Advanced analysis

Julien Vovelle

Contents

1 References **2**

2 Integration theory **3**
2.1 Complex measures . . . 3
2.2 F. Riesz Representation Theorem by measures . . . 4
2.3 Integration of Banach-valued functions . . . 10
2.4 Strong compactness in L^p . . . 16
2.4.1 Criteria for strong compactness . . . 16
2.4.2 Obstructions to strong convergence . . . 21
2.5 Differentiation, maximal functions . . . 27
2.5.1 Maximal functions . . . 27
2.5.2 Lebesgue's differentiation Theorem . . . 30
2.5.3 A brief reminder on the theory of ordinary differential equations . . . 32
2.5.4 Solving ODEs with vector fields of Sobolev regularity . . . 35

3 Hahn-Banach theorem **41**
3.1 Preliminaries on hyperplanes and linear functional . . . 41
3.1.1 Riesz' lemma . . . 41
3.1.2 Quotient spaces . . . 41
3.1.3 Hyperplanes . . . 42
3.2 Hahn-Banach theorem . . . 44
3.2.1 Analytic form of the Hahn-Banach theorem: extension of linear functionals 44
3.2.2 Geometric form of the Hahn-Banach theorem: separation . . . 46
3.2.3 Applications of the Hahn-Banach theorem in finite dimension . . . 49
3.2.4 Uniform convexity, projection on convex sets . . . 52

4 Fréchet spaces **55**

5 Weak topologies **57**
5.1 Definition of the weak topology, weakly open, closed, bounded sets . . . 57
5.2 Bounded sets in Fréchet spaces . . . 61
5.3 Topological dual spaces of some standard Banach and Fréchet spaces . . . 63
5.3.1 Additive set functions . . . 63
5.3.2 A list of topological dual spaces . . . 65
5.3.3 A list of topological dual spaces: spaces built on $\mathbb{R}^d$ . . . 66
5.3.4 The bounded-open topology and the final/inductive topology . . . 69
5.3.5 Separable spaces . . . 72

5.4 Weak-star topology . . . 73
5.4.1 Some applications of the Banach-Alaogu theorem . . . 75
5.4.2 Reflexive spaces . . . 77
5.4.3 Milman-Pettis Theorem . . . 78
5.4.4 Minimization of functionals . . . 80

6 Distribution theory **81**
6.1 Distribution: definition, elementary operations . . . 82
6.1.1 Space of distributions . . . 82
6.1.2 Bump functions, partition of unity, support . . . 84
6.1.3 Order of a distribution, derivation of a distribution . . . 86
6.1.4 Some operations on distributions . . . 88
6.1.5 Fundamental solutions of PDEs - I . . . 93
6.2 Tempered distributions . . . 94
6.2.1 Definition of tempered distributions . . . 94
6.2.2 Operations on tempered distributions . . . 98
6.2.3 Fundamental solutions of PDEs - II . . . 101

7 Sobolev spaces **104**
7.1 Definition and first properties . . . 105
7.2 Sobolev spaces defined on the whole space . . . 108
7.2.1 Proof of Morrey's inequality . . . 109
7.2.2 Proof of the Gagliardo-Nirenberg-Sobolev Inequality . . . 112
7.2.3 Sobolev's embeddings . . . 114
7.3 Sobolev spaces defined on bounded domains . . . 116
7.3.1 Geometry of the boundary of bounded open sets . . . 116
7.3.2 Local parametrization in a tubular neighbourhood of ∂U . . . 117
7.3.3 Measure on the boundary . . . 118
7.3.4 Approximation by smooth functions . . . 120
7.3.5 Extension . . . 121
7.3.6 Trace . . . 123
7.3.7 Kernel of the trace operator . . . 126
7.4 Operation of Lipschitz maps on Sobolev spaces . . . 127
7.4.1 Lipschitz functions . . . 127
7.4.2 Operation of Lipschitz functions on Sobolev spaces . . . 131

A Surface measure **132**

1 References

Here are some references for the course.

- *Real analysis*, by G.B. Folland [Fol99]
- *Functional analysis*, by P.D. Lax, [Lax02]
- *Functional analysis, Sobolev spaces and partial differential equations*, by H. Brezis, [Bre11]
- *Analysis*, by E.H. Lieb and M. Loss, [LL01]
- *Functional Analysis,* by W. Rudin, [Rud73]

- *Éléments de distributions et d'équations aux dérivées partielles,* by C. Zuily, [Zui02]
- *Analyse II*, by C. Villani, [Vil03]

With the color cyan are given some references to the first-year courses, for those who attended them.

2 Integration theory

The first sections, Section 2.1 and Section 2.2 are presented as complements only. We will not discuss these topics in the course, but you may find it useful later to have them here. In particular, the formulation of the Riesz representation theorem (by measure), Theorem 2.3, is adapted to a treatment of functions of bounded variations.

2.1 Complex measures

Definition 2.1 (Complex measure)**.** Let $(X, \mathcal{A})$ be a measure space. A complex measure over $(X, \mathcal{A})$ is a set function $\mu\colon \mathcal{A} \to \mathbb{C}$ such that, for all $A \in \mathcal{A}$, one has

$$\mu(A) = \sum_{i=1}^{\infty} \mu(A_i), \tag{2.1}$$

for all countable partition $(A_i)_{i\geq 1}$ of A, the sum in (2.1) being absolutely convergent.

When a complex measure takes values in $\mathbb{R}$, one sometimes speaks of *real* measure, or *signed* measure (however, certain authors authorize signed measures to take the value $+\infty$, *i.e.* $\mu\colon \mathcal{A} \to (-\infty, +\infty]$, *e.g.* [SS05, Chapter 6.4]). In this notes, "signed measure" are complex measures such that $\mu(A) \in \mathbb{R}$ for all $A \in \mathcal{A}$, so $|\mu(A)| < +\infty$ for every A.

Proposition 2.1 (Total variation)**.** *Let μ be a complex measure over the measure space $(X, \mathcal{A})$. The formula*

$$|\mu|(A) = \sup \left\{ \sum_{i=1}^{\infty} |\mu(A_i)| \right\}, \tag{2.2}$$

where the supremum is taken over all countable partitions $(A_i)_{i\geq 1}$ of A, defines a non-negative finite measure $|\mu|$ on $\mathcal{A}$ called the total variation *of μ.*

The most difficult point in Proposition 2.1 is the proof that $|\mu|$ is a finite measure, see [Rud87, Theorem 6.4] for instance, or "Intégration et mesures", Chapter VII-2.

Proposition 2.2 (Total variation, some properties)**.** *Let μ be a signed measure over the measure space $(X, \mathcal{A})$. We have the following properties:*

a) Jordan decomposition: $\mu = \mu_+ - \mu^-$, *where*

$$\mu^+ := \frac{1}{2}(|\mu| + \mu) \text{ and } \mu^- := \frac{1}{2}(|\mu| - \mu) \tag{2.3}$$

are finite positive measure,

b) Factorization: *there exists a measurable function $\sigma\colon X \to \mathbb{R}$ such that*

$$|\sigma(x)| = 1 \text{ for a.e. } x \in X \text{ and } \mu(A) = \int_A \sigma(x) d|\mu|(x), \tag{2.4}$$

for all $A \in \mathcal{A}$,

c) Hahn decomposition: *there exists some measurable sets A_- and A_+ such that $A_- \cap A_+ = \emptyset$ and the positive and negative parts defined in (2.3) are concentrated on A_+ and A_- respectively:*

$$\mu^+(E) = \mu(E \cap A_+) \text{ and } \mu^-(E) = \mu(E \cap A_-), \tag{2.5}$$

for all $E \in \mathcal{A}$.

Proof of Proposition 2.2. Taking $A_1 = A$ and $A_i = \emptyset$ if $i > 1$ in (2.2), we see that

$$|\mu(A)| \leq |\mu|(A), \tag{2.6}$$

for all $A \in \mathcal{A}$. The measures $\mu_\pm$ in (2.3) are positive therefore, and it is obvious that $\mu = \mu^+ - \mu^-$ (note also that $|\mu| = \mu^+ + \mu^-$). The factorization (2.4) is a consequence of the Radon-Nikodym theorem. To obtain the Hahn decomposition (2.5), we use (2.4) and set $A_\pm = \{\sigma = \pm 1\}$. □

The integral of a measurable function $f\colon X \to \mathbb{R}$ against μ is defined as

$$\int_X f d\mu = \int_X f d\mu^+ - \int_X f d\mu^-, \tag{2.7}$$

provided

$$\int_X |f| d|\mu| < +\infty. \tag{2.8}$$

Under (2.8), both terms in the right-hand side of (2.7) are well defined and finite, in particular we do not have to consider an indeterminate form $\infty - \infty$. Using the "notation" (2.7), the relation (2.4) can be written

$$\int_X f d\mu = \int_X f\sigma d|\mu|, \tag{2.9}$$

for $f = \mathbf{1}_A$, the characteristic function of A. By linearity, (2.9) holds true when f is a simple function. If $f\colon X \to \mathbb{R}$ is positive and satisfies (2.8), approximation by simple functions shows that (2.9) remains true for f. Finally, using the decomposition $f = f^+ - f^-$ into positive and negative parts, we obtain (2.9) for any f satisfying (2.8).

2.2 F. Riesz Representation Theorem by measures

For $d, m \in \mathbb{N} \setminus \{0\}$, the set of continuous, compactly supported functions $\mathbb{R}^d \to \mathbb{R}^m$ is denoted by $C_c(\mathbb{R}^d; \mathbb{R}^m)$. If $m = 1$, we use the simpler notation $C_c(\mathbb{R}^d)$. Let $\mathcal{A}$ be a σ-algebra on $\mathbb{R}^d$ containing all the Borel sets, let μ be a Borel positive measure on $\mathcal{A}$ such that $\mu(K) < +\infty$ for all compact subset K of $\mathbb{R}^d$, let $\sigma\colon \mathbb{R}^d \to \mathbb{R}^m$ be a bounded measurable function. For each $f \in C_c(\mathbb{R}^d; \mathbb{R}^m)$, we set

$$T(f) = \int_{\mathbb{R}^d} f(x) \cdot \sigma(x) d\mu(x), \tag{2.10}$$

where $f(x) \cdot \sigma(x)$ is the canonical scalar product of $f(x)$ and $\sigma(x)$ in $\mathbb{R}^m$ (we also denote by $|\cdot|$ the associated euclidean norm in (2.11) below). Then (2.10) defines a linear functional $C_c(\mathbb{R}^d; \mathbb{R}^m) \to \mathbb{R}$ with the following property: for all compact K of $\mathbb{R}^d$,

$$\sup \left\{ |T(f)|; f \in C_c(\mathbb{R}^d; \mathbb{R}^m), |f(x)| \leq 1 \text{ for all } x \in \mathbb{R}^d, f \text{ supported in } K \right\} < +\infty. \tag{2.11}$$

Indeed, the sup in (2.11) is bounded by $\|\sigma\|_{L^\infty(\mathbb{R}^d)} \mu(K)$. Conversely, we have the following representation theorem.

Theorem 2.3 (F. Riesz Representation Theorem by measures)**.** *Let $T\colon C_c(\mathbb{R}^d;\mathbb{R}^m)\to\mathbb{R}$ be a linear functional satisfying (2.11) for all compact K of $\mathbb{R}^n$. Then, there exists a σ-algebra $\mathcal{A}$ on $\mathbb{R}^d$ containing all the Borel sets, a regular Borel positive measure μ on $\mathcal{A}$ such that $\mu(K)<+\infty$ for all compact subset K of $\mathbb{R}^n$, a measurable function $\sigma\colon\mathbb{R}^d\to\mathbb{R}^m$ such that $|\sigma(x)|$ for a.e. $x\in\mathbb{R}^n$, for which the identity (2.10) is satisfied for all $f\in C_c(\mathbb{R}^d;\mathbb{R}^m)$.*

This statement of the F. Riesz Representation Theorem is taken from [EG15, Chapter 1.8] (note that we do not refer to [EG92], but instead to the revised version [EG15]). Beware that the term “measure” in [EG15] means[1] “outer measure”. You may be more familiar with the versions of the Riesz Representation Theorem (by measures) which can be found in *Real and complex analysis* by Rudin (see [Rud87, p.40 and p.130] for instance, or “Intégration et mesures”, Chapter VII-1). These versions can be deduced from Theorem 2.3, see Theorem 2.4 and Theorem 2.5 below. The measure μ in Theorem 2.3 is regular. We recall that this means that it is both outer regular: for all $A\in\mathcal{A}$,

$$\mu(A)=\inf\{\mu(U); A\subset U, U \text{ open}\}, \tag{2.12}$$

and inner regular: for all $A\in\mathcal{A}$,

$$\mu(A)=\sup\{\mu(K); K\subset A, K \text{ compact}\}. \tag{2.13}$$

We make one more additional remark (this will be used in the proof of Theorem 2.5): the measure μ in Theorem 2.3 is defined as follows: first,

$$\mu(V)=\sup\left\{T(f); |f(x)|\leq 1 \text{ for all } x\in\mathbb{R}^d, f \text{ supported in } V\right\}, \tag{2.14}$$

if V is open and then

$$\mu(A)=\inf\left\{\mu(V); A\subset V, V \text{ open}\right\}, \tag{2.15}$$

for any set A.

Theorem 2.4 (F. Riesz Representation Theorem, positive linear functionals)**.** *Let $T\colon C_c(\mathbb{R}^d)\to\mathbb{R}$ be a positive linear functional:*

$$f\in C_c(\mathbb{R}^d), f(x)\geq 0 \textit{ for all } x\in\mathbb{R}^d\Rightarrow T(f)\geq 0. \tag{2.16}$$

Then, there exists a σ-algebra $\mathcal{A}$ on $\mathbb{R}^d$ containing all the Borel sets, a regular Borel positive measure μ on $\mathcal{A}$ such that $\mu(K)<+\infty$ for all compact subset K of $\mathbb{R}^n$, for which the identity

$$T(f)=\int_X f d\mu. \tag{2.17}$$

is satisfied for all $f\in C_c(\mathbb{R}^d)$.

Proof of Theorem 2.4. Consider first T satisfying (2.16). Using the linearity of T, we obtain the following monotony property

$$f\leq g\Leftrightarrow 0\leq g-f\Rightarrow 0\leq T(f-g)=T(f)-T(g)\Leftrightarrow T(f)\leq T(g). \tag{2.18}$$

Let K be a compact subset of $\mathbb{R}^d$ and let $F_K\in C_c(\mathbb{R}^d)$ satisfy $0\leq F_K\leq 1$, $F_K\equiv 1$ on K. Then every function $f\in C_c(\mathbb{R}^d)$ such that $|f|\leq 1$, f supported in K can be bounded as follows:

$$-F_K\leq f\leq F_K. \tag{2.19}$$

[1] see “Intégration et mesures”, Chapter III-2.

Using (2.18), we deduce that $|T(f)| \le T(F_K)$: the condition (2.11) is satisfied. We deduce that T has the representation

$$T(f) = \int_X f\sigma(x)d\mu, \tag{2.20}$$

where $\sigma(x) \in \{-1, +1\}$. Suppose by contradiction that $A := \{\sigma = -1\}$ has a positive measure ε. By regularity of μ, there is a compact K with $\mu(K) \ge \varepsilon/2 > 0$ such that $K \subset A$. Let $f_K \in C_c(\mathbb{R}^d)$ be a function such that $0 \le f_k \le 1$ and $f_K \equiv 1$ on K. We have then

$$-T(f_K) \ge \int_K f_K d\mu \ge \mu(K) > 0, \tag{2.21}$$

a contradiction. So $\sigma = 1$ a.e., and we obtain the representation (2.17). □

Theorem 2.5 (F. Riesz Representation Theorem, bounded linear functionals)**.** *Let $\mathcal{M}(\mathbb{R}^d)$ denote the set of signed Borel measures μ on $\mathbb{R}^d$ endowed with the norm $\|\mu\|_{\mathcal{M}(\mathbb{R}^d)} = |\mu|(\mathbb{R}^d)$. Let $C_0(\mathbb{R}^d)$ denote the space of continuous functions $\mathbb{R}^d \to \mathbb{R}$ which tends to 0 at $+\infty$, endowed with the sup norm $\|f\|_{C(\mathbb{R}^d)} = \sup_{x\in\mathbb{R}^d} |f(x)|$. Then the map*

$$\mu \mapsto T_\mu, \quad T_\mu(f) = \int_{\mathbb{R}^d} f d\mu, \tag{2.22}$$

defines an isomorphism between $\mathcal{M}(\mathbb{R}^d)$ and the topological dual space to $C_0(\mathbb{R}^d)$ such that

$$\|T_\mu\| = \|\mu\|_{\mathcal{M}(\mathbb{R}^d)} \tag{2.23}$$

for all $\mu \in \mathcal{M}(\mathbb{R}^d)$.

Before giving the proof of Theorem 2.5, we must explain some notations in (2.23) and make some comments.

1. The integral in (2.22) has been defined in (2.7). Any $f \in C_0(\mathbb{R}^d)$ is measurable and bounded, so (2.8) is satisfied since

$$\int_{\mathbb{R}^d} |f|d|\mu| \le \|f\|_{C(\mathbb{R}^d)}|\mu|(\mathbb{R}^d). \tag{2.24}$$

2. The space $C_0(\mathbb{R}^d)$ endowed with the sup norm is a Banach space (it is a closed subspace of the Banach space $C_b(\mathbb{R}^d)$ of bounded continuous functions). The natural norm on the topological dual space to $C_0(\mathbb{R}^d)$ is the dual norm

$$\|T\| = \sup\{|T(f)|; f \in C_0(\mathbb{R}^d), \|f\|_{C(\mathbb{R}^d)} \le 1\}. \tag{2.25}$$

   This is the norm used in (2.23).

3. We have asserted that $\|\mu\|_{\mathcal{M}(\mathbb{R}^d)} := |\mu|(\mathbb{R}^d)$ defines a norm on $\mathcal{M}(\mathbb{R}^d)$. This is indeed the case and is rather easy to check: the positive homogeneity and the triangular inequality follow from (2.2). By (2.6), we have $\mu \equiv 0$ if $|\mu|(\mathbb{R}^d) = 0$.

4. We will establish that the map (2.22) is injective, the identity $\mu = \nu$, for $\mu, \nu \in \mathcal{M}(\mathbb{R}^d)$, being understood in the sense that $\mu(A) = \nu(A)$ for all Borel sets A. See also Remark 2.1 below on that aspect.

