx_k this linear form. It turns out that $\{dx_1, dx_2, \ldots, dx_n\}$ is a basis of $(\mathbb{R}^n)^*$. Therefore, any 1-form ω can be expressed in terms of the basis as

$$\omega = f_1 dx_1 + f_2 dx_2 + \ldots f_n dx_n$$

where f_k with $k = 1 \ldots n$ are scalar functions. After that, we can express the differential of a scalar function in this way

$$df = \frac{\partial f}{\partial x_1}\, dx_1 + \frac{\partial f}{\partial x_2}\, dx_2 + \cdots + \frac{\partial f}{\partial x_n}\, dx_n.$$

If $n \leq 3$ we will use the set of variables x, y, z rather than the numeration.

10.2 Integration of 1-forms on paths

Given a parameterized curve $\gamma : [a, b] \to D \subset \mathbb{R}^n$ and a continuous 1-form ω defined on D, the integral of ω on (or along) γ is the number

$$\int_\gamma \omega = \int_a^b \omega(\gamma(t))(\gamma'(t))\, dt.$$

The first natural question is to know in what extent the definition depends on γ as a function rather than on the "shape" $\gamma([a, b])$. Actually the integral will depend on the set $\gamma([a, b])$ and the order in which the points are placed, the sense the curve is walked. To decide between the two possible ways to walk the curve is to set an *orientation*. To be more precise, the integral of 1-forms does not change by reparameterizations which preserve the orientation.

Proposition 10.2.1. *Let $\gamma : [a, b] \to D \subset \mathbb{R}^n$ piecewise C^1, ω a continuous 1-form defined on D and $j : [c, d] \to [a, b]$ an increasing piecewise C^1 bijection. Then $\tilde{\gamma} : [c, d] \to D$ is a piecewise C^1 curve and*

$$\int_{\tilde{\gamma}} \omega = \int_\gamma \omega$$

Proof. The following equalities are not bothered by finite set of points where the derivatives are not defined

$$\int_{\tilde{\gamma}} \omega = \int_c^d \omega(\tilde{\gamma}(t))(\tilde{\gamma}'(t))\, dt = \int_c^d \omega(\gamma(j(t)))(\gamma'(j(t))j'(t))\, dt$$

$$= \int_c^d \omega(\gamma(j(t)))(\gamma'(j(t)))\, j'(t)\, dt = \int_a^b \omega(\gamma(\tau))(\gamma'(\tau))\, d\tau = \int_\gamma \omega$$

where we have used the chain rule and the linearity of the form. ■

The previous proposition motivates the notion of "path". A path is the class of equivalence of all the curves that can be obtained one form the other through regular reparameterizations that no reverse the orientation.

We may perform some operations with paths: if a curve γ can be considered as the concatenation of two of them γ_1, γ_2, up to a suitable reparameterization , we write $\gamma = \gamma_1 + \gamma_2$; if some parameterization γ^* of $\gamma([a,b])$ goes backwards we write $\gamma^* = -\gamma$. Now we have this obvious result.

Proposition 10.2.2. *For paths $\gamma, \gamma_1, \gamma_2$ with values on the domain of a continuous 1-form ω the following equalities hold:*

$$\int_{\gamma_1+\gamma_2} \omega = \int_{\gamma_1} \omega + \int_{\gamma_2} \omega \ ; \qquad \int_{-\gamma} \omega = -\int_\gamma \omega.$$

The chain rule is also the trick behind the following result.

Proposition 10.2.3. *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be C^1 and let $\gamma : [a,b] \to D$ piecewise C^1. Then*

$$\int_\gamma df = f(\gamma(b)) - f(\gamma(a)).$$

In particular, if γ is closed, that is, $\gamma(a) = \gamma(b)$, then $\int_\gamma df = 0$.

Proof. Only one line

$$\int_\gamma df = \int_a^b df(\gamma(t))(\gamma'(t))\, dt = \int_a^b \frac{d}{dt}(f(\gamma(t)))\, dt = f(\gamma(b)) - f(\gamma(a)).$$

■

Note that to say that the integral depends only on the starting and ending points (from now on "endpoints" with a distinction between them) means that it does not matter how the path is joining them, which is actually equivalent to say that the integral along a *closed* path is always 0. We have the following important result.

Theorem 10.2.4. *Let ω be an 1-form defined on $D \subset \mathbb{R}^n$. There there exists a C^1 function $f : D \to \mathbb{R}$ such that $\omega = df$ if and only if $\int_\gamma \omega$ depends only on the endpoints of γ (equivalently, if $\int_\gamma \omega = 0$ for every closed $\gamma \subset D$).*

Proof. Clearly one implication is a consequence of the previous proposition. Assume now that the line integral depends only on the endpoints of the curve. We may assume that D is connected, so the for a general open set it is enough to make the following construction on every connected component. Fix a point $x_0 \in D$ and for any other point $x \in D$ fix a C^1 curve $\gamma_x \subset D$ starting at x_0 and ending at x (the parameter interval is not relevant). Define now $f(x) := \int_{\gamma_x} \omega$. Note that this definition is not ambiguous by the hypothesis of independence on how the points x_0 and x are joined.
Now fix $x \in D$ and $h \in \mathbb{R}^n$ and take $\delta > 0$ small enough to have $x + h \in D$ for $\|h\| < \delta]$. Observe that

$$f(x+h) - f(x) = \int_{\gamma_{x+h}} \omega - \int_{\gamma_x} \omega = \int_{\gamma_{x+h}-\gamma_x} \omega$$

and the fact that the curve $\gamma_{x+h} - \gamma_x$ joins the points x and $x + h$, so it can be replaced by the segment $x + th$ with $t \in [0, 1]$. We have then

$$f(x+h) - f(x) = \int_0^1 \omega(x+th)(h)\, dt.$$

Since ω is continuous at x, given $\varepsilon > 0$ we can take a smaller $\delta > 0$ so $\|\omega(x + th) - \omega(x)\| < \varepsilon$ for all $t \in [0, 1]$, what implies

$$|\omega(x+th)(h) - \omega(x)(h)| < \varepsilon\|h\|.$$

As obviously $\int_0^1 \omega(x)(h)\, dt = \omega(x)(h)$, putting it all together

$$|f(x+h) - f(x) - \omega(x)(h)| = |\int_0^1 (\omega(x+th)(h) - \omega(x)(h))\, dt| \leq \varepsilon\|h\|$$

what means that f is differentiable at x and its differential is precisely $\omega(x)$ as wanted. ∎

An 1-form ω is called *exact* if it has a *primitive* f, that is, if $\omega = df$. Note that the primitive is not unique and two primitives of the same form should differ in a function whose differential is 0 and therefore constant on every connected component of the domain.

The characterization of exact 1-forms in terms of integrals is not a practical one. First of all note that if an exact 1-form $\omega = \sum_{k=1}^n f_k dx_k = df$ is C^1 then

its coefficients satisfy that $f_k = \frac{\partial f}{\partial x_k}$ and therefore

$$\frac{\partial f_k}{\partial x_j} = \frac{\partial^2 f}{\partial x_k \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_k} = \frac{\partial f_j}{\partial x_k}$$

since f is C^2 the commutation of derivatives holds. Amazingly, this necessary condition turns out to be almos a sufficient condition for exactness. For that reason we are going to name it: a 1-form $\omega = \sum_{k=1}^n f_k dx_k$ is called *closed* if $\frac{\partial f_k}{\partial x_j} = \frac{\partial f_j}{\partial x_k}$ on all its domain and for every pair of indices i, j. We say that an open subset D is *star-shaped* if there is a point $x_0 \in D$ such that for any other point $x \in D$ the segment joining x_0 and x is contained in D. In particular, star-shaped sets are connected, and convex open sets are star-shaped and

Theorem 10.2.5 (Lemma of Poincaré)**.** *Let $D \subset \mathbb{R}^n$ be a star-shaped open set and let ω be a C^1 closed form. Then ω is exact.*

Proof. By a translation we may assume that D is star-shaped with respect to 0, so the segment joining 0 and $(x_1, \ldots, x_n)$ is simply $(tx_1, \ldots, tx_n)$ with $t \in [0, 1]$. The candidate to primitive is obviously

$$f(x_1, \ldots, x_n) = \int_0^1 \sum_{k=1}^n f_k(tx_1, \ldots, tx_n) x_k \, dt$$

Now, for the partial derivative with respect to x_j we have

$$\frac{\partial f}{\partial x_j}(x_1, \ldots, x_n) = \int_0^1 f_j(tx_1, \ldots, tx_n) + \sum_{k=1}^n \frac{f_k}{\partial x_j}(tx_1, \ldots, tx_n) t x_k \, dt$$

$$= \int_0^1 f_j(tx_1, \ldots, tx_n) + \sum_{k=1}^n \frac{f_j}{\partial x_k}(tx_1, \ldots, tx_n) t x_k \, dt$$

$$= \int_0^1 f_j(tx_1, \ldots, tx_n) + \frac{d}{dt}(f_j(tx_1, \ldots, tx_n)) \, dt$$

$$= \int_0^1 \frac{d}{dt}(t f_j(tx_1, \ldots, tx_n)) \, dt = t f_j(tx_1, \ldots, tx_n) \,|_0^1 = f_j(x_1, \ldots, x_n)$$

where for the first equality we may exchange derivation and integration by the regularity of the functions, the second equality makes use of the hypothesis and the following equalities are based on the chain rule and derivation of products.

That shows that $df = \omega$ as wanted. ∎

A explicit form for the primitive of ω is given in the proof, however for small dimension is better to proceed in this way: suppose that $\omega = p\,dx + q\,dy$ is closed, that is $\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$. Then find a *partial primitive* g of q with respect to y, that is, $\frac{\partial g}{\partial y} = q$. The function we are looking for can be written as $f = g + h$ where h does not depend on y as

$$\frac{\partial h}{\partial y} = \frac{\partial f}{\partial y} - \frac{\partial g}{\partial y} = q - q = 0.$$

Then $h(x)$ is a function of a single variable and

$$h'(x) = \frac{\partial f}{\partial x} - \frac{\partial g}{\partial x} = p - \frac{\partial g}{\partial x}.$$

The last term should be only function of x in order to find h. Actually, it is

$$\frac{\partial}{\partial y}(p - \frac{\partial g}{\partial x}) = \frac{\partial p}{\partial y} - \frac{\partial^2 g}{\partial x \partial y} = \frac{\partial p}{\partial y} - \frac{\partial q}{\partial x} = 0$$

by the hypothesis. Now find h and we have $f = g + h$ explicitly.

Poincaré's results points out that closed forms are exact on special domains. If the domain is not start-shaped then the result may fail. Indeed, the form

$$\frac{-y\,dx}{x^2 + y^2} + \frac{x\,dy}{x^2 + y^2}$$

on $\mathbb{R}^2 \setminus \{(0,0)\}$ is closed (easily checkable) and not exact (consider the integral around the unit circle). The formula $f(x,y) = \arctan(y/x)$ provides a primitive valid on a halfplane that can be extended to any domain of the form $\mathbb{R}^2 \setminus R$ where R is a halfline departing from $(0,0)$ (being $f(x,y)$ is a measure of the angle between (x,y) and R). The characterization of the domains where all closed form is exact is actually a topological matter, however the notions and proofs involved are beyond the scope of these notes.

10.3 The Green-Riemann formula

This section is concerned with integration of 1-forms in $\mathbb{R}^2$ over *simple closed paths*, or more generally, the boundary of bounded open domain provided that

it is piecewise C^1. We will start with simpler elementary domains. We say that a domain is *elemental with respect to the X axis* if it is limited by two vertical lines $x = a$ and $x = b$ and two graphs of C^1 functions $f, g : [a, b] \to \mathbb{R}$ such that $f > g$. The boundary ∂D of an elementary domain D is supposed to be oriented anticlockwise, that is the graph of f is walked from right to left and the segment on $x = a$ is walked down, for instance. A domain *elemental with respect to the Y axis* is defined similarly, or we can think of a how looks like a domain elemental with respect to the X axis after a rotation of $\pi/2$.

Lemma 10.3.1. *Let D be a elemental domain with respect to the X axis and $p(x, y)$ a C^1 function defined on a domain containing $\overline{D}$. Then*

$$\int_{\partial D} p\,dx = -\iint_D \frac{\partial p}{\partial y}\,dxdy.$$

Proof. Note that the vertical segments do not add to the line integral because it does not contain dy. The graphs are parameterized by $(x, g(x))$ and $(x, f(x))$, but this last one must be walked backwards. Thus

$$\int_{\partial D} p\,dx = \int_a^b p(x, g(x))\,dx - \int_a^b p(x, f(x))\,dx =$$

$$-\int_a^b (p(x, f(x)) - p(x, g(x)))\,dx = -\int_a^b \left(\int_{g(x)}^{f(x)} \frac{\partial p}{\partial y}(x, y)\,dy\right) dx =$$

$$-\iint_D \frac{\partial p}{\partial y}(x, y)\,dxdy$$

as wanted. ∎

If we had considered an elemental domain with respect to the Y axis the proof would have followed the same lines but the anticlockwise orientation of the domain has an opposite relation to the orientation of the Y axis. Therefore the analogous result does not contain the sign minus.

Lemma 10.3.2. *Let D be a elemental domain with respect to the Y axis and $q(x, y)$ a C^1 function defined on a domain containing $\overline{D}$. Then*

$$\int_{\partial D} q\,yx = \iint_D \frac{\partial q}{\partial x}\,dxdy.$$

It is easy to prove that a convex domain with C^1 boundary is elemental with respect to both the X and Y axis, and many other domains can be expressed as a non overlapping union of domains which are elemental with respect to the two axis. Putting together the previous lemmas and the observation we have the following.

Proposition 10.3.3 (Green-Riemann, elemental domains)**.** *Let D be a domain which is elemental with respect to both the X and Y axis, and let $\omega = p\,dx + q\,dy$ an 1-form which is C^1 on a domain that contains $\overline{D}$. Then*

$$\int_{\partial D} \omega = \int_{\partial D} p\,dx + q\,dy = \iint_D \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) dxdy.$$

Moreover, the same formula hold if D is a domain such that it can be decomposed into a finite non overlapping union of domains with C^1 boundaries which are elemental with respect to both the X and Y axis.

Proof. The first statement is just the sum of the formulas provided by the lemmas. For the second statement we have only to remark that the double integras are additive for non overlapping unions of domains. Namely, if $D = \bigcup_{k=1}^n D_k$ then

$$\sum_{k=1}^n \iint_{D_k} \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) dxdy = \iint_{\bigcup_{k=1}^n D_k} \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) dxdy.$$

On the other hand, any C^1 non trivial (not reduced to a single point) piece of curve contained $\bigcup_{k=1}^n \partial D_k \setminus \partial D$ is contained into a shared boundary $\partial D_i \cap \partial D_j$ where $i \neq j$ are unique. This curve does not contribute to $\int_{\partial D}$ since it is walked in opposite directions when computing $\int_{\partial D_i}$ and $\int_{\partial D_j}$ with the subsequent cancelation. That can be expressed with the "arithmetics of paths" as $\sum_{k=1}^n \partial D_k = \partial D$. Therefore

$$\sum_{k=1}^n \int_{\partial D_k} p\,dx + q\,dy = \int_{\sum_{k=1}^n \partial D_k} p\,dx + q\,dy = \int_{\partial D} p\,dx + q\,dy$$

which completes the proof of the proposition. ∎

The previous result cast more light on the relation between closed and exact 1-forms. Indeed, if the 1-form is closed then the function inside the double integral vanishes, so the line integral along the boundary of any elemental

domain is zero. Any Jordan closed curve is the boundary of a region. However, provided that such a curve is C^1, it may be impossible to decompose it into finitely many elemental domains. We will improve the previous proposition for more general domains.

Theorem 10.3.4 (Green-Riemann, general)**.** *Let D be a bounded open domain with C^1 boundary and let $p\,dx + q\,dy$ be an 1-form which is C^1 on a domain that includes $\overline{D}$. Then*

$$\int_{\partial D} p\,dx + q\,dy = \iint_D \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) dxdy.$$

Proof. It is enough to prove the result for 1-forms of type $p\,dx$ as we have seen before. Let $F_1 \subset \partial D$ be the finite set of points where the boundary is not smooth. Let $K \subset \partial D$ be the compact set of the points of the boundary where the tangent line is vertical. Indeed, the set has closed intersection with each C^1 piece, explicitly $\{\gamma(t)\cdot(1,0) : \gamma'(t)\cdot(1,0) = 0\}$. The orthogonal projection of K onto the X axis K_X is also compact and has measure 0 by Morse-Sard's theorem. The set K_X can be covered with finitely many open intervals whose total length can be taken smaller than a number δ we will precise later. Let F_2 denote the finite set of their butts. The vertical lines built on the points of $F = F_1 \cup F_2$ define closed strips S_k with $k = 1, \ldots, n$ of two types: stripes of type B (bad) if they contain points of ∂D at which the tangent is vertical; stripes of type A (alright) just the others. Note that at any point of $\partial D \cap S$ with S of type A is possible to express locally the boundary as the graph of a function $y = f(x)$. Indeed, the points where the implicit function theorem is not applicable are contained in the strips of type B. Adding a finite set of points F_3 to F we may moreover assume that any connected part of $\partial D \cap S$ for S of type A is a graph of the sort $y = f(x)$. Under the hypotheses, the number of connected parts of $\partial D \cap S$ must be finite and thus $D \cap S$ can be decomposed into finitely elemental domains. Therefore

$$\int_{\partial(D\cap S)} p\,dx = -\iint_{D\cap S} \frac{\partial p}{\partial y}\,dxdy$$

for any S of type A. Now we are going to deal with the stripes of type B. Take $\varepsilon > 0$. Since the set $\overline{D}$ is bounded and $\frac{\partial p}{\partial y}$ continuous we may take $\delta > 0$ small enough to guarantee that

$$\sum_{k\in B} \iint_{D\cap S_k} \left|\frac{\partial p}{\partial y}\right| dxdy < \varepsilon.$$

On the other hand, if the covering by open intervals of K_X is tight enough we may assume that first component of the derivative $(\gamma'(t)\cdot(1,0)$ for γ a C^1 piece of $\partial D)$ is small than $\varepsilon/\text{length}(\partial D)$. That implies

$$\sum_{k\in B}\left|\int_{\partial(D\cap S_k)} p\,dx\right|<\varepsilon.$$

Now we have

$$\left|\sum_{k\in A\cup B}\iint_{D\cap S_k}\frac{\partial p}{\partial y}\,dxdy+\sum_{k\in A\cup B}\int_{\partial(D\cap S)}p\,dx\right|<2\varepsilon.$$

Since the cancelation happens for the vertical segments added by the sripe, we get

$$\left|\int_{\partial D}p\,dx+\iint_D\left(\frac{\partial p}{\partial y}\right)dxdy\right|<2\varepsilon$$

and being $\varepsilon>0$ arbitrary we arrive to the formula in the statement. ∎

It is necessary to remark that the anticlockwise orientation of D not always goes "anticlockwisely". Consider the square $[0,3]^2$ which is made up of 9 smaller squares of side 1. Remove the central square $[1,2]^2$ and call D the remainder. Then ∂D has two connected components, and the inner one is walked clockwise. Indeed, endow the boundary of the 8 squares whose union is D with the anticlockwise orientation drawing the arrows and then look. Another explanation based in Green-Riemann, if we add $[1,2]^2$ to D with its anticlockwise orientation on the border, the shared boundary of D should be walked the other way around for the annihilation of the line integrals.

10.4 Forms of degree 2

We will begin by the definition the algebraic forms of degree 2. Consider a bilinear continuous map $b: X\times X\to\mathbb{R}$. We say that b is *symmetric* if $b(x,y)=b(y,x)$ and *antisymmetric* if $b(x,y)=-b(x,y)$ for every $x,y\in X$. Note that the bilinear continuous forms on X made up a vector space and the sets or symmetric and antisymmetric are linear subspaces. Moreover, the space of continuous bilinear forms is *direct sum* of the symmetric and antisymmetric subspaces. That is consequence of the unique decomposition of a bilinear form b as a symmetric and antisymmetric forms

$$b(x,y)=\frac{1}{2}\big(b(x,y)+b(y,x)\big)+\frac{1}{2}\big(b(x,y)-b(y,x)\big).$$

If the dimension of X is finite, say n, then the dimension of the space of bilinear forms (all them are continuous in this case) is n^2. Indeed, if $\{e_1, \ldots, e_n\}$ is basis of X the value of a bilinear form b is determined by the n^2 values of the coefficients $\{b(e_i, e_j) : 1 \leq i, j \leq n\}$ as

$$b(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j b(e_i, e_j).$$

Now, observe that a symmetric form is determined by $n(n+1)/2$ coefficients since $b(e_i, e_j) = b(e_j, e_i)$, therefore that number is the dimension of the subspace of symmetric forms. Analogously, the dimension of the subspace of antisymmetric forms is $n(n-1)/2$. This is a lesser number because for all $1 \leq i \leq n$ we have $b(e_i, e_i) = 0$.

Let $\mathcal{A}_2(X)$ denote the set of all continuous antisymmetric bilinear forms on X. This space is naturally endowed with the topology of the supremum norm (on $B_X \times B_X$). In order to keep coherence with the general theory of *alternate forms*, which we are no discussing here, $\mathcal{A}_0(X)$ will denote the scalars and $\mathcal{A}_1(X) = X^*$. We will introduce the *exterior product* in a very restricted version

$$\wedge : \mathcal{A}_1(X) \times \mathcal{A}_1(X) \to \mathcal{A}_2(X)$$

defined as $(\alpha \wedge \beta)(x, y) = \alpha(x)\beta(y) - \alpha(y)\beta(x)$ for $\alpha, \beta \in \mathcal{A}_1(X)$. There is not difficulty in checking that $\wedge$ is bilinear and antisymmetric.

In case of finite dimension, with $X = \mathbb{R}^n$ we will often consider a basis of $\mathcal{A}_2(\mathbb{R}^n)$ which is built of from the standard basis of $\mathcal{A}_1(\mathbb{R}^n)$ with the help of the exterior product. Namely, the basis made up of the following elements

$$dx_i \wedge dx_j(e_k, e_l) = \begin{cases} 1 & if \quad k = i, l = j \\ -1 & if \quad k = j, l = i \\ 0 & \text{otherwise} \end{cases}$$

The basis of $\mathcal{A}_2(\mathbb{R}^n)$ can be enumerated as follows

$$\{dx_i \wedge dx_j : 1 \leq i < j \leq n\},$$

however for reasons that will be clear later in the case of $\mathbb{R}^3$, being the dimension of $\mathcal{A}_2(\mathbb{R}^3) = 3$, we prefer the cyclic ordering $\{dx_2 \wedge dx_3, dx_3 \wedge dx_1, dx_1 \wedge dx_2\}$,

or alternatively $\{dy \wedge dz, dz \wedge dx, dx \wedge dy\}$.

After the algebraic part we are ready for the analytic definition.

Definition 10.4.1. *A differential form ω of degree 2 (also called 2-form) is a function defined on a open subset of X with values on $\mathcal{A}_2(X)$.*

As in the case of 1-forms we are interested in 2-forms which are at least continuous, and often differentiable or with further regularity. This regularity is revealed by the scalar coefficient functions in the case of finite dimension

$$\omega = \sum_{1 \leq i < j \leq n} f_{ij}\, dx_i \wedge dx_j.$$

Given a differentiable 1**-form** $\omega = \sum_{i=1}^{n} f_i\, dx_i$ its *exterior derivative* is the 2-form defined as

$$d\omega = \sum_{1 \leq i < j \leq n} \left(\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} \right) dx_i \wedge dx_j.$$

It is not difficult to see that the operation of exterior differentiation is linear, and moreover it satisfies the following analogous of the Leibniz differentiation rule: given a function f and ω a 1-differential form then

$$d(f\omega) = df \wedge \omega + f d\omega$$

were the properties of the operation $\wedge$ are used when it comes to simplification.

Note that with the help of exterior differentiation we can rewrite the condition of closeness for 1-forms: the 1-form ω is closed if and only if $d\omega = 0$.