Proof of Theorem 2.5. Let $\nu \in \mathcal{M}(\mathbb{R}^d)$. By (2.9), we have

$$T_\nu(f) = \int_{\mathbb{R}^d} f(x)\sigma(x)d\mu(x), \quad \mu := |\nu|, \tag{2.26}$$

where $\sigma = 1$ μ-a.e. It follows that

$$|T_\nu(f)| \leq \|f\|_{C(\mathbb{R}^d)}\mu(\mathbb{R}^d). \tag{2.27}$$

This shows that T_ν is in the topological dual to $C_0(\mathbb{R}^d)$ and that $\|T_\nu\| \leq \|\nu\|_{\mathcal{M}(\mathbb{R}^d)}$. Let us apply Theorem 2.3, to show that (2.22) is surjective: if T is in the topological dual space to $C_0(\mathbb{R}^d)$, then T is characterized by its values on $C_c(\mathbb{R}^d)$ since $C_c(\mathbb{R}^d)$ is dense in $C_0(\mathbb{R}^d)$ for the sup norm. It is sufficient to establish that $T(f) = T_\nu(f)$ for a given ν, for all $f \in C_c(\mathbb{R}^d)$. Let K be a compact subset of $\mathbb{R}^d$ and let $f \in C_c(\mathbb{R}^d)$ be supported in K, with $|f| \leq 1$. Then

$$|T(f)| \leq \|T\|\|f\|_{C(\mathbb{R}^d)} \leq \|T\|, \tag{2.28}$$

so (2.11) is satisfied. By Theorem 2.3, there exists a σ-algebra $\mathcal{A}$ on $\mathbb{R}^d$ containing all the Borel sets, a regular Borel positive measure μ on $\mathcal{A}$ such that $\mu(K) < +\infty$ for all compact subset K of $\mathbb{R}^n$, a measurable function $\sigma\colon \mathbb{R}^d \to \mathbb{R}^m$ such that $|\sigma(x)|$ for a.e. $x \in \mathbb{R}^n$. The bound (2.11) that we have obtained by (2.28) is independent of K, and this induces μ to be finite. This is clear if we come back to (2.14): taking $V = \mathbb{R}^d$, we obtain

$$\mu(\mathbb{R}^d) \leq \|T\|. \tag{2.29}$$

For $A \in \mathcal{A}$, let us set

$$\nu(A) = \int_A \sigma d\mu. \tag{2.30}$$

Then ν is a signed measure: if (A_i) is a countable partition of a set A, then

$$\sum_i |\nu(A_i)| \leq \sum_i \mu(A_i) = \mu(A) < +\infty, \tag{2.31}$$

and $\nu(A) = \sum_i \nu(A_i)$ as a consequence of the identity

$$\sum_i \int_{\mathbb{R}^d} \sigma \mathbf{1}_{A_i} d\mu = \int_{\mathbb{R}^d} \sum_i \sigma \mathbf{1}_{A_i} d\mu, \tag{2.32}$$

which follows from the dominated convergence theorem (or Fubini-Tonelli theorem) since

$$\int_{\mathbb{R}^d} \sum_i |\sigma \mathbf{1}_{A_i}| d\mu \leq \mu(A) < +\infty. \tag{2.33}$$

We claim that $|\nu| = \mu$. Let $A \in \mathcal{A}$ and let $B_\pm = \{\sigma = \pm 1\}$. Since $\nu(A \cap B_\pm) = \pm\mu(A \cap B_\pm)$ by (2.30), the definition (2.2) gives us

$$|\nu|(A \cap B_\pm) = \mu(A \cap B_\pm), \tag{2.34}$$

and thus $|\nu| = \mu$ since

$$|\nu|(A) = |\nu|(A \cap B_+) + |\nu|(A \cap B_-). \tag{2.35}$$

We also notice that

$$\nu^+(A \cap B_-) = \frac{1}{2}\left(|\nu|(A \cap B_-) + \nu(A \cap B_-)\right) = \frac{1}{2}\left(\mu(A \cap B_-) - \mu(A \cap B_-)\right) = 0, \tag{2.36}$$

so

$$\nu^+(A) = \nu^+(A \cap B_+) = \frac{1}{2}\left(|\nu|(A \cap B_+) + \nu(A \cap B_+)\right) = \mu(A \cap B_+). \tag{2.37}$$

Similarly $\nu^-(A) = \mu(A \cap B_-)$. Passing from characteristic functions to simple functions and then to integrable functions, we obtain

$$\int_{\mathbb{R}^d} f d\nu^+ = \int_{B_+} f d\mu, \quad \int_{\mathbb{R}^d} f d\nu^- = \int_{B_-} f d\mu, \tag{2.38}$$

for any $f \in C_c(\mathbb{R}^d)$. Finally, by means of (2.38), we can gather (2.10) and the definition (2.7) to obtain

$$T(f) = \int_{\mathbb{R}^d} f\sigma d\mu = \int_{B_+} f d\mu - \int_{B_-} f d\mu = \int_{\mathbb{R}^d} f d\nu^+ - \int_{\mathbb{R}^d} f d\nu^- = T_\nu(f). \tag{2.39}$$

Let us now show the injective character of the map (2.22). Let $\nu_1, \nu_2 \in \mathcal{M}(\mathbb{R}^d)$ be such that $T_{\nu_1} = T_{\nu_2} =: T$. Let us say that a signed measure ν is *regular* if both ν^+ and ν^- are regular (it is easy to see that this is equivalent to the fact that $|\nu|$ is regular). Since the map (2.22) is surjective, as proved already, we have $T = T_\nu$, where ν is a *regular* signed measure on a σ-algebra which contains the Borel sets. Since $\nu_1 = \nu$ & $\nu_2 = \nu$ implies $\nu_1 = \nu_2$, we may as well assume that one of the two measures, say ν_2, is regular. Can we use this extra information to help us to conclude? We will see that this is helpful indeed if ν_1 has a sign (one of the measure $\nu_1^\pm$ is trivial). This is of course a restrictive case, but still, this is interesting to consider it. For the moment, let us go on with the proof of uniqueness, without assuming that ν^1_+ or $\nu^1_- = 0$. Let K be a compact set. Set $f(x) = \mathbf{1}_K(x)$ and consider the sequence of functions

$$f_n(x) = \sup_{y \in \mathbb{R}^d} \left[f(y) - n|x - y|\right]. \tag{2.40}$$

If $x \in K$, the sup in (2.40) is 1. If x is at distance greater than n^{-1} of K, then the sup is 0. In both cases, the sup is reached for $y = x$. Suppose that $0 < d(x, K) < n^{-1}$ and let $y_K \in K$ such that $|x - y_K| = d(x, K)$. If $y \in K$, then

$$f(y) - n|x - y| = 1 - n|x - y| \leq 1 - nd(x, K) = f(y_K) - n|x - y_K|. \tag{2.41}$$

If $y \notin K$, then

$$f(y) - n|x - y| = -n|x - y| \leq 0 \leq 1 - nd(x, K) = f(y_K) - n|x - y_K|. \tag{2.42}$$

Consequently, $f_n(x) = 1 - nd(x, K)$. Finally, we obtain the alternative expression

$$f_n(x) = (1 - nd(x, K))^+. \tag{2.43}$$

We may have defined $f_n(x)$ directly by (2.43), but it is instructive to use the general regularization formula (2.40), which is known as[2] "sup-convolution". This is indeed a way to regularize f since each f_n is n-Lipschitz continuous (see Lemma 2.6 below). We obtain also a non-increasing sequence such that $f_n(x) \downarrow f(x)$ for all x. Note also that each f_n is compactly supported. Recall that the identity $T_{\nu_1}(f_n) = T_{\nu_2}(f_n)$ means

$$\int_{\mathbb{R}^d} f_n d\nu_1^+ - \int_{\mathbb{R}^d} f_n d\nu_1^- = \int_{\mathbb{R}^d} f_n d\nu_2^+ - \int_{\mathbb{R}^d} f_n d\nu_2^-. \tag{2.44}$$

[2] the "inf-convolution" of f is $f_n(x) = \inf_{y \in \mathbb{R}^d} \left[f(y) + n|x - y|\right]$

By monotone convergence in each of the four terms of (2.44), we deduce that

$$\nu_1(K) = \nu_2(K). \tag{2.45}$$

Let us prove that f_n is n-Lipschitz continuous. This can be deduced from the expression (2.43) since $x \mapsto d(x, K)$ is 1-Lipschitz continuous, but we can as well prove the following general lemma.

Lemma 2.6 (Regularity of the sup-convolution)**.** *Let $f\colon \mathbb{R}^d \to \mathbb{R}$ be a bounded function. Then the sup-convolution f_n defined by (2.40) is n-Lipschitz continuous.*

Proof of Lemma 2.6. Fix M such that $|f| \leq M$ on $\mathbb{R}^d$. We show first that the sup is a max. Set $F(y, x) = f(y) - n|y - x|$. Note that $F(y, x) \leq M$ for all y, x, so the sup defining $f_n(x)$ is finite. For each $M \in \mathbb{N} \setminus \{0\}$, there exists $y_m \in \mathbb{R}^d$ such that

$$F(y_m, x) > f_n(x) - m^{-1}. \tag{2.46}$$

The sequence (y_m) is a maximizing sequence. Let us show that it is bounded. We have

$$f_n(x) = \sup_{y \in \mathbb{R}^d} F(y, x) \geq F(x, x) \geq -M. \tag{2.47}$$

With (2.46), this implies

$$n|y_m - x| = -F(y_m, x) + f(y_m) \leq -f_n(x) + m^{-1} + M \leq 2M + m^{-1}. \tag{2.48}$$

Up to extraction of a subsequence, we can assume that (y_m) is converging to a given y_* in $\mathbb{R}^d$. We can pass to the limit in (2.46) then, to deduce that $f_n(x) = F(y_*, x)$. If $x_1, x_2 \in \mathbb{R}^d$ now, and if $y_1, y_2 \in \mathbb{R}^d$ are such that $F(y_i, x_i) = f_n(x_i)$, $i = 1, 2$, then

$$f_n(x_1) - f_n(x_2) = f(y_1) - n|y_1 - x_1| - f_n(x_2) \tag{2.49}$$

The triangular inequality $|y_1 - x_1| \leq |y_1 - x_2| + |x_1 - x_2|$ gives

$$f_n(x_1) - f_n(x_2) \geq -n|x_1 - x_2| + F(y_1, x_2) - f_n(x_2) \geq -n|x_1 - x_2|. \tag{2.50}$$

By symmetry of x_1 and x_2 in (2.50), we obtain the result. □

Let V be an open set. We will use (2.45) to show that $\nu_1(V) = \nu_2(V)$. The set $K_n = V^c \cap \bar{B}(0, n)$ is compact, so $\nu_1(K_n) = \nu_2(K_n)$, which we can rewrite as

$$\nu_1^+(K_n) + \nu_2^-(K_n) = \nu_2^+(K_n) + \nu_1^-(K_n). \tag{2.51}$$

The sequence (K_n) is non-decreasing with $\cup K_n = V^c$. By taking the limit $n \to +\infty$ in (2.51) and using the continuity for increasing limits of positive measures, we obtain

$$\nu_1^+(V^c) + \nu_2^-(V^c) = \nu_2^+(V^c) + \nu_1^-(V^c), \tag{2.52}$$

which is equivalent to $\nu_1(V^c) = \nu_2(V^c)$, and thus to

$$\nu_1(V) = \nu_2(V). \tag{2.53}$$

At this point, let us assume that $\nu_1^- = 0$. We will show that ν^1 is regular. This will give the conclusion $\nu_1 = \nu_2$, by (2.12)-(2.13) and (2.45)-(2.53). Let A be a Borel set and let $\varepsilon > 0$. There exists some compact sets K and some open set V such that

$$K \subset A \subset V, \quad \nu_2 + (V \setminus K) < \varepsilon. \tag{2.54}$$

From (2.45) and (2.53), we deduce then that

$$\nu_1^+(V \setminus K) = \nu_1(V \setminus K) = \nu_2(V) - \nu_2(K) = \nu_2(V \setminus K) \leq \nu_2 + (V \setminus K) < \varepsilon. \tag{2.55}$$

This shows that $\nu_1 = \nu_1^+$ is regular. Let us now show that $\nu_1 = \nu_2$ without the extra assumption that $\nu_1^- = 0$. Let $\mu = \nu_+^1 + \nu_2^-$ and $\tilde{\mu} = \nu_2^+ + \nu_1^-$. These are two positive measures, for which we want to prove that $\mu(A) = \tilde{\mu}(A)$, $A \in \mathcal{B}(\mathbb{R}^d)$, the class of Borel sets of $\mathbb{R}^d$. There are some general results, that allow to infer, from the equality of two measures on a class $\mathcal{C}$, the equality on the σ-algebra $\sigma(\mathcal{C})$ generated by $\mathcal{C}$. In particular, we have the following result ([Bil95, Theorem 10.4], or "Intégration et mesures", Chapter III-1):.

Theorem 2.7. *Assume*

i) $\mathcal{C}$ *is a* π*-system (it is stable by finite intersection),*

ii) $\mu = \tilde{\mu}$ *on* $\mathcal{C}$*,*

iii) $\mathbb{R}^d$ *is* $\mathcal{C}$*-*σ*-finite: there exists a non-decreasing sequence of sets* $A_n \in \mathcal{C}$ *such that* $\mathbb{R}^d = \cup A_n$ *and* $\mu(A_n) = \tilde{\mu}(A_n) < +\infty$ *for each* n*.*

Then $\mu = \tilde{\mu}$ *on* $\sigma(\mathcal{C})$*.*

The finiteness hypothesis *iii)* is necessary, as shown by the example $\mu =$ one-dimensional Lebesgue measure, $\tilde{\mu} = 2\mu$, $\mathcal{C}$ being the class of sets $(a, +\infty)$, $a \in \mathbb{R}$. In our case, the measures are finite: using (2.53), we can directly apply the theorem with $\mathcal{C}$ being the class of open sets of $\mathbb{R}^d$. To conclude, we still have to prove the identity (2.23). It is a direct consequence of (2.27) (where $\mu = |\nu|$) and (2.29)-(2.39). □

Remark 2.1 (Regularity of measures). The proof of Theorem 2.5 shows that any $\mu \in \mathcal{M}(\mathbb{R}^d)$ is regular.

Remark 2.2 (F. and M. Riesz). Frigyes and Marcel Riesz are two Hungarian brothers, both mathematicians. Frigyes Riesz is considered as one of the founder of functional analysis. Marcel Riesz has worked on different topics, the theory of partial differential equations in particular. He was the mentor of Lars Hörmander. We have encountered one (or three, depending on the way we count) theorem by F. Riesz. Theorem 2.15 below involves M. Riesz.

2.3 Integration of Banach-valued functions

Let E be a Banach space with norm $\|\cdot\|$, and let $(X, \mathcal{A}, \mu)$ be a measure space with μ finite: $\mu(X) < +\infty$. In this section, we will discuss the integration of functions $u\colon X \to E$. This theory is needed in various contexts, for instance:

- when one wants to solve partial differential equations involving both a time variable t and a space variable x. The solution u is often viewed as a map $[0, T] \to E$, where E is a "functional space", a space of functions depending on the variable x, the integration space being $X = [0, T]$, $\mu =$Lebesgue measure,
- when one considers random variables taking values in an infinite-dimensional space. Then $(X, \mathcal{A}, \mu)$ is a probability space. A renowned example is given by the Wiener measure, which gives a way to draw elements in $E = C([0, 1]; \mathbb{R}^d)$ at random, modelling the Brownian motion.

For references, see [Eva10, Appendix E.5], [Yos80, Chapter V.5], [DPZ14, Chapter 1.1].

Definition 2.2 (Measurable functions)**.** A function $u\colon X \to E$ is said to be

- *measurable* if $u^{-1}(B) \in \mathcal{A}$ for all Borel set B of E,
- *weakly measurable* if, for all φ in the topological dual E^* of E, the map $\varphi \circ u\colon X \to \mathbb{R}$ is measurable.

Proposition 2.8 (Approximation by simple functions)**.** *Assume that E is* separable. *If $u\colon X \to E$ is measurable, then there is a sequence of simple*[3] *functions u_n such that $\|u_n(x) - u(x)\| \downarrow 0$ for all $x \in X$.*

Proof of Proposition 2.8. Let (v_k) be a sequence of points in E such that $\{v_k; k \in \mathbb{N}\}$ is dense in E. Given a "level" $n \geq 0$, we construct u_n as a function taking values in the finite set $D_n = \{v_k; k \leq n\}$. For each $x \in X$, we set $u_n(x) = v_k$, where v_k is the closest element to $u(x)$ in D_n. Let $\varepsilon > 0$. There exists N such that $\|u(x) - v_N\| < \varepsilon$. By construction then, $\|u(x) - u_n(x)\| < \varepsilon$ for $n \geq N$. So $u_n(x) \to u(x)$ when $n \to +\infty$. It is clear also by construction that $n \mapsto \|u_n(x) - u(x)\|$ is non-increasing. Let us check that u_n is measurable. We have

$$u_n(x) = v_k \iff \forall j \in \{1, \dots, n\}, \|u(x) - v_k\| \leq \|u(x) - v_j\|, \tag{2.56}$$

so

$$\{u_n = v_k\} = \bigcap_{j=1}^{n} \{x \in X; \|u(x) - v_j\| - \|u(x) - v_k\| \geq 0\}. \tag{2.57}$$

To conclude it is sufficient to remark that, for all $v \in E$,

$$f\colon x \mapsto \|u(x) - v\| \tag{2.58}$$

is measurable since $\{f \geq \alpha\} = u^{-1}(\bar{B}(v, \alpha))$. □

Remark 2.3 (Control on $\|u_n\|$)**.** Let us define

$$\tilde{u}_n = u_n \mathbf{1}_{\{\|u_n\| \leq 2\|u\|\}}, \tag{2.59}$$

where (u_n) is the sequence defined in Proposition 2.8. The set $\{\|u_n\| \leq 2\|u\|\}$ is measurable so $(\tilde{u}_n)$ is a sequence of simple functions. For each x such that $u(x) \neq 0$, for each $\varepsilon > 0$ with $\varepsilon < \|u(x)\|$, there is a $n = n(x)$ such that $\|u(x) - u_m(x)\| < \varepsilon$ for all $m \geq n(x)$. Then

$$\|u_m(x)\| \leq \|u(x) - u_m(x)\| + \|u(x)\| \leq 2\|u(x)\|, \tag{2.60}$$

and $\tilde{u}_m(x) = u_m(x)$. This shows that $(\tilde{u}_n)$ is converging point-wise to u. This gives us a sequence of simple functions which converges to u and such that $\|u_n(x)\| \leq 2\|u(x)\|$ for all $x \in X$. This is used in the proof of Proposition-Definition 2.12.