Now we are going to discuss exactness and closeness for 2-forms. Let ω be a 2-form. We say that ω is *exact* if there is a 1-form α such that $\omega = d\alpha$. As in the case of exact 1-forms, the 2-forms which are exact satisfy a sort identity with partial derivatives of the coefficients. Firstly, consider an exact 2-form in $\mathbb{R}^3$, whose primitive $f_1 dx_1 + f_2 dx_2 + f_3 dx_3$ is C^2, and so the 2-form can be written as

$$\left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3} \right) dx_2 \wedge dx_3 + \left(\frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1} \right) dx_3 \wedge dx_1 + \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right) dx_1 \wedge dx_2.$$

Note now that the following identity holds

$$\frac{\partial}{\partial x_1}\left(\frac{\partial f_3}{\partial x_2}-\frac{\partial f_2}{\partial x_3}\right)+\frac{\partial}{\partial x_2}\left(\frac{\partial f_1}{\partial x_3}-\frac{\partial f_3}{\partial x_1}\right)+\frac{\partial}{\partial x_3}\left(\frac{\partial f_2}{\partial x_1}-\frac{\partial f_1}{\partial x_2}\right)=0.$$

In other words (and standard notation in $\mathbb{R}^3$), if $\omega = Ady\wedge dz + Bdz\wedge dx + Cdx\wedge dy$ is exact then

$$\frac{\partial A}{\partial x}+\frac{\partial B}{\partial y}+\frac{\partial C}{\partial z}=0.$$

The previous equality satisfied by exact 2-forms can be replaced by $\binom{n}{3}$ equalities in dimension n which are the key of the next definition. We say that a 2-form

$$\omega=\sum_{1\le i<j\le n} f_{ij}\,dx_i\wedge dx_j$$

is *closed* if the following equality is satisfied

$$\frac{\partial f_{ij}}{\partial x_k}+\frac{\partial f_{jk}}{\partial x_i}+\frac{\partial f_{ki}}{\partial x_j}=0$$

whenever i,j,k are different integers between 1 and n with the convention that $f_{rs}=-f_{sr}$ in case $r>s$. The expressions above are actually the coefficients of a 3-form, the exterior differential of ω. We are not going into it, but we may infer that $d^2(\alpha)=d(d\alpha)=0$ for every 1-form α, as it was for scalar functions, namely $d^2(f)=d(df)=0$ (always under the hypothesis of being C^2). Therefore the iteration of the operation of exterior differential is always 0 (it is not the case for the standard differential).

The Lemma of Poincaré is true also for 2-forms, that is closed 2-forms defined on star-shaped domains are exact. Instead of proving that we will provide a method to compute primitives of 2-forms in $\mathbb{R}^3$. Consider the form $\omega = Ady\wedge dz+Bdz\wedge dx+Cdx\wedge dy$ where A,B,C are functions of x,y,z. The objetive is to eliminate z form both the functions and the basis of 2-forms. Firstly consider an 1-form $\alpha = pdx+qdy$ (p,q are functions of x,y,z). Its exterior differential is

$$d\alpha=\left(\frac{\partial q}{\partial x}-\frac{\partial p}{\partial y}\right)dx\wedge dy-\frac{\partial q}{\partial z}dy\wedge dz+\frac{\partial p}{\partial z}dz\wedge dx.$$

Now we are going to compute p,q so $\omega-d\alpha$ only contains the $dx\wedge dy$ term. For that it is necessary these two equations be fulfilled

$$A=-\frac{\partial q}{\partial z};\quad B=\frac{\partial p}{\partial z}$$

what turns out possible with partial primitives. From now on the functions p, q are supposed known. We have

$$\omega - d\alpha = \left(C - \frac{\partial q}{\partial x} + \frac{\partial p}{\partial y} \right) dx \wedge dy.$$

We claim that the function between brackets does not depend on z. We will compute its partial derivative with respect to z

$$\frac{\partial}{\partial z}\left(C - \frac{\partial q}{\partial x} + \frac{\partial p}{\partial y} \right) = \frac{\partial C}{\partial z} - \frac{\partial^2 q}{\partial x \partial z} + \frac{\partial^2 p}{\partial y \partial z} = \frac{\partial C}{\partial z} + \frac{\partial A}{\partial x} + \frac{\partial B}{\partial y} = 0$$

where the hypothesis of being ω closed is used by the first time. Once we know that the function between brackets does not contain z the problem is reduced to dimension 2 where to find a primitive is not difficult by partial integration as we may assume that the primitive is of the form $f(x, y)dx$ (or $g(x, y)dy$).

10.5 Integration of 2-forms on surfaces

We will consider in first place parameterized C^1 surfaces with boundary embedded into $\mathbb{R}^n$ which can be described by an injective function $\Gamma : D \to \mathbb{R}^n$ and the following conditions:

1. $D \subset \mathbb{R}^2$ is compact with C^1 boundary;
2. there is an open set $D \subset C \subset \mathbb{R}^2$ where Γ extends as C^1 function;
3. $\frac{\partial \Gamma}{\partial u}(u, v)$ and $\frac{\partial \Gamma}{\partial v}(u, v)$ are linearly independent for every $(u, v) \in D$.

The integral of a 2-form of parameterized C^1 surface with boundary Γ is defined when $\Gamma(D)$ is contained into the domain of ω by the formula

$$\int_\Gamma \omega = \iint_D \omega\big(\Gamma(u, v)\big)\Big(\frac{\partial \Gamma}{\partial u}(u, v), \frac{\partial \Gamma}{\partial v}(u, v)\Big)\, du dv.$$

Firstly, note that if we interchange the role of the variables (u, v) taking $\tilde{\Gamma}(v, u) = \Gamma(u, v)$ defined on $\tilde{D} = \{(v, u) : (u, v) \in D\}$ then the value of the integral change multiplied by -1. Indeed, this is consequence of the antisymmetry of ω. This phenomenon is the analogous of the change of sign in the integral of 1-forms when the path is walked backwards. The principle behind is that parameterized surfaces (the ones we are considering) can given

an "orientation" that plays a role similar to the orientation of curves. In the case of surfaces embedded into $\mathbb{R}^3$ having an orientation is simply to distinguish between the to "faces" of the surface, as for instance we can distinguish between up and down when the surface is given as the graph of a function of two variables.

Another issue we have to deal with is the to prove that the notion integral for 2-forms does not depend on the particular choice of the parameterization but on the shape $\Gamma(D)$ of the surface together with the orientation, that means, a similar statement to Proposition 10.2.1.

Proposition 10.5.1. *Let $\Gamma : D \to \mathbb{R}^n$ a C^1 surface with boundary, ω a continuous 2-form defined on a set containing $\Gamma(D)$ and $h : \tilde{D} \to D$ a C^1 bijection with positive jacobian. Then $\tilde{\Gamma} : \tilde{D} \to \mathbb{R}^n$ is a piecewise C^1 surface and*

$$\int_{\tilde{\Gamma}} \omega = \int_{\Gamma} \omega$$

Proof. Writing the change of variables as $h(s,t) = (u(s,t), v(s,t))$ and substituting into the expression of the first integral (some variables are omitted for the sake of readability) we have

$$\iint_{\tilde{D}} \omega\big(\tilde{\Gamma}(s,t)\big)\Big(\frac{\partial\tilde{\Gamma}}{\partial s}(s,t), \frac{\partial\tilde{\Gamma}}{\partial t}(s,t)\Big)\, dsdt =$$

$$\iint_{\tilde{D}} \omega\big(\tilde{\Gamma}(s,t)\big)\Big(\frac{\partial\Gamma}{\partial u}\frac{\partial u}{\partial s} + \frac{\partial\Gamma}{\partial v}\frac{\partial v}{\partial s}, \frac{\partial\Gamma}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial\Gamma}{\partial v}\frac{\partial v}{\partial t}\Big)\, dsdt =$$

$$\iint_{\tilde{D}} \omega\big(\tilde{\Gamma}(s,t)\big)\Big(\frac{\partial\Gamma}{\partial u}, \frac{\partial\Gamma}{\partial u}\Big)\Big(\frac{\partial u}{\partial s}\frac{\partial v}{\partial t} - \frac{\partial v}{\partial s}\frac{\partial u}{\partial t}\Big)\, dsdt =$$

$$\iint_{\tilde{D}} \omega\big(\Gamma(u(s,t),v(s,t))\big)\Big(\frac{\partial\Gamma}{\partial u}(u(s,t),v(s,t)), \frac{\partial\Gamma}{\partial v}(u(s,t),v(s,t))\Big)\left|\frac{\partial(u,v)}{\partial(s,t)}\right| dsdt$$

$$= \iint_{D} \omega\big(\Gamma(u,v)\big)\Big(\frac{\partial\Gamma}{\partial u}(u,v), \frac{\partial\Gamma}{\partial v}(u,v)\Big)\, dudv$$

where in the first equality we have used the chain rule for derivatives, in the second equality the bilinearity and antisymmetry of the 2-form, third equality is just to make explicit the involved variables and finally the last equality is due to the change of variables formula for the integral. ∎

Now we will consider a more general type of surfaces. We say that connected set in $\mathbb{R}^n$ is an *oriented piecewise C^1 surface* if it is the union of the images of finitely many parameterized C^1 surfaces with border, those surfaces can only intersect on points of their borders and the intersection when happens is a non trivial curve, and finally the orientations induced by the parameterizations on each piece are compatible. This is something with a clear meaning for surfaces in $\mathbb{R}^3$ thinking of orientation with the help of the normal vector field. In this 3-dimensional setting there is an important example. Asume that a compact set with nonempty interior has a boundary which is made up of finitely many parameterized surfaces (with boundary). Then there is a natural standard orientation: the normal field points to the exterior of the set. Thus, in case of an oriented piecewise C^1 surface $\Gamma = \Gamma_1 + \cdots + \Gamma_m$ where Γ_k are parameterized C^1 pieces we define

$$\int_\Gamma \omega = \sum_{k=1}^{m} \int_{\Gamma_k} \omega$$

for any 2 form defined on a domain containing Γ. It is not difficult to check that the definition does not depends on how Γ is decomposed into C^1 parameterized pieces. For instance, the sphere needs such a decomposition and it can be done of infinitely many fashions. Moreover, removing one point of the sphere, the remainder is a parameterized surface and one point less does not bother when it comes to integration.

10.6 Gauss and Stokes

This section will be developed in $\mathbb{R}^3$ providing analogous and related results to the Green-Riemann formula. We say that a bounded open domain $E \subset \mathbb{R}^3$ with C^1 boundary is elemental with respect to the XY plane if its boundary is contained in the "cylinder" $\{(x, y, z) : (x, y) \in \partial D\}$ where D is the orthogonal projection of E onto the XY plane, and the graphs of two C^1 functions $f, g : D \to \mathbb{R}$ with $f > g$. Analogously elemental domains with respect to the YZ and the XZ planes are defined.

Lemma 10.6.1. *Let $E \subset \mathbb{R}^3$ be an elemental domain with respect to the XY plane and $R(x, y, z)$ a C^1 function defined on a domain containing $\overline{E}$. Then*

$$\int_{\partial E} R\, dx \wedge dy = \iiint_E \frac{\partial R}{\partial z}\, dxdydz$$

Proof. The integral of the 2-form on those parts of ∂E contained in the "cylinder" is null (from the geometrical point of view the field $R\,dx\wedge dy$ is vertical meanwhile the normal vectors of the cylinder are horizontal). The upper and lower parts of the domain are given by the parameterizations $\Gamma_1(x,y) = (x,y,f(x))$ and $\Gamma_2(x,y) = (x,y,g(x,y))$ with $(x,y)\in D$, where the second has to be reversed to be according to the orientation (towards the exterior). As we have

$$(dx\wedge dy)\Big(\frac{\partial\Gamma_1}{\partial x},\frac{\partial\Gamma_1}{\partial y}\Big) = (dx\wedge dy)\Big(\frac{\partial\Gamma_2}{\partial x},\frac{\partial\Gamma_2}{\partial y}\Big) = 1$$

then

$$\int_{\partial E} R\,dx\wedge dy = \iint_D R(x,y,f(x,y))\,dxdy - \iint_D R(x,y,g(x,y))\,dxdy =$$

$$\iint_D \big(R(x,y,f(x,y)) - R(x,y,g(x,y))\big)\,dxdy = \iint_D \left(\int_{g(x,y)}^{f(x,y}\frac{\partial R}{\partial z}\,dz\right)dxdy$$

$$= \iiint_E \frac{\partial R}{\partial z}\,dxdydz$$

as wanted. ∎

Since the analogous results are true for elemental domains with respect to the YZ and XZ planes we have the following.

Theorem 10.6.2 (Gauss-Ostrogradsky)**.** *Let $E\subset\mathbb{R}^3$ be a domain which is elemental with respect to the three planes XY, YZ and XZ and let $Pdy\wedge dz + Qdz\wedge dx + Rdx\wedge dy$ a 2-form which is C^1 on a domain containing $\overline{E}$. Then*

$$\int_{\partial E} P\,dy\wedge dz + Q\,dz\wedge dx + R\,dx\wedge dy =$$

$$\iiint_E \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right)dxdydz$$

Moreover, the same formula hold if E is a domain such that it can be decomposed into a finite non overlapping union of domains with C^1 boundaries which are elemental with respect to the three coordinate planes.

Proof. The result is just the sum of the three equalities

$$\int_{\partial E} P\,dy\wedge dz = \iiint_E \frac{\partial P}{\partial x}\,dxdydz$$

$$\int_{\partial E} Q\,dz \wedge dx = \iiint_E \frac{\partial Q}{\partial y}\,dxdydz$$
$$\int_{\partial E} R\,dx \wedge dy = \iiint_E \frac{\partial R}{\partial z}\,dxdydz$$

where the last one come from the previous lemma and the two other ones are the analogous that can be obtained switching the coordinate planes. ■

Remark 10.6.3. *The theorem of Gauss-Ostrogradsky can be proved with a similar degree of generality that the Green-Riemann theorem, but the extra work is not worth at all.*

Now we will obtain a result which relates the integration of an 1-form along the relative boundary (the "free points" of the boundaries of the pieces) of an oriented piecewise C^1 surface and the integral of its exterior differential over that surface. Firstly, given an oriented piecewise C^1 surface S we have to assign an orientation to the relative boundary ∂S. That will be the anticlockwise orientation when we look at the surface from "above", that is, from the part the normal vectors points towards. We will say that a piece of the surface is flat with respect to the plane XY if it can be represented as the graph of a function $z = f(x, y)$.

Theorem 10.6.4 (Stokes)**.** *Let S be an oriented piecewise C^2 oriented surface and let ω an 1-form which C^1 on a domain containing S. Then*

$$\int_{\partial S} \omega = \int_S d\omega.$$

Proof. Decomposing the surface into C^2 pieces we just have to prove the result for each piece which are C^2 surfaces with boundary. Indeed, the surface integrals are additive and the integral of the 1-form vanishes on the shared parts of the relative boundary. With the help of the implicit function theorem we may decompose the surface into smaller flat pieces (remember that the function representing the surface can be enlarged smoothly beyond its domain). Therefore we may assume that S is C^2 flat with respect to XY (the other two orientations are obtained likewise). Assume now that S is represented as $z = f(x, y)$ with $(x, y) \subset D$ a domain with C^1 boundary. In order to proof the result we are going to develop both members of the equality. Put $\omega = Pdx + Qdy + Rdz$ and $(X(t), Y(t))$ with $t \in [a, b]$ a parameterization of the border. Thus we have

$$\int_{\partial S} \omega =$$

$$\int_a^b \big(P(\cdot)X'(t) + Q(\cdot)Y'(t) + R(\cdot)(\frac{\partial f}{\partial x}(\cdot\cdot)X'(t) + \frac{\partial f}{\partial x}(\cdot\cdot)Y'(t))\big)\,dt =$$

$$\int_a^b \big((P(\cdot) + R(\cdot)\frac{\partial f}{\partial x}(\cdot\cdot))X'(t) + (Q(\cdot) + R(\cdot)\frac{\partial f}{\partial y}(\cdot\cdot))Y'(t)\big)\,dt =$$

$$\int_a^b \big(p(\cdot)X'(t) + q(\cdot)Y'(t)\big)dt = \int_{\partial D} p\,dx + q\,dy$$

where $(\cdot) = (X(t), Y(t), f(X(t), Y(t)))$, $(\cdot\cdot) = (X(t), Y(t))$ and

$$p(x,y) = P(x,y,f(x,y)) + R(x,y,f(x,y))\frac{\partial f}{\partial x}(x,y),$$

$$q(x,y) = Q(x,y,f(x,y)) + R(x,y,f(x,y))\frac{\partial f}{\partial y}(x,y).$$

Therefore

$$\int_{\partial S} \omega = \int_{\partial D} p\,dx + q\,dy = \iint_D \Big(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\Big)\,dxdy$$

where the last equality is thanks to the Green-Riemann formula. Now we have to compute

$$\frac{\partial p}{\partial y} = \frac{\partial P}{\partial y} + \frac{\partial P}{\partial z}\frac{\partial f}{\partial y} + \left(\frac{\partial R}{\partial y} + \frac{\partial R}{\partial z}\frac{\partial f}{\partial y}\right)\frac{\partial f}{\partial x} + R\frac{\partial^2 f}{\partial x \partial y}$$

$$\frac{\partial q}{\partial x} = \frac{\partial Q}{\partial x} + \frac{\partial Q}{\partial z}\frac{\partial f}{\partial x} + \left(\frac{\partial R}{\partial x} + \frac{\partial R}{\partial z}\frac{\partial f}{\partial x}\right)\frac{\partial f}{\partial y} + R\frac{\partial^2 f}{\partial y \partial x}$$

where the variables have been removed for sake of better readability. Therefore

$$\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} = \frac{\partial Q}{\partial x} + \frac{\partial Q}{\partial z}\frac{\partial f}{\partial x} + \frac{\partial R}{\partial x}\frac{\partial f}{\partial y} - \frac{\partial P}{\partial x} - \frac{\partial P}{\partial z}\frac{\partial f}{\partial y} - \frac{\partial R}{\partial y}\frac{\partial f}{\partial x} = (*)$$

Now we are going to compute the surface integral of the statement. Firstly

$$d\omega = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) dy \wedge dz + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) dz \wedge dx + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx \wedge dy$$

that implies

$$d\omega(U,V) = -\frac{\partial R}{\partial y}\frac{\partial f}{\partial x} + \frac{\partial Q}{\partial z}\frac{\partial f}{\partial x} - \frac{\partial P}{\partial z}\frac{\partial f}{\partial y} + \frac{\partial R}{\partial x}\frac{\partial f}{\partial y} + \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial x} = (**)$$

where $U = (1, 0, \frac{\partial f}{\partial x})$ and $V = (0, 1, \frac{\partial f}{\partial y})$. The equality $(*) = (**)$ completes the proof of the theorem. ■

Remark 10.6.5. *The hypothesis C^2 in the last theorem contrasts with the C^1 assumption in previous results. This is a consequence of the chosen method of proof. And the result can be proved under more relaxed hypotheses.*

Many of the previous results can be expressed in terms of the relation between the integrals of a $(k-1)$-form and it exterior differential, which is a k-form, on the $(k-1)$-dimensional smooth boundary of a k-dimensional object, respectively. Now we are ready for this new point of view:

1. Let $\gamma : [a, b] \to \mathbb{R}^n$ be a parameterized piecewise C^1 curve (injective). Its relative boundary is $\partial\gamma = \{\gamma(a), \gamma(b)\}$. The orientation of γ induces an orientation on that two points set, that is a distinction. Given a 0-form, that is a scalar function, f we define $\int_{\partial\gamma} f = f(\gamma(b)) - f(\gamma(a))$. With this notation Proposition 10.2.3 becomes

$$\int_{\gamma} df = \int_{\partial\gamma} f.$$

2. If $\omega = p\,dx + q\,dy$ is an 1-form on $\mathbb{R}^2$ which is C^1, its exterior differential is the 2-form

$$d(p\,dx + q\,dy) = dp \wedge dx + dq \wedge dy = \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) dx \wedge dy.$$

   Since 2-forms in $\mathbb{R}^2$ have dimension 1 they are assimilable to scalar functions. Taking $\int_D f dx \wedge dy = \int_D f\,dxdy$, the Green-Riemann formula Theorem 10.3.4 becomes

$$\int_{\partial D} \omega = \int_D d\omega.$$

3. Let $\omega = P\,dy \wedge dz + Q\,dz \wedge dx + R\,dx \wedge dy$ be a 2-form in $\mathbb{R}^3$. Its exterior differential ia 3-form, which is assimilable to a scalar function because there is only one basic element in dimension 3, namely $dx \wedge dy \wedge dz$. The formula for the exterior differential was insinuated in Section 4

$$d\omega = \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dx \wedge dy \wedge dz.$$

   After this, the Gauss-Ostrogradsky Theorem 10.6.2 becomes

$$\int_{\partial E} \omega = \int_E d\omega.$$

4. And, of course, Stokes Theorem 10.6.4 itself follows the same scheme.

All these results can be summed up into a general Cartan-Stokes Theorem proved in the frame of the theory of differential forms on manifolds.

10.7 Rationale and remarks

According to the syllabus, we will consider only the integration of 1-forms and 2-forms. Eventually, scalar functions will be considered 0-forms, and 3-forms appear implicitly in the necessary conditions for a 2-form to be exact. Therefore, the exterior multiplication is only defined for 1-forms and the definition of exterior derivative is a restricted one.

The development of the theory is quite standard. Only two comments: the topological facts are not much stressed (no mention of simply connected domains, nor homotopy neither homology), and the proof of the Green-Riemann formula is a real one, that means, it is not based on an *a priori* existence of a nice decomposition of the domain. Such a struggle is not repeated for the $\mathbb{R}^3$ theorems, though.

An interesting comment could be that some differential equations that have local solutions may have or not global solutions depending on the domain where they are considered, and that is a purely topological matter (suggest they look for information on the *Rham cohomology*).

10.8 Exercises

1. Calculate the integral
$$\int_\gamma y\,dx - x\,dy$$
being γ the triangle with vertices $(0,0)$, $(1,0)$ and $(0,1)$ orientated likewise.

2. Calculate the integral
$$\int_\gamma (y-z)dx + (z-x)dy + (x-y)dz$$

being γ e the triangle with vertices $(a,0,0)$, $(0,b,0)$ and $(0,0,c)$ with $a,b,c>0$ orientated likewise.

3. Calculate the integral

$$\int_\gamma z^2dx + x^2dy + y^2dz,$$

being γ the spherical triangle with vertices $(a,0,0)$, $(0,a,0)$ and $(0,0,a)$, on the sphere centred at $(0,0,0)$ with radius $a>0$.

4. Consider the differential form

$$\omega(x,y,z) = \frac{2x}{z}\,dx + \frac{2y}{z}\,dy + \Big(1 - \frac{x^2+y^2}{z^2}\Big)\,dz,$$

defined on the set

$$A = \{(x,y,z)\in\mathbb{R}^3 \colon z\neq 0\}.$$

Show that it is exact and find all its primitives.

5. Find all the functions $\varphi,\psi\in C^1(\mathbb{R})$ with $\psi(0)=0$ such that the differential form

$$\omega(x,y,z) = (z+z^2)\,dx + \varphi(y)\,\psi(z)\,dy + \big(x + 2z(x+y^2)\big)\,dz,$$

is exact. Then, find all its primitives.

6. Interpret geometrically the integral of the differential form

$$\omega = \frac{-ydx + xdy}{x^2+y^2},$$

along a closed curve counterclockwise that encloses the origin. Then deduce the value of the following integral

$$\int_0^{2\pi} \frac{dt}{a^2\cos^2 t + b^2\sin^2 t}.$$

7. Calculate

$$\iint_\phi z\,dx\wedge dy,$$

where ϕ is the parameterized surface

$$\{\phi(u,v) := (u+v, u^2+v^2, u-v) : u,v\in[-1,1]\}.$$

8. Prove that on a oriented surface M there is a 2-form ω such that for every $N \subset M$ surface with boundary, then the area of N is the absolute value of $\int_N \omega$.

9. Given the 1-form on $\mathbb{R}^3$

$$\omega_1(x_1, x_2, x_3) = x_1\, dx_1 - dx_3$$

$$\omega_2(x_1, x_2, x_3) = 2x_3^2\, dx_1$$

$$\omega_3(x_1, x_2, x_3) = dx_1 - x_2 x_3\, dx_2$$

find $\omega = (2\omega_1 - x_2\omega_3) \wedge \omega_2$.

10. Given $f(x, y) = x^2 y$, $g(x, y) = xy^2$, find the simplest expression for

$$\omega(x, y) = df(x, y) \wedge dg(x, y).$$

11. Given two differentiable functions f and g on $\mathbb{R}^3$, assume that

$$df \wedge dg = \lambda\, dx \wedge dy$$

where λ is a non null function on $\mathbb{R}^3$. Prove that f and g depends only on x, y.

12. Compute the exterior derivative of the form

$$\omega(x, y, z) = yz^2(\cos xy)\, dx + xz^2(\cos xy)\, dy + (x + y)\, dz.$$

13. Compute the exterior derivative of the form

$$\omega(x, y, z) = (2xy + y^2)\, dx + (x^2 + 2xy)\, dy + 3z^2\, dz,$$

and interpret the result.

14. Find all the primitives of the 2-form on $\mathbb{R}^2$

$$\omega(x, y) = (2x + y - 3x^2y^2)\, dx \wedge dy.$$

15. Find all the primitives of the 2-form on $\mathbb{R}^3$

$$\omega(x, y, z) = (y^3 - x^3)dx \wedge dy + (x - 2z)dy \wedge dz + (2z - y)dz \wedge dx.$$

16. Given the 1-form

$$\omega(x,y,z) = y\,dx - x\,dy + dz,$$

find conditions on the C^1 functions $u(x,y,z)$ y $v(x,y,z)$ in order to the form $\omega - v\,du$ be closed. Show that u and v do not depend on z.

17. Let $m \geq 0$ and take $r = (x^2+y^2+z^2)^{1/2}$ as usual. Find a function $f(r)$ such that if $\vec{F} = f(r)(x,y,z)$, then $\operatorname{div}(\vec{F}) = r^m$. Apply that to express the integral

$$\iiint_D r^m\,dxdydz$$

in terms of a surface integral on ∂D (regularity is assumed). Find the value for $D = B(0,1)$.

18. Let $f : \mathbb{R}^2 \to \mathbb{R}$ an arbitrary C^1 function and consider the 2-form

$$\omega = f(x,y)\,dx \wedge dy + x^2 y\,dy \wedge dz - xy^2\,dz \wedge dx.$$

Show that ω is exact and find a primitive.

19. Let ω be a 1-form C^1 defined on $\mathbb{R}^3 \setminus \{0\}$. Show that if ω is closed, then it is exact. What is the role of the dimension?

20. Let ω be a 1-form C^1 defined on $\mathbb{R}^n \setminus \{0\}$. Let $R_0 \subset \mathbb{R}^n$ be a half-line starting at 0 such that ω has primitives on the set $\mathbb{R}^n \setminus R_0$. Prove that ω has primitives on $\mathbb{R}^n \setminus R$ for R any half-line starting at 0. Does ω have primitive on $\mathbb{R}^n \setminus \{0\}$?

21. Prove that the following 2-form is closed

$$\frac{x\,dy \wedge dz}{(x^2+y^2+z^2)^{3/2}} + \frac{y\,dz \wedge dx}{(x^2+y^2+z^2)^{3/2}} + \frac{z\,dx \wedge dy}{(x^2+y^2+z^2)^{3/2}},$$

and find a primitive defined on a maximal open subset of $\mathbb{R}^3$.

Chapter 11

Classic Vector Analysis

11.1 Operations with vectors in $\mathbb{R}^3$

The geometrical interpretation in $\mathbb{R}^2$ of the arithmetical operations in the field of complex numbers $\mathbb{C}$ thanks to *Argand's diagram* is of great help in plane geometry. That motivated Hamilton to seek a similar arithmetic structure for $\mathbb{R}^3$. After several barren tries, in 1843 he came out with the idea of defining the operation in $\mathbb{R}^4$ instead of $\mathbb{R}^3$ and giving up with commutativity. He defined the *quaternions* like an extension of the complex numbers (and so of the real ones) as the formal expressions

$$a + bi + cj + dk$$

where i, j, k are imaginary units that interact among them accordingly to these rules

$$i^2 = j^2 = k^2 = -1;\ ij = -ji = k;\ jk = -kj = i,\ ki = -ik = j.$$

The real numbers are interpreted as those such that $b = c = d = 0$, so a is called the *scalar part* and $bi + cj + dk$ is called the *vector part* of the quaternion. The arithmetic operations between quaternions are performed using distributivity, implicitly assumed, in order to reduce the result to the canonical form above with the help of the relationships between units, minding that the order matters. Following this rules, the product turns out to be associative and every non zero element has an inverse (both left and right). Namely, the inverse of $a + bi + cj + dk$ is

$$\frac{a - bi - cj - dk}{a^2 + b^2 + c^2 + d^2}.$$

There is an obvious analogy with complex numbers. The term on the numerator is called *conjugate* and the real number on the denominator is the square of the *modulus*. As it happens with the modulus of complex numbers, the modulus is multiplicative.

Now we will consider quaternions with *real* part zero, which are called *purely imaginary* and can be interpreted into $\mathbb{R}^3$. The product of two purely imaginary quaternions is not purely imaginary in general

$$(x_1 i + y_1 j + z_1 k)(x_2 i + y_2 j + z_2 k) =$$

$$-(x_1 x_2 + y_1 y_2 + z_1 z_2) + (y_1 z_2 - z_1 y_2)i + (z_1 x_2 - x_1 z_2)j + (x_1 y_2 - y_1 x_2)k$$

The scalar part of the result, after changing the sign can be identified with the Euclidean scalar product of vectors. The vector part of the product is called the *vector product*. If u, v are quaternions with null scalar part its product is can be represented as

$$uv = -u \cdot v + u \times v$$

being $u \times v$ the vector product. Since the modulus is multiplicative we have

$$|u|^2 |v|^2 = |u \cdot v|^2 + |u \times v|^2$$

It is not difficult to check that

$$u \cdot (u \times v) = v \cdot (v \times u) = 0$$

which means that $u \times v$ is orthogonal to both u and v, so its direction is well determined in $\mathbb{R}^3$ if u and v are independent. We also have a consequence of the non commutativity: $u \times v = -v \times u$.

Some time time after the discovery of the quaternions it was clear that in order to deal with the Euclidean geometry of $\mathbb{R}^3$ it is not necessary the full power of its the algebraic structure. We can work more easily in that frame just keeping the scalar and vector products once we know their properties. Let us note that the easiest method to compute the vector product without appealing to quaternions is the following symbolic determinant

$$u \times v = \begin{vmatrix} i & j & k \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

where $u(u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$. Since most of the vectors will be crowned with a little arrow in the following sections, we will denote from now on the basis of $\mathbb{R}^3$ derived from the quaternions by $\{\vec{i}, \vec{j}, \vec{k}\}$.

11.2 Differential forms on $\mathbb{R}^3$

Along this chapter we will consider *scalar* and *vector fields* in $\mathbb{R}^3$. These are simply real functions and functions with values in $\mathbb{R}^3$ defined on some open domain of $\mathbb{R}^3$, often the whole space. We will follow this terminology (fields) in order to stress the fact the different nature of the domain, which is made of points, and the range which can be made of either numbers or vectors. As to vector fields comes, it is worth noticing that it can be interpreted both as differential 1-forms or 2-forms. Firstly we will establish the identification between the vectors of $\mathbb{R}^3$ and the *alternate forms* of degrees 1 and 2 on $\mathbb{R}^3$, whose respective spaces on have dimension 3. Such an identification can be done with the help of basis, namely

$$a\vec{i} + b\vec{j} + c\vec{k} \longleftrightarrow a\,dx + b\,dy + c\,dz;$$

$$a\vec{i} + b\vec{j} + c\vec{k} \longleftrightarrow a\,dy \wedge dz + b\,dz \wedge dx + c\,dx \wedge dy.$$

These associations are canonical in the sense that the actions of the forms on vectors $\vec{u}, \vec{v} \in \mathbb{R}^3$, being $\vec{u} = (u_1, u_2, u_3), \vec{v} = (v_1, v_2, v_3)$, can be represented by the previous vector products for 1-forms as

$$(a\,dx + b\,dy + c\,dz)(u) = (a\vec{i} + b\vec{j} + c\vec{k}) \cdot u = au_1 + bu_2 + cu_3$$

and for 2-forms as follows

$$(a\,dy \wedge dz + b\,dz \wedge dx + c\,dx \wedge dy)(u, v) =$$

$$(a\vec{i} + b\vec{j} + c\vec{k}) \cdot (\vec{u} \times \vec{v}) = \begin{vmatrix} a & b & c \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}.$$

The proof of the these equalities can reduced to check them on pairs or triplets of basic vectors thanks to the linearity. For instance, the second one on the triplet $(\vec{i}, \vec{j}, \vec{k})$ we have $(dy \wedge dz)(\vec{j}, \vec{k}) = 1$ and $\vec{i} \cdot (\vec{j}, \vec{k}) = \vec{i} \cdot \vec{i} = 1$.

Once we know how to switch from differential forms language to vector analysis language we can establish a relation between the integration of forms and the integration on curves and surfaces (see the corresponding part in previous chapters). Using the relations above any differential form ω of degrees

1 or 2 defined on a domain of $\mathbb{R}^3$ can be transformed into a vector field $\overrightarrow{F}$. If $\gamma(t)$ is a parameterized curve we have

$$\int_\gamma \omega = \int_a^b \omega(\gamma(t))(\gamma'(t))\,dt = \int_a^b \overrightarrow{F}(\gamma(t)) \cdot \gamma'(t)\,dt = \int_\gamma \overrightarrow{F} \cdot d\overrightarrow{\ell}\,.$$

Now, if Γ is a parameterized C^1 surface with boundary (with domain D) and ω is a 2 form that is identified with the a vector fiel $\overrightarrow{F}$ then

$$\int_\Gamma \omega = \iint_D \omega(\Gamma)\left(\frac{\partial\Gamma}{\partial u}, \frac{\partial\Gamma}{\partial v}\right) du dv = \iint_D \overrightarrow{F} \cdot \left(\frac{\partial\Gamma}{\partial u} \times \frac{\partial\Gamma}{\partial v}\right) du dv = \iint_\Gamma \overrightarrow{F} \cdot d\overrightarrow{S}\,.$$

These transformations show that the integration of forms can be expressed actually as integration with respect to the intrinsic measures either for curves or surfaces. That implies the already proven result that the integration of forms is invariant by change of parameterization but an eventually change of sign in case the orientation is reversed.

11.3 Vector operators

The *exterior differential* acting on differential forms on $\mathbb{R}^3$ can adopt several forms, after the identification with fields of the previous section, called *vector (differential) operators*. Since the correspondence is done through a choice of an orthonormal basis, the appearance of the differential operators strongly relies on the associate coordinates "x, y, z" and the associated partial derivations which may give the false impression that the canonical basis and the cartesian coordinates are privileged. For that reason, we will stress the fact that the vector operators are *intrinsic*, that is, they do not depend on the choice of coordinates. This could be done by direct computation on an orthogonal change of coordinates, or proving that the exterior differentiation is intrinsic. However we will choose alternative methods which moreover cast light on the geometrical or physical meaning of the vector operators which is basic for the applications.

Under the hypothesis of some regularity, there are some operations involving differentiation that can be performed to scalar and vector fields. These operations can be labelled with the help of a symbolic operator named *nabla*

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) = \frac{\partial}{\partial x}\,\vec{i} + \frac{\partial}{\partial y}\,\vec{j} + \frac{\partial}{\partial z}\,\vec{k}.$$

The first operation we will consider is well known: the gradient. Let us recall that gradient of a scalar field (function) f is the vector field defined by

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right).$$

Despite the fact that the gradient is defined in term of the cartesian coordinates associated, it has an intrinsic meaning. Indeed, its modulus is the maximum value of the directional derivative over all the norm one vectors, and provided it is not zero, the gradient points in the maximizing direction.

Given a vector field $\overrightarrow{F} = (f_1, f_2, f_3)$, its *divergence* is the scalar field defined by

$$\nabla \cdot \overrightarrow{F} = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}.$$

It is not obvious from this definition that the divergence is an intrinsic notion. That can be deduced by alternative methods, as straight computation. It is easier to remark that if we identify the vector field with a (differential) 2-form, its divergence can be identified with its exterior differential. Therefore, if we know that the exterior differential is an intrinsic notion independent from the coordinate system then the same is true for the divergence. The third method we will give also provides an interpretation of the divergence. The Gauss-Ostrogradsky theorem says wit this notation that if $D \subset \mathbb{R}^3$ is a bounded open domain and $\overrightarrow{F}$ is C^1 on a domain which includes $\overline{D}$ then

$$\iint_{\partial D} \overrightarrow{F} \cdot d\overrightarrow{S} = \iiint_D \nabla \cdot \overrightarrow{F} \, dV.$$

For that reason divergence can be interpreted as the net rate of the flux leaving/entering a small volume around the point

$$\nabla \cdot \overrightarrow{F}(x) = \lim_{\varepsilon \to 0^+} \frac{\iint_{\partial B(x,\varepsilon)} \overrightarrow{F} \cdot d\overrightarrow{S}}{\mathrm{Vol}(B(x,\varepsilon))}.$$

Suppose that the vector field represents the speeds of a fluid. If the fluid is incompressible that implies that the net rate of the flux is 0 as the amount of fluid entering the ball equals that one getting out, so the divergence is 0. If the is not incompressible then the divergence represents variations in density at that point. In other interpretations of vector fields the divergence represents a magnitude related to the field that is created/destroyed at the point. For

instance, the divergence of the static electric force field represents the charge per unit volume.

Given a vector field $\overrightarrow{F} = (f_1, f_2, f_3)$, its *rotational* or *curl* is the vector field defined by

$$\nabla \times \overrightarrow{F} = \left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}, \ \frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x}, \ \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right).$$

Note that if we identify the vector field $\overrightarrow{F}$ with a n 1-form ω then $\nabla \times \overrightarrow{F}$ can be identify with $d\omega$. That shows that the definition of the rotational is intrinsic. In order to have an interpretation we have to appeal to Stokes' theorem, which can be rewritten in those terms. Let Γ be an oriented parameterized C^2 surface with boundary in $\mathbb{R}^3$ then

$$\int_{\partial\Gamma} \overrightarrow{F} \cdot d\overrightarrow{\ell} = \iint_{\Gamma} \nabla \times \overrightarrow{F} \cdot d\overrightarrow{S}.$$

Consider a disc $D(x, \varepsilon, \overrightarrow{n})$ with center at x, radius ε and contained in a plane perpendicular to a norm one vector $\overrightarrow{n}$. Then provided that the orientation of the disc is the one of $\overrightarrow{n}$ we have

$$\nabla \times \overrightarrow{F}(x) \cdot \overrightarrow{n} = \lim_{\varepsilon \to 0^+} \frac{\int_{\partial D(x,\varepsilon,\overrightarrow{n})} \overrightarrow{F} \cdot d\overrightarrow{\ell}}{\pi \varepsilon^2}.$$

If $\overrightarrow{F}$ is a force field, then $\int_{\partial D(x,\varepsilon,\overrightarrow{n})} \overrightarrow{F} \cdot d\overrightarrow{r}$ is the work done along a closed circuit. In case the field is conservative then this number is 0,which means that a mass moving along the circle by effect of the force do not gain kinetic energy after a doing a turn. If we place a still tiny wheel at x whose axis is align with $\overrightarrow{n}$ it will not turn in a conservative field. However, if the field is not a conservative one, the wheel will turn with an impulse proportional to $\nabla \times \overrightarrow{F}(x) \cdot \overrightarrow{n}$.

Now we will consider an operator that involves second order derivatives, the *Laplacian*, which is actually a combination of gradient and divergence. Given a twice differentiable scalar field take

$$\Delta f = \nabla \cdot (\nabla f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$$

The Laplacian is also represented as $\nabla^2 f$. The intrinsicness of the Laplacian is clear, however the combination of the interpretations of the gradient and the

divergence do not cast light on the what the Laplacian means. For that reason we are going to build a straight one. Assume that f has a Taylor development around $0 = (0, 0, 0)$, for simplicity, of second order of the form

$$f(x, y, z) = f(0) + \frac{\partial f}{\partial x}(0)x + \frac{\partial f}{\partial y}(0)y + \frac{\partial f}{\partial z}(0)z +$$

$$\frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}(0)x^2 + \frac{\partial^2 f}{\partial y^2}(0)y^2 + \frac{\partial^2 f}{\partial z^2}(0)z^2 +\right.$$

$$\left. 2\frac{\partial^2 f}{\partial x \partial y}(0)xy + 2\frac{\partial^2 f}{\partial x \partial z}(0)xz + 2\frac{\partial^2 f}{\partial y \partial z}(0)yz\right) + o(\|(x, y, z)\|^2).$$

The integration over the sphere $\partial B(0, \varepsilon)$ leads to the cancellation of the first order terms and the mixed ones (those containing xy,xz and yz), so it remains

$$\iint_{\partial B(0,\varepsilon)} f \, dS = 4\pi\varepsilon^2 f(0) + \frac{4\pi\varepsilon^4}{6}\frac{\partial^2 f}{\partial x^2}(0) + \frac{4\pi\varepsilon^4}{6}\frac{\partial^2 f}{\partial y^2}(0) + \frac{4\pi\varepsilon^4}{6}\frac{\partial^2 f}{\partial z^2}(0) + o(\varepsilon^4)$$

$$= 4\pi\varepsilon^2 f(0) + \frac{4\pi\varepsilon^4}{6}\Delta f(0) + o(\varepsilon^4).$$

The trick to compute easily the integrals was the following: by symmetry we obviously have

$$\iint_{\partial B(0,\varepsilon)} x^2 \, dS = \iint_{\partial B(0,\varepsilon)} y^2 \, dS = \iint_{\partial B(0,\varepsilon)} z^2 \, dS$$

but

$$\iint_{\partial B(0,\varepsilon)} x^2 \, dS + \iint_{\partial B(0,\varepsilon)} y^2 \, dS + \iint_{\partial B(0,\varepsilon)} z^2 \, dS = \iint_{\partial B(0,\varepsilon)} \varepsilon^2 \, dS = 4\pi\varepsilon^4.$$

Therefore the average of $f(x, y, z) - f(0, 0, 0)$ over the sphere for $\varepsilon > 0$ small is

$$\frac{1}{4\pi\varepsilon^2}\iint_{\partial B(0,\varepsilon)} (f - f(0)) \, dS = \frac{\varepsilon^2}{6}\Delta f(0) + o(\varepsilon^2).$$

Therefore, the Laplacian measures the difference between the value of the function on a point and its average around the point. The functions whose Laplacian is null are called *harmonic* and we will see later that the value at a given point is actually the average of the values on centred spheres.

The last operator we will consider is the Laplacian of a vector field, that appears in applications to Electromagnetism. If $\overrightarrow{F} = (f_1, f_2, f_3)$ then we define

$$\Delta \overrightarrow{F} = (\Delta f_1, \Delta f_2, \Delta f_3).$$

The fact that this definition is intrinsic is consequence of the following identity whose proof is left to the reader

$$\Delta \overrightarrow{\mathbf{A}} = \nabla(\nabla \cdot \overrightarrow{\mathbf{A}}) - \nabla \times (\nabla \times \overrightarrow{\mathbf{A}}).$$

11.4 Newtonian potential

The function

$$f(x, y, z) = \frac{1}{\sqrt{x^2 + y^2 + z^2}}$$

plays a very important role in the study of fields in $\mathbb{R}^3$, not only because it appears relate to classic physical fields as the gravitational and electromagnetic, but also because its study will provide us with theoretical tools to deal with very general mathematical problems.

We will write its derivatives of first and iterated second order

$$\frac{\partial f}{\partial x} = \frac{-x}{(x^2 + y^2 + z^2)^{3/2}}; \quad \frac{\partial f}{\partial y} = \frac{-y}{(x^2 + y^2 + z^2)^{3/2}}; \quad \frac{\partial f}{\partial z} = \frac{-z}{(x^2 + y^2 + z^2)^{3/2}};$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{3(2x^2 - y^2 - z^2)}{(x^2 + y^2 + z^2)^{5/2}}; \quad \frac{\partial^2 f}{\partial y^2} = \frac{3(2y^2 - z^2 - x^2)}{(x^2 + y^2 + z^2)^{5/2}}; \quad \frac{\partial^2 f}{\partial z^2} = \frac{3(2z^2 - x^2 - y^2)}{(x^2 + y^2 + z^2)^{5/2}}.$$

These equalities show that the decreasing rate at infinity is r^{-2} for the first derivatives and r^{-3} for the second order ones (even the non computed). Moreover, clearly $\nabla f = 0$ thus it is harmonic on $\mathbb{R}^3 \setminus \{0\}$. Now we will consider more complicated functions build from the function above. In order to keep some simplicity we will introduce the notation $\overrightarrow{p}$ and $\overrightarrow{r}$ for points of $\mathbb{R}^3$, mainly the first one will be a "free" variable and the second one an "integration" variable. Assume we are given points $\overrightarrow{r}_1, \ldots, \overrightarrow{r}_n$ and numbers $m_1, \ldots, m_n$ that we will call "charges" (or "masses"). The *potential* produced at $\overrightarrow{p}$ by the charges m_k's placed at the points $\overrightarrow{r}_k$'s is

$$\Phi(\overrightarrow{p}) = \sum_{k=1}^{n} \frac{m_k}{\| \overrightarrow{r}_k - \overrightarrow{p} \|}.$$

Note that this function is harmonic except at the singularities $\overrightarrow{r}_k$'s. If we think of the potential produced by many small charges we arrive naturally to a generalization of the potential with the help of integration. Let μ be a finite signed σ-additive Borel measure with compact support. Under these assumptions the potential can be defined for points $\overrightarrow{p}$ out the support of μ as

$$\Phi(\overrightarrow{p}) = \int \frac{d\mu(\overrightarrow{r})}{\|\overrightarrow{r} - \overrightarrow{p}\|}.$$

This function is harmonic on the complement of the support of the measure as the interchange of derivation and integration is not a problem in absence of singularities. However, the potential Φ could be defined at more points if the integral is convergent, although we cannot say anything of the regularity of Φ unless we make some assumptions on the measure μ. As we will see later, for measures compactly supported which are continuous with respect to the Lebesgue measure (*continuous densities* from now on appealing to the physical origin) the potential Φ is defined everywhere. For not compactly supported measures, if we impose special decay conditions to μ at infinity we may even have the potential defined everywhere. Nevertheless some special cases are treated by analogy with physical situations.