Remark 2.4 (Alternative construction of approximating sequence of simple functions)**.** In Proposition 2.8, we start from the point $u(x) \in E$ and build $u_n(x)$ for $n = 1, 2, \dots$ simply by adding points to the "clouds" of points $D_1, D_2, \dots$ and "jumping" to a new value if it gets closer to $u(x)$. The correction $\tilde{u}_n(x)$ in Remark 2.3 follows the same process, except that we first wait at the point 0, until there arrives a point v_k in the ball $\bar{B}(0, 2\|u(x)\|)$.
To construct the approximating sequence (u_n), there is an other way to proceed, more similar to what is done in the case of real-valued functions. For a fixed size $\varepsilon > 0$, we construct a "grid" of the whole space E by considering some balls $B_{k,\varepsilon}$ centred on the elements v_k, then replace

[3] a function $X \to E$ is said *simple* if it is measurable and takes only a finite number of values

locally u by the constant value given by the center of the ball, and truncate in some way to get a simple function. One can consider the collection of balls $B_{k,\varepsilon} = B(v_k, \varepsilon)$ for instance. They do not form a grid, since there may be some overlaps, so first we correct this by considering the collection $\{B'_{k,\varepsilon}; k \geq 1\}$, where

$$B'_{k,\varepsilon} = B_{k,\varepsilon} \setminus \cup_{i=1}^{k-1} B_{i,\varepsilon}. \tag{2.61}$$

Then we set $u_\varepsilon(x) = v_k$ if $u(x) \in B'_{k,\varepsilon}$ and $k < \varepsilon^{-1}$, and set $u_\varepsilon(x) = 0$ otherwise:

$$u_\varepsilon = \sum_{k<\varepsilon^{-1}} v_k \mathbf{1}_{A_{k,\varepsilon}}, \quad A_{k,\varepsilon} = u^{-1}(B'_{k,\varepsilon}). \tag{2.62}$$

We obtain a sequence of simple function (u_ε) such that $u_\varepsilon(x) \to u(x)$ for all x. Again, appears the problem that we do not control $\|u_\varepsilon(x)\|$ by $\|u(x)\|$ in any satisfactory way. We can correct this as in Remark 2.3, or simply adapt the choice of the "grid" from the very beginning. Indeed, the problem comes from the inadequate choice of the initial grid, which is not sufficiently fine around the origin. So, instead of starting from the collection of balls $B(v_k, \varepsilon)$ with fixed radius ε, we can start from the collection of balls $B_{k,\varepsilon} = B(v_k, \varepsilon\|v_k\|)$, which get smaller as one gets closer to the origin. One can check then that (2.61)-(2.62) defines a sequence (u_ε) of simple functions converging point-wise to u , with the control $\|u_\varepsilon\| \leq 2\|u\|$ as soon as $\varepsilon < 1/2$.

Proposition 2.9 (Notion of measurable function)**.** *Assume that E is* separable. *Then it is equivalent to be measurable and weakly measurable.*

Proof of Proposition 2.9. The composition of two measurable maps is measurable, so measurability implies weak-measurability, independently on the fact that E is separable. Let us assume now that $u\colon X \to E$ is weakly-measurable. To prove that u is measurable, it is sufficient to establish the point-wise convergence of a sequence of simple functions to u. If we take a look at the proof of Proposition 2.8, we see that it is sufficient to show that the function f in (2.58) is measurable. We can assume that $v = 0$ (otherwise, consider $\tilde{u}(x) = u(x) - v$). Our purpose therefore, is to show that $x \mapsto \|u(x)\|$ is measurable. We use the following result.

Lemma 2.10. *Let E be a separable Banach space. There exists a countable set $\{\varphi_n; n \in \mathbb{N}\} \subset E^*$ such that $\|u\| = \sup_{n\in\mathbb{N}} |\varphi_n(u)|$ for all $u \in E$.*

Using Lemma 2.10, we can describe the function $f\colon x \mapsto \|u(x)\|$ as the countable supremum of the measurable functions $x \mapsto |\varphi_n(u(x))|$, so f is measurable, which is the desired conclusion. □

Proof of Lemma 2.10. We admit the following result, that will be proved later (see Theorem 3.8), as a consequence of the Hahn-Banach theorem. For all $u \in E$,

$$\|u\| = \sup \{|\varphi(u)|; \varphi \in E^*, \|\varphi\| \leq 1\}. \tag{2.63}$$

Let $\{v_k; k \in \mathbb{N}\}$ be a dense subset of E. To each k and to each $m \in \mathbb{N}$, $m \geq 1$, we can associate a $\varphi_{k,m} \in E^*$ with $\|\varphi_{k,m}\| \leq 1$, such that

$$\|v_k\| \leq |\varphi_{k,m}(v_k)| + m^{-1}. \tag{2.64}$$

Let $u \in E$ and let $\varepsilon > 0$. We choose k, m such that $\|u - v_k\| < \varepsilon$, $m^{-1} < \varepsilon$. Then (2.64) gives

$$\|u\| \leq \|v_k\| + \|u - v_k\| < |\varphi_{k,m}(v_k)| + 2\varepsilon. \tag{2.65}$$

Since $\|\varphi_{k,m}\| \leq 1$, we also have

$$|\varphi_{k,m}(v_k)| \leq |\varphi_{k,m}(v_k - u)| + |\varphi_{k,m}(u)| \leq \|v_k - u\| + |\varphi_{k,m}(u)| < \varepsilon + |\varphi_{k,m}(u)|, \tag{2.66}$$

and thus $\|u\| < |\varphi_{k,m}(u)| + 3\varepsilon$, which shows that the set

$$\{\varphi_{k,m}; k \in \mathbb{N}, m \in \mathbb{N}, m \geq 1\} \tag{2.67}$$

is a countable subset of E^* which has the desired property. $\square$

We will define now the integral of simple functions. A simple function u is a function of the form

$$u = \sum_{i=1}^{m} u_i \mathbf{1}_{A_i}, \tag{2.68}$$

where $u_i \in E$, $A_i \in \mathcal{A}$. The decomposition (2.68) is said to be *canonical*, [SS05, Chapter 2.1], if the sets A_i are disjoint and the values u_i distinct and non-zero 0: for all $i, j \in \{1, \ldots, n\}$,

$$u_i \neq 0, \quad i \neq j \Rightarrow A_i \cap A_j = \emptyset, u_i \neq u_j. \tag{2.69}$$

The canonical form is unique. If $u_1^*, \ldots, u_n^*$ are the values taken by u, excepting 0 if 0 is in the range of u, then

$$u = \sum_{i=1}^{n} u_i^* \mathbf{1}_{A_i^*}, \quad A_i^* = \{x \in X; u(x) = u_i^*\}, \tag{2.70}$$

is the canonical decomposition of u.

Proposition-Definition 2.11 (Integral of simple functions)**.** *Let*

$$u = \sum_{i=1}^{m} u_i \mathbf{1}_{A_i} \tag{2.71}$$

be a simple measurable function $X \to E$. Assume that (2.71) *is the canonical decomposition of u. Then, the integral of u with respect to μ is the element of E defined as the combination*

$$\int_X u d\mu = \sum_{i=1}^{m} \mu(A_i) u_i. \tag{2.72}$$

It has the following properties:

(i) (2.72) *remains true if* (2.71) *is not canonical,*

(ii) we have the triangular inequality

$$\left\| \int_X u d\mu \right\| \leq \int_X \|u\| d\mu. \tag{2.73}$$

(iii) the following linearity relation is satisfied

$$\int_X u d\mu + \int_X v d\mu = \int_X (u + v) d\mu, \tag{2.74}$$

where u and v are simple functions.

Proof of Proposition-Definition 2.11. Note that the right-hand side of (2.72) is well defined since $\mu(A_i) < +\infty$ for all i. Let us prove *(i)*. We consider the decomposition (2.71), not necessary in canonical form. We will study how the quantity

$$S = \sum_{i=1}^{m} \mu(A_i) u_i \tag{2.75}$$

is affected (or rather *not affected*), when we reduce (2.71) to a canonical form. Assume first that there are no overlaps between the sets A_i: $A_i \cap A_j = \emptyset$ if $i \neq j$. The canonical form of u is then obtained by regrouping the sets A_i on which u takes the same value, and eliminating the value 0: with the notations in (2.70),

$$u = \sum_{j=1}^{n} u_j^* \mathbf{1}_{A_j^*}, \quad A_j^* = \bigcup_{i/u_i = u_j^*} A_i. \tag{2.76}$$

Then

$$\mu(A_j^*) = \sum_{i/u_i = u_j^*} \mu(A_i) \Rightarrow S = \int_X u d\mu. \tag{2.77}$$

Our task therefore is to eliminate the possible overlaps between the sets A_i in (2.71). Assume without loss of generality that $\cup_{i=1}^m A_i = X$. At a given point x, there are k of the sets A_i which overlap, $1 \leq k \leq m$. Keeping the notation k for the number of overlaps, we can write therefore

$$u = \sum_{k=1}^{m} \sum_{J_k \subset \{1,\ldots,m\}, \# J_k = k} u_{J_k} \mathbf{1}_{A_{J_k}}, \quad A_{J_k} := \left(\bigcap_{i \in J_k} A_i \right) \setminus \left(\bigcup_{i \notin J_k} A_i \right), \quad u_{J_k} := \sum_{i \in J_k} u_i. \tag{2.78}$$

Now, we have $A_J \cap A_{J'} = \emptyset$ if $J \neq J'$, so there remains to show that S in (2.75) satisfies

$$S = \sum_{k=1}^{m} \sum_{J \in \mathcal{J}_k} \mu(A_J) u_J, \quad \mathcal{J}_k := \{J \subset \{1,\ldots,m\}, \# J = k\}. \tag{2.79}$$

We obtain (2.79) by fixing an index $i \in \{1,\ldots,m\}$ and repeating the discussion on the number of overlaps of the sets A_j, but on A_i instead of X: we write the partition

$$A_i = \bigcup_{k=1}^{m} \bigcup_{J \in \mathcal{J}_{k,i}} A_i \cap A_J, \quad \mathcal{J}_{k,i} := \{J \in \mathcal{J}, i \in J\}. \tag{2.80}$$

By additivity of the measures (and since $A_i \cap A_J = A_J$ if $i \in J$), this gives

$$S = \sum_{i=1}^{m} \sum_{k=1}^{m} \sum_{J \in \mathcal{J}_{k,i}} \mu(A_J) u_i = \sum_{i=1}^{m} \sum_{k=1}^{m} \sum_{J \in \mathcal{J}_k} \mu(A_J) u_i \mathbf{1}_{i \in J}. \tag{2.81}$$

There remains to exchange the sums over i, k, J: since

$$u_J = \sum_{1 \leq i \leq m, i \in J} u_i = \sum_{i=1}^{m} u_i \mathbf{1}_{i \in J}, \tag{2.82}$$

we obtain (2.79). Proving *(i)* was necessary to obtain the linear relation (2.74), which is now straightforward. The triangular inequality (2.73) follows from the triangular inequality for finite sums in E. □

Proposition-Definition 2.12 (Bochner integrable function)**.** *let $(X, \mathcal{A}, \mu)$ be a finite measure space. Let E be a separable Banach space. A measurable function $u\colon X \to E$ is said to be* Bochner integrable *(or, more simply,* integrable*) if*

$$\int_X \|u\| d\mu < +\infty. \tag{2.83}$$

By Remark 2.3, there is a sequence of simple functions u_n such that $\|u_n(x) - u(x)\| \to 0$ for all $x \in X$ and $\|u_n - u\| \leq 3\|u\|$ everywhere in X. The integral of u against μ is defined as

$$\int_X u d\mu = \lim_{n\to+\infty} \int_X u_n d\mu. \tag{2.84}$$

This limit exists and is independent on the choice of the approximating sequence (u_n) of simple functions. The commutation relation

$$\varphi\left(\int_X u d\mu\right) = \int_X \varphi(u) d\mu \tag{2.85}$$

is satisfied for all continuous linear form $\varphi \in E^*$.

Proof of Proposition-Definition 2.12. By the linear relation (2.74) and the triangular inequality (2.73), we have

$$\|S_n - S_m\| \leq \int_X \|u_n - u_m\| d\mu \leq \int_X \|u - u_m\| d\mu + \int_X \|u - u_m\| d\mu, \tag{2.86}$$

where

$$S_n := \int_X u_n d\mu \tag{2.87}$$

By the Lebesgue dominated convergence theorem, we have

$$\lim_{m\to+\infty} \int_X \|u - u_m\| d\mu = 0, \tag{2.88}$$

and (2.86) shows that (S_n) is Cauchy, so the limit (2.84) exists. Replacing u_m in (2.86) by u'_n, where (u'_n) is an other sequence of simple functions such that $\|u - u'_n\| \to 0$ with $\|u - u'_n\| \leq C\|u\|$, shows that $\lim S_n$ is independent on the choice of the approximating sequence. To obtain (2.85), it is sufficient to check the identity when u is a simple function (we use the fact that φ is continuous), but this is a direct consequence of the linearity of φ then. □

Proposition 2.13 (Action of bounded linear operator)**.** *Let $(X, \mathcal{A}, \mu)$ be a finite measure space. Let E, F be some separable Banach spaces, let $u\colon X \to E$ be integrable and $T\colon E \to F$ be a bounded linear operator. Then $T \circ u\colon X \to F$ is integrable and*

$$T\left(\int_X u d\mu\right) = \int_X T \circ u d\mu. \tag{2.89}$$

Proof of Proposition 2.13. Since T is continuous, $T \circ u$ is measurable. We have also

$$\int_X \|T(u)\|_F d\mu \leq \|T\|_{\mathcal{L}(E,F)} \int_X \|u\|_E d\mu < +\infty, \tag{2.90}$$

so $T \circ u$ is integrable. Let (u_n) be a sequence of simple functions such that $\|u - u_n\| \to 0$ and $\|u - u_n\| \leq 3\|u\|$. Then $T \circ u_n$ is a simple function. By linearity of T and (2.72) in Proposition-Definition 2.12 we have the commutation relation

$$T\left(\int_X u_n d\mu\right) = \int_X T \circ u_n d\mu, \tag{2.91}$$

for all n. By continuity of T, the left-hand side of (2.91) is converging to

$$T\left(\int_X u d\mu\right). \tag{2.92}$$

By continuity of T and the triangular inequality, we also have

$$\left\|\int_X T \circ u d\mu - \int_X T \circ u_n d\mu\right\|_F d\mu \leq \int_X \|T\circ u - T\circ u_n\|_F d\mu \leq \|T\|_{\mathcal{L}(E,F)} \int_X \|u - u_n\|_E d\mu, \tag{2.93}$$

so

$$\lim_{n\to+\infty} \int_X T \circ u_n d\mu = \int_X T \circ u d\mu, \tag{2.94}$$

and (2.89) is obtained in the limit $n \to +\infty$ in (2.91). □

2.4 Strong compactness in L^p

2.4.1 Criteria for strong compactness

We give the following compactness result in $L^p(\mathbb{R}^d)$. For $z \in \mathbb{R}^d$, we denote by τ_z the action of translation $\tau_h u(x) = u(x - z)$. Denote by

$$\omega_{L^p}(u;\eta) = \sup_{|z|<\eta} \|\tau_z u - u\|_{L^p(\mathbb{R}^d)}. \tag{2.95}$$

the modulus of continuity (with respect to translations) of a function $u \in L^p(\mathbb{R}^d)$.

Proposition 2.14 (Modulus of continuity in L^p). *We have*

$$\lim_{\eta\to 0} \omega_{L^p}(u;\eta) = 0, \tag{2.96}$$

for any given $u \in L^p(\mathbb{R}^d)$.

Proof of Proposition 2.14. We establish first the following inequality: if $u, v \in L^p(\mathbb{R}^d)$, then

$$\omega_{L^p}(u;\eta) \leq \omega_{L^p}(v;\eta)| + 2\|u - v\|_{L^p(\mathbb{R}^d)}. \tag{2.97}$$

Let $z \in \mathbb{R}^d$ with $|z| < \eta$. By the triangular inequality in L^p and the identity $\|\tau_z w\|_{L^p(\mathbb{R}^d)} = \|w\|_{L^p(\mathbb{R}^d)}$, we have

$$\|\tau_z u - u\|_{L^p(\mathbb{R}^d)} \leq \|\tau_z v - v\|_{L^p(\mathbb{R}^d)} + 2\|u - v\|_{L^p(\mathbb{R}^d)} \leq \omega_{L^p}(v;\eta)| + 2\|u - v\|_{L^p(\mathbb{R}^d)}. \tag{2.98}$$

Taking the sup over $|z| < \eta$ gives (2.97). Next, let $\varepsilon > 0$ and let v be a continuous compactly supported function such that $\|u - v\|_{L^p(\mathbb{R}^d)} < \varepsilon$. Let K be a compact set such that v is supported in K, and let $\omega_\infty(v;\eta)$ denote the classical modulus of continuity

$$\omega_\infty(v;\eta) = \sup_{x\in K, |z|<\eta} |v(x + z) - v(x)|. \tag{2.99}$$

If $|z| < \eta \leq 1$, we have

$$\left[\int_{\mathbb{R}^d} |v(x+z) - v(x)|^p dx\right]^{1/p} \leq |K_1|^{1/p} \omega_\infty(v;\eta), \tag{2.100}$$

where $K_1 = K + \bar{B}(0,1)$, which, using (2.97), gives

$$\omega_{L^p}(u;\eta) \leq |K_1|^{1/p} \omega_\infty(v;\eta) + 2\varepsilon. \tag{2.101}$$

A continuous function on a compact set is uniformly continuous, so we can take η small enough to ensure that $|K_1|^{1/p}\omega_\infty(v;\eta) < \varepsilon$, and thus $\omega_{L^p}(u;\eta) < 3\varepsilon$. ☐

Theorem 2.15 (Kolmogorov-M.Riesz-Fréchet Theorem)**.** *Let $1 \leq p < +\infty$, let $H \subset L^p(\mathbb{R}^d)$ satisfy*

(i) H is bounded: *there exists $M \geq 0$ such that $\|u\|_{L^p(\mathbb{R}^d)} \leq M$ for all $u \in H$,*

(ii) the convergence (2.96) is uniform on H: *for all $\varepsilon > 0$, there exists $\eta > 0$ such that $|z| < \eta$ implies $\|\tau_z u - u\|_{L^p(\mathbb{R}^d)} < \varepsilon$ for all $u \in H$.*

Let $\chi \in C_c^\infty(\mathbb{R}^d)$ satisfy

$$0 \leq \chi \leq 1, \quad \chi \equiv 1 \textit{ on } B(0,1), \quad \mathrm{supp}(\chi) \subset B(0,2), \tag{2.102}$$

and, for $R > 0$, define the rescaled truncate function $\chi_R(x) = \chi(R^{-1}x)$. Then, for all $R > 0$,

$$H_R = \{u\chi_R; u \in H\} \tag{2.103}$$

is relatively compact in $L^p(\mathbb{R}^d)$.