We will obtain a result which is basic in order to understand the behaviour of the potential of continuous distributions, but firstly we have to consider a particular distribution on a particular nice surface: the sphere. Consider the potential produced by a homogeneous charge located on a sphere of radius $R > 0$ centred at 0. Assume that the density is ρ, that is the quotient of the charge by the measure area. By symmetry reasons the potential must depend only on the distance of the point $\overrightarrow{p}$ to the origin, that is, the norm $\|\overrightarrow{p}\|$. Therefore we may assume that the point lies on the positive part of the Z axis and so $\overrightarrow{p} = (0, 0, z_0)$ with $z_0 \geq 0$. Consider the following parameterization of the sphere

$$\overrightarrow{r}(\theta, \phi) = \begin{cases} x = R\cos\theta\cos\phi; \\ y = R\sin\theta\cos\phi; \\ z = R\sin\phi. \end{cases}$$

The factor of area transformation is $R^2\cos\phi$ and the distance to $\overrightarrow{p}$ is

$$\|\overrightarrow{r} - \overrightarrow{p}\|^2 = R^2\cos^2\theta\cos^2\phi + R^2\sin^2\theta\cos^2\phi + (R\sin\phi - z_0)^2$$

$$= R^2 - 2Rz_0\sin\phi + z_0^2.$$

The potential at $\overrightarrow{p}$ is given by

$$\Phi(\overrightarrow{p}) = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \frac{\rho R^2 \cos\phi \, d\theta d\phi}{\sqrt{R^2 - 2Rz_0 \sin\phi + z_0^2}}$$

$$= 2\pi \int_{-\pi/2}^{\pi/2} \frac{\rho R^2 \cos\phi \, d\phi}{\sqrt{R^2 - 2Rz_0 \sin\phi + z_0^2}} = -\frac{2\pi\rho R}{z_0} \sqrt{R^2 - 2Rz_0 \sin\phi + z_0^2} \Big|_{-\pi/2}^{\pi/2}$$

$$= \frac{2\pi\rho R}{z_0}\sqrt{(R+z_0)^2} - \frac{2\pi\rho R}{z_0}\sqrt{(R-z_0)^2} = \frac{2\pi\rho R}{z_0}(R + z_0 - |R - z_0|).$$

The expression depends on the relative position of the point with respect to the sphere

$$\Phi(\overrightarrow{p}) = \begin{cases} 4\pi R\rho & \text{if} \quad \|\overrightarrow{p}\| < R; \\ \frac{4\pi R^2 \rho}{\|\overrightarrow{p}\|} & \text{if} \quad \|\overrightarrow{p}\| \geq R. \end{cases}$$

Note that the integral converges at the singular points (the sphere of radius R) making the potential continuous on all $\mathbb{R}^3$, although not differentiable. Moreover, for exterior points the potential of the charged sphere behaves as if all the charge ($4\pi R^2\rho$) was concentrated at the origin.

Now we will consider a homogeneous charge on a ball where the volumetric density is denoted ρ too. The ball of radius $R > 0$ can be considered as made up of spheres of radiuses $0 \leq s \leq R$. The potential produced by a sphere of radius s and thickness δs at the point $\overrightarrow{p}$ is

$$\Phi_s(\overrightarrow{p}) = \begin{cases} \frac{4\pi s^2 \rho}{\|\overrightarrow{p}\|} \, \delta s & \text{if} \quad 0 \leq s \leq \min\{\|\overrightarrow{p}\|, R\}; \\ 4\pi s\rho \, \delta s & \text{if} \quad \min\{\|\overrightarrow{p}\|, R\} < s \leq R. \end{cases}$$

In case $\|\overrightarrow{p}\| \geq R$ only the first formula is necessary and therefore

$$\Phi(\overrightarrow{p}) = \int_0^R \frac{4\pi s^2 \rho}{\|\overrightarrow{p}\|} \, ds = \frac{4\pi R^3 \rho}{3\|\overrightarrow{p}\|}$$

which means that the homogeneously charged sphere behaves like a punctual charge concentred at the origin. In case that $\|\overrightarrow{p}\| < R$ we have

$$\Phi(\overrightarrow{p}) = \int_0^R \Phi_s(\overrightarrow{p}) \, ds = \int_0^{\|\overrightarrow{p}\|} \frac{4\pi s^2 \rho}{\|\overrightarrow{p}\|} \, ds + \int_{\|\overrightarrow{p}\|}^R 4\pi s\rho \, ds =$$

$$\frac{4\pi\|\overrightarrow{p}\|^3\rho}{3\|\overrightarrow{p}\|}+2\pi R^2\rho-2\pi\|\overrightarrow{p}\|^2\rho=2\pi R^2\rho-\frac{2\pi}{3}\|\overrightarrow{p}\|^2\rho.$$

Putting both expressions together we have

$$\Phi(\overrightarrow{p})=\begin{cases}2\pi R^2\rho-\frac{2\pi\|\overrightarrow{p}\|^2\rho}{3} & \text{if } \|\overrightarrow{p}\|<R;\\ \frac{4\pi R^3\rho}{3\|\overrightarrow{p}\|} & \text{if } \|\overrightarrow{p}\|\geq R.\end{cases}$$

In spite of the complicated formula this function is C^1. At points not belonging to the sphere the regularity is C^∞, however the potential is not harmonic in the interior of the ball

$$\frac{\partial^2\Phi}{\partial x^2}=\frac{\partial^2\Phi}{\partial y^2}=\frac{\partial^2\Phi}{\partial z^2}=-\frac{4\pi\rho}{3}$$

and so $\Delta\Phi=-4\pi\rho$ in the interior of the ball.

Consider now a variable volumetric density ρ with compact support. If we assume ρ is measurable and bounded, the potential created by the charge $\mu=\rho\,dV$, where dV represents the 3-dimensional Lebesgue measure, is

$$\Phi(\overrightarrow{p})=\iiint\frac{\rho(\overrightarrow{r})\,dV}{\|\overrightarrow{r}-\overrightarrow{p}\|}$$

which is defined everywhere. Indeed, the support of ρ can be included into a ball and the function ρ is bounded, therefore the convergence of the integral on balls above implies the convergence in this situation. A direct argument is possible too: if $\overrightarrow{p}$ belongs to the support of ρ it is a singular point, however the volume of a ball of radius $\varepsilon>0$ is proportional to ε^3 meanwhile the function goes to ∞ proportionally to ε^{-1}, which implies the convergence. Actually, this argument works with a function inside the integral whose growth proportionally to ε^{-2}. That implies the convergence of

$$\iiint\frac{|\rho(\overrightarrow{r})|\,dV}{\|\overrightarrow{r}-\overrightarrow{p}\|^2}=\iiint|\rho(\overrightarrow{r})|\|\nabla(\|\overrightarrow{r}-\overrightarrow{p}\|^{-1})\|\,dV$$

and therefore the differentiability of $\Phi(\overrightarrow{p})$ everywhere and the validity of the following formula

$$\nabla\Phi(\overrightarrow{p})=\iiint\rho(\overrightarrow{r})\nabla(\|\overrightarrow{r}-\overrightarrow{p}\|^{-1})\,dV.$$

In order to have second order derivatives we need to ask some regularity to ρ. If ρ were differentiable at some point $\overrightarrow{p}$ then it would be possible to compensate the growing of rate ε^{-3} near the singularity with the "balanced" difference $\rho(\overrightarrow{r}) - \rho(\overrightarrow{p})$, implying that we may change locally ρ by the constant value $\rho(\overrightarrow{p})$. This is a delicate task that we are no going to detail here, however the interpretation is easy: the Laplacian of Φ at $\overrightarrow{p}$ can be calculated decomposing the charge into two parts: the part inside a small ball of radius ε where we may assume that ρ is constant and the charge out the small ball which produces a potential whose Laplacian is 0 at $\overrightarrow{p}$ since this point is not in the support. The consequence of that argument provided the extra regularity of ρ is the possibility of being recovered from the potential it generates

$$\Delta\Phi(\overrightarrow{p}) = -4\pi\rho(\overrightarrow{p}).$$

This remarkable formula is known as *Poisson's equation*. It is natural to think the feasibility of the inverse problem: given a function f defined on a domain D with some regularity hypotheses. is possible to express f as a potential? In general the answer is negative. Indeed, the function to be a candidate for the charge is evidently

$$\rho = \frac{-1}{4\pi}\Delta f.$$

If f is C^3 and D is bounded then the potential

$$\Phi(\overrightarrow{p}) = \frac{-1}{4\pi}\iiint_D \frac{\Delta f(\overrightarrow{r})\, dV}{\|\overrightarrow{r} - \overrightarrow{p}\|}$$

is a function such that $\Delta\Phi = \Delta f$ on D, but $\Phi \neq f$ in general. For instance, if $f(x, y, z) = x^2 + y^2 + z^2$ and $D = B(0, 1)$ the integral formula will produce the potential of a homogenous charged ball (as above in this section) which differs from f in a constant. In general, the difference $\Phi - f$ will be a harmonic function on D. If f and its derivatives satisfy some particular decay conditions we could enforce the equality as a consequence of the properties of harmonic functions we are going to study in next section.

The previous arguments have a nice application. Every C^2 vector field can be decomposed locally as the sum of a gradient of a scalar function and the rotational of a vector field. Indeed, given $\overrightarrow{F}$ take $\rho = \nabla \cdot \overrightarrow{F}$. If $D \subset \mathbb{R}^3$ is a ball (or more generally, a bounded star shaped domain) then we may consider the potential Φ generated by the density ρ on D and take $f = -(4\pi)^{-1}\Phi$. Now we have

$$\nabla \cdot (\overrightarrow{F} - \nabla f) = \rho - \rho = 0.$$

Therefore, the field $\overrightarrow{F} - \nabla f$ is *closed* regarded as a 2-form. Then there exists a *primitive* $\overrightarrow{G}$ on D such that $\nabla \times \overrightarrow{G} = \overrightarrow{F} - \nabla f$ and thus

$$\overrightarrow{F} = \nabla f + \nabla \times \overrightarrow{G}$$

as we wanted. Is worth noticing that the decomposition above does not make any sense from the point of view given by the theory of differential forms. Indeed, we are adding a 1-form and a 2-form.

11.5 Harmonic functions

In the previous section we have seen that potential functions are harmonic away their support, that is, their Laplacian is zero. Harmonic functions appear in many applications, so that we will prove some additional properties they have. Let start by this straightforward application of the Gauss-Ostrogradsky theorem: if f is harmonic in a domain that contains the ball $B[\overrightarrow{p}, R]$ with $R > 0$ then

$$\iint_{\partial B[\overrightarrow{p},R]} \nabla f \cdot d\overrightarrow{S} = \iiint_{B[\overrightarrow{p},R]} \Delta f \, dV = 0.$$

We can rewrite this equality

$$\iint_{\partial B[\overrightarrow{p},R]} \nabla f \cdot \overrightarrow{N} \, dS = 0$$

and the term $\nabla f \cdot \overrightarrow{N}$ can be interpreted as a *normal derivative*, usually denoted by $\frac{\partial f}{\partial n}$ (in this particular case is a radial derivative $\frac{\partial f}{\partial r}$). We may parameterize the sphere $\partial B[\overrightarrow{p}, r]$ by means of the unit sphere $\partial B[0,1]$ as $\overrightarrow{p} + r\overrightarrow{x}$. In such a case we have $\overrightarrow{x} = \overrightarrow{N}$ as well. Since the sizes of the spheres differ in a r^2 factor we have

$$\iint_{\partial B[0,1]} \nabla f(\overrightarrow{p} + r\overrightarrow{x}) \cdot \overrightarrow{x} \, dS(\overrightarrow{x}) = \frac{1}{r^2} \iint_{\partial B[\overrightarrow{p},r]} \nabla f \cdot \overrightarrow{N} \, dS = 0.$$

Then, integrating with respect to r we get

$$0 = \int_0^R \iint_{\partial B[0,1]} \nabla f(\overrightarrow{p} + r\overrightarrow{x}) \cdot \overrightarrow{x} \, dS(\overrightarrow{x}) \, dr =$$

$$\iint_{\partial B[0,1]} \int_0^R \frac{d}{dr}\Big(f(\overrightarrow{p} + r\overrightarrow{x})\Big) \, dr \, dS(\overrightarrow{x}) = \iint_{\partial B[0,1]} f(\overrightarrow{p} + r\overrightarrow{x})\Big|_{r=0}^{R} dS(\overrightarrow{x})$$

$$= \iint_{\partial B[0,1]} (f(\overrightarrow{p} + R\overrightarrow{x}) - f(\overrightarrow{p}))\, dS(\overrightarrow{x})$$

which implies after rescaling (integration over the ball of radius R) that

$$0 = \iint_{\partial B[\overrightarrow{p},R]} (f - f(\overrightarrow{p}))\, dS = \iint_{\partial B[\overrightarrow{p},R]} f\, dS - 4\pi R^2 f(\overrightarrow{p})$$

and so

$$f(\overrightarrow{p}) = \frac{1}{4\pi R^2} \iint_{\partial B[\overrightarrow{p},R]} f\, dS.$$

This remarkable identity is the so called *mean value property* of the harmonic functions, that is the value at any point can be expressed as an average of the values over any sphere around that point. The mean value property is true in any dimension with the corresponding adaptation. Note that in dimension 1 is evident because the harmonic functions are exactly the *affine* functions, so the mean value property just says that the value at the middle of a segment is the arithmetic mean of the values at the butts. In dimension 2 the above proof can be adapted with the use of the Green-Riemann theorem and the ideas to be discussed in the next section. Nevertheless, in dimension 2 the theory of harmonic functions have strong bonds with complex analysis which provide alternative techniques.

We will go on with the 3-dimensional frame to state and prove the results although the they are valid in any dimension. The mean value property has a surprising consequence.

Theorem 11.5.1. *Let $D \subset \mathbb{R}^3$ be a connected domain and f a harmonic function defined on D. Then*

(a) f does not have relative strictly extremum values;

(b) if f attains an absolute extreme value on D then f is constant;

(c) if D is bounded and f can be extended continuously to $\overline{D}$ then f attains its extreme values on ∂D.

Proof. We will argue with maximums, being the argument with minimums similar.
Assume that f has a relative strict maximum at $\overrightarrow{p}$. Then there is $\varepsilon > 0$

such that $f|_{\partial B(\overrightarrow{p},\varepsilon)} < f(\overrightarrow{p})$. By the continuity of f (a strict inequality at a particular point integrated remains strict) we get that

$$\frac{1}{4\pi\varepsilon^2}\iint f\,dS < f(\overrightarrow{p})$$

which is a contradiction.
Now, if the function attains a maximum the previous argument shows that actually we have $f|_{\partial B(\overrightarrow{p},\varepsilon)} = f(\overrightarrow{p})$ for any $\overrightarrow{p}$ where the maximum is attained and any $\varepsilon > 0$ such that $B[\overrightarrow{p},\varepsilon] \subset D$. That shows that the set

$$\{\overrightarrow{p} \in D : f(\overrightarrow{p}) = \max(f)\}$$

is open. As it is clearly closed, then it must be all D by connection.
Finally, the last statement is a consequence that the maximum has to be attained somewhere. If attained on D, then the function is constant and so the maximum is also attained on ∂D. ■

A consequence of the previous result is the uniqueness of the solution of the Dirichlet's problem on bounded domains. In case, of unbounded domains we have the following.

Corollary 11.5.2. *A harmonic function defined on $\mathbb{R}^3$ which vanishes at ∞ must be null.*

This result can be applied to prove that a function vanishing at ∞ such that its derivatives also vanishes at ∞ with a suitable rate of decay is actually a potential with charge given by Poisson's equation.

11.6 Vector Analysis in $\mathbb{R}^2$

So far we have discussed results and topics for $\mathbb{R}^3$. The adaptation of the results to $\mathbb{R}^2$ is not merely to take $z = 0$ everywhere. Let start with the following observation: the Green-Riemann formula

$$\int_{\partial D} p\,dx + q\,dy = \iint_D \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) dx\,dy$$

can be deduced from Stokes' formula just doing $z = 0$. However, what is the analogous of Gauss-Ostrogradsky? In other words, how to express the

flux integral on the plane and what is its corresponding "divergence" formula? In order to answer that question, suppose firstly that the boundary of D is given by some C^1 curve $\overrightarrow{\gamma}(s)$ with $s \in [0, L]$ the arc-length. That implies $\|\overrightarrow{\gamma}'(s)\| = 1$. We put $\overrightarrow{\gamma}(s) = (x(s), y(s))$ the integral above over ∂D can be expressed as

$$\int_{\partial D} p\,dx + q\,dy = \int_0^L (p(x(s), y(s))\,x'(s) + q(x(s), y(s))\,y'(s))\,ds$$

whose interpretación has beed discussed previously. If we desire a 2-dimensional flux integral we need the to work with the unitary normal vector (pointing outside) $\overrightarrow{n}(s) = (y'(s), -x'(s))$. The corresponding flux integral is

$$\int_{\partial D} \overrightarrow{F} \cdot d\overrightarrow{\nu} := \int_{\partial D} \overrightarrow{F} \cdot \overrightarrow{n}\,ds = \int_0^L (p(x(s), y(s))\,y'(s) - q(x(s), y(s))\,x'(s))\,ds$$

where $\overrightarrow{F} = (p, q)$. Note that the last member can be interpreted as a standard line integral

$$\int_0^L (-q(x(s), y(s))\,x'(s) + p(x(s), y(s))\,y'(s))\,ds = \int_{\partial D} -q\,dx + p\,dy.$$

Therefore, applying Green-Riemann we get

$$\int_{\partial D} \overrightarrow{F} \cdot d\overrightarrow{\nu} = \iint_D \left(\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y}\right) dx\,dy$$

which is a genuine version of the Gauss-Ostrogradsky theorem. It is clear, that this formula can be extended to the same hypotheses of Green-Riemann. Moreover, we may interpret the expression inside the plane dimensional integral as a *2-dimensional divergence* and the meaning of this divergence is the same that in 3 dimensions for two dimensional fluids (this idealization is also studied in Fluid Dynamics).

The application of the previous result to the gradient of a scalar function $f(x, y)$ gives

$$\int_{\partial D} \nabla f \cdot d\overrightarrow{\nu} = \iint_D \Delta f\,dx\,dy.$$

In particular, if the function is harmonic (in 2 dimensions) then the integrals are zero. In the theory of Newtonian potential in 3 dimensions an important

role was played by the function $(x^2+y^2+z^2)^{-1/2}$ which the unique non trivial harmonic function with spherical symmetry, turning a blind eye on the fact that it s not defined at 0. The analogous role in $\mathbb{R}$ is played by the function $\phi(x,y) = (-1/2)\log(x^2+y^2)$ for which

$$\int_{\partial D} \nabla\phi \cdot d\overrightarrow{\nu} = -2\pi$$

on any domain D containing 0. The function ϕ can be interpreted as a 3-dimensional potential produced by a charge placed on the Z axis with linear density 1, being necessary some "physical trick" in order to obtain the result. Also, the proof of the mean value theorem in the previous section can adapted without trouble to obtain

$$f(\overrightarrow{p}) = \frac{1}{2\pi R}\int_{\partial B[\overrightarrow{p},R]} f\, d\ell$$

for any harmonic function defined on a domain that contains $B[\overrightarrow{p}, R]$.

Finally, we will discuss the application of Green-Riemann to the computation of areas. It is clear, that a choice of functions p, q such that $\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} = 1$, as for instance $q = x, p = 0$ or $p = -y, q = 0$, will imply

$$\text{area}(D) = \int_{\partial D} p\, dx + q\, dy.$$

The advantage that this method offers is based that it is often easier to parameterize the boundary rather than to express as a graph. Another choice of functions p, q which offers some symmetry is the following

$$\text{area}(D) = \frac{1}{2}\int_{\partial D} -y\, dx + x\, dy.$$

In case of not having a natural parameterization of the curve, the part contained in one of the half-planes $x > 0$ or $x < 0$ taking as parameter $t = y/x$ on each piece where it can be done uniquely. The relation among the differentials $dy = t\, dx + x\, dt$ carried to the last area formula gives

$$\text{area}(D) = \frac{1}{2}\int_{\partial D} -tx\, dx + tx\, dx + x^2 dt = \frac{1}{2}\int_{\partial D} x^2 dt$$

which is a remarkable formula that keeps some resemblance with the well known *polar* formula for the area

$$\text{area}(D) = \frac{1}{2}\int_\alpha^\beta \rho^2 d\theta$$

if $D = \{(r\cos\theta, r\sin\theta) : 0 \le r \le \rho(\theta), \alpha \le \theta \le \beta\}$.

11.7 Assorted applications

We include some applications coming from different branches of Physics.

11.7.1 Mechanics

Assume that a particle of mass m follows a path $\overrightarrow{r}(t)$ as a consequence of a force field $\overrightarrow{F}$ acting on it. Newton's second law says that

$$\overrightarrow{F} = m\,\frac{d^2\overrightarrow{r}}{dt^2}.$$

Assume that the force is *conservative*, that is, it is the gradient of some scalar function. Take a function V such that $\overrightarrow{F} = -\nabla V$ and call it *potential energy*. We already know that

$$\int_{t_1}^{t_2} \overrightarrow{F}\cdot d\overrightarrow{r} = V(\overrightarrow{r}(t_1)) - V(\overrightarrow{r}(t_2))$$

where the first integral is a line integral along $\overrightarrow{r}(t)$ with $t_1 \le t \le t_2$ (we follow the standard notation labelling the integral with the time interval instead of the path). It is worth noticing that the Leibniz rule for differentiation of products also works for scalar and vector products. In particular

$$\frac{d}{dt}\Big(\frac{d\overrightarrow{r}}{dt}\cdot\frac{d\overrightarrow{r}}{dt}\Big) = \frac{d^2\overrightarrow{r}}{dt^2}\cdot\frac{d\overrightarrow{r}}{dt} + \frac{d\overrightarrow{r}}{dt}\cdot\frac{d^2\overrightarrow{r}}{dt^2} = 2\,\frac{d\overrightarrow{r}}{dt}\cdot\frac{d^2\overrightarrow{r}}{dt^2}.$$

We will apply that to the line integral above

$$\int_{t_1}^{t_2} \overrightarrow{F}\cdot d\overrightarrow{r} = \int_{t_1}^{t_2} m\,\frac{d^2\overrightarrow{r}}{dt^2}\cdot\frac{d\overrightarrow{r}}{dt}\,dt = \frac{m}{2}\,\frac{d\overrightarrow{r}}{dt}\cdot\frac{d\overrightarrow{r}}{dt}\Big|_{t_1}^{t_2} = \frac{mv^2(t_2)}{2} - \frac{mv^2(t_1)}{2}$$

where $v = \|\frac{d\overrightarrow{r}}{dt}\|$. The magnitude $mv^2/2$ is called the *kinetic energy*. Now from the equality

$$V(\overrightarrow{r}(t_1)) - V(\overrightarrow{r}(t_2)) = \frac{mv^2(t_2)}{2} - \frac{mv^2(t_1)}{2}$$

we get

$$V(\overrightarrow{r}(t_1)) + \frac{mv^2(t_1)}{2} = V(\overrightarrow{r}(t_2)) + \frac{mv^2(t_2)}{2}$$

which is called the *conservation of energy law*: the sum of the kinetic and the potential energy remains constant along the time.