To prove Theorem 2.15, we will use the following result.

Theorem 2.16 (Compact sets in Banach spaces)**.** *A set A in a Banach space E is relatively compact if, and only if, for all given radius $r > 0$, it can be covered by a finite number of balls of radius r. In particular, if for every $\varepsilon > 0$, every $u \in A$ can be decomposed as $u = v + w$, where $v \in K_\varepsilon$, $\|w\| < \varepsilon$, and K_ε is relatively compact, then A is relatively compact.*

Proof of Theorem 2.16. Let us begin with the proof of the corollary, which may be summarized as

"small+relatively compact $\Rightarrow$ relatively compact"

in a Banach space. Let $r > 0$. Take $\varepsilon = r/2$. Cover K_ε by a finite number of ball $B(x_i, r/2)$, $i \in I$, I finite. Then A is covered by the balls $B(x_i, r)$, $i \in I$. For the first equivalence statement, look at [Fol99, p. 15]. ☐

Proof of Theorem 2.15. It is clear that H_R satisfies *(i)*. Let us show that it also satisfies *(ii)*. If $z \in \mathbb{R}^d$ with $|z| < \eta$, then

$$\|\tau_z(u\chi_R) - u\chi_R\|_{L^p(\mathbb{R}^d)} \leq \omega_\eta(u) + \|u(\tau_z\chi_R - \chi_R)\|_{L^p(\mathbb{R}^d)} \leq \omega_\eta(u) + \|u\|_{L^p(\mathbb{R}^d)}\|\tau_z\chi_R - \chi_R\|_{L^\infty(\mathbb{R}^d)}. \tag{2.104}$$

This shows that

$$\omega_{L^p}(u\chi_R;\eta) \leq \omega_{L^p}(u;\eta) + M\omega_{L^\infty}(\chi_R;\eta), \tag{2.105}$$

where M is the bound in *(i)*. We have $\nabla\chi_R(x) = R^{-1}(\nabla\chi)(R^{-1}x)$, so $\|\nabla\chi_R\|_{L^\infty(\mathbb{R}^d)} \leq R^{-1}\|\nabla\chi\|_{L^\infty(\mathbb{R}^d)} =: C(R,\chi)$ and

$$\omega_{L^\infty}(\chi_R;\eta) \leq C(R,\chi)\eta. \tag{2.106}$$

We obtain then

$$\omega_{L^p}(u\chi_R;\eta) \leq \omega_{L^p}(u;\eta) + MC(R,\chi)\eta, \tag{2.107}$$

so H_R satisfies *(ii)*. Let (ρ_η) be an approximation of the unit,

$$\rho_\eta(x) = \eta^{-d}\rho_1(\eta^{-1}x), \quad \rho_1 \in C_c^\infty(\mathbb{R}^d), \quad \operatorname{supp}(\rho_1) \subset \bar{B}(0,1). \tag{2.108}$$

Each $u \in H_R$ can be decomposed as $u = v + w$, where

$$v = \rho_\eta * u, \quad w = u - \rho_\eta * u, \tag{2.109}$$

We have

$$w(x) = \int_{\bar{B}(0,\eta)} \rho_\eta(y)(u - \tau_y u)(x)dy. \tag{2.110}$$

The triangular inequality gives

$$\|w\|_{L^p(\mathbb{R}^d)} \leq \omega_\eta(u)\int_{\bar{B}(0,\eta)} \rho_\eta(y)dy = \omega_\eta(u). \tag{2.111}$$

Let $\varepsilon > 0$, choose $\eta > 0$ such that $\omega_\eta(u) < \varepsilon$ for all $u \in H$. Then $\|w\|_{L^p(\mathbb{R}^d)} < \varepsilon$. We will prove now that the set

$$B_\eta := \{\rho_\eta * u; u \in H_R\} \tag{2.112}$$

is relatively compact in $L^p(\mathbb{R}^d)$. Each function $v \in B_\eta$ is supported in $K := \bar{B}(0, 2R+1)$. By compact injection of $C(K)$ in $L^p(K)$ when K is compact in $\mathbb{R}^d$, it will be sufficient to show that B_η is relatively compact in $C(K)$. To obtain this result, we will use Ascoli's Theorem. By *(i)*, we have first the uniform bound

$$\|\rho_\eta * u\|_{L^\infty(K)} = \|\rho_\eta * u\|_{L^\infty(\mathbb{R}^d)} \leq \|\rho_\eta\|_{L^{p'}(\mathbb{R}^d)}\|u\|_{L^p(\mathbb{R}^d)} \leq \|\rho_\eta\|_{L^{p'}(\mathbb{R}^d)}M, \tag{2.113}$$

for all $u \in H_R$. Since

$$|\rho_\eta * u(x) - \rho_\eta * u(y)| \leq \|\nabla(\rho_\eta * u)\|_{L^\infty(\mathbb{R}^d)}|x-y|, \tag{2.114}$$

the equi-continuity of B_η follows from the bound

$$\|\nabla(\rho_\eta * u)\|_{L^\infty(\mathbb{R}^d)} = \|\nabla(\rho_\eta) * u\|_{L^\infty(\mathbb{R}^d)} \leq \|\nabla\rho_\eta\|_{L^{p'}(\mathbb{R}^d)}\|u\|_{L^p(\mathbb{R}^d)} \leq \|\nabla\rho_\eta\|_{L^{p'}(\mathbb{R}^d)}M, \tag{2.115}$$

for all $u \in H_R$. We conclude by Theorem 2.16. □

Remark 2.5 (Relative compactness of H). If, to *(i)* and *(ii)*, we add the condition

(iii) *the elements in H are uniformly small at infinity*: for all $\varepsilon > 0$, there exists $R > 0$ such that

$$\int_{|x|>R} |u(x)|^p dx < \varepsilon, \tag{2.116}$$

for all $u \in H$, then we can conclude that H itself is relatively compact in $L^p(\mathbb{R}^d)$.

Proof of Remark 2.5. We apply Theorem 2.16 again. Let $\varepsilon > 0$ and let $R > 0$ be such that

$$\int_{|x|>R} |u(x)|^p dx < \varepsilon^p, \tag{2.117}$$

for all $u \in H$. Each u in H can be decomposed as $u = v + w$, where $v = u\chi_R \in H_R$, $w = u(1 - \chi_R)$. We have shown that H_R is relatively compact in $L^p(\mathbb{R}^d)$ so it is sufficient to prove that $\|w\|_{L^p(\mathbb{R}^d)} < \varepsilon$. This is a direct consequence of (2.117). □

The control on the modulus of continuity for translations (2.95) is often given by a control of the derivatives of the functions u for $u \in H$.

Proposition 2.17. *We have*

$$\omega_{L^p}(u; \eta) \leq \|\nabla u\|_{L^p(\mathbb{R}^d)}\eta, \tag{2.118}$$

for all functions $u \in C^1(\mathbb{R}^d)$ such that $u \in L^p(\mathbb{R}^d)$ and $|\nabla u| \in L^p(\mathbb{R}^d)$.

Proof of Proposition 2.17. Let $x \in \mathbb{R}^d$, $z \in \mathbb{R}^d$ with $|z| < \eta$. We apply the identity

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(t)dt, \quad \varphi \in C^1([0, 1]), \tag{2.119}$$

to $\varphi(t) = u(tx + (1 - t)(x - z)) = u(x - (1 - t)z)$ to obtain

$$|u(x) - \tau_z u(x)| = \left|\int_0^1 \nabla u(x - (1 - t)z) \cdot z\right| \leq \eta \int_0^1 |\nabla u(x - (1 - t)z)|dt. \tag{2.120}$$

By Jensen's inequality, we obtain

$$|u(x) - \tau_z u(x)|^p \leq \eta^p \int_0^1 |\nabla u(x - (1 - t)z)|^p dt. \tag{2.121}$$

Since

$$\int_{\mathbb{R}^d} |\nabla u(x - (1 - t)z)|^p dx = \int_{\mathbb{R}^d} |\nabla u(x)|^p dx = \|\nabla u\|^p_{L^p(\mathbb{R}^d)}, \tag{2.122}$$

integration with respect to x in (2.121) and Fubini's theorem gives (2.118). □

We give now a second criterion for compactness in L^p.

Theorem 2.18 (Vitali's theorem)**.** *Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space. Let $1 \leq p < +\infty$ and let (u_n) be a sequence of $L^p(X)$ satisfying*

(i) a.e. convergence: *$u_n \to u$ a.e. in X, where $u\colon X \to \mathbb{R}$ is measurable,*

(ii) equi-integrability: *for all $\varepsilon > 0$, there exists $M > 0$ such that*

$$\int_{\{|u_n|>M\}} |u_n|^p d\mu < \varepsilon, \tag{2.123}$$

for all n,

(iii) uniform smallness out of a set of finite measure: *for all $\varepsilon > 0$, there exists a measurable set Γ of finite measure such that*

$$\int_{X\setminus\Gamma} |u_n|^p d\mu < \varepsilon, \tag{2.124}$$

for all n.

Then $u \in L^p(X)$ and $u_n \to u$ in $L^p(X)$.

Before proving Theorem 2.18, let us give an alternative definition of the equi-integrability property.

Proposition 2.19 (Equi-integrability)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $1 \leq p < +\infty$. Let (u_n) be a sequence of $L^p(X)$. If (u_n) satisfies (2.123)-(2.124) of Theorem 2.18 then*

$$\sup_n \|u_n\|_{L^p(X)} < +\infty, \tag{2.125}$$

and (u_n) satisfies also (2.126)-(2.124), where (2.126) is given by

(iv) for all $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\int_A |u_n|^p d\mu < \varepsilon, \tag{2.126}$$

for all n and for all measurable set A with $\mu(A) < \delta$.

Conversely, (2.125) and (2.126) imply (ii).

Corollary 2.20 (Equi-integrability of a single function)**.** *Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space and let $u \in L^p(X)$. Then, for all $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\int_A |u|^p d\mu < \varepsilon, \tag{2.127}$$

for all measurable set A with $\mu(A) < \delta$.

Proof of Corollary 2.20. Let us show that $\{u\}$ satisfies (2.123)-(2.124), then we can apply Proposition 2.19 to conclude. Recall that $|u| < +\infty$ a.e. since, using Markov's inequality,

$$\mu(\{|u| > M\}) \leq \frac{1}{M^p} \int_X |u|^p d\mu \to 0 \text{ when } M \to +\infty.$$

By the dominated convergence theorem, we have therefore

$$\int_{\{|u|>M\}} |u|^p d\mu = \int_X \mathbf{1}_{\{|u|>M\}} |u|^p d\mu \to 0 \text{ when } M \to +\infty. \tag{2.128}$$

This gives (2.123). To obtain (2.124), we can apply the dominated convergence theorem again, to $\mathbf{1}_{\Gamma_n} |u|^p$, where (Γ_n) is an increasing sequence of sets of finite measure such that $X = \cup_{n \in \mathbb{N}} \Gamma_n$. $\square$

Proof of Proposition 2.19. If (u_n) satisfies (2.123)-(2.124), then

$$\int_A |u_n|^p d\mu = \int_{A \cap \{|u_n|>M\}} |u_n|^p d\mu + \int_{A \cap \{|u_n| \leq M\}} |u_n|^p d\mu \tag{2.129}$$

gives, for a fixed $\varepsilon > 0$ and for M large enough,

$$\int_A |u_n|^p d\mu \leq \frac{\varepsilon}{2} + M^p \mu(A), \tag{2.130}$$

and we obtain (2.126) by choosing δ sufficiently small. Next, let us apply (2.123)-(2.124) with $\varepsilon = 1$. There exists $M > 0$ and Γ of finite measure such that

$$\int_{\{|u_n|>M\}} |u_n|^p d\mu \leq 1, \quad \int_{X \setminus \Gamma} |u_n|^p d\mu \leq 1, \forall n. \tag{2.131}$$

This gives

$$\int_X |u_n|^p d\mu = \int_{\{|u_n|>M\}} |u_n|^p d\mu + \int_{\{|u_n|\leq M\}\cap\Gamma^c} |u_n|^p d\mu + \int_{\{|u_n|\leq M\}\cap\Gamma} |u_n|^p d\mu \leq 2 + M^p\mu(\Gamma), \tag{2.132}$$

and (u_n) is indeed bounded in $L^p(X)$, as asserted in (2.125). Conversely, assume (2.126)-(2.124) and (2.125). Let C be a bound for the left-hand side in (2.125). By the Markov inequality, we have

$$|\{|u_n| > M\}| = |\{|u_n|^p > M^p\}| \leq \frac{1}{M^p}\int_X |u_n|^p d\mu \leq \frac{C^p}{M^p}. \tag{2.133}$$

This shows that, given $\delta > 0$, the measure of $|\{|u_n| > M\}|$ is smaller than δ for M large enough independent on n, and then (2.123) follows from (2.126) applied to $|\{|u_n| > M\}|$. □

Proof of Theorem 2.18. Using Fatou's Lemma in (2.125), we obtain

$$\int_A |u|^p d\mu < +\infty, \tag{2.134}$$

so $u \in L^p(X)$. We will now show that we can reduce the problem to the case where $u = 0$. This amounts to show that the sequence of general term $w_n := u_n - u$ has the same properties as the original sequence (u_n). Using the characterization of equi-integrability given in Proposition 2.19, the bound

$$\int_A |w_n|^p d\mu \leq 2^p \int_A |u_n|^p d\mu + 2^p \int_A |u|^p d\mu \tag{2.135}$$

and Corollary 2.20, we obtain the equi-integrability property for (w_n). Similar arguments also give the uniform smallness condition for (w_n). Let us now prove the theorem, assuming $u = 0$. Let $\varepsilon > 0$, let $M > 0$ and let Γ be a set of finite measure such that, respectively (2.123) and (2.124) are satisfied. We have

$$\int_X |u_n|^p d\mu < 2\varepsilon + \int_X v_n d\mu, \tag{2.136}$$

where $v_n = \mathbf{1}_\Gamma \mathbf{1}_{\{|u_n|\leq M\}}|u_n|^p$ is dominated by the integrable function $\mathbf{1}_\Gamma M^p$. By the dominated convergence theorem, we obtain

$$\int_X |u_n|^p d\mu < 3\varepsilon, \tag{2.137}$$

for n large enough, and conclude that $u_n \to 0$ in $L^p(X)$. □

2.4.2 Obstructions to strong convergence

We present three obstructions to strong convergence: loss of mass, concentration, oscillation. Each phenomenon is independent on the others in the examples we provide, but of course they may combine in some situation. The loss of mass may be considered as minor, compared to the phenomena of concentration and oscillation, but it is central in the more general question of convergence of probability measures on Polish spaces ("tightness criterion").

Loss of mass Let ψ be a positive continuous compactly supported function on $\mathbb{R}^d$ and let $\psi_n(x) = \psi(x - n\nu)$, where $\nu \in \mathbb{R}^d$, $|\nu| = 1$. The sequence of functions (ψ_n) is vanishing at infinity in the direction ν. Fix $R > 0$ such that the support of ψ is contained in $\bar{B}(0, R)$. Then, for all $x \in \mathbb{R}^d$, $\psi_n(x) = 0$ for all n larger than $|x| + R$. So ψ_n is converging to 0 a.e. It is not converging to 0 in $L^p(\mathbb{R}^d)$ however, since $\|\psi_n\|_{L^p(\mathbb{R}^d)} = \|\psi\|_{L^p(\mathbb{R}^d)}$ for all n (invariance by translation of the Lebesgue measure). If we examine Theorem 2.15, then $H := \{\psi_n; n \in \mathbb{N}\}$ satisfies the two criteria *(i)*-*(ii)* and so also the conclusion, which is only local however (and it is true that $\|psi_n\|_{L^p(K)} \to 0$ for any compact set K. Actually, $\psi_n \chi_R \equiv 0$ as soon as $n > 2R$, so H_R is finite. If we examine Theorem 2.18 now, where $X = \mathbb{R}^d$, $\mathcal{A}$ is the Borel σ-algebra and μ the Lebesgue measure, then the situation is the following one: the criteria *(i)*-*(ii)* are satisfied (why?), the default is in the criterion *(iii)*. Indeed, if Γ has finite measure, then we can apply the Lebesgue dominated convergence theorem to $\mathbf{1}_\Gamma |\psi_n|^p$, which is dominated by $\mathbf{1}_\Gamma \|\psi\|^p_{L^\infty(\mathbb{R}^d)}$. We have in particular

$$\int_{\mathbb{R}^d \setminus \Gamma} |\psi_n(x)|^p dx = \|\psi\|^p_{L^p(\mathbb{R}^d)} - \int_\Gamma |\psi_n(x)|^p dx \geq \frac{1}{2} \|\psi\|^p_{L^p(\mathbb{R}^d)}, \tag{2.138}$$

for n large enough, which is not compatible with (2.124).

Concentration Consider the approximation of the unit (ρ_n) defined on $\mathbb{R}^d$ by

$$\rho_n(x) = n^d \rho_1(nx), \tag{2.139}$$

where $\rho_1 \in C_c^\infty(\mathbb{R}^d)$ is a smooth, compactly supported function such that

$$\rho_1 \geq 0, \quad \int_{\mathbb{R}^d} \rho_1(x)dx = 1, \quad \mathrm{supp}(\rho_1) \subset B(0,1). \tag{2.140}$$

Then (ρ_n) satisfies the criteria *(i)* and *(iii)* in the Vitali's theorem 2.18. Indeed, for all $x \neq 0$, we have $\rho_n(x) = 0$ for $n > |x|^{-1}$, so $\rho_n \to 0$ a.e. We also have

$$\int_{|x|>1} |\rho_n(x)|dx = 0, \tag{2.141}$$

for all n. The criterion *(ii)* is not satisfied however, and (ρ_n) is not converging to 0 in $L^1(\mathbb{R}^d)$, Indeed,

$$\|\rho_n\|_{L^1(\mathbb{R}^d} = 1, \tag{2.142}$$

by homogeneity and (2.140). The criterion *(ii)* is not satisfied because there is concentration of (ρ_n) around 0. Indeed, we have

$$\int_{B(0,r)} |\rho_n(x)|dx = 1 \tag{2.143}$$

when $n > r^{-1}$, which is not compatible with the equi-integrability criterion (2.126).