11.7.2 Hydrostatics

On a still fluid the pressure p is a scalar field that at any point represents the magnitude of the force per unit area applied on the a face of a tiny plane surface under the assumption that the fluid is removed from the other side. The experimental knowledge says that the force is always normal to the surface and its magnitude does not depend on the orientation of the surface, at the same point. Assume that a non porous body D with C^1 boundary is subdued to a pressure field. The total force applied on D is given by the surface integral

$$-\iint_{\partial D} p\,\overrightarrow{N}\,dS = -\iint_{\partial D} p\,d\overrightarrow{S}$$

where the sign "$-$" is necessary because the pressure by its very definition is positive, the normal $\overrightarrow{N}$ points out the outside meanwhile the force is exerted towards the body. The trick to compute this integral, which is vector-valued, is to reduce it to a flux integral. Consider the field $\overrightarrow{F} = p\,\vec{i}$. Then

$$\vec{i}\cdot\iint_{\partial D} p\,d\overrightarrow{S} = \iint_{\partial D} \overrightarrow{F}\cdot d\overrightarrow{S} = \iiint_D \nabla\cdot\overrightarrow{F}\,dV = \iiint_D \frac{\partial p}{\partial x}\,dV.$$

by Gauss-Ostrogradsky at the last step. That can be done likewise also for $\vec{j}$ and $\vec{k}$ with obvious consequences that can be written simultaneously for each coordinate as the integral of a vector function

$$-\iint_{\partial D} p\,\overrightarrow{N}\,dS = -\iiint_D \Big(\frac{\partial p}{\partial x}, \frac{\partial p}{\partial y}, \frac{\partial p}{\partial z}\Big)\,dV = -\iiint_D \nabla p\,dV.$$

If we consider the *hydrostatic pressure* at ground level given by

$$p(x, y, z) = c - \rho g z$$

where c is constant (pressure at $z = 0$), ρ the density of the fluid (that also may depend on the point, but we are considering constant at our scale) and g the standard gravity constant at ground level. Since we have $\nabla p = -\rho g\vec{k}$, the total force exerted on D is

$$-\iiint_D \nabla p\, dV = \iiint_D \rho g\vec{k}\, dV = \mathrm{vol}(D)\rho g\vec{k}$$

that is the so called *Archimedes' principle*: the total force is exerted vertically upright and is equivalent to the weight of the mass of fluid that the body D displaces. We may complete this result calculating the *line of action* the resultant force. Indeed, the physical forces are not completely represented by vectors of $\mathbb{R}^3$. It is necessary to specify the line through this force acts or equivalently its *moment* with respect a given point. The moment of a resultant force is the sum of the individual moments. Let $\overrightarrow{r} = (x, y, z)$ be the position of a point of ∂D. The total moment of the force exerted by the pressure is

$$-\iint_{\partial D} p\, \overrightarrow{r} \times \overrightarrow{N}\, dS.$$

Since the result is a vector we will repeat the previous trick multiplying scalarly by $\vec{i}$ and so

$$-\vec{i} \cdot \iint_{\partial D} p\, \overrightarrow{r} \times \overrightarrow{N}\, dS = -\iint_{\partial D} p\, \vec{i} \cdot (\overrightarrow{r} \times \overrightarrow{N})\, dS =$$

$$-\iint_{\partial D} p \cdot (\vec{i} \times \overrightarrow{r}) \cdot \overrightarrow{N}\, dS = -\iint_{\partial D} p \cdot (\vec{i} \times \overrightarrow{r}) \cdot d\overrightarrow{S}$$

that can be calculated by the Gauss-Ostrogradsky theorem. Using that $p = c - \rho g z$ we have

$$-p \cdot (\vec{i} \times \overrightarrow{r}) = (cz - \rho g z^2)\vec{j} + (\rho g y z - cy)\vec{k}$$

and so

$$\nabla \cdot (-p \cdot (\vec{i} \times \overrightarrow{r})) = \rho g y.$$

Therefore we have

$$-\vec{i} \cdot \iint_{\partial D} p\, \overrightarrow{r} \times \overrightarrow{N}\, dS = \rho g \iiint_D y\, dV.$$

Analogous computations show that

$$-\vec{j} \cdot \iint_{\partial D} p\, \overrightarrow{r} \times \overrightarrow{N}\, dS = -\rho g \iiint_D x\, dV;$$

$$-\vec{k}\cdot\iint_{\partial D} p\,\overrightarrow{r}\times\overrightarrow{N}\,dS=0.$$

Knowing that the *center of mass* of D (with uniform density) is the point of coordinates

$$\overrightarrow{C_M}=\frac{1}{\mathrm{vol}(D)}\left(\iiint_D x\,dV,\iiint_D y\,dV,\iiint_D z\,dV\right)$$

our result can be written as

$$-\iint_{\partial D} p\,\overrightarrow{r}\times\overrightarrow{N}\,dS=\overrightarrow{C_M}\times\mathrm{vol}(D)\rho g\vec{k}$$

which means that the resultant force $\mathrm{vol}(D)\rho g\vec{k}$ is exerted along the line passing through $\overrightarrow{C_M}$, exactly at the weight of D as if it was filled with fluid. In despite of the technical difficulty of our calculations, it is possible a much simpler way to reach the same conclusions: the volume D filled with fluid would be in equilibrium so its weight applied on $\overrightarrow{C_M}$ compensates all the external forces and moments over ∂D exerted by the rest of the fluid.

11.7.3 Hydrodynamics

Assume we have a moving fluid in such a way that at every point we have a speed $\overrightarrow{v}$, a pressure p and a density ρ that also may depend on time. If we delimit a region D within the fluid, the conservation of the mass implies that the mass flux through the boundary ∂D per time unit must balance the variation of mass inside D, that is

$$\iint_{\partial D}\rho\,\overrightarrow{v}\cdot d\overrightarrow{S}=-\frac{d}{dt}\iiint_D\rho\,dV=-\iiint_D\frac{\partial\rho}{\partial t}\,dV$$

where the sign "$-$" is due to the fact that fluid going out counts positively and the last equality is just standard derivation of integrals with respect to parameters. Applying Gauss-Ostrogradsky we have

$$0=\iiint_D\nabla\cdot(\rho\overrightarrow{v})\,dV+\iiint_D\frac{\partial\rho}{\partial t}\,dV=\iiint_D\left(\nabla\cdot(\rho\,\overrightarrow{v})+\frac{\partial\rho}{\partial t}\right)dV.$$

As this has to be true for every domain D we deduce

$$\nabla\cdot(\rho\,\overrightarrow{v})+\frac{\partial\rho}{\partial t}=0$$

which is the so called *continuity equation* of fluids. As we said at the beginning this equation just expresses the mass conservation. If the density is constant (e.g. liquids) we obtain $\nabla \cdot \overrightarrow{v} = 0$: volumen entering equals volume going out. Assume now that ∂D encompasses a part of the fluid and moves along with it. Newton's second and third laws combined say that the acceleration observed on D (the mass inside) is due to the external forces, notably the effect of the pressure and the weight if we are studying the problem at ground level. Assuming that D is small enough to consider $\overrightarrow{v}$ homogeneous on it we have

$$\frac{d\overrightarrow{v}}{dt} \iiint_D \rho \, dV = -\iint_{\partial D} p \, d\overrightarrow{S} + \iiint_D \rho \, \overrightarrow{F} \, dV$$

being $\overrightarrow{F}$ a force by mass unit ($\overrightarrow{F} = g\vec{k}$ for the ordinary weight). After Gauss-Ostrogradsky and replacing terms

$$0 = \iiint_D \frac{d\overrightarrow{v}}{dt} \, \rho \, dV + \iiint_D \nabla p \, dV - \iiint_D \rho \, \overrightarrow{F} \, dV$$

where ∇p comes from our previous study of the Archimedes' principle. Since the equality has to be true for D arbitrarily small we get

$$\frac{d\overrightarrow{v}}{dt} \, \rho + \nabla p - \rho \, \overrightarrow{F} = 0.$$

This equation has an important handicap for the applications. In practise it is easier to determine the speed at a given point and then its variation along time $\frac{\partial \overrightarrow{v}}{\partial t}$ which is different from $\frac{d\overrightarrow{v}}{dt}$ that represents the acceleration of a part of the moving fluid. In order to obtain the relation between both derivatives assume $\overrightarrow{v} = (v_x, v_y, v_z)$ being all of them functions of x, y, z, t. Now for v_x we have

$$\frac{dv_x}{dt} = \frac{\partial v_x}{\partial t} + \frac{\partial v_x}{\partial x}\frac{dx}{dt} + \frac{\partial v_x}{\partial y}\frac{dy}{dt} + \frac{\partial v_x}{\partial z}\frac{dz}{dt} =$$

$$\frac{\partial v_x}{\partial t} + \frac{\partial v_x}{\partial x} v_x + \frac{\partial v_x}{\partial y} v_y + \frac{\partial v_x}{\partial z} v_z = \frac{\partial v_x}{\partial t} + \nabla v_x \cdot \overrightarrow{v}.$$

The same can ve done for v_y, v_z and putting all together we get

$$\frac{d\overrightarrow{v}}{dt} = \frac{\partial \overrightarrow{v}}{\partial t} + (\overrightarrow{v} \cdot \nabla)\overrightarrow{v}$$

where the term $\overrightarrow{v} \cdot \nabla$ acts like a differential operator on each coordinate of $\overrightarrow{v}$. The previous equations of fluids takes now the form

$$\frac{\partial \overrightarrow{v}}{\partial t} + (\overrightarrow{v} \cdot \nabla)\overrightarrow{v} + \frac{1}{\rho}\nabla p - \overrightarrow{F} = 0$$

which known as *Euler's equation*. Note that if the fluid is *stationary*. that is, the speed at any point remains constant in time, then $\frac{\partial \overrightarrow{v}}{\partial t} = 0$.
Euler's equation is still much complicated, however some reasonable assumptions can lead to simpler forms. Assuming that the fluid is *irrotational* which means free of whirlpools and in terms of equations $\nabla \times \overrightarrow{v} = 0$ then

$$\nabla v_x \cdot \overrightarrow{v} = \frac{\partial v_x}{\partial x} v_x + \frac{\partial v_x}{\partial y} v_y + \frac{\partial v_x}{\partial z} v_z = \frac{\partial v_x}{\partial x} v_x + \frac{\partial v_y}{\partial x} v_y + \frac{\partial v_z}{\partial x} v_z = \frac{\partial \overrightarrow{v}}{\partial x} \cdot \overrightarrow{v}$$

and the same is true for y, z. Going above to the place where $(\overrightarrow{v} \cdot \nabla)\overrightarrow{v}$ first appeared we get

$$(\overrightarrow{v} \cdot \nabla)\overrightarrow{v} = \frac{1}{2}\nabla v^2$$

(remember, under the hypothesis that $\nabla \times \overrightarrow{v} = 0$). Using this for Euler's equation gives us

$$\frac{\partial \overrightarrow{v}}{\partial t} + \frac{1}{2}\nabla v^2 + \frac{1}{\rho}\nabla p - \overrightarrow{F} = 0$$

still not quite practical. Assume moreover that the fluid is stationary ($\frac{\partial \overrightarrow{v}}{\partial t} = 0$), the external force is conservative ($\overrightarrow{F} = -\nabla V$) and the density ρ constant (liquid). Then we have

$$0 = \frac{1}{2}\nabla v^2 + \frac{1}{\rho}\nabla p + \nabla V = \nabla\Big(\frac{v^2}{2} + \frac{p}{\rho} + V\Big)$$

which implies the important equation

$$\frac{v^2}{2} + \frac{p}{\rho} + V = \text{constant}$$

known as *Bernoulli's equation* which is a form of the law of conservation of energy for fluids. Note that for V constant or being its variation negligible with respect to the pressure (e.g. at ground level the fluid moves approximatively at the same height) then the raising of the speed implies a lowering of the pressure, which is the so called *Venturi effect*. Introducing the effect of the *viscosity*, which implies that the speed of the fluid near immobile objects is zero for instance, is possible to obtain a set of formulas which model much better real fluids: the *Navier-Stokes equations*.

11.7.4 Electromagnetic fields

The well known *Coulomb's law* says that to point charges are repelled (or attracted if they are of different sign) in the empty space with a force proportional to their magnitudes and inversely proportional to the square of the distance between them. Following standard conventions the intensity of this force is written

$$F = \frac{q_1 q_2}{4\pi\epsilon_0\, r^2}$$

where the constant ϵ_0 depends on the unit system. The air, or matter in general, between the charges has some effect that could be included in the formula but will not consider. Note that the field produced by a single charge is of Newtonian type, thus the theory of Newtonian potential can be applied to study the field produced by a charge density. Indeed, let $\overrightarrow{\mathbf{E}}$ the electric field field produced by a continuous electric density ρ. Poisson equation, after the correction of the 4π term is

$$\nabla \cdot \overrightarrow{\mathbf{E}} = \frac{\rho}{\epsilon_0}.$$

Note that there is not "$-$" because the repulsion is the effect suffered by a positive test charge placed in a field produced by a positive density $\rho > 0$. Static electric fields have a potential that simplifies their description and the mechanical effects on charges.

However, a main ingredient here is the great mobility of charges, notably through certain substances called *conductors*. Therefore, the variation with time of ρ, and so that of $\overrightarrow{\mathbf{E}}$, must be taken into account. The *law of conservation of charge* implies that we may apply the fluid model to the *electric current* $\overrightarrow{J}$ (flux of charge per time and surface unit) to obtain

$$-\frac{d}{dt}\iiint_D \rho\, dV = \iint_{\partial D} \overrightarrow{J} \cdot d\overrightarrow{S}$$

whose meaning must be obvious at this stage. Commuting derivation and integral together the Gauss-Ostrogradsky theorem and the fact that D is arbitrary leads to

$$-\frac{\partial \rho}{\partial t} = \nabla \cdot \overrightarrow{J}.$$

Note that so far we are considering the charge to be either positive or negative and the current is interpreted in a positive sense, that is, if a region receives a positive flux of charge then the charge "increases". Actually, what moves inside conductors are *electrons* which have negative charge. This fact is not

relevant form the point of view of classic electrodynamics.
When an ordinary conductor is placed into an electric field the charges move inside and after a while the movement stops because of the *electrical resistance*. In the reached equilibrium, the electrostatic potential is constant on the conductor and so the electric field $\overrightarrow{\mathbf{E}}$ is normal to the surface of the conductor. The potential could be computed solving the Laplace equation $\Delta V = 0$ with the boundary conditions imposed by the charges distributed on the conductors (V is constant on the boundaries). The former argument does not apply to *superconductors* which are substances that under certain conditions (extremely low temperatures) possess null electrical resistance.
A system of two point charges of equal magnitude and different signs placed at "short" distance is called a *dipole*. Far from a dipole the intensity of the field decreases faster than for a single point charge because there is an almost cancelation of effects: the sum of two nearly oposite vectors with nearly the same modulus. However, near the dipole things are obviously different, and the effect of the field on another dipole not only will include attraction/repulsion but also a torque (rotational force).
This brief discussion on dipoles was aimed to present the magnetic force. *Magnets* behave between them like electric dipoles. The magnetic field is represented by a vector field $\overrightarrow{\mathbf{B}}$ whose effect is not only felt on magnetic materials but also on moving electric charges according to the formula

$$\overrightarrow{F} = q\overrightarrow{v} \times \overrightarrow{\mathbf{B}}$$

where q and $\overrightarrow{v}$ are respectively the charge and the speed. Since the magnetic force is perpendicular to the trajectory it does not modify the kinetic energy (the work done by $\overrightarrow{F}$ is 0), however the magnetic field bends the trajectory and eventually drives trapped ionized moving particles to the poles following a helix path (look for the explanation of the *auroras*). It is well known that a piece of magnet is again a magnet with two different poles, so it is impossible to isolate "magnetic monopoles". That leads to consider a larger magnet as composed of a density of "magnetic dipoles" instead of charges and so in any arbitrarily (small) volume there is always a compensation written as

$$\nabla \cdot \overrightarrow{\mathbf{B}} = 0.$$

There are equations linking $\overrightarrow{\mathbf{E}}$ and $\overrightarrow{\mathbf{B}}$ notably when they vary with the time. Charges in movement (currents) produce a magnetic field, for instance electromagnets, and variations in the magnetic field induce currents, that is, variations on the electric field, for instance dynamos. Both phenomenons are

modelled by the laws of Faraday-Henry and Ampère-Maxwell

$$\nabla \times \overrightarrow{\mathbf{E}} = -\frac{\partial \overrightarrow{\mathbf{B}}}{\partial t};$$

$$\nabla \times \overrightarrow{\mathbf{B}} = \mu_0 \overrightarrow{J} + \epsilon_0 \mu_0 \frac{\partial \overrightarrow{\mathbf{E}}}{\partial t}.$$

Note that the application of the divergence $(\nabla\cdot)$ to the first equation says nothing new meanwhile for the second we recover the conservation of charge. The set of for equations: these two last ones together $\nabla\cdot\overrightarrow{\mathbf{E}} = \rho/\epsilon_0$ and $\nabla\cdot\overrightarrow{\mathbf{B}} = 0$ is called *Maxwell equations* which totally describes the electromagnetic field. The constant μ_0 plays for the magnetism in vacuum a role analogue to ϵ_0, and we are deliberately omitting the modifications that happens inside the matter. We will do more tricky manipulations on Maxwell equations. Since $\nabla\cdot\overrightarrow{\mathbf{B}} = 0$ there is $\overrightarrow{\mathbf{A}}$ such that $\nabla\times\overrightarrow{\mathbf{A}} = \overrightarrow{\mathbf{B}}$. Now we have

$$\nabla \times \overrightarrow{\mathbf{E}} = -\frac{\partial}{\partial t}(\nabla \times \overrightarrow{\mathbf{A}}) = -\nabla \times \frac{\partial \overrightarrow{\mathbf{A}}}{\partial t}$$

and so

$$\nabla \times \left(\overrightarrow{\mathbf{E}} + \frac{\partial \overrightarrow{\mathbf{A}}}{\partial t}\right) = 0.$$

Therefore there exists a function ϕ such that

$$-\nabla\phi = \overrightarrow{\mathbf{E}} + \frac{\partial \overrightarrow{\mathbf{A}}}{\partial t}$$

that would allow us to consider that $\overrightarrow{\mathbf{E}}$ has a scalar potential (like in the static case) but corrected by a term $\frac{\partial\overrightarrow{\mathbf{A}}}{\partial t}$ coming from the magnetic counterpart of the field.
Note that the property of $\overrightarrow{\mathbf{A}}$ remains by adding the gradient of a scalar function. Let us assume for instance that $\nabla\cdot\overrightarrow{\mathbf{A}} = 0$ (technically that would require a special decay of $\overrightarrow{\mathbf{B}}$ at infinity, but we will turn a blind eye on it). Under that hypothesis we would have

$$\Delta\phi = -\nabla\cdot\overrightarrow{\mathbf{E}} = -\frac{\rho}{\epsilon_0}.$$

This is a very nice consequence, however we have to forget it because there is a choice for $\overrightarrow{\mathbf{A}}$ that was better for the development of the theory. The vector

potential $\overrightarrow{\mathbf{A}}$ can be chosen to satisfy an apparently more strange condition due to Lorenz: take $\overrightarrow{\mathbf{A}}$ such that $\nabla \times \overrightarrow{\mathbf{A}} = \overrightarrow{\mathbf{B}}$ and

$$\nabla \cdot \overrightarrow{\mathbf{A}} + \epsilon_0 \mu_0 \frac{\partial \phi}{\partial t} = 0.$$

Working with the Lorenz's condition we have

$$-\Delta\phi = \nabla \cdot \Big(\overrightarrow{\mathbf{E}} + \frac{\partial \overrightarrow{\mathbf{A}}}{\partial t} \Big) = \frac{\rho}{\epsilon_0} + \frac{\partial}{\partial t}(\nabla \cdot \overrightarrow{\mathbf{A}}) = \frac{\rho}{\epsilon_0} - \epsilon_0 \mu_0 \frac{\partial^2 \phi}{\partial t^2}$$

that can be rewritten as

$$\Delta\phi - \epsilon_0 \mu_0 \frac{\partial^2 \phi}{\partial t^2} = -\frac{\rho}{\epsilon_0}.$$

On the other hand, the Ampère-Maxwell equation can be transform as

$$\nabla \times (\nabla \times \overrightarrow{\mathbf{A}}) = \mu_0 \overrightarrow{J} + \epsilon_0 \mu_0 \frac{\partial}{\partial t} \Big(-\nabla\phi - \frac{\partial \overrightarrow{\mathbf{A}}}{\partial t} \Big) =$$

$$= \mu_0 \overrightarrow{J} - \epsilon_0 \mu_0 \Big(\nabla \frac{\partial \phi}{\partial t} \Big) - \epsilon_0 \mu_0 \frac{\partial^2 \overrightarrow{\mathbf{A}}}{\partial t^2} = \mu_0 \overrightarrow{J} + \nabla(\nabla \cdot \overrightarrow{\mathbf{A}}) - \epsilon_0 \mu_0 \frac{\partial^2 \overrightarrow{\mathbf{A}}}{\partial t^2}.$$

Having in mind that $\Delta \overrightarrow{\mathbf{A}} = \nabla(\nabla \cdot \overrightarrow{\mathbf{A}}) - \nabla \times (\nabla \times \overrightarrow{\mathbf{A}})$ we have

$$\Delta \overrightarrow{\mathbf{A}} - \epsilon_0 \mu_0 \frac{\partial^2 \overrightarrow{\mathbf{A}}}{\partial t^2} = -\mu_0 \overrightarrow{J}.$$

The couple of equations we have just obtained

$$\Delta\phi - \epsilon_0 \mu_0 \frac{\partial^2 \phi}{\partial t^2} = -\frac{\rho}{\epsilon_0},$$

$$\Delta \overrightarrow{\mathbf{A}} - \epsilon_0 \mu_0 \frac{\partial^2 \overrightarrow{\mathbf{A}}}{\partial t^2} = -\mu_0 \overrightarrow{J}$$

provide a simpler and quite symmetric version of Maxwell equations in terms of the potentials ϕ and $\overrightarrow{\mathbf{A}}$ which are more convenient for theoretical studies. In absence of charges and currents (that is, far away from them in practise), the equations became homogeneous

$$\Delta\phi - \epsilon_0 \mu_0 \frac{\partial^2 \phi}{\partial t^2} = 0,$$

$$\Delta \overrightarrow{\mathbf{A}} - \epsilon_0 \mu_0 \frac{\partial^2 \overrightarrow{\mathbf{A}}}{\partial t^2} = 0.$$

It is a very remarkable fact that still have non trivial solutions which are of *wave* type. Note that in the same conditions of absence of charges and currents $\overrightarrow{\mathbf{E}}$ and $\overrightarrow{\mathbf{B}}$ satisfy similar equations that can be obtained more easily from Maxwell equations. The speed of the waves is the number

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}}$$

which amazingly coincides with the speed of light. This is even more shocking if we consider that ϵ_0 and μ_0 could be determined working with batteries and wires in a modest laboratory. That leads to Maxwell to think that light is actually an electromagnetic wave. Moreover, the system of equations can be written as an only equation in $\mathbb{R}^4$ for the pair $(\overrightarrow{\mathbf{A}}, \phi)$ and a 4-dimensional Laplacian-like operator

$$\Box = \Big(\frac{\partial^2}{\partial x^2}, \frac{\partial^2}{\partial y^2}, \frac{\partial^2}{\partial z^2}, -\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \Big).$$

All these goes to the *Theory of Relativity*, but our way ends here.

11.8 Rationale and remarks

The dimension 3 played an important role in the chapter of differential forms. It was so because the statement of the classic results. It is not difficult to see that the general formulation of in terms of differential forms allows the formulation in higher dimensions. However, there are aspects of the dimension 3 that exclusive. For instance, 3 is also the dimension of the 2-forms, so vector fields can be thought of as either 1-forms or 2-forms.

That peculiarity is related to the so called "vector product", incidentally used in the computation of the area of a surface. For that reason, we start with the quaternion origin of that notion. We show the correspondence between differential forms and fields.

Vector operators are the classic interpretation of the exterior differential. The geometrical-physical interpretation is valuable for the applications later. Then we introduce the Newtonian potential as a way to found a function with a

prescribed Laplacian. The computations are somehow informal but complete. The problem of uniqueness leads naturally to the study of harmonic functions.

Among the applications we have choose some basic hydrostatics (the theoretical results compares to heuristic ones), hydrodynamics and electro-magnetic field, where we get the 4-dimensional form of the Maxwell's equations at the end.

11.9 Exercises

1. Let f and g be scalar functions, $\overrightarrow{F}$ and $\overrightarrow{G}$ be vectorial fields, all defined on $\mathbb{R}^3$. Prove the following formulas:

   (a) $\nabla(fg) = g\,\nabla f + f\,\nabla g$.

   (b) $\nabla \cdot (f\overrightarrow{F}) = \nabla f \cdot \overrightarrow{F} + f\,\nabla \cdot \overrightarrow{F}$.

   (c) $\nabla \times (f\overrightarrow{F}) = \nabla f \times \overrightarrow{F} + f\,\nabla \times \overrightarrow{F}$.

   (d) $\nabla \cdot (\overrightarrow{F} \times \overrightarrow{G}) = (\nabla \times \overrightarrow{F}) \cdot \overrightarrow{G} + \overrightarrow{F} \cdot (\nabla \times \overrightarrow{G})$.

2. Show that
$$\Delta \overrightarrow{F} = \nabla(\nabla \cdot \overrightarrow{F}) - \nabla \times (\nabla \times \overrightarrow{F}).$$

3. Prove using orthogonal coordinate transformations (positive for the rotacional) in $\mathbb{R}^3$ that the vector operators are intrinsic.

4. Find the expression of the vector operators in cylindrical and spherical coordinates.

5. Show that the following operations do not define intrinsic operators:
$$f \longrightarrow \Big(\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial y^2}, \frac{\partial^2 f}{\partial z^2}\Big);$$
$$\overrightarrow{F} \longrightarrow \Big(\frac{\partial f_1}{\partial x} + \frac{\partial f_1}{\partial y} + \frac{\partial f_1}{\partial z}, \frac{\partial f_2}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_2}{\partial z}, \frac{\partial f_3}{\partial x} + \frac{\partial f_3}{\partial y} + \frac{\partial f_3}{\partial z}\Big);$$
   where $\overrightarrow{F} = (f_1, f_2, f_3)$.