Oscillations-1 Let $\psi \in C_c^\infty(\mathbb{R})$ be a positive function supported in $(0,1)$ such that $0 \leq \psi(x) \leq M$ for all $x \in \mathbb{R}$. We rescale the function ψ in the x-direction by a factor $1/n$ and consider n copies of this rescaled functions distributed regularly on $(0,1)$, so we define $\psi_n(x)$ by

$$\psi_n(x) = \sum_{k=0}^{n-1} \psi(nx-k), \quad x \in \mathbb{R}. \tag{2.144}$$

Then, for all $\varphi \in C([0,1])$, we have

$$\int_0^1 \psi_n(x)\varphi(x)dx \to \int_0^1 A_\psi \varphi(x)dx, \quad A_\psi := \int_0^1 \psi(x)dx. \tag{2.145}$$

Indeed, a change of variable gives

$$\int_0^1 \psi_n(x)\varphi(x)dx = \frac{1}{n}\sum_{k=0}^{n-1}\int_0^1 \psi(x)\varphi\left(\frac{x+k}{n}\right)dx. \tag{2.146}$$

Let

$$\omega_{L^\infty}(\varphi;\eta) = \sup\{|\varphi(x)-\varphi(y)|; x,y\in[0,1], |x-y|<\eta\} \tag{2.147}$$

denote the modulus of continuity of φ. Since

$$\left|\varphi\left(\frac{x+k}{n}\right)-\varphi\left(\frac{k}{n}\right)\right| \leq \omega(\varphi;1/n), \tag{2.148}$$

for all $x\in[0,1]$, (2.146) gives

$$\int_0^1 \psi_n(x)\varphi(x)dx = \frac{A_\psi}{n}\sum_{k=0}^{n-1}\varphi\left(\frac{k}{n}\right) + \mathcal{O}_\psi(\omega(\varphi;1/n)), \tag{2.149}$$

and (2.145) follows when $n\to+\infty$. Assume that (ψ_n) is converging strongly in $L^1(0,1)$ to a limit function ψ_∞. Then (2.145) implies that ψ is constant, equal to A_ψ. Straightforward calculations give

$$\|\psi_n - A_\psi\|_{L^1(0,1)} = \|\psi - A_\psi\|_{L^1(0,1)}, \tag{2.150}$$

hence (ψ_n) cannot converge strongly in $L^1(0,1)$.

Oscillations-2 : Fourier series, Fourier transform Let us use the following conventions for Fourier series and Fourier Transform: if $u\in L^1(\mathbb{T}^d)$ (where $\mathbb{T}^d$ is the d-dimension torus) and $v\in L^1(\mathbb{R}^d)$,

$$\hat{u}(n) = \int_{\mathbb{T}^d} u(x)e^{-2\pi i x\cdot n}dx, \quad \hat{v}(\xi) = \int_{\mathbb{R}^d} v(x)e^{-2\pi i x\cdot\xi}dx, \tag{2.151}$$

where $n\in\mathbb{Z}^d$, $\xi\in\mathbb{R}^d$ respectively. Informally, using the inverse Fourier formula, we have the decompositions

$$u(x) = \sum_{n\in\mathbb{Z}^d}\hat{u}(n)e^{2\pi i x\cdot n}, \quad v(x) = \int_{\mathbb{R}^d}\hat{v}(\xi)e^{2\pi i x\cdot\xi}d\xi, \tag{2.152}$$

which give u or v as a superposition of functions oscillating at the frequency n or ξ: one should have a good control on the oscillating behaviour of u, once the decay of the coefficients is known. This decay can be estimated in various manners. For instance we have the following statements or functional inequalities (where $A\lesssim B$ indicates that $A\leq C(d)B$ for a constant $C(d)$ which depends on d only, and $A\approx B$ means that both $A\lesssim B$ and $B\lesssim A$ are satisfied). Regularity C^k: for $k\in\mathbb{N}$,

$$\|u\|_{C^k(\mathbb{T}^d)} \lesssim (2\pi)^k\sum_{n\in\mathbb{Z}^d}(1+|n|^2)^{k/2}|\hat{u}(n)|, \quad \|v\|_{C^k(\mathbb{R}^d)} \lesssim (2\pi)^k\int_{\mathbb{R}^d}(1+|\xi|^2)^{k/2}|\hat{u}(\xi)|d\xi. \tag{2.153}$$

Regularity H^k: for $k\in\mathbb{N}$,

$$\|u\|^2_{H^k(\mathbb{T}^d)} \approx \sum_{n\in\mathbb{Z}^d}(1+|n|^2)^k|\hat{u}(n)|^2, \tag{2.154}$$

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

$$\|v\|^2_{H^k(\mathbb{R}^d)} \approx \int_{\mathbb{R}^d} (1+|\xi|^2)^k |\hat{u}(\xi)|^2 d\xi. \tag{2.155}$$

Hölder regularity: for $\alpha \in (0,1)$, and $Q := [0,1)^d$,

$$\|u\|_{C^{0,\alpha}(\mathbb{T}^d)} = \sup_{x\in Q)} |u(x)| + \sup_{x\neq y\in Q} \frac{|u(x)-u(y)|}{|x-y|^\alpha} \lesssim C(\alpha) \sum_{n\in\mathbb{Z}^d} (1+|n|^2)^{\alpha/2} |\hat{u}(n)|, \tag{2.156}$$

and

$$\|v\|_{C^{0,\alpha}(\mathbb{R}^d)} = \sup_{x\in\mathbb{R}^d)} |u(x)| + \sup_{x\neq y\in \mathbb{R}^d} \frac{|u(x)-u(y)|}{|x-y|^\alpha} \lesssim C(\alpha) \int_{\mathbb{R}^d} (1+|n|^2)^{\alpha/2} |\hat{v}(\xi)| d\xi. \tag{2.157}$$

Proof of (2.156) *and* (2.157). Use

$$|e^{2\pi i x\cdot\zeta} - 1| \leq C(\alpha)|\zeta|^\alpha |x|^\alpha, \quad C(\alpha) := \sup_{t\in\mathbb{R}^*} \frac{|e^{2\pi i t}-1|}{t^\alpha}. \tag{2.158}$$

□

Definition 2.3 (Schwartz space)**.** The Schwartz space $\mathscr{S}(\mathbb{R}^d)$ is the space of infinitely differentiable functions whose derivatives decay faster at infinity than any polynomial: $v \in \mathscr{S}(\mathbb{R}^d)$ if v is of class C^∞ and all the semi-norms

$$p_{m,k}(v) = \sup_{x\in\mathbb{R}^d, |\alpha|\leq k} (1+|x|^2)^{m/2} |\partial_x^\alpha v(x)| \tag{2.159}$$

are finite.

We have or will use the following notations: for $\alpha \in \mathbb{N}^d$, $x^\alpha = x_1^{\alpha_1}\cdots x_d^{\alpha_d}$, $\partial_x^\alpha = \partial_{x^1}^{\alpha_1}\cdots\partial_{x_d}^{\alpha_d}$. If $v \in \mathscr{S}(\mathbb{R}^d)$, $\alpha,\beta \in \mathbb{N}^d$, then $x \mapsto x^\beta \partial_x^\alpha v(x)$ is in $L^p(\mathbb{R}^d)$ for all $1 \leq p \leq \infty$, simply because

$$|x^\beta \partial_x^\alpha v(x)| \leq \frac{p_{m,k}(v)}{(1+|x|^2)^{(d+1)/2}} \in L^p(\mathbb{R}^d), \tag{2.160}$$

where $m = |\beta| + d + 1$ and $k = |\alpha|$. In particular,

$$\|\hat{v}\|_{C_0(\mathbb{R}^d)} \leq \|v\|_{L^1(\mathbb{R}^d)} \leq p_{d+1,0}(v). \tag{2.161}$$

Using(2.161) and the formulas

$$(\partial_x^\alpha v)\hat{}(\xi) = (2\pi i\xi)^\alpha \hat{v}(\xi), \quad [(-2\pi i X)^\alpha v]\hat{}(\xi) = \partial_\xi^\alpha \hat{v}(\xi), \tag{2.162}$$

where $X^\alpha v$ denotes the function $x \mapsto x^\alpha v(x)$, we can establish the following result.

Proposition 2.21 (Fourier and Schwartz space)**.** *If $v \in \mathscr{S}(\mathbb{R}^d)$, then $\hat{v} \in \mathscr{S}(\mathbb{R}^d)$ and we have $p_{m,k}(\hat{v}) \lesssim_{m,k} p_{k+d+1,m}(v)$.*

Proposition 2.22 (A form of the uncertainty principle of Heisenberg)**.** *We have*

$$\|(X-x_0)v\|_{L^2(\mathbb{R}^d)} \|(\nabla_x - 2\pi i\xi_0)v\|_{L^2(\mathbb{R}^d)} \geq \frac{d}{2}\|v\|^2_{L^2(\mathbb{R}^d)}, \tag{2.163}$$

for all $v \in \mathscr{S}(\mathbb{R}^d)$, for all $x_0, \xi_0 \in \mathbb{R}^d$.

Remark 2.6 (Localisation in x and ξ). since

$$\|(\nabla_x - 2\pi i\xi_0)v\|_{L^2(\mathbb{R}^d)} = 2\pi\|(\xi - \xi_0)\hat{v}\|_{L^2(\mathbb{R}^d)}, \tag{2.164}$$

by (2.162) and the Plancherel Theorem, (2.163) says, in certain quantitative way, that we cannot have both v localized around x_0 and $\hat{v}$ localized around ξ_0.

Proof of Proposition 2.22. Using

$$\|(X - x_0)v\|^2_{L^2(\mathbb{R}^d)} = \|X(\tau_{-x_0}v)\|^2_{L^2(\mathbb{R}^d)}, \tag{2.165}$$

and

$$\|(\nabla_x - 2\pi i\xi_0)v\|^2_{L^2(\mathbb{R}^d)} = \|\nabla_x(ve_{-\xi_0})\|^2_{L^2(\mathbb{R}^d)}, \quad e_\xi(x) := \exp(2\pi i x\cdot\xi), \tag{2.166}$$

and the fact that the transformations $v \mapsto \tau_{-x_0}v$ and $v \mapsto ve_{-\xi_0}$ affect only one of the three terms in (2.163), it is sufficient to consider the cases $x_0 = \xi_0 = 0$. Then we integrate the identity

$$\mathrm{div}_x(xv^2/2) = \mathrm{div}(x)v^2/2 + x\cdot\nabla_x(v^2/2) = dv^2/2 + xv\cdot\nabla_x v \tag{2.167}$$

to obtain, by the Cauchy-Schwarz inequality,

$$\frac{d}{2}\|v\|^2_{L^2(\mathbb{R}^d)} = -\int_{\mathbb{R}^d} xv\cdot\nabla_x v \le \|Xv\|_{L^2(\mathbb{R})}\|\nabla_x v\|_{L^2(\mathbb{R})}. \tag{2.168}$$

□

Proposition 2.23 (Nash Inequality). *We have*

$$\|u\|^{1+2/d}_{L^2(\mathbb{T}^d)} \lesssim \|\nabla u\|_{L^2(\mathbb{T}^d)}\|u\|^{2/d}_{L^1(\mathbb{T}^d)}, \quad \|v\|^{1+2/d}_{L^2(\mathbb{R}^d)} \lesssim \|\nabla v\|_{L^2(\mathbb{R}^d)}\|v\|^{2/d}_{L^1(\mathbb{R}^d)}. \tag{2.169}$$

for all $u \in H^1(\mathbb{T}^d)$, $v \in H^1(\mathbb{R}^d)$.

Proof of Proposition 2.23. We do the proof in the case of the Torus only. First we use Parseval's identity:

$$\|u\|^2_{L^2(\mathbb{T}^d)} = \sum_{n\in\mathbb{Z}^d} |\hat{u}(n)|^2, \tag{2.170}$$

and then make the distinction between the sum over $|n| < R$ and $|n| \ge R$, where R will be chosen later. Use the estimate $|\hat{u}(n)| \le \|u\|_{L^1(\mathbb{T}^d)}$ in the range $|n| < R$ and the estimate $|\hat{u}(n)|^2 \le R^{-2}|n|^2|\hat{u}(n)|^2$ in the range $|n| \ge R$ to obtain the bound

$$\|u\|^2_{L^2(\mathbb{T}^d)} \lesssim R^d\|u\|^2_{L^1(\mathbb{T}^d)} + R^{-2}\|\nabla u\|^2_{L^2(\mathbb{T}^d)}. \tag{2.171}$$

By optimization over R, (2.171) gives (2.169). □

Theorem 2.24 (Averaging lemma). *Let $(\Omega, \mathbb{P})$ be a probability space, let $a\colon \Omega \to \mathbb{R}^d$ be a random variable such that $|a| \le M$ a.s. and let $u\colon \Omega\times\mathbb{R}^d \to \mathbb{R}$ be a measurable function such that, a.s., $x \mapsto u(x) \in \mathscr{S}(\mathbb{R}^d)$ and $\mathbb{E}\left[p_{m,k}(u)\right] < +\infty$ for all $m, k \in \mathbb{N}$. We assume that a satisfies the following condition: there exists a constant $C(a) \ge 0$ and $\alpha \in (0, 1]$ such that*

$$\mathbb{P}(|a\cdot\nu| < \varepsilon) \le C(a)\varepsilon^\alpha \tag{2.172}$$

for all directions $\nu \in \mathbb{R}^d$, $|\nu| = 1$. Then the average $\rho := \mathbb{E}\left[u\right]$ satisfies the estimate

$$\|\rho\|^2_{H^{\alpha/2}(\mathbb{R}^d)} := \int_{\mathbb{R}^d}(1 + |\xi|^2)^{\alpha/2}|\hat{\rho}(\xi)|^2 d\xi \le C\mathbb{E}\left[\|u\|^2_{L^2(\mathbb{R}^d)} + \|a\cdot\nabla_x u\|^2_{L^2(\mathbb{R}^d)}\right], \tag{2.173}$$

where the constant C depends on d, $C(a)$ and α only.

Remark 2.7 (Non-degeneracy condition). The non-degeneracy condition (2.172) is a quantitative version of the property $\mathbb{P}(a\cdot\nu=0)=0$ for all ν in the unit sphere, which means that the term $a\cdot\nabla_x u$ will give a sufficient amount of information on directional derivatives to give a definite gain of regularity for the average ρ.

Proof of Theorem 2.24. Set $w=a\cdot\nabla_x u$. We have: a.s., for all $\xi\in\mathbb{R}^d$, $\hat{w}(\xi)=-2\pi i a\cdot\xi\hat{u}(\xi)$, which gives: a.s., for all $\xi\in\mathbb{R}^d$,

$$-2\pi i a\cdot\xi\hat{u}(\xi)+2\pi\hat{u}(\xi)=\hat{w}(\xi)+2\pi\hat{u}(\xi)\Rightarrow\hat{u}(\xi)=\frac{1}{2\pi}\frac{1}{1-ia\cdot\xi}(\hat{w}(\xi)+2\pi\hat{u}(\xi)). \tag{2.174}$$

By the Fubini Theorem, we deduce that, for all $\xi\in\mathbb{R}^d$,

$$\hat{\rho}(\xi)=\frac{1}{2\pi}\mathbb{E}\left[\frac{1}{1-ia\cdot\xi}(\hat{w}(\xi)+2\pi\hat{u}(\xi))\right]. \tag{2.175}$$

By the Cauchy-Schwarz inequality, it follows that

$$|\hat{\rho}(\xi)|^2\leq\frac{1}{4\pi^2}\mathbb{E}\left[\frac{1}{1+|a\cdot\xi|^2}\right]\mathbb{E}\left[|\hat{w}(\xi)+2\pi\hat{u}(\xi)|^2\right]. \tag{2.176}$$

We will show that

$$\mathbb{E}\left[\frac{1}{1+|a\cdot\xi|^2}\right]\leq\frac{C_1}{|\xi|^\alpha},\quad\forall|\xi|\geq 1. \tag{2.177}$$

Since

$$\mathbb{E}\left[|\hat{w}(\xi)+2\pi\hat{u}(\xi)|^2\right]\leq 2\mathbb{E}\left[|\hat{w}(\xi)|^2+4\pi^2|\hat{u}(\xi)|^2\right], \tag{2.178}$$

the estimate (2.177) and the theorem of Plancherel will give the desired conclusion (2.173). To prove (2.177), we introduce the non-increasing function $\theta(t)=(1+t^2)^{-1}$. We have

$$\theta(t)=-\int_t^\infty\theta'(s)ds=\int_0^\infty\mathbf{1}_{\{s>t\}}|\theta'(s)|ds. \tag{2.179}$$

By Fubini's theorem

$$\mathbb{E}\left[\frac{1}{1+|a\cdot\xi|^2}\right]=\int_0^\infty\mathbb{E}\left[\mathbf{1}_{\{s>|a\cdot\xi|\}}\right]|\theta'(s)|ds=\int_0^\infty\mathbb{P}\left(s>|a\cdot\xi|\right)|\theta'(s)|ds. \tag{2.180}$$

Write $\xi=|\xi|\nu$, where $|\nu|=1$. We use the non-degeneracy hypothesis (2.172) to get

$$\mathbb{E}\left[\frac{1}{1+|a\cdot\xi|^2}\right]\leq C(a)\int_0^\infty s^\alpha|\xi|^{-\alpha}|\theta'(s)|ds=C(a)\left[\int_0^\infty s^\alpha|\theta'(s)|ds\right]|\xi|^{-\alpha}, \tag{2.181}$$

which is the desired result. □

Fourier and Laplace operator: on the Torus, the Fourier orthonormal basis of $L^2(\mathbb{T}^d)$ is a spectral basis of the Laplace operator: setting $e_n(x)=\exp(2\pi i x\cdot n)$, we have

$$-\Delta e_n=\lambda_n e_n,\quad\lambda_n:=4\pi^2|n|^2. \tag{2.182}$$

The Bessel-Parseval identity then shows that, for $k\in\mathbb{N}$, and up to the multiplicative factor $4\pi^2$, the right-hand side of (2.154) is

$$\langle(\mathrm{Id}-\Delta)^k u,u\rangle_{L^2(\mathbb{T}^d)}=\|A^k u\|^2_{L^2(\mathbb{T}^d)},\quad A:=(\mathrm{Id}-\Delta)^{1/2}. \tag{2.183}$$

In the end of the course, we will study the spectral theorem for compact self-adjoint operators, and see how it can be applied to exhibit a spectral basis (e_n,λ_n) of the Laplace operator with homogeneous Dirichlet conditions on a bounded open subset U of $\mathbb{R}^d$.