6. Compute the flux of the field (ax, by, cz) on the sphere centred at the origin and radius $R > 0$. Compute the flux of the same field on any other sphere.

7. Assume that the bounded domain $D \subset \mathbb{R}^3$ has piecewise C^1 border and $\vec{F}$ is a C^2 field defined on $\mathbb{R}^3$. Find the flux of $\mathrm{rot}(\vec{F})$ through ∂D.

8. Consider on $\mathbb{R}^3$ the vector field $\vec{F} = (x^3/a^4, y^3/b^4, z^3/c^4)$ where $a, b, c > 0$.

   (a) Describe geometrically the surfaces that are orthogonal to $\vec{F}$.

   (b) Compute the flux of $\vec{F}$ through the sphere centred at $(0, 0, 0)$ and radius 1.

9. Find the flux of the field $\overrightarrow{F} = (x + \alpha y, y - \alpha x, \beta)$, where $\alpha, \beta \in \mathbb{R}$ are parameters, through the portion of paraboloid $z = 1 - x^2 - y^2$ above the plane $z = 0$ and exterior orientation.

10. Consider the pyramid on $[-1, 1]^2$ with vertex at $(0, 0, 3)$ and let L e the surface made up from the lateral faces with exterior orientation. Let $\vec{F} = e^{x+y-2z}(1, 1, 1)$. Compute

$$\iint_L \vec{F} \cdot d\vec{S}.$$

11. Show that the integral for the Newtonian potential generated by a constant linear density $\rho > 0$ on the Z axis diverge. However, the analogous integral for the Newtonian force converges on $\mathbb{R}^3$ except the Z axis. Find a function Φ such that $\overrightarrow{F} = \nabla\Phi$ and explain why that is compatible with the first statement of this exercise.

12. Consider the ellipsoid S with formula $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ and let $\rho(x, y, z)$ be the distance to the origin from the tangent plane to the ellipsoid at (x, y, z). Show that

$$\iint_S \rho \, dS = 4\pi abc;$$

$$\iint_S \frac{1}{\rho} dS = \frac{4\pi}{3} abc \left(\frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2} \right).$$

   Hint: $1/\rho(x, y, z) = \sqrt{\frac{x^2}{a^4} + \frac{y^2}{b^4} + \frac{z^2}{c^4}}$.

13. Show that the following function is harmonic

$$f(x, y, z) = \frac{xy}{(x^2 + y^2 + z^2)^{5/2}}.$$

Calcule lthe flux integral

$$I = \iint_{\partial D} \nabla f \, d\vec{S}$$

where $D \subset \mathbb{R}^3$ has C^1 border and $(0,0,0) \notin \partial D$.

14. Compute the area limited on the first quadrant by the curve

$$x^3 + y^3 = 3xy.$$

15. Let f a harmonic function defined on an open set of $\mathbb{R}^2$. Let D be a compact disc centred at p with $r > 0$ contained on the domain of f. Prove that

$$f(p) = \frac{1}{2\pi r} \int_{\partial D} f \, ds.$$

16. Let $\alpha \geq 1$ and consider the function $f_\alpha(x,y,z) = (x^2 + y^2 + z^2)^\alpha$.

    (a) Find a simplified expression for Δf_α.

    (b) Compute the flux of ∇f_1 through the sphere

$$x^2 + y^2 + z^2 = 2z.$$

17. Let $D \subset \mathbb{R}^3$ be compact with C^1 border and let f a scalar C^1 function defined on a neighbourhood of D. Prove that

$$\iint_{\partial D} f d\overrightarrow{S} = \iiint_D \nabla f \, dV.$$

18. Let $D \subset \mathbb{R}^3$ be a compact with C^1 border. Show that the integral

$$\iint_{\partial D} \nabla \left(\frac{1}{\rho} \right) \cdot d\overrightarrow{S}$$

where $\rho = \sqrt{x^2 + y^2 + z^2}$ takes the values -4π or 0 depending on 0 being interior or exterior to D. Interpret the integral in terms f the solid angle and make a guess on the values in case $0 \in \partial D$.

19. Prove that if f, g are enough regular in $\overline{D}$ with $D \subset \mathbb{R}^3$ (or $\mathbb{R}^2$) open, then

$$\iiint_D g\Delta f \, dV + \iiint_D \nabla g \cdot \nabla f \, dV = \iint_{\partial D} g\nabla f \cdot d\overrightarrow{S}.$$

Prove that a continuous function $f : \overline{D} \to \mathbb{R}$ which is an harmonic on D minimices the "energy integral" among all the regular functions taking the same values on ∂D, that is,

$$\iiint_D \|\nabla f\|^2 dV = \min\{\iiint_D \|\nabla g\|^2 dV : g|_{\partial D} = f|_{\partial D}\}.$$

20. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2 function whose level sets coincide with the level sets of a harmonic function. Show that

$$\frac{\Delta f}{\|\nabla f\|^2}$$

depends only on the value of f.

21. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a C^2 functino such that

$$\iint_{\partial D} \nabla f dS \geq 0$$

for every bounded open set D with C^1 border. Prove that f cannot have strict relative maximums. What about strict minimums?

22. Find the conditions that should be satisfied by the functions ϕ, ψ in order to

$$f(x, y, z) = \phi(\sqrt{x^2 + y^2})\psi(z)$$

be harmonic. (Hint: use $\phi(r)$ with $r = \sqrt{x^2 + y^2}$).

Chapter 12

Appendix A: The Stone-Weierstrass theorem

12.1 General Topology

Topology of metric spaces it is not enough to cover the needs of tha Analysis of functions of real variables. Indeed, we know that the uniform convergence can be understood in terms of a metric topology, although pointwise convergence cannot. For that reason we need *general topological spaces*.

A *topology* on a set X is a family $\tau \subset \mathcal{P}(X)$ such that:

1. $\emptyset, X \in \tau$;
2. if $U, V \in \tau$, then $U \cap V \in \tau$;
3. if $(U_i)_{i \in I} \subset \tau$, then $\bigcup_{i \in I} U_i \in \tau$.

Note that the family of open sets in a metric space is a topology, so we will refer to the elements of τ as *open sets*, and their complements will be called *closed sets*. A set X endowed with a topology is called a *topological space*. The topological space (X, τ) is said to be *Hausdorff* if for every $x, y \subset X$ with $x \neq y$ there exists $A, B \in \tau$ with $x \in U$, $y \in V$ and $U \cap V = \emptyset$. Clearly, metric topologies are Hausdorff. Continuity can be defined locally, nut we are only interested in global continuity: a mapping $f : (X_1, \tau_1) \to (X_2, \tau_2)$ between topological spaces is continuous if $f^{-1}(V) \in \tau_1$ whenever $U \in \tau_2$.

A topological space (X, τ) is said to be *compact* if for every family $(U_i)_{i\in I} \subset \tau$ with $X = \bigcup_{i\in I} U_i$ there exists $J \subset I$ finite such that $X = \bigcup_{i\in J} U_i$. Compactness is a very important property in Analysis since it works quite well together continuity.

Theorem 12.1.1 (Weierstrass)**.** *Let K be a compact space and $f : K \to \mathbb{R}$ a continuous function. Then f is bounded and its maximum and its minimum are reached at some points of K.*

Proof. It is not difficult to prove that compactness is stable by continuous images. Therefore, $f(K)$ would be a compact subset of $\mathbb{R}$, so it is bounded and it contains its maximum and minimum. ∎

As in the metric case, for a compact space K, we will denote $C(K)$ the set of real continuous functions defined on K, eventually endowed with the supremum norm.

Theorem 12.1.2 (Urysohn)**.** *Let K be a compact Hausdorff space and let $A, B \subset K$ be disjoint closed subsets. Then there exists $f : K \to [0, 1]$ such that $f|_A = 0$ and $f|_B = 1$.*

Proof. Consider a maximal family $\mathfrak{F}$ of open sets such if $U \in \mathfrak{F}$ then $A \subset U \subset X \setminus B$ and $\overline{U} \subset V$ for all $V \in \mathfrak{F}$ with $U \subset V$. The family $\mathfrak{F}$ either contains a *clopen* subset (simultaneously open and closed) or for any $U, V \in \mathfrak{F}$ with $U \subsetneq V$ there is $W \in \mathfrak{F}$ such that $U \subsetneq W \subsetneq V$. In the first case the construction of f is obvious, so we will assume the second case holds. By induction we may take $U_t \in \mathfrak{F}$ for every dyadic $t \in [0, 1]$ in such a way that $t \leq s$ implies $U_t \subset U_s$. Define now

$$f(x) = \inf\{t \in [0, 1] : x \in U_t\}$$

This function satisfies $f|_A = 0$ and $f_B = 1$. We have to check its continuity. By construction we have

$$f^{-1}([0, s)) = \bigcup_{t<s} U_t; \quad f^{-1}((s, 1]) = \bigcup_{t>s} (X \setminus \overline{U_t})$$

which implies the continuity of f. ∎

12.2 Approximation by continuous functions

The results in this section exploits two additional structures on $C(K)$, namely the algebra structure, i.e. $C(K)$ is closed for the standard product of functions, and the *lattice* structure, which means that $C(K)$ is closed for the boolean operations max and min.

Proposition 12.2.1. *A linear subspace $X \subset C(K)$ is dense if and only there is $\varepsilon \in (0,1)$ such that for every disjoint closed sets $A, B \subset K$ there is $f \in X$ with values in $[-1,1]$ such that $f|_A \leq \varepsilon - 1$ and $f|_B \geq 1 - \varepsilon$.*

Proof. Assume that X is dense and take a continuous function g such that $g|_A = -1 + \varepsilon/2$ is its minimum and $g|_B = 1 - \varepsilon/2$ is its maximum. Then find $f \in X$ such that $\|f - g\|_\infty < \varepsilon/2$. Clearly f fulfils the required conditions. The reverse implication is more delicate. Firstly note that by scaling it is enough to prove the approximation by elements from X for functions with values in $[-1,1]$. Consider $g \in C(K)$ with values in $[-1,1]$ and take $A = g^{-1}([-1,-2/3])$ and $B = g^{-1}([2/3,1])$. Apply the hypothesis to find f with its values in $[-1,1]$, $f(A) \subset [-1, \varepsilon - 1]$ and $f(B) \subset [1-\varepsilon, 1]$. Take $f_1 = 3^{-1} f$ and set $\lambda = (2+\varepsilon)/3 < 1$. A elemental computation shows that $\|g - f_1\|_\infty < \lambda$, that is $(g-f)(K) \subset [-\lambda, \lambda]$. We have now that $\lambda^{-1}(g - f_1)$ is a function taking values in $[-1,1]$, so we can repeat the previous argument to find f_2 such that

$$-\lambda \ \leq \ \lambda^{-1}(g - f_1) - f_2 \ \leq \ \lambda$$

what implies

$$\|g - f_1 - \lambda f_2\|_\infty \leq \lambda^2.$$

Inductively we can find a sequence (f_n) such that

$$\left\|g - f_1 - \lambda f_2 - \cdots - \lambda^{n-1} f_n\right\|_\infty \leq \lambda^n$$

which prove that g is approximated by elements from X as λ^n goes to 0. ∎

Theorem 12.2.2. *Let $X \subset C(K)$ be a vector lattice that contains the constants and assume that X tells apart on points from K. Then X is dense in $C(K)$.*

Proof. The linearity and the possibility of adding constants allows to find a function $f_{x,y} \in X$ such that $f_{x,y}(x) = -2$ and $f_{x,y}(y) = 2$ whenever $x, y \in K$ with $x \neq y$. Let $A \subset K$ be closed and $y \in K \setminus A$. The family of sets $(U_x)_{x \in A}$

where $U_x = f_{x,y}^{-1}((-\infty,-1))$ is an open cover of A. Let $x_1,\dots,x_n \in A$ such that the corresponding sets cover A. Then

$$f_{A,y} = \max\{-1, \min\{f_{x_1,y}, f_{x_2,y},\dots,f_{x_n,y}\}\}$$

is a function such that $f_{A,y}|_A = -1$ and $f_{A,y}(y) = 2$. Now, if $B \subset K$ is closed and $A \cap B = \emptyset$ we may consider the open cover of B given by the sets $V_y = f_{A,y}^{-1}((1,+\infty))$ and take a finite subcover given by points $y_1,\dots,y_m$. Thus the function

$$f_{A,B} = \min\{1, \max\{f_{A,y_1}, f_{A,y_2},\dots,f_{A,y_m}\}\}$$

satisfies $f_{A,B}|_A = -1$ and $f_{A,B}|_B = 1$. The proof finishes by applying the denseness criterion Proposition 12.2.1. ∎

Lemma 12.2.3. *Let $(p_n(t))$ the sequence of polynomials defined on $[0,1]$ inductively by $p_1(t) = 0$ and $p_{n+1}(t) = p_n(t) + 2^{-1}(t - p_n(t)^2)$. Then $(p_n(t))$ converges uniformly to $\sqrt{t}$ and consequently the function $|t|$ can be uniformly approximated by polynomials on any bounded interval of $\mathbb{R}$.*

Proof. Firstly show that $p_n(t) \leq \sqrt{t}$. Indeed, assuming $p_n(t) \leq \sqrt{t}$ by induction and keeping in mind that $t \in [0,1]$ we have

$$p_{n+1}(t) = p_n(t) + 2^{-1}(\sqrt{t} + p_n(t))(\sqrt{t} - p_n(t))$$

$$\leq p_n(t) + \sqrt{t}(\sqrt{t} - p_n(t)) \leq p_n(t) + (\sqrt{t} - p_n(t)) = \sqrt{t}.$$

Now we have $p_{n+1}(t) \geq p_n(t)$ so the sequence is increasing and bounded. The limit $p(t)$ is the only solution of the functional equation $p(t) = p(t) + 2^{-1}(t - p(t)^2)$. The convergence is uniform by Dini's theorem 1.5.2.
It is clear that the composition $p_n(t^2)$ converges to $|t|$ uniformly on $[-1,1]$. In order to approximate $|t|$ on any bounded interval we may suppose that it is of the form $[-M,M]$ with $M > 0$. It is not difficult to see that sequence of polynomials

$$P_n(t) = Mp_n\left(\frac{t^2}{M^2}\right)$$

converges uniformly to $|t|$ on $[-M,M]$. ∎

Theorem 12.2.4. *Let $X \subset C(K)$ be a subalgebra that contains the constants and assume that X tells apart on points from K. Then X is dense in $C(K)$.*

Proof. Take any $f \in X$ and let $(P_n(t))$ be a sequence of polynomials which converges uniformly to $|t|$ on the set $f(K) \subset \mathbb{R}$. Note that $P_n \circ f \in X$ by the hypotheses, so $|f| \in \overline{X}$, which implies the lattice property for $\overline{X}$. Theorem 12.2.2 says that $\overline{X}$ is dense in $C(K)$ and so $\overline{X} = C(K)$. ∎

As a consequence we recover the classical theorems of Weierstrass.

Corollary 12.2.5. *The following statements are due to K. Weierstrass:*

1. *The algebraic polynomials in one variable are dense in $C[a, b]$.*
2. *The polynomials in n variables are dense in $C(K)$ with $K \subset \mathbb{R}^n$ compact.*
3. *The trigonometric polynomials are dense in $C[0, 2\pi]$.*

Proof. Only the last statement needs to be addressed. The fact that the trigonometric polynomials are an actual algebra follows from these well known equalities

$$\cos(\alpha)\cos(\beta) = 2^{-1}(\cos(\alpha+\beta) + \cos(\alpha-\beta))$$
$$\sin(\alpha)\sin(\beta) = 2^{-1}(\cos(\alpha-\beta) - \cos(\alpha+\beta))$$
$$\sin(\alpha)\cos(\beta) = 2^{-1}(\sin(\alpha+\beta) + \sin(\alpha-\beta))$$

and separation of points in $[0, \pi]$ is done just by $\sin t$ and $\cos t$. ∎

Chapter 13

Appendix B: Some properties of L^p spaces

13.1 Basic properties

Along the chapter (Ω, Σ, μ) will be a complete measure space. Eventually we could require the measure to be finite or σ-finite, however completeness is important in order not to care about what happens on a null measure set. The spaces $\mathcal{L}^p(\mu)$ for $0 < p < +\infty$ are defined as

$$\mathcal{L}^p(\mu) = \{f \text{ measurable} : \int |f|^p \, d\mu < \infty\}.$$

The case $p = \infty$ is treated in a different way

$$\mathcal{L}^\infty(\mu) = \{f \text{ measurable} : \exists M > 0, \mu(|f| > M) = 0\}.$$

We call that property being *essentially bounded*. We may complete the scale of spaces by taking $\mathcal{L}^0(\mu)$ the set of measurable real-valued functions, that was named $\mathcal{M}$ in the chapter of Measure Theory. Firstly note the following fact.

Proposition 13.1.1. *$\mathcal{L}^p(\mu)$ is a vector space for $0 \leq p \leq +\infty$.*

Proof. The case $\mathcal{L}^0(\mu)$ was already studied in Measure Theory, and $\mathcal{L}^\infty(\mu)$ is quite obvious. Note that if $a, b \geq 0$ then

$$\left(\frac{a+b}{2}\right)^p \leq \max\{a^p, b^p\} \leq a^p + b^p.$$

Therefore

$$\int |f+g|^p \, d\mu \leq 2^p \int |f|^p d\mu + 2^p \int |g|^p d\mu < \infty$$

if $f, g \in \mathcal{L}^p(\mu)$. Homogeneity is evident. ∎

On these spaces we define a distinguished functional for $p >' 0$

$$\|f\|_p = \left(\int |f|^p \, d\mu \right)^{1/p} \text{ if } f \in \mathcal{L}^p(\mu) \text{ and } 0 < p < \infty \, ;$$

$$\|f\|_\infty = \inf\{M \geq 0 : \mu(|f| > M) = 0\} \text{ if } f \in \mathcal{L}^\infty(\mu).$$

It is easy to see that $\|\lambda f\|_p = |\lambda| \|f\|_p$ for any $\lambda \in \mathbb{R}$ and $\|f\|_p = 0$ if and only if $f = 0$ almost everywhere. The following is not so obvious.

Proposition 13.1.2. *$\| \cdot \|_p$ is a seminorm for $1 \leq p \leq \infty$.*

Proof. The case $p = \infty$ is easy, so we will assume $1 \leq p < \infty$. After we know the functional is homogeneous, being a norm is equivalent to prove the convexity of the "unit ball"

$$B_p = \{f \in \mathcal{L}^p(\mu) : \|f\|_p \leq 1\} = \{f \in \mathcal{L}^p(\mu) : \int |f|^p \, d\mu \leq 1\}.$$

Let $f, g \in B_p$ and $\lambda \in [0,1]$. The convexity of $t \to t^p$ implies

$$|\lambda f(x) + (1-\lambda) g(x)|^p \leq \lambda |f(x)|^p + (1-\lambda) |g(x)|^p$$

Integrating we get

$$\int |\lambda f + (1-\lambda) g|^p \, d\mu \leq \lambda \int |f|^p \, d\mu + (1-\lambda) \int |g|^p \, d\mu \leq 1$$

which is the desired convexity of B_p.
For the sake of completeness we will prove that the convexity of B_p implies the triangle property for $\| \cdot \|_p$. Indeed, we may assume $f, g \in \mathcal{L}^p(\mu)$ are such that $\|f\|_p, \|g\|_p \neq 0$. Then $f/\|f\|_p, g/\|g\|_p \in B_p$ and thus

$$\frac{\|f\|_p}{\|f\|_p + \|g\|_p} \frac{f}{\|f\|_p} + \frac{\|g\|_p}{\|f\|_p + \|g\|_p} \frac{g}{\|g\|_p} \in B_p$$

and thus

$$\left\| \frac{f+g}{\|f\|_p + \|g\|_p} \right\|_p \leq 1$$

implying $\|f+g\|_p \leq \|f\|_p + \|g\|_p$ as desired. ∎

Note that the key inequality for convexity goes the oposite way if $p < 1$, however we have the following.

Proposition 13.1.3. *The formula $d(f,g) := \|f-g\|_p^p$ defines an invariant translation pseudometric on $\mathcal{L}^p(\mu)$ for $0 < p < 1$.*

Proof. Note that for $a, b \geq 0$ and $0 < p \leq 1$ we have $(a+b)^p \leq a^p + b^p$. Therefore

$$\int |f-g|^p \, d\mu \leq \int |f-h|^p \, d\mu + \int |h-g|^p \, d\mu$$

and the fact that the metric is translation invariant is obvious. ∎

Another useful inequality.

Theorem 13.1.4. *Assume that $\mu(\Omega) = 1$, we are given a convex function $\phi : (a,b) \to \mathbb{R}$ and $f \in \mathcal{L}^1$ such that $f(x) \in (a,b)$ for almost every point. Then*

$$\phi\left(\int f \, d\mu\right) \leq \int \phi \circ f \, d\mu$$

where the integral is taken with value ∞ in case $\phi \circ f \notin \mathcal{L}^1$.

Proof. The inequality is evident for simple functions provided that they are taken into the "canonical form", that is, on a partition of Ω. The general statement follows by taking limits: if $s_n \to f$ pointwise then $\phi \circ s_n \to \phi \circ f$ for a sequence of simple functions such that $s_n \in (a,b)$ because a convex function is continuous. In order to ensure the convergence of the integral note that for $(\phi \circ f)^-$ is posible to use dominated convergence (the function ϕ is bounded below by a liner one) and for $(\phi \circ f)^+$ monotone convergence. ∎

Consider $\sim$ the equivalence relation $f \sim g$ if $f = g$ almost everywhere. The quotient spaces $L^p(\mu) = \mathcal{L}^p(\mu)/\sim$ are still vector spaces and the functional $\|\cdot\|_p$ is well defined on them. Note that now $\|\cdot\|_p$ is a norm on $L^p(\mu)$ if $1 \leq p \leq \infty$ and $\|\cdot\|_p^p$ defines a translation invariant metric on $L^p(\mu)$ if $0 < p < 1$.

Theorem 13.1.5. *$(L^p(\mu), \|\cdot\|_p)$ is complete for $1 \leq p \leq \infty$.*

Proof. Assume $p < \infty$. It is enough to prove that an absolutely convergent series is convergent, so assume $(f_n) \in L^p(\mu)$ with $\sum_{n=1}^{\infty} \|f_n\|_p < \infty$. Take

$g_n = \sum_{k=1}^n |f_k|$ which is an increasing sequence of positive functions. The triangle property of the norm implies $\sup_n \|g_n\|_p < \infty$. In particular we have

$$\int \lim_n g_n^p = \lim_n \int g_n^p \, d\mu < \infty$$

that implies $\lim_n g_n < \infty$ almost everywhere. Therefore the series $\sum_{n=1}^\infty f_n$ is almost everywhere pointwise convergent. Let f its limit whereas its convergent and take 0 otherwise. Clearly f is a measurable function and we have

$$\int |f|^p \, d\mu \leq \int (\sum_{n=1}^\infty |f_n|)^p \, d\mu = \int \lim_n g_n^p \, d\mu < \infty$$

meaning that $f \in L^p(\mu)$. The convergence of the series to f will be consequence of the $\|\cdot\|_p$-boundedness of the tails

$$\|f - \sum_{k=1}^n f_k\|_p^p \leq \|\sum_{k=n+1}^\infty |f_k|\|_p^p \leq (\sum_{k=n+1}^\infty \|f_k\|_p)^p \to 0$$

where the last inequality comes from the monotone convergence theorem applied to the triangle inequality. The case $p = \infty$ can be handled with the same standard ideas as the proof of the completeness of ℓ_∞ or $C(K)$. ∎

13.2 Convergence

Here we will compare several types of convergence. Firstly we will introduce the notion of *convergence in measure*. We say that a sequence $(f_n) \subset \mathcal{L}^0(\mu)$ converge in measure to f if

$$\lim_n \mu(\{|f_n - f| > \varepsilon\}) = 0$$

for every $\varepsilon > 0$. Clearly, the limit in measure is determined almost everywhere. Note that Chebyshev inequality says that if $f \in L^p(\mu)$ and $\varepsilon > 0$ then

$$\mu(\{|f| \geq \varepsilon\}) \leq \varepsilon^{-p} \int |f|^p \, d\mu = \varepsilon^{-p} \|f\|_p^p$$

implying the following.