Conclusion The obstruction to compactness illustrated above are frequently encountered in the study of PDEs (the sequence of function can be a sequence of solutions, when stability of solutions is studied, or a sequence of approximate solution, built with the purpose to study existence of a solution). For instance, in homogenization theory, where one tries to substitute an homogeneous object to an object having a lot of oscillations at small scales, understanding and controlling oscillations of functions is fundamental. It is often desirable to go beyond the mere observation that a sequence is not compact and to exhibit a limiting object to pursue the study (in particular to use the algebra that goes with the PDE under study). For the sequence (ρ_n) in (2.139), we can say that

$$\rho_n \to \delta_0, \tag{2.184}$$

where δ_0 is the "Dirac mass" at 0. The convergence (2.184) can be understood in the sense of measures or in the sense of distributions (in any case, this is a kind of weak convergence). For the sequence (ψ_n) in (2.144), we can prove that, for all $\Phi \in C([0,M])$, for all $\varphi \in C([0,1])$ we have

$$\lim_{n\to+\infty} \int_0^1 \varphi(x)\Phi(\psi_n(x))dx = \int_0^1 \varphi(x)dx \int_0^1 \Phi(\psi(y))dy, \tag{2.185}$$

which again gives us a measure as a limit object, which is the product measure $\lambda \otimes \psi_\sharp\lambda$ on $[0,1]\times[0,M]$, where λ denotes the Lebesgue measure on $[0,1]$ and $\psi_\sharp\lambda$ is the push-forward of λ by ψ. There are many different ways to measure of the regularity/integrability of functions. We will study in particular the Sobolev spaces, which are normed vector spaces. Weak convergence in normed space will be the object of the next chapter, once we have proved the Hahn-Banach theorem and its important corollaries. Some other instances of noticeable functional spaces in analysis are Fréchet Spaces and we will briefly consider them before beginning the theory of distributions.

2.5 Differentiation, maximal functions

2.5.1 Maximal functions

We denote by $|A|$ the Lebesgue measure of a Borel set $A \subset \mathbb{R}^d$. To $u \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, we associate the maximal function Mu defined by

$$Mu(x) = \sup_{B\ni x} \frac{1}{|B|}\int_B |u(y)|dy, \tag{2.186}$$

where the supremum is taken over all balls B such that $x \in B$. Considering open or closed balls makes no difference (why?) so we will assume that all balls are open. If $B = B(z,r)$ is such a ball and $\theta > 0$, we denote by θB the ball $B(x,\theta r)$. We will establish the following result.

Theorem 2.25 (Estimating the size of Maximal functions)**.** *The function Mu defined in (2.186) is measurable. We have*

$$|\{Mu > \alpha\}| \leq \frac{3^d}{\alpha}\|u\|_{L^1(\mathbb{R}^d)}, \tag{2.187}$$

for all $\alpha > 0$ if $u \in L^1(\mathbb{R}^d)$ and

$$\|Mu\|_{L^\infty(\mathbb{R}^d)} \leq \|u\|_{L^\infty(\mathbb{R}^d)}, \tag{2.188}$$

if $u \in L^\infty(\mathbb{R}^d)$ and

$$\|Mu\|_{L^p(\mathbb{R}^d)} \leq 2(3^d p')^{1/p}\|u\|_{L^p(\mathbb{R}^d)}, \quad p' := \frac{p}{p-1}, \tag{2.189}$$

if $u \in L^p(\mathbb{R}^d)$ and $1 < p < +\infty$.

Proof of Theorem 2.25. The proof of (2.187) is taken from [SS05, Chapter 3.1.1]. Let $\alpha > 0$. If $Mu(x) > \alpha$, then there is an open ball B_x such that $x \in B_x$ and

$$\int_{B_x} |u(y)|dy > \alpha|B_x|, \tag{2.190}$$

which implies $Mu > \alpha$ on B_x. Therefore $\{Mu > \alpha\}$ is open. Assume that $u \in L^1(\mathbb{R}^d)$. We will show

$$|A_\alpha| \leq \frac{C}{\alpha}\|u\|_{L^1(A_\alpha)}, \quad A_\alpha = \{Mu > \alpha\}, \tag{2.191}$$

from which (2.187) follows. If $x \in A_\alpha$, there is a ball B_x such that $x \in B_x$ and (2.190) is satisfied. Let $K \subset A_\alpha$ be a compact set. From the covering of K by the balls B_x, $x \in K$, we extract a covering by balls B_i, $i \in I$ with I finite, satisfying

$$|B_i| \leq \frac{1}{\alpha}\int_{B_i} |u(y)|dy, \tag{2.192}$$

for all $i \in I$. Note that

$$K \subset \bigcup_{i\in I} B_i \subset A_\alpha, \tag{2.193}$$

the second inclusion being a consequence of the inclusion $B_x \subset A_\alpha$ for all $x \in A_\alpha$. We use an elementary version of the Vitali's covering theorem (see Theorem 2.26 below) to find a subset $J \subset I$ such that $\{B_i; i \in J\}$ is a collection of *disjoint* balls and

$$\left|\bigcup_{i\in I} B_i\right| \leq 3^d \left|\bigcup_{i\in J} B_i\right| = 3^d \sum_{i\in J} |B_i|. \tag{2.194}$$

From (2.192), (2.193) and (2.194), we deduce that

$$|K| \leq \frac{3^d}{\alpha}\sum_{i\in J}\int_{B_i} |u(y)|dy = \frac{3^d}{\alpha}\int_{\cup_{i\in J}B_i} |u(y)|dy \leq \frac{3^d}{\alpha}\int_{A_\alpha} |u(y)|dy. \tag{2.195}$$

Since $K \subset A_\alpha$ is arbitrary, (2.191) follows by inner regularity of the Lebesgue measure. Assume now that $u \in L^p(\mathbb{R}^d)$ with $1 < p \leq +\infty$. The inequality (2.188) is clear, so let us consider the case $1 < p < +\infty$. By Fubini-Tonelli's theorem, for every positive measurable function v, we have

$$\|v\|^p_{L^p(\mathbb{R}^d)} = \int_{\mathbb{R}^d}\int_0^\infty \mathbf{1}_{\{v(x)>\alpha\}} p\alpha^{p-1} d\alpha dx = \int_0^\infty p\alpha^{p-1}|\{v > \alpha\}|d\alpha. \tag{2.196}$$

We use the notation $a \wedge b := \min(a, b)$. If we truncate the sum over α in (2.196), we obtain

$$\|v \wedge R\|^p_{L^p(\mathbb{R}^d)} = \int_0^R p\alpha^{p-1}|\{v > \alpha\}|d\alpha. \tag{2.197}$$

Assume $u \in L^1 \cap L^p(\mathbb{R}^d)$. We can apply (2.191) then, to obtain, for $R > 0$,

$$\|Mu\wedge R\|^p_{L^p(\mathbb{R}^d)} = \int_0^R p\alpha^{p-1}|\{Mu > \alpha\}|d\alpha \leq 3^d\int_0^R p\alpha^{p-2}\int_{\mathbb{R}^d} |u(x)|\mathbf{1}_{\{Mu(x)>\alpha\}} dx d\alpha. \tag{2.198}$$

If we bound $\mathbf{1}_{\{Mu(x)>\alpha\}}$ by 1 in (2.198), we obtain

$$\|Mu \wedge R\|^p_{L^p(\mathbb{R}^d)} \leq 3^d p'\|u\|_{L^1(\mathbb{R}^d)} R^{p-1} < +\infty. \tag{2.199}$$

Using Fubini-Tonelli's theorem and integrating with respect to α in (2.198), we deduce

$$\|Mu \wedge R\|_{L^p(\mathbb{R}^d)}^p \leq 3^d p' \int_{\mathbb{R}^d} |u(x)|[u(x) \wedge R]^{p-1} dx. \tag{2.200}$$

By Hölder's inequality, it follows that

$$\|Mu \wedge R\|_{L^p(\mathbb{R}^d)}^p \leq 3^d p' \|u\|_{L^p(\mathbb{R}^d)} \|Mu \wedge R\|_{L^p(\mathbb{R}^d)}^{p-1}. \tag{2.201}$$

Since $\|Mu \wedge R\|_{L^p(\mathbb{R}^d)} < +\infty$, we have

$$\|Mu \wedge R\|_{L^p(\mathbb{R}^d)} \leq 3^d p' \|u\|_{L^p(\mathbb{R}^d)}. \tag{2.202}$$

We pass to the limit $R \to +\infty$ then. By monotone convergence, (2.202) gives us the estimate

$$\|Mu\|_{L^p(\mathbb{R}^d)} \leq 3^d p' \|u\|_{L^p(\mathbb{R}^d)}. \tag{2.203}$$

Although (2.203) has not the same constant as (2.189), it is still a satisfactory estimate $\|Mu\|_{L^p} \lesssim \|u\|_{L^p}$, but there remains to relax the hypothesis $u \in L^1 \cap L^p(\mathbb{R}^d)$. We cannot simply invoke an argument of density. Indeed (this is precisely our main issue), the map $u \mapsto Mu$ is not easy to apprehend. If $u \in L^p(\mathbb{R}^d)$, the possible obstruction to the integrability property $u \in L^1(\mathbb{R}^d)$ comes from the behaviour of u on "large" sets (more precisely: on sets of infinite measure), where $|u|$ may be much larger than $|u|^p$. Let us therefore consider the truncation u_α defined for $\alpha > 0$ by $u_\alpha = u\mathbf{1}_{|u|>\alpha}$. In this truncation procedure "from below", we replace u by 0 where $|u| \leq \alpha$, so $u_\alpha \in L^1 \cap L^p(\mathbb{R}^d)$. Since $|u| \leq |u_\alpha| + \alpha$, we have $Mu \leq Mu_\alpha + \alpha$, and thus

$$\{Mu > 2\alpha\} \subset \{Mu_\alpha > \alpha\}. \tag{2.204}$$

We use (2.204) to revisit the proof of (2.202). In (2.198), we have

$$\begin{aligned}
\|Mu \wedge R\|_{L^p(\mathbb{R}^d)}^p &= \int_0^R p\alpha^{p-1} |\{|Mu| > \alpha\}| d\alpha = 2^p \int_0^{R/2} p\alpha^{p-1} |\{Mu > 2\alpha\}| d\alpha \\
&\leq 2^p \int_0^{R/2} p\alpha^{p-1} |\{Mu_\alpha > \alpha\}| d\alpha \leq 2^p 3^d \int_0^{R/2} p\alpha^{p-2} \|u_\alpha\|_{L^1(\mathbb{R}^d)} d\alpha \\
&= 2^p 3^d \int_0^{R/2} p\alpha^{p-2} \int_{\mathbb{R}^d} |u(x)| \mathbf{1}_{\{|u(x)|>\alpha\}} d\alpha = 2^p 3^d p' \||u| \wedge (R/2)\|_{L^p(\mathbb{R}^d)}^p.
\end{aligned} \tag{2.205}$$

We conclude by taking the limit $R \to +\infty$. □

Theorem 2.26 (Vitali's covering theorem with a finite number of balls)**.** *Let $\{B_i; i \in I\}$ be a collection of balls with I finite. Then there exists a subcollection $\{B_i; i \in J\}$ of* disjoint *balls such that*

$$\left|\bigcup_{i\in I} B_i\right| \leq 3^d \left|\bigcup_{i\in J} B_i\right|. \tag{2.206}$$

Proof of Theorem 2.26. To construct J from I, we will need to eliminate some balls of the original collection. This loss will be compensated for the following reason: if B and B' are two balls with a non-empty intersection, with B having the largest radius,then $B' \subset 3B$. Let $\mathcal{B}_0 = \{B_i; i \in I\}$. Among the balls of largest radius in $\mathcal{B}_0$, we choose a given ball B_{i_1}. Let $\mathcal{B}_1 \subset \mathcal{B}_0$ be the subcollection obtained by eliminating from $\mathcal{B}_0$ all the balls intersecting B_{i_1} and B_{i_1} itself. Let us repeat this operation on $\mathcal{B}_1$, $\mathcal{B}_2, \dots$ until all balls in the considered collection are disjoint. We obtain a sequence of balls $B_{i_1}, \dots B_{i_k}$ and a sequence of collections

$$\mathcal{B}_k \subset \mathcal{B}_{k-1} \subset \cdots \subset \mathcal{B}_1 \subset \mathcal{B}_0, \tag{2.207}$$

in k steps, $k \leq \#I$. The conclusion follows by considering the collection

$$\{B_j; j \in J\} = \{B_{i_1}, \dots, B_{i_k}\} \bigcup \mathcal{B}_k, \tag{2.208}$$

and noting that $|3B| = 3^d|B|$. □

We will give an application of the L^p-estimate (2.189) to the study of ordinary differential equations with vector fields which are not Lipschitz continuous (but have Sobolev regularity), see Section 2.5.4. For the moment, we will see how the *Lebesgue's differentiation Theorem* can be deduced from the estimate (2.187).

2.5.2 Lebesgue's differentiation Theorem

Theorem 2.27 (Lebesgue's differentiation Theorem)**.** *If $u \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, then*

$$\lim_{r \to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} u(y)dy = u(x), \tag{2.209}$$

for a.e. $x \in \mathbb{R}^d$.

Proof of Theorem 2.27. We will establish the stronger result

$$\lim_{r \to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |u(y) - u(x)|dy = 0, \tag{2.210}$$

for a.e. $x \in \mathbb{R}^d$. Indeed, the inequality

$$\left| \int_{B(x,r)} u(y)dy - u(x) \right| = \left| \int_{B(x,r)} (u(y) - u(x))dy \right| \leq \frac{1}{|B(x,r)|} \int_{B(x,r)} |u(y) - u(x)|dy \tag{2.211}$$

shows that (2.210) implies (2.209). Denote by u^* the Hardy-Littlewood maximal function, defined by

$$u^*(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |u(y)|dy. \tag{2.212}$$

We will use the bound

$$\frac{1}{|B(x,r)|} \int_{B(x,r)} |u(y)|dy \leq u^*(x) \leq Mu(x). \tag{2.213}$$

We put forward two additional facts:

- the result (2.210) is local: it is sufficient to prove (2.210) for a.e. $x \in K$ where K is an arbitrary compact set. To establish (2.210) a.e. on K, we can replace u by $u\mathbf{1}_U$, where U is an open set with compact closure such that $K \subset U$. Otherwise speaking, we can assume without loss of generality that $u \in L^1(\mathbb{R}^d)$.
- (2.210) is satisfied for all $x \in \mathbb{R}^d$ if u is continuous (if $|u(y) - u(x)| < \varepsilon$ for all y in a neighbourhood of x, then the integral in (2.210) is smaller than ε for r small enough).

To transfer the property (2.210) for continuous functions to general integrable functions, we will use a density argument: given $\varepsilon > 0$, there exists $\tilde{u}$ a continuous compactly supported function such that

$$\|w\|_{L^1(\mathbb{R}^d)} < \varepsilon, \quad w := u - \tilde{u}. \tag{2.214}$$

For $x \in \mathbb{R}^d$, $r > 0$, we have, by triangular inequality,

$$\begin{aligned}\frac{1}{|B(x,r)|}\int_{B(x,r)} |u(y) - u(x)|dy \leq \frac{1}{|B(x,r)|}\int_{B(x,r)} |\tilde{u}(y) - \tilde{u}(x)|dy \\ + \frac{1}{|B(x,r)|}\int_{B(x,r)} |u(y) - \tilde{u}(y)|dy + |\tilde{u}(x) - u(x)|.\end{aligned} \tag{2.215}$$

Set

$$\delta[u;r](x) = \frac{1}{|B(x,r)|}\int_{B(x,r)} |u(y) - u(x)|dy. \tag{2.216}$$

We have $\limsup_{r\to 0} \delta[\tilde{u};r](x) = 0$ at all point x, so (2.215) and (2.213) (with $w = u - \tilde{u}$ instead of u) give

$$\limsup_{r\to 0} \delta[u;r] \leq Mw + |w|. \tag{2.217}$$

Let $\alpha > 0$. Using (2.217), we have

$$\left|\left\{\limsup_{r\to 0} \delta[u;r] > \alpha\right\}\right| \leq |\{Mw > \alpha/2\}| + |\{|w| > \alpha/2\}|\,. \tag{2.218}$$

By (2.187) and by the Markov inequality, the right-hand side of (2.218) is bounded by

$$2\frac{3^d+1}{\alpha}\|w\|_{L^1(\mathbb{R}^d)} < 2\frac{3^d+1}{\alpha}\varepsilon. \tag{2.219}$$

Taking the limit $\varepsilon \to 0$, we deduce that $|\{\limsup_{r\to 0} \delta[u;r] > \alpha\}| = 0$. Since α is arbitrary, it follows by continuity of the Lebesgue measure with respect to increasing limit that

$$\left|\left\{\limsup_{r\to 0} \delta[u;r] = 0\right\}\right| = \lim_{n\to+\infty}\left|\left\{\limsup_{r\to 0} \delta[u;r] > n^{-1}\right\}\right| = 0. \tag{2.220}$$

This shows that (2.210) is satisfied for a.e. $x \in \mathbb{R}^d$. □

Remark 2.8 (Lebesgue differentiation Theorem: L^1-convergence). Let $u \in L^1(\mathbb{R}^d)$. By the homogeneity property $|B(x,r)| = |B(0,1)|r^d$, the ratio $\delta[u;r]$ in (2.216) can be written

$$\delta[u;r](x) = \frac{1}{|B(0,1)|}\int_{B(0,1)} |u(x+rz) - u(x)|dz. \tag{2.221}$$

We have simply used the change of variable $z \mapsto y = x + rz$, of Jacobian r^d. By the Fubini theorem, we have then

$$\|\delta[u;r]\|_{L^1(\mathbb{R}^d)} = \frac{1}{|B(0,1)|}\int_{B(0,1)} \|\tau_{-rz}u - u\|_{L^1}dz \leq \omega_{L^1}(u;r), \tag{2.222}$$

and (2.96) gives us immediately the L^1 convergence

$$\lim_{r\to 0} \delta[u;r] = 0 \text{ in } L^1(\mathbb{R}^d). \tag{2.223}$$

If $(r_n) \downarrow 0$, it follows from (2.223) that there is a subsequence (r_{n_k}) such that $\delta[u, r_{n_k}] \to 0$ a.e. when $k \to +\infty$. For all $r = r_n$, there exists then k such that $n_k \leq n < n_{k+1}$, and we can try to estimate $\delta[u; r_n]$ by

$$\delta[u; r_n] \leq \frac{|B(0, r_{n_k})|}{|B(0, r_{n_{k+1}})|} \times \frac{1}{|B(0, r_{n_k})|} \int_{B(0, r_{n_k})} |u(y) - u(x)| dy = \left[\frac{r_{n_k}}{r_{n_{k+1}}}\right]^d \delta[u; r_{n_k}]. \quad (2.224)$$

It is possible to exploit (2.224) to prove $\delta[u; r_n] \to 0$ a.e. only if the sequence (r_{n_k}) decays slowly enough, and there is no reason for this. Even the original sequence may decay quickly and satisfy $r_{n+1} = o(r_n)$. There are other situation in mathematics where it requires some particular efforts to establish a result a.e. We can mention for instance:

- the law of large numbers in probability theory,
- the ergodic theorem in the study of dynamical systems (Birkhoff's ergodic theorem).