Proposition 13.2.1. *The convergence in norm $\|\cdot\|_p$ implies the convergence in measure.*

We also have.

Proposition 13.2.2. *If $\mu(\Omega) < \infty$ then the convergence almost everywhere implies the convergence in measure.*

Proof. If (f_n) converges to f almost everywhere then for every $\varepsilon > 0$ we have

$$\mu(\bigcap_{n=1}^{\infty} \bigcup_{k \geq n} \{|f_k - f| > \varepsilon\}) = 0.$$

Thus, if $\mu(\Omega) < \infty$, then

$$\mu(\{|f_n - f| > \varepsilon\}) \leq \mu(\bigcup_{k \geq n} \{|f_k - f| > \varepsilon\}) \to 0$$

as wished. ∎

The convergence in measure is a topological one.

Proposition 13.2.3. *If $\mu(\Omega) < \infty$ then the convergence in measure is metrized by*

$$d(f, g) = \int \min\{|f - g|, 1\}\, d\mu.$$

Proof. Since the notion is translation invariant we will consider neighbourhoods of 0. If $0 < \varepsilon < 1$ then

$$\mu(|f| > \varepsilon) \leq \varepsilon^{-1} \int \min\{|f|, 1\}\, d\mu$$

and thus the d-convergence implies the convergence in measure. On the other hand, if (f_n) converges in measure to 0 the right-hand side of the following formula can be done as smaller as we wish

$$\int \min\{|f_n|, 1\}\, d\mu \leq \varepsilon\mu(\Omega) + \mu(\{|f_n| > \varepsilon\})$$

which implies the convergence in the metric d. ∎

The argument in the proof of the following proposition was already used to prove Theorem 8.6.5.

Proposition 13.2.4. *If a sequence is convergent in measure, then it has a subsequence which converges almos everywhere.*

Proof. Let (f_n) converging in measure to f. Then it is possible to find n_1 such that

$$\mu(\{|f_{n_1} - f| > 1\}) \leq 1/2.$$

Inductive it is possible to build an increasing sequence $n_1 < n_2 < \dots$ such that the sets

$$A_k = \{|f_{n_k} - f| > 1/k\}$$

satisfy $\mu(A_k) \leq 2^{-k}$. Take $A = \bigcap_{k=1}^{\infty} \bigcup_{j \geq k} A_j$. And note that $\mu(A) = 0$. By construction we have for any $x \in A^c$ that $|f_{n_k}(x) - f(x)| \leq 1/k$ from a certain k on, and so the theorem is proven. ∎

We also have the following.

Theorem 13.2.5 (Egoroff's theorem)**.** *Assume $\mu(\Omega) < \infty$ and let (f_n) be a sequence of measurable functions that converges to f almost everywhere. Then for every $\varepsilon > 0$ there is a set $\Omega_\varepsilon \in \Sigma$ such that $\mu(\Omega \setminus \Omega_\varepsilon) < \varepsilon$ and (f_n) converges to f uniformly on Ω_ε.*

Proof. Consider the sequence $g_n = \sup\{|f_k - f| : k \geq n\}$ which converges to 0 almost everywhere. By Proposition 13.2.2, (g_n) converges to 0 in measure. Therefore, for every $n \in \mathbb{N}$ we can find $k_n \in \mathbb{N}$ such that

$$\mu\left(\{g_{k_n} > 1/n\}\right) < 2^{-n}\varepsilon.$$

Take $A_n = \{g_{k_n} \leq 1/n\}$ and $\Omega_\varepsilon = \bigcap_{n=1}^{\infty} A_n$. By construction, it is easy to check that $\mu(\Omega \setminus \Omega_\varepsilon) < \varepsilon$. For any $n \in \mathbb{N}$ and $x \in \Omega_\varepsilon \subset A_n$ we have

$$|f_m(x) - f(x)| \leq g_{k_n}(x) \leq 1/n$$

for any $m \geq k_n$, that implies the uniform convergence of (f_n) on Ω_ε. ∎

13.3 Classification of L^p spaces and examples

Let start with two examples with totally opposite behaviour. For a set Γ we will denote $\ell^p(\Gamma)$ the space L^p built on the measure space $(\Gamma, \mathcal{P}(\Gamma), \#)$. If $\Gamma = \mathbb{N}$ we will write simply ℓ^p. For these spaces we have $\ell_{p_1} \subset \ell_{p_2}$ if $p_1 \leq p_2$, and moreover the inclusion is continuous with norm 1. On the other hand,

if (Ω, Σ, μ) is a finite measure space the inclusion happens in the reverse way $L^{p_2}(\mu) \subset L^{p_1}(\mu)$ if $p_1 \leq p_2$. Indeed, assume $f \in L^{p_1}(\mu)$, then

$$\int |f|^{p_2}\, d\mu \leq \int_{|f|\leq 1} |f|^{p_2}\, d\mu + \int_{|f|>1} |f|^{p_2}\, d\mu$$

$$\leq \mu(\Omega) + \int_{|f|>1} |f|^{p_1}\, d\mu < \infty.$$

The norm of the inclusion can sharply estimated with the help of Hölder inequality, see next section.

The general case happens to be a blend of the two previous ones, although we will state the general result under the hypothesis of σ-finiteness.

Proposition 13.3.1. *Let (Ω, Σ, μ) be a σ-finite measure space and $1 \leq p \leq \infty$. Then $L^p(\mu)$ is isometric to a direct sum of $\ell^p(\Gamma)$ and $L^p(\nu)$ where $\Gamma \subset \mathbb{N}$ and ν is an atom-free finite measure (eventually void).*

Proof. By a result of measure theory we know that $\Omega = \Omega_a \cup \Omega_f$ with $\Omega_a \cap \Omega_f = \emptyset$ where Ω_a is atomic and Ω_f is atom-free. Clearly

$$\int |f|^p\, d\mu = \int_{\Omega_a} |f|^p\, d\mu + \int_{\Omega_f} |f|^p\, d\mu$$

what implies L^p is ℓ^p-sum of $L^p(\Omega_a)$ and $L^p(\Omega_f)$. Now, let $(A_\gamma)_{\gamma\in\Gamma}$ an enumeration of the atoms. The map

$$T : \ell^p(\Gamma) \to L^p(\Omega_a)$$

defined by $T((x_\gamma)) = \sum_\gamma x_\gamma \mu(A_\gamma)^{-1} \chi_{A_\gamma}$ is an isometry (details are left to the reader). For the atom-free part, if it is not of finite measure already, we may consider a decomposition $\Omega_f = \bigcup_n P_n$ where $\mu(P_n) = 1$. Consider the measure $\nu(A) = 2^{-n}\nu(P_n \cap A)$ which is finite on Ω_f and the map

$$S : L^p(\Omega_f) \to L^p(\nu)$$

defined by $T(f) = \sum_n 2^n \chi_{P_n} f$ is an isometry. ∎

In case that $L^p(\mu)$ is separable for some $1 \leq p < \infty$ (equivalently, $L^p(\mu)$ is separable for all $1 \leq p < \infty$ or (Σ, d_μ) is separable, see Proposition 8.6.2) it is posible to chose ν to be the Lebesgue measure on $[0, 1]$.

13.4 Duality

We will use the following arithmetical identity: if $1 < p, q < \infty$ satisfies $1/p + 1/q = 1$ and $a, b \geq 0$ then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

and the identity only happens if $a^p = b^q$. The proof of this identity can be obtained geometrically by interpreting the summands on the right-hand side as areas limited by the curve $y = x^{p-1}$ (or equivalently $x = y^{q-1}$).

Theorem 13.4.1 (Hölder inequality)**.** *Let $1 \leq p, q \leq \infty$ satisfy $1/p+1/q = 1$. If $f \in L^p(\mu)$ and $g \in L^q(\mu)$ then $fg \in L^1(\mu)$, $\|fg\|_1 \leq \|f\|_p \|g\|_q$. Moreover, if the functions are normalized then the equality holds if and only if $|f|^p = |g|^q$ almost everywhere.*

Proof. The case $1 \in \{p, q\}$ is easy. If $f \in L^\infty$ then the inequality $|f| \leq \|f\|_\infty$ holds almost everywhere, so for every integrable function g we have

$$\int |fg| \, d\mu \leq \int \|f\|_\infty g \, d\mu = \|f\|_\infty \|g\|_1.$$

Assume $1 < p, q < \infty$. We also may assume $\|f\|_p, \|g\|_q > 0$ otherwise both members of the inequality turn to be 0. Consider the norm-one functions $f/\|f\|_p$ and $g/\|g\|_q$ and apply the arithmetic inequality

$$\frac{|fg|}{\|f\|_p \|g\|_q} \leq \frac{|f|^p}{p\|f\|_p^p} + \frac{|g|^q}{q\|f\|_q^q}.$$

Integration gives

$$\int \frac{|fg|}{\|f\|_p \|g\|_q} \, d\mu \leq \int \frac{|f|^p}{p\|f\|_p^p} \, d\mu + \int \frac{|g|^q}{q\|g\|_q^q} \, d\mu = \frac{1}{p} + \frac{1}{q} = 1.$$

Therefore

$$\int |fg| \, d\mu \leq \|f\|_p \|g\|_q$$

as wanted. The statement about when the equality holds follows easily. ∎

With the help of Hölder inequality we can now obtain a precise bound for the inclusion operator between L^p-spaces when $\mu(\Omega) < \infty$. Assume $p_1 \leq p_2$ and $f \in L^{p_2} d\mu$. Then

$$\int |f|^{p_1} d\mu = \int 1 \cdot |f|^{p_1} \leq \left(\int 1^q d\mu \right)^{1/q} \left(\int (|f|^{p_1})^{p_2/p_1} d\mu \right)^{p_1/p_2}$$

where q is the conjugate exponent to p_2/p_1, that is, $1/q = 1 - p_1/p_2$ and thus we have

$$\|f\|_{p_1} \leq \mu(\Omega)^{1/p_2 - 1/p_1} \|f\|_{p_2}.$$

The sharpness of this bound can be tested on the function $f = 1$.

Theorem 13.4.2. *Let (Ω, Σ, μ) a measure space and $1 \leq p, q \leq \infty$. Then the map $\mathfrak{J} : L^q(\mu) \to L^p(\mu)^*$ defined by $\mathfrak{J}(g)(f) = \int fg \, d\mu$ for $g \in L^q(\mu)$ and $f \in L^p(\mu)$ is an (injective) isometry. Moreover*

1. *$L^p(\mu)^* = \mathfrak{J}(L^q(\mu))$ if $1 < p < \infty$;*
2. *$L^1(\mu)^* = \mathfrak{J}(L^\infty(\mu))$ if μ is σ-finite or a cardinal measure;*
3. *$\mathfrak{J}(L^1(\mu)) \subsetneq L^\infty(\mu)^*$ when they are infinite dimensional.*

Proof. Hölder's inequality implies that $\mathfrak{J}$ is well defined and $\|\mathfrak{J}(g)\| \leq \|g\|_p$. In order to check that $\mathfrak{J}$ is actually an isometry, if $1 < p < \infty$ take $f = \text{sign}(g)|g|^{q-1}$ and note that $\|f\|_p = \|g\|_q^{q-1}$ We have

$$\int fg \, d\mu = \int |g|^q \, d\mu = \|g\|_q^q = \|f\|_p \|g\|_q$$

In case, $p = \infty$ take just $f = \text{sign}(g)$ with same proof, and for $p = 1$ it is possible to find "almost norming" functions taking $\varepsilon > 0$ and $f = \chi_A$ where $A \subset \{|g| \geq \|g\|_\infty - \varepsilon\}$ with $0 < \mu(A) < \infty$.
Firstly assume μ is finite and $1 \leq p < \infty$. Let F be a continuous linear functional defined on $L^p(\mu)$. Then the formula $\nu(A) = F(\chi_A)$ for $A \in \Sigma$ defines a σ-additive signed measure. Indeed, the formula is well defined as $\chi_A \in L^p(\mu)$ and finite additivity is quite obvious. For the σ-additivity firstly note that $|\nu(A)| \leq (\mu(A))^{1/p}$. If we have a disjoint sequence $(A_n) \subset \Sigma$ then

$$\|\chi_{\bigcup_{k=1}^\infty A_k} - \chi_{\bigcup_{k=1}^n A_k}\|_p = \|\chi_{\bigcup_{k=n+1}^\infty A_k}\|_p = \mu(\bigcup_{k=n+1}^\infty A_k)^{1/p} \to 0$$

as n goes to ∞. That implies

$$\nu(\bigcup_{k=1}^{\infty} A_k) = F(\chi_{\bigcup_{k=1}^{\infty} A_k}) = \lim_n F(\chi_{\bigcup_{k=1}^{n} A_k}) = \lim_n \sum_{k=1}^{n} \nu(A_k) = \sum_{k=1}^{\infty} \nu(A_k).$$

We also have ν is of bounded variation

$$\sum_{n=1}^{\infty} |\nu(A_n)| = \sum_{n=1}^{\infty} \pm F(\chi_{A_n}) = F(\sum_{n=1}^{\infty} \pm \chi_{A_n}) \leq \|F\| \, \mu(\Omega)^{1/p}.$$

The measure ν is absolutely continuous with respect to μ, so the Radon-Nikodym theorem gives us a function $g \in \mathcal{L}^1(\mu)$ such that

$$F(\chi_A) = \nu(A) = \int_A g \, d\mu = \int \chi_A g \, d\mu.$$

The extreme equality extends to simple functions naturally and then to any $f \in \mathcal{L}^\infty(\mu)$

$$F(f) = \int f g \, d\mu$$

because of the uniform denseness of simple functions among the bounded measurable functions. Now we are going to check that g lies actually in $\mathcal{L}^q(\mu)$. Indeed, if $p = 1$ we claim that g is essentially bounded. Otherwise, for every n it would be possible to find $A \in \Sigma$ with $\mu(A) > 0$ and $|g(x)| > n$ for $x \in A$. Taking $f = \text{sign}(g)\chi_A$ we will have

$$|F(f)| = \left| \int f g \, d\mu \right| \geq n \, \mu(A) = n \, \|f\|_1$$

which violates the continuity of F as $f \neq 0$. This argument also gives $\|g\|_\infty \leq \|F\|$. If $p > 1$, assume that A is a set where g is bounded. Take $f = \text{sign}(g)\chi_A |g|^{q-1}$ which is also bounded. Note that $|f|^p = |g|^q$ on A. We have

$$\int_A |g|^q d\mu = \int f g \, d\mu = F(f) \leq \|F\| \, \|f\|_p = \|F\| \, \left(\int_A |g|^q d\mu \right)^{1/p}$$

and thus

$$\left(\int_A |g|^q d\mu \right)^{1/q} = \left(\int_A |g|^q d\mu \right)^{1-1/p} \leq \|F\|.$$

Since the bound does not depend on A, taking $A_n = \{|g| \leq n\}$ and applying the monotone convergence theorem we get $g \in \mathcal{L}^q(\mu)$ and $\|g\|_q \leq \|F\|$.
The cases $p = 1$ and $p > 1$ can be put together in the following way: if (Ω, Σ, μ) is a general measure space, then for every $A \in \Sigma$ there is $g_A \in \mathcal{L}^q(\mu)$ supported by A such that $F(f) = \int f g_A \, d\mu$ for every $f \in \mathcal{L}^p(\mu)$ supported by A and $\|g_A\|_q \leq \|F\|$. In case, (Ω, Σ, μ) was σ-finite it is clear how to extend the result. Assume (A_n) are disjoint, cover Ω and have finite measure. Put $g_n = g_{A_n}$ and define $g(x) = g_n(x)$ if $x \in A_n$. The function g is measurable and for every $f \in \mathcal{L}^p(\mu)$ and $n \in \mathbb{N}$ we have

$$F(\chi_{\bigcup_{k=1}^n A_k} f) = \sum_{k=1}^n F(\chi_{A_k} f) = \sum_{k=1}^n \int_{A_n} f g \, d\mu = \int_{\bigcup_{k=1}^n A_k} f g \, d\mu$$

which means that g represents F on $\bigcup_{k=1}^n A_k$. The previous observation gives

$$\int_{\bigcup_{k=1}^n A_k} |g|^q \, d\mu \leq \|F\|^q.$$

The monotone convergence theorem implies $g \in \mathcal{L}^q(\mu)$. Now g represents F on the whole space because $\chi_{\bigcup_{k=1}^n A_k} f \to f$ in the $\|\cdot\|_p$ norm.
So far, representation theorem has been proved for $1 \leq p < \infty$ and (Ω, Σ, μ) a σ-finite measure space. The general statement for $1 < p < \infty$ follows from the fact that a continuous operator defined of $L^p(\mu)$ is supported on a σ-finite subset of Ω. Let $F : L^p(\mu) \to \mathbb{R}$ a continuous operator such that is not σ-finitely supported in the following sense: for every countably many sets of finite positive measure (A_n) there is A disjoint with them all and $f \in L^p(\mu)$ supported in A such that $F(f) > 0$. Using transfinite induction is possible to find a disjoint collection $(A_\alpha)_{\alpha=1}^{\omega_1} \subset \Sigma$ with positive measure such that A_α supports a norm one element $f_\alpha \in L^p(\mu)$ with $F(f_\alpha) \neq 0$. A standard argument with non countable cardinals shows that there $\delta > 0$ such that $F(f_\alpha) > \delta$ for infinitely many α's. We may relabel some of these elements with $\mathbb{N}$. Now note that $\|\sum_{k=1}^n f_n\| = n^{1/p}$ and $F(\sum_{k=1}^n f_n) > n\delta$. That implies $\|F\| > \delta n^{1-1/p}$ that goes to ∞ as n does which is a contradiction. We have now F is supported on a σ-finite set so we can apply the previous part to get $g \in L^q(\mu)$ supported on this same set representing f.
The last fact follows form the existence of finitely additive measures on $\mathcal{P}(\mathbb{N})$ which are not σ-additive, which in terms of Banach spaces is just the fact that $(\ell^1)^{**} \neq \ell^1$. A Hahn-Banach extension of a functional suitably defined by a finitely additive measure on (Ω, Σ) will do the work. ∎

13.5 Uniform convexity of $L^p(\mu)$ for $1 < p < \infty$

We say that a Banach space is *uniformly convex* if

$$\delta_X(t) = 1 - \sup\{\left\|\frac{x+y}{2}\right\| : \|x\| = \|y\| = 1, \|x-y\| \geq t\} > 0$$

for all $t \in (0, 2]$. The function $\delta_X(t)$ is called the *modulus of uniform convexity* of X. The main aim now is to prove the uniform convexity of $L^p(\mu)$ spaces for $1 < p < \infty$. We will distinguish between two cases with different proofs.

Theorem 13.5.1. *$L^p(\mu)$ is uniformly convex for $1 < p < \infty$.*

Proof. The norm $\|\cdot\|_p$ in $\mathbb{R}^2$ is strictly convex for $1 < p < \infty$. A compactness argument shows that for every $t > 0$ then

$$\delta(t) = 1 - \sup\{\left|\frac{x+y}{2}\right|^p : |x|^p + |y|^p = 1, |x-y| \geq t\} > 0.$$

Using homogeneity we get that if $|a-b|^p \geq t^p(|a|^p + |b|^p)$ then

$$\left|\frac{a+b}{2}\right|^p \leq (1 - \delta(t)) \left(\frac{|a|^p + |b|^p}{2}\right).$$

Assume $f, g \in L^p(\mu)$ with $\|f\|_p = \|g\|_p = 1$ and $\|f - \|_p \geq t$. Consider the set

$$A = \{t^p(|f|^p + |g|^p) \leq 4|f-g|^p\}$$

Note that

$$\int_{A^c} |f-g|^p d\mu \leq \frac{t^p}{4} \int_{A^c} (|f|^p + |g|^p)\, d\mu \leq \frac{t^p}{2}$$

and therefore

$$\int_A |f-g|^p d\mu \geq \frac{t^p}{2}.$$

We get the following inequality we will use soon

$$\int_A \frac{|f|^p + |g|^p}{2}\, d\mu \geq \int_A \left|\frac{|f| + |g|}{2}\right|^p d\mu \geq \int_A \left|\frac{f-g}{2}\right|^p d\mu \geq \frac{t^p}{2^p 2}.$$

The uniform convexity follows easily now

$$1 - \left\|\frac{f+g}{2}\right\|_p^p \geq \int \left(\frac{|f|^p + |g|^p}{2} - \left|\frac{f+g}{2}\right|^p\right) d\mu \geq$$

$$\int_A \left(\frac{|f|^p + |g|^p}{2} - \left| \frac{f+g}{2} \right|^p \right) d\mu \geq \delta(t) \int_A \frac{|f|^p + |g|^p}{2}\, d\mu \geq \frac{t^p \delta(t)}{2^{p+1}}.$$

Indeed,

$$\sup\{ \left\| \frac{f+g}{2} \right\|_p : \|f\|_p = \|g\|_p, \|f-g\|_p \geq t \} \leq \left(1 - \frac{t^p \delta(t)}{2^{p+1}} \right)^{1/p} < 1$$

as wished. ∎

In the case $p \geq 2$ is possible to obtain a sharper inequality by simpler means. We will use an arithmetical inequality. If $a, b \in \mathbb{R}$ and $2 \leq p < \infty$ then

$$\left(\left| \frac{a+b}{2} \right|^p + \left| \frac{a-b}{2} \right|^p \right)^{1/p} \leq \left(\left(\frac{a+b}{2} \right)^2 + \left(\frac{a-b}{2} \right)^2 \right)^{1/2} = \left(\frac{a^2+b^2}{2} \right)^{1/2}$$

and so

$$\left| \frac{a+b}{2} \right|^p + \left| \frac{a-b}{2} \right|^p \leq \left(\frac{a^2+b^2}{2} \right)^{p/2} \leq \frac{|a^p| + |b|^p}{2}$$

because of the convexity of $t \to t^{p/2}$. Now, if $f, g \in L^p(\mu)$ with $\|f\|_p = \|g\|_p = 1$ then

$$\int \left| \frac{f+g}{2} \right|^p d\mu + \int \left| \frac{f-g}{2} \right|^p d\mu \leq \frac{1}{2} \left(\int |f|^p \, d\mu + \int |g|^p \, d\mu \right) = 1$$

and thus

$$\left\| \frac{f+g}{2} \right\|_p^p \leq 1 - \left\| \frac{f-g}{2} \right\|_p^p.$$

It follows

$$\left\| \frac{f+g}{2} \right\|_p \leq \left(1 - \left(\frac{t}{2} \right)^p \right)^{1/p}$$

if $\|f-g\|_p \geq t$. Therefore $L^p(\mu)$ is uniformly convex with modulus of uniform convexity

$$\delta_{L^p(\mu)}(t) \geq 1 - \left(1 - \left(\frac{t}{2} \right)^p \right)^{1/p} \sim \frac{t^p}{2^p p}$$

when $t \sim 0$.
The case $1 < p < 2$ is trickier and it is known that $\delta_{L^p(\mu)}(t) \sim c_p t^2$.

Chapter 14

Appendix C: Introduction to Lagrangian and Hamiltonian mechanics

14.1 Coordinates and speeds

The standard way of presenting Newtonian mechanics is with the help of vector notation. For instance, the position of a system made up of N particles is depicted by N spatial vectors $(\vec{r}_1, \ldots, \vec{r}_N)$. The equations in Lagrangian mechanics are scalar, so we need to change to the quite unnatural notation

$$(x_1, x_2, x_3, \ldots, x_{3N-2}, x_{3N-1}, x_{3N}) := (\vec{r}_1, \ldots, \vec{r}_N)$$

In many applications the positions of the particles cannot be totally arbitrary, that means, the variables $\{x_i : 1 \leq i \leq 3N\}$, and perhaps the time t, satisfy dependence relations, which can be interpreted in the language of smooth manifolds (see Section 5.2). Assume that locally the position can be described by $n \leq 3N$ independent variables $\{q_j : 1 \leq j \leq n\}$ which are calles *generalized coordinates*. In Newtonian mechanics the equations of movement are elegantly expressed in terms of the cartesian coordinates and their derivatives. Lagrangian mechanics seeks a form of the equations of movement valid for every set of coordinates. In particular, expressing the equations of the movement in terms of the generalized coordinates would simplify their resolution.