All these examples are not strictly uncorrelated: the law of large numbers can be deduced from Birkhoff's ergodic theorem, and the classical proof of the latter uses maximal functions, [SS05, Chapter 6.5.2]).

2.5.3 A brief reminder on the theory of ordinary differential equations

Let us first recall some elementary facts about ordinary differential equations (ODEs). If $t \mapsto x(t) \in \mathbb{R}^d$ is seen as the parametrization of a curve γ in $\mathbb{R}^d$, an ODE $\dot{x}(t) = a(x(t))$ gives a way to follow γ by assigning the tangent vector $\dot{x}(t)$ at each position $x(t)$. The curve γ may be a conic for instance: this is what happens in the fundamental case of the resolution of the two-body problem in the theory of gravitation of Newton. An ODE $\dot{x}(t) = a(x(t))$ in dimension d can also be seen as a system of d equations in d unknowns: the modelling of a vast number of phenomena in chemistry, biology, physics... uses such an approach based on ODEs. In this section, we will discuss only autonomous ODEs, where the vector field $a\colon \mathbb{R}^d \to \mathbb{R}^d$ only depends on the state variable x, and not on the time variable, although the generalization of the theory to the non-autonomous case is possible and important.

Global existence, blow-up in finite time. Consider first the following elementary examples, where $x(t)$, $y(t)$, $z(t)$ may denote the size of a given population at time t:

1. $d = 1$, $\dot{x}(t) = x(t)$,
2. $d = 1$, $\dot{y}(t) = y(t)\ln(|y(t)|)$,
3. $d = 1$, $\dot{z}(t) = (z(t))^2$.

These three examples can be written as equations in separate variables

$$\frac{dx}{G(x)} = dt.$$

By computing a primitive function of G^{-1}, we obtain the following explicit expressions ($t > 0$, $x_0, y_0, z_0 > 0$):

$$x(t) = e^t x_0, \quad y(t) = \exp(e^t \ln(y_0)), \quad z(t) = \frac{z_0}{1 - z_0 t}.$$

The expression for $z(t)$ is valid as long as $t < z_0^{-1}$. In example *1.*, the growth of $a(x) = x$ is linear, and $t \mapsto x(t)$ as en exponential growth. This indicates in itself that the solution of an

ODE with a vector field that grows superlinearly may blow up in finite time. The example *3*. illustrates this fact (we may as well consider the ODE $\dot{z}(t) = |z(t)|^p$ with $p > 1$). The example *2*. shows that it is possible to go a little bit beyond the sublinear case, regarding the growth of a, and still get solutions defined globally in time (see Osgood's lemma for a general result). Now, consider a vector field $a\colon \mathbb{R}^d \to \mathbb{R}^d$ which is *globally Lipschitz continuous*: there exists $L \geq 0$ such that

$$|a(x) - a(y)| \leq L|x - y|, \tag{2.225}$$

where $|\cdot|$ is the euclidean norm on $\mathbb{R}^d$. If we apply (2.225) to $y = 0$, we obtain

$$|a(x)| \leq C(|x| + 1), \quad C = L + |a(0)|,$$

and this shows that a has a *sublinear growth*. In view of the three examples discussed above, we expect the ODE $\dot{x}(t) = a(x(t))$ to admit some solutions which are defined globally in time.

Uniqueness. Under the Lipschitz condition (2.225), we also have the estimate

$$|x_1(t) - x_2(t)| \leq (e^{Lt} - 1)|x_1(0) - x_2(0)|, \tag{2.226}$$

where $t \in [0, T]$ and $x_1(t)$ and $x_2(t)$ are two solutions of the ODE $\dot{x}(t) = a(x(t))$ on $[0, T]$. The estimate (2.226) is a consequence of the Grönwall Lemma. It gives a result of continuous dependence on the data and, in particular, a result of uniqueness, for the solutions of the ODE. The example of the functions $x\colon t \mapsto [(t - t_0)^+]^m$, $t_0 \geq 0$, which give infinitely many solutions to the ODE $\dot{x}(t) = a(x(t))$, $a(x) = m|x|^{1-1/m}$ satisfying $x(0) = 0$, shows that uniqueness will not hold if a Lipschitz condition is replaced by a Hölder condition. To sum up these discussions, we consider the following two cases: globally Lipschitz continuous vector field, locally Lipschitz continuous vector field.

Globally Lipschitz continuous vector field. If a is globally Lipschitz continuous, we expect to be able to solve $\dot{x}(t) = a(x)$ in a unique manner (once prescribed a given point at time $t = 0$), and globally in time. The classical proof of this result uses the integral formulation of the ODE (with initial condition $x(0) = x_0$):

$$x(t) = x_0 + \int_0^t a(x(s))ds, \tag{2.227}$$

where x is *a priori* considered to be a continuous function of t.

Theorem 2.28 (Cauchy-Lipschitz global)**.** *Let $a\colon \mathbb{R}^d \to \mathbb{R}^d$ be a globally Lipschitz continuous vector field. Then, for all $x_0 \in \mathbb{R}^d$, the ODE $\dot{x}(t) = a(x(t))$ has a unique solution $t \mapsto x(t; x_0)$ satisfying* (2.227) *for all $t \in \mathbb{R}$.*

We can see on the integral equation (2.227) that x is C^1 in t. If a has some additional regularity, for instance a is of class C^k, $k \geq 1$, then x is of class C^{k+1} in t (proof by recursion on k). The *flow* associated to the vector field a is the collection of trajectories $t \mapsto x(t; x_0)$, usually denoted $X_t(x_0)$ or $\Phi_t(x_0)$.

Locally Lipschitz continuous vector field. Assume now that $a\colon \mathbb{R}^d \to \mathbb{R}^d$ is locally Lipschitz continuous. We have a good (local) dependence on the data by (2.226), but blow up is possible, as illustrated by the example *3*.

Theorem 2.29 (Cauchy-Lipschitz local, blow-up)**.** *Let $a\colon \mathbb{R}^d \to \mathbb{R}^d$ be a locally Lipschitz continuous function and let $x_0 \in \mathbb{R}^d$. A* solution (x, J) *to the Cauchy Problem*

$$\begin{cases} \dot{x}(t) = a(x(t)), \\ x(0) = x_0, \end{cases} \tag{2.228}$$

is a continuous function $x\colon J \to \mathbb{R}^d$, where J is an interval of $\mathbb{R}$ containing 0, satisfying (2.227) for all $t \in J$. We have the following results.

1. ***Compatibility.*** *Two solutions (x_1, J_1) and (x_2, J_2) to (2.228) satisfy $x_1 = x_2$ on $J_1 \cap J_2$.*
2. ***Maximal solution.*** *There is a unique maximal solution (x^*, J^*) to (2.228): if (x, J) is an other solution, then $J \subset J^*$ and $x = x^*$ on J.*
3. ***Blow-up in finite time.*** *If $\sup J^* < +\infty$, then $x^*(t)$ is unbounded in the neighbourhood of $\sup J^*$ (same result in the neighbourhood of $\inf J^*$ if $\inf J^* > -\infty$).*

Solving transport equation (omit on first reading). Let u be a (smooth) solution to the transport equation

$$\partial_t u(x,t) + a(x) \cdot \nabla_x u(x,t) = 0, \tag{2.229}$$

with initial condition $u(x, 0) = u_0(x)$. By differentiation of u along the flow Φ_t of a, we obtain

$$\frac{\partial}{\partial t} [u(\Phi_t(x), t)] = (\partial_t u)(\Phi_t(x), t) + \dot{\Phi}_t(x) \cdot \nabla_x u(\Phi_t(x), t).$$

Since $\dot{\Phi}_t(x) = a(\Phi_t(x))$, we get

$$\frac{\partial}{\partial t} [u(\Phi_t(x), t)] = (\partial_t u + a \cdot \nabla_x u)(\Phi_t(x), t) = 0.$$

Consequently u is given by

$$u(x,t) = u_0 \circ \Phi^t(x), \tag{2.230}$$

where Φ^t denotes the inverse of $x \mapsto \Phi_t(x)$ (actually $\Phi^t = \Phi_{-t}$ here, since the equation is autonomous). This link between the partial differential equation (2.229) and the ODE $\dot{x}(t) = a(x(t))$ is both a reason to extend the theory of the Cauchy Problem for ODEs to the cases of vector fields a with a regularity less than Lipschitz, and an help to achieve this goal. It happens frequently that the transport equation (2.229) is part of a larger system of equations, where a depends of some the unknowns of the problems and where natural estimates give a Sobolev regularity for a. Consider for instance the resolution of the Vlasov-Poisson system ($x, v \in \mathbb{R}^3$)

$$\partial_t f(t,z) + a(t,z) \cdot \nabla_z f(t,z) = 0, \quad z = \begin{pmatrix} x \\ v \end{pmatrix}, \quad a(t,z) = \begin{pmatrix} v \\ E(x,t) \end{pmatrix},$$

where the field E is given by $E(x,t) = -\nabla_x V(x,t)$ and the potential V is deduced from the Poisson Equation $-\Delta_x V(x,t) = \rho(x,t)$, where ρ is the density, which depends on the solution f itself:

$$\rho(x,t) = \int_{\mathbb{R}^3} f(t,x,v)dv.$$

To generalize the resolution of (2.229) to the case of non-smooth functions, it is also natural to consider weak solutions to the equation (2.229). Then, in view of (2.230), we will need to give a meaning to an integral

$$\int_{\mathbb{R}^d} v \circ \Phi^t(x) dx, \tag{2.231}$$

where v is a non-negative integrable function (v may be $|u_0|$ for instance). Recall that $\Phi^t = \Phi_{-t}$ since we consider the autonomous case, so (2.231) will be well defined if $\Phi_t(x)$ is defined for a.e. $x \in \mathbb{R}^d$ and if Φ_t satisfies the condition: there exists $L \geq 0$ (called a compressibility constant) such that, for all t,

$$\int_{\mathbb{R}^d} v \circ \Phi_t(x)dx \leq L \int_{\mathbb{R}^d} v(x)dx, \tag{2.232}$$

for all integrable non-negative function v on $\mathbb{R}^d$. The estimate (2.234) in Theorem 2.30 below is an essential step in the construction of a generalized flow satisfying (2.232) (in the approach developed by Crippa and de Lellis, [CDL08]).

2.5.4 Solving ODEs with vector fields of Sobolev regularity

In this section, we will establish the following result, where B_r denotes the ball $B(0, r)$.

Theorem 2.30 (Stability of the flow, a quantitative estimate)**.** *Let a and $\tilde{a}$ be two smooth, bounded, vector fields on $\mathbb{R}^d$. Let Φ_t and $\tilde{\Phi}_t$ be the associated flows. We assume that there are some constants $L, \tilde{L} \geq 0$ such that*

$$\int_{\mathbb{R}^d} \varphi \circ \Phi_t(x)dx \leq L \int_{\mathbb{R}^d} \varphi(x)dx, \quad \int_{\mathbb{R}^d} \varphi \circ \tilde{\Phi}_t(x)dx \leq \tilde{L} \int_{\mathbb{R}^d} \varphi(x)dx, \tag{2.233}$$

for all non-negative, continuous and compactly supported function $\varphi\colon \mathbb{R}^d \to \mathbb{R}_+$. Let $p \in (1, +\infty)$ and let $T > 0$ be fixed. Let $K = \max(\|a\|_{L^\infty(\mathbb{R}^d)}, \|\tilde{a}\|_{L^\infty(\mathbb{R}^d)})$. Then we have the estimate

$$\sup_{t\in[0,T]} \|\Phi_t - \tilde{\Phi}_t\|_{L^1(B_1)} \leq C|\log(\|a - \tilde{a}\|_{L^1(B_R)})|^{-1}, \tag{2.234}$$

where $R = 1 + KT$ and the constant C depends only on T, p, K, and on the quantities $\|D_x a\|_{L^p(B_{3+KT})}$, L, $\tilde{L}$.

Notations: we will use the following notations: for $r > 0$ and $u \in L^1_{\text{loc}}(\mathbb{R}^d)$ we define the local maximal function $M_r[u]$ by

$$M_r[u](x) = \sup_{0<t<r} \frac{1}{|B(x, t)|} \int_{B(x,t)} |u(y)|dy. \tag{2.235}$$

We will also consider, for $\delta > 0$, the function

$$G_\delta(t) = \int_{B_1} \log\left(1 + \frac{|\Phi_t(x) - \tilde{\Phi}_t(x)|}{\delta}\right) dx, \tag{2.236}$$

where log denote the inverse of exp. We denote by $D_x a$ the matrix-valued function $x \mapsto (\partial_{x_j} a^i(x))_{1\leq i,j\leq d}$ and, for $E \subset \mathbb{R}^d$, also denote by $\|D_x a\|_{L^p(E)}$ the norm

$$\|D_x a\|_{L^p(E)} = \left(\int_E |D_x a(x)|^p dx\right)^{1/p}, \tag{2.237}$$

where, given $A \in \mathcal{M}_d(\mathbb{R})$, $|A|$ is the norm subordinated to the euclidean norm on $\mathbb{R}^d$:

$$|A| = \sup\left\{|Ax|; x \in \mathbb{R}^d; |x| \leq 1\right\}.$$

We also use the notation $\mathrm{div}(a)$ for the divergence operator

$$\mathrm{div}(a)(x) = \sum_{i=1}^{d} \frac{\partial a^i}{\partial x_i}(x) = \mathrm{tr}(D_x a)(x). \tag{2.238}$$

If $s \in \mathbb{R}$, $s^- = \max(-s, 0)$ is the negative part of s. At last, we denote by $C_1, C_2, \dots$ any constant that depends on T, p, K, and on the quantities $\|D_x a\|_{L^p(B_{3+KT})}$, L, $\tilde{L}$ only.

Before giving the proof of Theorem 2.30, we do some preliminary remarks.

Remark 2.9 (Rescaling). Let $r \geq 1$. Let $\Psi_t(x) = r^{-1}\Phi_{rt}(rx)$. We have $\Psi_0(x) = x$ and

$$\dot{\Psi}_t(x) = \dot{\Phi}_{rt}(rx) = a(\Phi_{rt}(rx)) = a(r\Psi_t(x)). \tag{2.239}$$

By uniqueness, $\Psi_t(x) = r^{-1}\Phi_{rt}(rx)$ is the flow associated to the vector field $x \mapsto a(rx)$. If we apply (2.234) to Ψ_t and $\tilde{\Psi}_t$ on the interval $[0, r^{-1}T]$, we obtain

$$r^{-(d+1)} \sup_{t\in[0,T]} \|\Phi_t - \tilde{\Phi}_t\|_{L^1(B_r)} \leq C|\log(r^{-d}\|a - \tilde{a}\|_{L^1(B_R)})|^{-1}, \tag{2.240}$$

where $R = r + \|a\|_{L^\infty(\mathbb{R}^d)}T$ and the constant C depends only on r, T, p, K, and on the quantities $\|D_x a\|_{L^p(B_{3r+KT})}$, L, $\tilde{L}$. Since $r \geq 1$, this gives the more general estimate

$$\sup_{t\in[0,T]} \|\Phi_t - \tilde{\Phi}_t\|_{L^1(B_r)} \leq C|\log(\|a - \tilde{a}\|_{L^1(B_R)})|^{-1}. \tag{2.241}$$

Remark 2.10 (Global flow). Since a is at least of class C^1, it is locally Lipschitz continuous. Let (x, J) be the maximal solution to (2.228). The integral equation (2.227) gives the bound

$$|x(t)| \leq |x_0| + t\|a\|_{L^\infty(\mathbb{R}^d)}, \tag{2.242}$$

for all $t \in J$. The blow-up criterion in Theorem 2.29 ensures that x is defined globally: $J = \mathbb{R}$. Consequently, the flow $\Phi_t(x)$ is defined for all $(t, x) \in \mathbb{R} \times \mathbb{R}^d$ and we have the finite propagation property

$$\Phi_t(B_r) \subset B_{r+tK}, \tag{2.243}$$

for all $r \geq 0$ and $t \geq 0$.