As the state of a system is determined by the positions and speeds of all the particles whose it is made up, the first step is to find the relations between

cartesian and generalized coordinates. Put $q = (q_1, \ldots, q_n)$ and $\dot{q} = (\dot{q}_1, \ldots, \dot{q}_n)$ is derivatives with respect time, that we will call *generalized speeds*. As we said above

$$x_i = x_i(q, t)$$

so the derivative with respect time will be of the form

$$\dot{x}_i = \dot{x}_i(q, \dot{q}, t).$$

The explicit expression can be find using the chain rule

$$\dot{x}_i = \sum_j \frac{\partial x_i}{\partial g_j}\, \dot{q}_j + \frac{\partial x_i}{\partial t}.$$

Note the difference between total derivative with respect to t which concerns to an actual movement, and the partial derivative with respect to t which expresses a constraint of the system that depends of the moment. On the other hand, we may consider theoretical variations of coordinates keeping t constant, which are called *virtual displacements*.

Lemma 14.1.1. *The following relations hold*

$$\frac{\partial \dot{x}_i}{\partial \dot{q}_j} = \frac{\partial x_i}{\partial q_j}\ ;$$

$$\frac{\partial \dot{x}_i}{\partial q_j} = \frac{d}{dt}\left(\frac{\partial x_i}{\partial q_j}\right).$$

Proof. We have

$$\dot{x}_i = \sum_k \frac{\partial x_i}{\partial q_k}\, \dot{q}_k + \frac{\partial x_i}{\partial t}$$

from which the first formula follows trivially. For the second one, just compare these expressions

$$\frac{\partial \dot{x}_i}{\partial q_j} = \sum_k \frac{\partial^2 x_i}{\partial q_k \partial q_j}\, \dot{q}_k + \frac{\partial^2 x_i}{\partial t \partial q_j}$$

and, by the chain rule,

$$\frac{d}{dt}\left(\frac{\partial x_i}{\partial q_j}\right) = \sum_k \frac{\partial^2 x_i}{\partial q_j \partial q_k}\, \dot{q}_k + \frac{\partial^2 x_i}{\partial q_j \partial t}$$

which are equal by commutativity of derivation as the coordinate functions are supposed regular enough. ∎

14.2 Forces, work and energy

The force acting on a particle is a quantitative description on how the interaction with the rest of the universe. We may distinguish between the interaction with other particles of the system and interactions with objects from outside the system, that is *internal forces* and *outer forces*. We assume that the *principle of superposition* holds, that is, that the interactions are additive and so the total force applied on a particle is the sum of all the individual forces (one for each interaction) applied on it.

The work done by a force, which only depends on the configuration of the system, along the displacement of a particle is independent of time in the sense that the speed of the particle does not matter for the computation. Therefore, we may consider the work done by forces in virtual displacements. As the force is variable, it is convenient to use a differential expression that physicist like to interpret in terms of infinitesimals. In order to tell apart of real displacements done as time goes by, we will use δ instead of d for differentials. The differential of work expressed in cartesian coordinates appears as

$$\delta W = \sum_i f_i \, \delta x_i$$

where $\{f_i : 1 \leq i \leq 3N\}$ are the components of the force suitably enumerated. In order to express δW in terms of the generalized coordinates, observe that

$$\delta x_i = \sum_j \frac{\partial x_i}{\partial q_j} \, \delta q_j$$

and then

$$\delta W = \sum_i f_i (\sum_j \frac{\partial x_i}{\partial q_j} \, \delta q_j) = \sum_j (\sum_i f_i \, \frac{\partial x_i}{\partial q_j}) \, \delta q_j = \sum_j Q_j \, \delta q_j$$

where $Q_j = \sum_i f_i \, \frac{\partial x_i}{\partial q_j}$ are called the *generalized components of the force*. One important task is to determine the forces acting on a given system.

In the very important case that the force is *conservative* (work done between two point does not depend on the trajectory) the differential form δW is exact and there exist a function $V = V(q)$ such that $Q_j = -\frac{\partial V}{\partial q_j}$ where the sign minus is taken for the sake of the physical interpretation of V as *potential*

energy. The equation we will prove for the movement admits more general potentials that eventually could depend on the speed too.

Another important case of forces, specially from the Lagrangian point of view, are the *constraint forces* which reduce the degrees of freedom of the system and so motivates the use of a suitable choice of generalized coordinates. In many cases practical cases the constraint forces do no work in virtual displacements (e.g. they are normal to the movement). The system where that holds are called *holonomic systems* and the generalized components of the constraint forces reduce to 0.

If the masses of the particles are suitably enumerated, the *kinetic energy* of the system is defined as

$$T = \frac{1}{2} \sum_i m_i \dot{x}_i^2$$

In despite that the energy depends only on the cartesian speeds, in generalized coordinates depends on $\{q, \dot{q}, t\}$. However, if t does not appear explicitly in the change of variables, then T is a quadratic function with respect to $\dot{q}$. Let us finish with the following easy fact.

Lemma 14.2.1.

$$m_i \ddot{x}_i = \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{x}_i} \right).$$

14.3 Equations of movement

From now on we will assume that the cartesian coordinates are taken with respect to an *inertial frame*. In such a case, Newton's law applies to every particle of the system and therefore

$$m_i \ddot{x}_i = f_i$$

for $1 \leq i \leq 3N$ and f_i being the i-th component of the total force applied to the correspondent particle of the system. We wish to find the form of those equations in generalized coordinates, the answer is the following.

Proposition 14.3.1. *The equations of the movement of the system of n degrees of freedom in generalized coordinates are*

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} = Q_j$$

for $1 \leq j \leq n$. Moreover, if the system is holomomic the computation of Q_j does not include the constraint forces.

Proof. Consider the following chain of equalities where the previous lemmata are applied together a trick based on the derivative of a product

$$Q_j = \sum_i f_i \frac{\partial x_i}{\partial q_j} = \sum_i m_i \ddot{x}_i \frac{\partial x_i}{\partial q_j} = \sum_i \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{x}_i}\right)\frac{\partial x_i}{\partial q_j} =$$

$$\sum_i \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{x}_i}\frac{\partial x_i}{\partial q_j}\right) - \sum_i \frac{\partial T}{\partial \dot{x}_i}\frac{d}{dt}\left(\frac{\partial x_i}{\partial q_j}\right) =$$

$$\sum_i \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{x}_i}\frac{\partial \dot{x}_i}{\partial \dot{q}_j}\right) - \sum_i \frac{\partial T}{\partial \dot{x}_i}\frac{\partial \dot{x}_i}{\partial q_j} = \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j}$$

which gives the desired equality. The observation about constraint forces is consequence that their generalized components with respect to such a set of coordinates is 0. ∎

14.4 The Lagrangian

Assume that our system is holonomic and the external forces derive from a potential $V(q,t)$ in the sense that $Q_j = -\frac{\partial V}{\partial q_j}$ (that includes the conservative case if there is no dependence on t) then the equations of movement can be rewritten as

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} + \frac{\partial V}{\partial q_j} = 0.$$

Define the *Langrangian* function of the system by

$$L(q,\dot{q},t) = T(q,\dot{q},t) - V(q,t)$$

and then the previous equality becomes

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0.$$

That equation could be obtained under more general assumptions meaning that we would consider the equation as the starting point for the foundations of Mechanics. Let us state this fact as a theorem.

Theorem 14.4.1. *Every mechanical system is characterized by a function $L(q, \dot{q}, t)$ in such a way that at any initial configuration the trajectories of the system in time satisfies the equations*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0$$

for $1 \leq j \leq n$, being n the degrees of freedom of the system.

From now on we will asume that the Lagrange equations characterize the evolution of the system once we know its Lagrangian, however the derivation from Newton's laws to this form was done under specific hypotheses (holonomy, potentials depending only on positions...).

In order to have a nicer foundation for the Mechanics, it is worth noticing that it is possible to change the practical but unnatural differential equation by an "universal minimization principle".

Theorem 14.4.2. *Every mechanical system is characterized by a function $L(q, \dot{q}, t)$ in such a way that the trajectory carried out by the system between two given configurations $(q_1, \dot{q}_1, t_1)$ and $(q_2, \dot{q}_2, t_2)$ an extremal value of the integral*

$$\int_{t_1}^{t_2} L(q, \dot{q}, t)\, dt$$

amongst all the other smooth trajectories between the same configurations.

Proof. For simplicity we will assume that the system has only a degree of freedom. Actually, this computation was done on Chapter 4, nevertheless we will repeat it with the notation from Mechanics. From now on $(q, \dot{q}, t)$ denotes the trajectory followed by the system, so $(q(t_i), \dot{q}(t_i)) = (q_i, \dot{q}_i)$ for $i = 1, 2$. In order to show the extremality of the real trajectory we will consider C^2 perturbations of the form $h(t)$ such that it and $\dot{h}(t)$ vanish at t_1, t_2. Extremality implies the directional derivative

$$\frac{d}{ds}\int_{t_1}^{t_2} L(q + sh, \dot{q} + s\dot{h}, t)\, dt$$

must be 0 at $s = 0$. The parametric derivation can be performed under the integral sign this way

$$\int_{t_1}^{t_2} \Big(\frac{\partial L}{\partial q}(q + sh, \dot{q} + s\dot{h}, t)\dot{h} + \frac{\partial L}{\partial \dot{q}}(q + sh, \dot{q} + s\dot{h}, t)\ddot{h}\Big)\, dt.$$

Integration by parts give that

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{q}}(q+sh, \dot{q}+s\dot{h}, t)\ddot{h}\, dt =$$

$$\frac{\partial L}{\partial \dot{q}}(q+sh, \dot{q}+s\dot{h}, t)\dot{h}\Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}}(q+sh, \dot{q}+s\dot{h}, t)\Big)\dot{h}\, dt$$

$$= -\int_{t_1}^{t_2} \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}}(q+sh, \dot{q}+s\dot{h}, t)\Big)\dot{h}\, dt.$$

Using this information above we get

$$\frac{d}{ds}\int_{t_1}^{t_2} L(q+sh, \dot{q}+s\dot{h}, t)\, dt =$$

$$\int_{t_1}^{t_2} \Big(\frac{\partial L}{\partial q}(q+sh, \dot{q}+s\dot{h}, t) - \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}}(q+sh, \dot{q}+s\dot{h}, t)\Big)\dot{h}\, dt$$

and the annulation at $s=0$ of this implies

$$\int_{t_1}^{t_2} \Big(\frac{\partial L}{\partial q}(q, \dot{q}, t) - \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}}(q, \dot{q}, t)\Big)\dot{h}\, dt = 0$$

for all the perturbations h satisfying the required assumptions. As it is possible to take h with support as small as we wish contained into (t_1, t_2) we deduce that

$$\frac{\partial L}{\partial q}(q, \dot{q}, t) - \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}}(q, \dot{q}, t)\Big) = 0$$

for $t \in (t_1, t_2)$ as we wanted. ∎

Conservation laws are always a first step for the integration, or at least simplification, of the equations of the movement. The conservation of linear moment and angular moment in Newtonian mechanics can be generalized in the following fashion. Associate to a generalized coordinate q_j we will consider the generalized mometum

$$p_j = \frac{\partial L}{\partial \dot{q}_j}\ .$$

We have the following.

Proposition 14.4.3. *If the Lagrangian L does not contain explicitly the coordinate q_j then p_j remains constant along the trajectory of the system.*

Proof.

$$\frac{dp_j}{dt} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) = \frac{\partial L}{\partial q_j} = 0.$$

■

14.5 The Hamiltonian

If the pass from cartesian to generalized coordinates does not involve the time the kinetic energy T is a quadratic form with respect to $\dot{q}$. Assuming as well that V does not depend on $\dot{q}$ we have

$$\sum_j p_j \dot{q}_j = \sum_j \frac{\partial T}{\partial q_j}\dot{q}_j = 2T$$

and so the total energy can be written as

$$T + V = 2T - L = \sum_j p_j \dot{q}_j - L.$$

We can prove the following principle which is somehow more general that the assumptions of the preceding computation.

Theorem 14.5.1. *If L does not contain explicitly the time, then the following magnitude remains constant along the trajectories of the system*

$$H = \sum_j p_j \dot{q}_j - L.$$

Proof.

$$\frac{dH}{dt} = \sum_j \frac{dp_j}{dt}\dot{q}_j + \sum_j p_j \ddot{q}_j - \frac{dL}{dt} =$$

$$\sum_j \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_j}\Big)\dot{q}_j + \sum_j \frac{\partial L}{\partial \dot{q}_j}\ddot{q}_j - \sum_j \frac{\partial L}{\partial q_j}\dot{q}_j - \sum_j \frac{\partial L}{\partial \dot{q}_j}\ddot{q}_j =$$

$$\sum_j \Big(\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_j}\Big) - \frac{\partial L}{\partial q_j}\Big)\dot{q}_j = 0$$

as claimed. ■

If the matrix whose coefficients are

$$\Big(\frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j}\Big)_{i,j}$$

has non null determinant which is a plausible hypothesis regarded from the point of view of the kinetic energy, then for the Jacobian we have

$$\frac{\partial(p_1, p_2, \ldots, p_n)}{\partial(\dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n)} \neq 0$$

because the matrix is the same. That implies the possibility of replacing the set of generalized speeds $\dot{q}$ by the generalized moments p.

The *Hamiltonian* is the function H of the preceding theorem when expressed in terms of the set of variables (q, p). The previous result establishes the invariance of the Hamiltonian along the trajectories of the system if it does not contain explicitly the time. Note that we can recover L from H, so they containt the same information about the system. However, the equations of the movement when expressed in terms of the Hamiltonian look a bit different.

Theorem 14.5.2. *The trajectories of the system with Hamiltonian $H(q, p)$ with n degrees of freedom are the solutions of the equations*

$$\frac{\partial H}{\partial q_j} = -\frac{dp_j}{dt}; \qquad \frac{\partial H}{\partial p_j} = \frac{dq_j}{dt}.$$

Allegedly the number of equations is double. The reason is that the relationships between variables p and q must be included. This is equivalent to consider the obvious relations $\dot{q} = dq/dt$ in Theorem 14.4.1.

Proof. Using the very definition of H we have

$$\frac{\partial H}{\partial q_j} = \sum_i p_i \frac{\partial \dot{q}_i}{\partial q_j} - \Big(\frac{\partial L}{\partial q_j}\Big)_{p=cte} = \sum_i p_i \frac{\partial \dot{q}_i}{\partial q_j} - \frac{\partial L}{\partial q_j} - \sum_i \frac{\partial L}{\partial \dot{q}_i}\frac{\partial \dot{q}_i}{\partial q_j} =$$

$$\sum_i p_i \frac{\partial \dot{q}_i}{\partial q_j} - \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_j}\Big) - \sum_i p_i \frac{\partial \dot{q}_i}{\partial q_j} = -\frac{dp_j}{dt}.$$

And for the other equation note that

$$\frac{\partial L}{\partial p_j} = \sum_i \frac{\partial L}{\partial q_i}\frac{\partial q_i}{\partial p_j} + \sum_i \frac{\partial L}{\partial \dot{q}_i}\frac{\partial \dot{q}_i}{\partial p_j} = 0 + \sum_i p_j \frac{\partial \dot{q}_i}{\partial p_j},$$

so

$$\frac{\partial H}{\partial p_j} = \sum_i p_i \frac{\partial \dot{q}_i}{\partial p_j} + \dot{q}_j - \frac{\partial L}{\partial p_j} = \sum_i p_i \frac{\partial \dot{q}_i}{\partial p_j} + \dot{q}_j - \sum_i \frac{\partial L}{\partial \dot{q}_j}\frac{\partial \dot{q}_i}{\partial p_j} = \frac{dq_j}{dt}$$

as wanted. ■

The solution of the Hamilton system produces trajectories in the *phase space* (q, p). The advantage with respect to the representation in $(q, \dot{q})$ is that in the phase space. For instance, if H does not contain the time then the level curves $H = cte$ are the trajectories of the system. Indeed, the orthogonal field to the level curves is $(\frac{\partial H}{\partial q}, \frac{\partial H}{\partial p})$ and so the field $(\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial q})$ which is orthogonal to the previous one must be tangent to the level curves of H.

Another reason, the $2n$-dimensional *divergence* of the Hamiltonian field is zero. Indeed,

$$\sum_j \frac{\partial}{\partial q_j}\Big(\frac{\partial H}{\partial p_j}\Big) + \sum_j \frac{\partial}{\partial p_j}\Big(-\frac{\partial H}{\partial q_j}\Big) = \sum_j \frac{\partial^2 H}{\partial p_j \partial q_j} - \sum_j \frac{\partial^2 H}{\partial q_j \partial p_j} = 0.$$

That implies the $2n$-dimensional volume is preserved by the flow of the system (as we did in dimension 3, see Section 11.3), that is, interpreting the Hamiltonian field $(\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial q})$ a the speed field of a fluid. That leads to the following result of Liouville.

Theorem 14.5.3. *Given a mechanical system, its Hamiltonian flow preserves volumes in the phase space. In particular, given an open set D in (q, p) which is composed of initial states, let be D_t the evolution of those states after a time $t > 0$. Then the $2n$-dimensional volume of D_t remains constant and equal to the volume of D.*

A well know application is the so called Poincaré's recurrence theorem: if the orbits of the system are confined in a bounded set then any initial state will be arbitrarily approximated by the evolution of the system after some time. We will finish with a version of the uncertainty principle for mechanical systems. For simplicity consider only a degree of freedom. Assume that the position q and momentum p are known with some errors $\Delta q(0)$ and $\Delta p(0)$ at the beginning. Then after some time the combined uncertainty does not decrease

$$\Delta q(t)\Delta p(t) \geq \Delta q(0)\Delta p(0).$$

Indeed, the product of the uncertainties $\Delta q(t)\Delta p(t)$ represents the area of a rectangle that contains the evolution through the flow of the system of the rectangle of sides $\Delta q(0)$ and $\Delta p(0)$ that contain all the possible initial states. It is somehow surprising that Classical Mechanics anticipates *Heissemberg's uncertainty principle*.

Bibliography

[1] M. Aigner, G. M. Ziegler, *Proofs form the Book*, (6th ed.), Springer, 2018.

[2] T. M, Apostol, *Análisis Matemático*, (1^a ed.), Ed. Reverté, Barcelona, 1960.

[3] T. M, Apostol, *Análisis Matemático*, (2^a ed.), Ed. Reverté, Barcelona, 1991.

[4] F. Bombal, L. Rodríguez, G. Vera *Problemas de Análisis Matemático*, (3 vol.), Editorial AC, 1994.

[5] I. Bronshtein, K. Semendiaev, *Manual de Matemáticas para ingenieros y estudiantes*, Editorial MIR, Moscú, 1973.

[6] G. Bruhat *Cours de Physique générale*, (4 vol.) Masson, 1963-1968.

[7] H. Cartan, *Formas Diferenciales*, Omega - Colección Métodos, Barcelona, 1972.

[8] F. del Castillo, *Análisis matemático II*, Alhambra, 1980.

[9] G. Choquet, *Cours de Topologie*, (2^{eme} ed.), Dunod, Paris, 2000.

[10] D. L. Cohn, *Measure Theory*, Birkhäuser, 2013.

[11] R. Courant, F. John, *Introducción al Cálculo y al Análisis Matemático*, (2 vol.), Limusa, México, 1982.

[12] P. J. Davis, R. Hersh, *Experiencia Matemática*, Ministerio de Educación y Ciencia, Ed. Labor, Barcelona, 1988.

[13] E. A. Desloge, *Classical Mechanics*, John Wiley & Sons Inc, 1982.

[14] J. Dieudonné, *Fundamentos de Análisis Moderno*, Ed. Reverté, Barcelona, 1979.

[15] B. A. Dubrovin, A. T. Fomenko, S. P. Novikov, *Métodos y aplicaciones de Geometría Moderna*, (2 vol.), Editorial URSS, Moscú, 2000.

[16] C. H. Edwards JR., *Advanced Calculus of Several Variables*, Dover Publ. Inc., New York, 1994.

[17] M. Fabian, P. Habala, P. Hájek, V. Montesinos and V. Zizler, *Banach Space Theory. The Basis for Linear and Nonlinear Analysis*, CMS Books in Mathematics, Springer, New York, 2011.

[18] K. Falconer, *Fractal Geometry : Mathematical Foundations and Applications*, Wiley, 2006.

[19] J. A. Fernández Viña, *Análisis Matemático*, (3 vol.), Tecnos, Madrid, 1986.

[20] J. A. Fernández Viña, E. Sánchez Mañes, *Ejercicios de Análisis Matemático*, (3 vol.), Tecnos, Madrid, 1994.

[21] V. Ilin, E. Pozniak, *Fundamentos del Análisis Matemático*, (3 vol.) Editorial MIR, Moscú, 1991.

[22] G. Jäger, *Física Teórica*, Labor, 1959.

[23] G. Joos, I. M. Freeman, *Theoretical Physics*, (3^{rd} ed.), Blackie & Son LMT, Glasgow, 1960.

[24] A. N. Kolmogórov, S. V. Fomín, *Elementos de la Teoría de Funciones y del Análisis Funcional*, Editorial MIR, Moscú, 1975.

[25] L. D. Kudriávtsev, *Curso de Análisis Matemático*, (2 vol.), Editorial MIR, Moscú, 1988.

[26] G. G. Lorenz, *Approximations of Functions*, (2^{nd} ed.), Chelsea, 1986.

[27] J. A. Marín Tejerizo, *Ampliación de Matemáticas para Técnicos*, S.A.E.T.A., Madrid, 1965.

[28] J. A. Marín Tejerizo, *Problemas de Cálculo Integral*, S.A.E.T.A., Madrid, 1968.

[29] A. Mishchenko, A. Fomenko, *A course of Differential Geometry and Topology*, MIR Publishers, Moscow, 1988.

[30] N. Piskunov, *Cálculo Diferencial e Integral*, Montaner y Simon, S.A., Barcelona, 1978.

[31] P. Puig Adam, *Curso Teórico-Práctico de Cálculo Integral*, Biblioteca Matemática S. L., Madrid, 1975.

[32] J. Rey Pastor, P. Pi Calleja, C. A. Trejo, *Análisis Matemático* (3 vol.) Kapelusz, Buenos Aires, 1968.

[33] W. Rudin, *Análisis Real y Complejo*, (3^a ed.) McGraw-Hill, 1987.

[34] S. Salas, E. Hille, G. J. Etgen, *Calculus Una y varias variables*, Ed. Reverté, 2006.

[35] L. A. Santaló, *Vectores y Tensores, con sus Aplicaciones*, EUDEBA, Buenos Aires, 1962.

[36] J. J. Scala, R. Riaza López, L. Ortiz Berrocal, *Cálculo Vectorial Aplicado*, Sección de Pub. E.T.S. de Ingenieros Industriales, Madrid, 1967.

[37] J. H. Shapiro, *A Fixed-Point Farrago*, Springer, 2016.

[38] M. Spivak, *Cálculo en Variedades*, Ed. Reverté, Barcelona, 1988.

[39] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, 1970.

[40] E. M. Stein, R. Shakarchi, *Princeton Lectures in Analysis*, (4 vol.), Princeton and Oxford University Press, 2003.

[41] G. Valiron, *Théorie des Fonctions*, (3^{eme} ed.) Masson, Paris, 1990.

[42] G. Vera, *Lecciones de Análisis Matemático II*, `https://webs.um.es/gvb/OCW/OCW-AM-II_files/PDF/AM-II.pdf`

[43] C. E. Weatherburn, *Advanced Vector Analysis*, G. Bell and Sons, LMT, London, 1943.