Remark 2.11 (Compressibility constant). Assume that

$$\|[\mathrm{div}(a)]^-\|_{L^\infty(\mathbb{R}^d)} < +\infty. \tag{2.244}$$

Then the first inequality in (2.233) is satisfied with $L = \exp(T\|[\mathrm{div}(a)]^-\|_{L^\infty(\mathbb{R}^d)})$. Indeed, note first that the integrals in (2.233) are well defined in virtue of the finite propagation property (2.243). The change of variable $z = \Phi_t(x)$ gives

$$\int_{\mathbb{R}^d} \varphi \circ \Phi_t(x)dx = \int_{\mathbb{R}^d} \varphi(z)|(J\Phi_t^{-1}(z))|dz. \tag{2.245}$$

Recall that

$$D_x\Phi_t(x) = \exp\left(\int_0^t D_x a(\Phi_s(x))ds\right), \tag{2.246}$$

for the reason that $t \mapsto D_x\Phi_t(x)$ solves the linear ODE on $\mathbb{R}^{d\times d}$ given by $\dot{A}_t = D_x a(\Phi_t)A_t$, with initial condition $A_0 = \mathrm{I}_d$, and that $\det(\exp(A)) = \exp(\mathrm{tr}(\mathrm{A}))$, so that the Jacobian determinant $J\Phi_t$ satisfies

$$J\Phi_t(x) = \exp\left(\int_0^t \mathrm{tr}(D_x a)(\Phi_s(x))ds\right) = \exp\left(\int_0^t (\mathrm{div}(a))(\Phi_s(x))ds\right)$$
$$\geq \exp\left(\int_0^t (-\,\mathrm{div}(a))^-(\Phi_s(x))ds\right) \geq \exp\left(-T\|[\mathrm{div}(a)]^-\|_{L^\infty(\mathbb{R}^d)}\right), \quad (2.247)$$

and thus

$$J\Phi_t^{-1}(z) = \left[J\Phi_t \circ \Phi_t^{-1}(z)\right]^{-1} \leq \exp(T\|[\mathrm{div}(a)]^-\|_{L^\infty(\mathbb{R}^d)}). \quad (2.248)$$

Proof of Theorem 2.30. First, given $r, \rho > 0$ we note that we have the local L^p-estimate

$$\|M_r[u]\|_{L^p(B_\rho)} \leq 2(3^d p')^{1/p}\|u\|_{L^p(B_{r+\rho})}, \quad p' := \frac{p}{p-1}, \quad (2.249)$$

for all $1 < p < +\infty$ and $u \in L^p_{\mathrm{loc}}(\mathbb{R}^d)$. This follows from the bounds $M_r[u](x) \leq M[u\mathbf{1}_{B(x,r)}](x)$ and $\mathbf{1}_{B(x,r)} \leq \mathbf{1}_{B_{r+\rho}}$ if $x \in B_\rho$. Then we apply (2.189) to $u\mathbf{1}_{B_{r+\rho}}$ to obtain (2.249). Next, we prove that, for $x \in \mathbb{R}^d$, $r > 0$, we have

$$\left|\frac{1}{|B(x,r)|}\int_{B(x,r)} a(z)dz - a(x)\right| \leq rM_r[|D_x a|](x). \quad (2.250)$$

We use the expansion

$$\frac{1}{|B(x,r)|}\int_{B(x,r)} a(z)dz - a(x) = \frac{1}{|B(x,r)|}\int_{B(x,r)}\int_0^1 D_x a(tz + (1-t)x)\cdot(z-x)dtdz, \quad (2.251)$$

and the triangular inequality to obtain the first estimate

$$\left|a(x) - \frac{1}{|B(x,r)|}\int_{B(x,r)} a(z)dz\right| \leq \frac{1}{|B(x,r)|}\int_{B(x,r)}\int_0^1 |D_x a|(tz+(1-t)x)|z-x|dtdz. \quad (2.252)$$

By Fubini's theorem and the change of variable $B(x,r) \ni z \mapsto x + t(z-x) \in B(x,rt)$ we can transform the right-hand side of (2.252) in

$$\frac{1}{|B(x,r)|}\int_0^1\int_{B(x,rt)} |D_x a|(z)|z-x|t^{-1}dzt^{-n}dt. \quad (2.253)$$

Since $t^n|B(x,r)| = |B(x,tr)|$, using the change of variable $t' = rt$ gives us

$$(2.253) = \int_0^1 \frac{1}{|B(x,rt)|}\int_{B(x,rt)} |D_x a|(z)|z-x|t^{-1}dzdt$$
$$= \int_0^r \frac{1}{|B(x,t)|}\int_{B(x,t)} |D_x a|(z)|z-x|t^{-1}dzdt, \quad (2.254)$$

which can be estimated by

$$\int_0^r M_r[|D_x a|](x)dt \leq rM_r[|D_x a|](x). \quad (2.255)$$

We complete (2.250) with the following estimate: for $x, y \in \mathbb{R}^d$,

$$\left|\frac{1}{|B(x,r)|}\int_{B(x,r)} a(z)dz - \frac{1}{|B(y,r)|}\int_{B(y,r)} a(z)dz\right| \leq C_d|x-y|M_{2r}[|D_x a|](x), \tag{2.256}$$

where $r = |x-y|$ and the constant C_d depends on d only. To establish (2.256), let us write first

$$\begin{aligned} A_{x,y} &:= \frac{1}{|B(x,r)|}\int_{B(x,r)} a(z)dz - \frac{1}{|B(y,r)|}\int_{B(y,r)} a(z)dz \\ &= \frac{1}{|B_r|}\int_{B_r} a(x+z)dz - \frac{1}{|B_r|}\int_{B_r} a(y+z)dz, \end{aligned} \tag{2.257}$$

to obtain

$$|A_{x,y}| \leq \frac{1}{|B_r|}\int_{B_r}\int_0^1 |D_x a|(tx+(1-t)y+z)|x-y|dtdz. \tag{2.258}$$

We use Fubini's theorem and the change of variable $z' = tx + (1-t)y + z$. For $t \in [0,1]$, we have

$$z' \in B(tx+(1-t)y, r) = B(x+(1-t)(y-x), r) \subset B(x, 2r), \tag{2.259}$$

so

$$|A_{x,y}| \leq \frac{1}{|B_r|}\int_0^1\int_{B(x,2r)} |D_x a|(z)|x-y|dzdt \leq \frac{|B_{2r}|}{|B_r|}|x-y|M_{2r}[|D_x a|](x), \tag{2.260}$$

and this gives the desired result since $|B_{2r}| = 2^d|B_r|$. Now let us decompose the difference $a(x) - a(y)$ into the three parts

$$a(x) - \frac{1}{|B(x,r)|}\int_{B(x,r)} a(z)dz, \quad \frac{1}{|B(x,r)|}\int_{B(x,r)} a(z)dz - \frac{1}{|B(y,r)|}\int_{B(y,r)} a(z)dz, \tag{2.261}$$

and

$$\frac{1}{|B(y,r)|}\int_{B(y,r)} a(z)dz - a(y). \tag{2.262}$$

If we apply (2.250) to the first and third term and apply (2.256) to the second term, we conclude to the following estimate:

$$|a(x) - a(y)| \leq C_d|x-y|(M_{2r}[|D_x a|](x) + M_{2r}[|D_x a|](y)), \tag{2.263}$$

where $r = |x-y|$ and the constant C_d depends on d only. We will now justify that

$$G_\delta(t) \leq \int_0^t\int_{B_1} \frac{|a(\Phi_s(x)) - \tilde{a}(\tilde{\Phi}_s(x))|}{\delta + |\Phi_s(x) - \tilde{\Phi}_s(x)|}dxds, \tag{2.264}$$

for all $t \in [0,T]$, where G_δ is defined by (2.236). The estimate (2.264) can be obtained formally by differentiation of G_δ. Since the norm $x \mapsto |x|$ is not differentiable at $x = 0$, we have to use a first step of regularization. If $J\colon \mathbb{R}^d \to \mathbb{R}_+$ is a smooth function and

$$G_{J,\delta}(t) = \int_{B_1} \log\left(1 + \frac{J(\Phi_t(x) - \tilde{\Phi}_t(x))}{\delta}\right)dx, \tag{2.265}$$

then $t \mapsto G_{J,\delta}$ is of class C^1 and

$$G'_{J,\delta}(t) = \int_{B_1} \frac{\nabla J[\Phi_t(x) - \tilde{\Phi}_t(x)] \cdot [a(\Phi_t(x)) - \tilde{a}(\tilde{\Phi}_t(x))]}{\delta + J(\Phi_t(x)) - \tilde{\Phi}_t(x))}dx. \tag{2.266}$$

To obtain (2.266), we use the ODEs $\dot{\Phi}_t(x) = a(\Phi_t(x))$, $\dot{\tilde{\Phi}}_t(x) = \tilde{a}(\tilde{\Phi}_t(x))$. To justify that $G_{J,\delta}$ is of class C^1, we use the theorem of differentiation under the integral sign. Indeed, setting

$$F(t,x) = \log\left(1 + \frac{J(\Phi_t(x) - \tilde{\Phi}_t(x))}{\delta}\right), \tag{2.267}$$

we have

1. being continuous, $x \mapsto F(t,x)$ is integrable (since measurable and bounded) for all, and hence at least one $t \in [0,T]$,
2. for all $x \in B_r$, $t \mapsto F(t,x)$ is of class C^1 on $[0,T]$,
3. there exists an integrable function $\bar{F}$ such that $|\partial_t F(t,x)| \leq \bar{F}(x)$ for all $x \in B_1$, $t \in [0,T]$.

For $\bar{F}(x)$, using (2.243), we can simply consider the constant function

$$x \mapsto \frac{M + \tilde{M}}{\delta} \sup_{z \in B_{2+T(M+\tilde{M})}} |\nabla J(z)|. \tag{2.268}$$

By integration with respect to t in (2.266), and provided $J(0) = 0$, we obtain

$$G_{J,\delta}(t) = \int_0^t \int_{B_1} \frac{\nabla J[\Phi_s(x) - \tilde{\Phi}_s(x)] \cdot [a(\Phi_s(x)) - \tilde{a}(\tilde{\Phi}_s(x))]}{\delta + J(\Phi_s(x)) - \tilde{\Phi}_s(x))} dx ds. \tag{2.269}$$

We consider then an approximation of $x \mapsto J_0(x) := |x|$ by some functions J_ε of class C^1, for instance

$$J_\varepsilon(x) = \sqrt{\varepsilon + |x|^2}. \tag{2.270}$$

With this choice (2.270) for J_ε, we have $\nabla J_\varepsilon(x) = \frac{x}{\sqrt{\varepsilon + |x|^2}}$, so $|\nabla J_\varepsilon(x)| \leq 1$ for all x. This has the consequence that

$$|G_{J_\varepsilon,\delta}(t)| \leq \int_0^t \int_{B_1} \frac{|a(\Phi_s(x)) - \tilde{a}(\tilde{\Phi}_s(x))|}{\delta + J_\varepsilon(\Phi_s(x)) - \tilde{\Phi}_s(x))} dx ds. \tag{2.271}$$

By dominated convergence, we can pass to the limit $[\varepsilon \to 0]$ in (2.271) to obtain (2.264). It follows from (2.264) that

$$G_\delta(t) \leq \frac{1}{\delta} \int_0^t \int_{B_1} |a(\tilde{\Phi}_s(x)) - \tilde{a}(\tilde{\Phi}_s(x))| dx ds + \int_0^t \int_{B_1} \frac{|a(\Phi_s(x)) - a(\tilde{\Phi}_s(x))|}{\delta + |\Phi_s(x) - \tilde{\Phi}_s(x)|} dx ds. \tag{2.272}$$

We focus on the first term in the right-hand side of (2.272). By (2.233), it is bounded by

$$\frac{\tilde{L}}{\delta} \int_0^t \int_{\tilde{\Phi}_s(B_1)} |a(y) - \tilde{a}(y)| dy. \tag{2.273}$$

Using (2.243) therefore gives

$$G_\delta(t) \leq \frac{\tilde{L}T}{\delta} \|a - \tilde{a}\|_{L^1(B_{1+KT})} + \int_0^t \int_{B_1} \frac{|a(\Phi_s(x)) - a(\tilde{\Phi}_s(x))|}{\delta + |\Phi_s(x) - \tilde{\Phi}_s(x)|} dx ds. \tag{2.274}$$

We apply then (2.263) with points $\Phi_s(x)$, $\tilde{\Phi}_s(x)$ to obtain a bound on the second term in the right-hand side of (2.274) by

$$C_d \int_0^t \int_{B_1} (M_2[|D_x a|](\Phi_s(x)) + M_2[|D_x a|](\tilde{\Phi}_s(x)))dxds. \tag{2.275}$$

By the "compressibility" condition (2.233) and the bound on the speed of propagation (2.243), we can estimate (2.275) by

$$2C_d(L+\tilde{L}) \int_0^t \int_{B_{1+KT}} M_2[|D_x a|](z)dzds \leq 2C_d(L+\tilde{L})T \int_{B_{1+KT}} M_2[|D_x a|](z)dz. \tag{2.276}$$

We use the Hölder inequality and the local L^p estimate (2.249) to estimate (2.276) from above by

$$2C_d(L+\tilde{L})T|B_{1+KT}|^{\frac{1}{p'}} 2(3^d p')^{1/p} \|D_x a\|_{L^p(B_{3+KT})} \leq C_1, \tag{2.277}$$

where C_1 is a generic constant depending on T, p, K, and on the quantities $\|D_x a\|_{L^p(B_{3+KT})}$, L, $\tilde{L}$, and conclude that

$$G_\delta(t) \leq \frac{\tilde{L}T}{\delta} \|a-\tilde{a}\|_{L^1(B_{1+KT})} + C_1. \tag{2.278}$$

We take now $\delta = \|a-\tilde{a}\|_{L^1(B_{1+KT})}$ so that (2.278) gives $G_\delta(t) \leq C_2$. Let $\eta > 0$ be a given parameter. Let

$$H(x) = \log\left(1 + \frac{|\Phi_t(x) - \tilde{\Phi}_t(x)|}{\delta}\right), \tag{2.279}$$

and let $A = \{x \in B_1; H(x) \leq \frac{C_2}{\eta}\}$. By the Markov inequality, we have

$$|B_1 \setminus K| = \left|\left\{x \in B_1; H > \frac{C_2}{\eta}\right\}\right| \leq \frac{\eta}{C_2} \int_{B_1} H(x)dx \leq \eta. \tag{2.280}$$

Therefore, by definition of the set A, we have

$$|B_1 \setminus A| \leq \eta, \quad |\Phi_t - \tilde{\Phi}_t| \leq \delta \exp(C_2/\eta) \text{ on } A. \tag{2.281}$$

We decompose then

$$\|\Phi_t - \tilde{\Phi}_t\|_{L^1(B_1)} = \|\Phi_t - \tilde{\Phi}_t\|_{L^1(B_1 \setminus A)} + \|\Phi_t - \tilde{\Phi}_t\|_{L^1(A)}. \tag{2.282}$$

On A, we use the bound

$$\|\Phi_t - \tilde{\Phi}_t\|_{L^1(A)} \leq |A|\delta \exp(C_2/\eta) \leq |B_1|\delta \exp(C_2/\eta) \leq C_3 \delta \exp(C_2/\eta). \tag{2.283}$$

On the complementary set $B_1 \setminus A$, we have, by (2.243)

$$\|\Phi_t - \tilde{\Phi}_t\|_{L^1(B_1 \setminus A)} \leq \left(\|\Phi_t\|_{L^\infty(B_1)} + \|\tilde{\Phi}_t\|_{L^\infty(B_1)}\right)|A| \leq C_4 \eta. \tag{2.284}$$

Without loss of generality, we can assume $\delta < 1$. Choosing $\eta = \frac{2C_2}{|\log(\delta)|}$ then gives

$$\|\Phi_t - \tilde{\Phi}_t\|_{L^1(B_1)} \leq C_5 |\log(\delta)|^{-1} + C_3 \delta^{1/2} \leq C_6 |\log(\delta)|^{-1}, \tag{2.285}$$

which is the desired result. □

3 Hahn-Banach theorem

3.1 Preliminaries on hyperplanes and linear functional

All the results in this section have a proof which does not use the Hahn-Banach theorem.

3.1.1 Riesz' lemma

Recall that, in a metric space (E, d), the distance $d(x, A)$ from a point $u \in E$ to a set $A \subset E$ is defined as

$$d(u, A) = \inf_{v \in A} d(u, v). \tag{3.1}$$

In particular $d(u, A) = 0$ if, and only if, $u \in \bar{A}$.

Proposition 3.1 (F. Riesz' lemma)**.** *Let E be a normed vector space and M a closed subspace of E, $M \neq E$. Then, for all $\varepsilon > 0$, there exists $u \in E$ of norm $\|u\| = 1$ such that $d(u, M) \geq 1 - \varepsilon$.*

Proof of Proposition 3.1. let d be the metric associated to the norm on E. Let $w \in E \setminus M$. We have $\delta := d(w, M) > 0$ and, given δ' with $\delta' > \delta$, there exists $v' \in M$ such that $0 < d(w, v') \leq \delta'$. Let $u_1 = w - v'$. Then

$$\|u_1 - v\| = d(w, v' + v) \geq \delta \tag{3.2}$$

for all $v \in M$, so $d(u_1, M) \geq \delta$. The point $u = u_1/\|u_1\|$ is of norm 1 and

$$d(u, M) \geq \delta/\|u_1\| \geq \delta/\delta'. \tag{3.3}$$

Choosing $\delta' = (1 - \varepsilon)^{-1}\delta$ will give the result. □

3.1.2 Quotient spaces

Let X be a vector space, and Y a subspace of E. We denote by X/Y the quotient space relatively to equivalence relation "$u \sim v$ if $u - v \in Y$" and also denote by $\pi\colon X \to X/Y$ the canonical surjection.

Theorem 3.2 (Isomorphism theorem)**.** *Let $T\colon X \to Z$ be a linear map between the vector spaces X and Z. Then T induces an isomorphism*

$$X/\mathrm{Ker(T)} \to \mathrm{Im}(T), \quad u + \mathrm{Ker(T)} \mapsto T(u). \tag{3.4}$$

If p is a semi-norm on a vector space X, then $Y = \{u \in E; p(u) = 0\}$ is a subspace of X. On X/Y, $\|\pi(u)\| := p(u)$ is well defined and is a norm. For instance, L^1 is obtained as such a quotient space for the semi-norm given by the integral. In the following proposition, we consider from the start a normed vector space.

Proposition 3.3 (Norm on the quotient space)**.** *Let E be a normed vector space, and M a* closed *subspace of E. Consider the map*

$$\pi(u) \mapsto \|\pi(u)\| := d(u, M) \tag{3.5}$$

on E/M. It defines a norm on E/M such that $\pi\colon E \to E/M$ has operator norm $\|\pi\| = 1$. The topology defined by the norm and the quotient topology coincide. If E is a Banach space, then E/M also.

Proof of Proposition 3.3. It is clear that (3.5) is well defined: it depends on the equivalence class of u only. It inherits the homogeneity property $\|\lambda\pi(u)\| = |\lambda|\|\pi(u)\|$ from the homogeneity property

$$\|\lambda u - v\| = |\lambda|\|u - \lambda^{-1}v\|, \lambda \neq 0.$$

The separation axiom $\|\pi(u)\| = 0 \Rightarrow \pi(u) = 0$ is satisfied because M is closed. The subadditivity property is also deduced from the subadditivity property for the original norm on E: given $\varepsilon > 0$, and $u, v \in E$, there exists $w_u, w_v \in M$ such that

$$\|u - w_u\| \leq \|\pi(u)\| + \varepsilon, \quad \|v - w_v\| \leq \|\pi(v)\| + \varepsilon.$$

Then

$$\|\pi(u+v)\| \leq \|(u+v) - (w_u + w_v)\| \leq \|u - w_u\| + \|v - w_v\| \leq \|\pi(u)\| + \|\pi(v)\| + 2\varepsilon.$$

The result follows by taking $\varepsilon \to 0$. Since $0 \in M$, we have $\|\pi(u)\| \leq \|u\|$ and Riesz' Lemma shows that $\|\pi\| = 1$. If $r > 0$, then

$$\pi^{-1}(B(\pi(u), r)) = B(u, r) + M = \bigcup_{v \in M} B(u, r) + \{v\} = \bigcup_{v \in M} B(u + v, r)$$

is open, so $\pi\colon (E, \|\cdot\|) \to (E/M, \|\cdot\|)$ is continuous. Since the quotient topology $\mathcal{T}$ is the finest topology that makes π continuous, it contains the topology $\mathcal{T}'$ associated to the norm on E/F. Conversely, let U be open for the quotient topology and let $\pi(u) \in U$. Since $u \in \pi^{-1}(U)$ open, there is a ball $B(u, r)$ of positive radius $r > 0$ included in $\pi^{-1}(U)$. Then $B(\pi(u), r)$ is included in U. Indeed, $\pi(v) \in B(\pi(u), r)$ is equivalent to the existence of $z \in M$ such that $\|u - v - z\| < r$, which means $v + z \in B(u, r)$, and implies $v \in \pi^{-1}(U) + M = \pi^{-1}(U)$. We have proved that, for each $\pi(u) \in U$, there is a non-trivial ball $B(\pi(u), r) \subset U$: U is $\mathcal{T}'$-neighbourhood of each of its points, so $U \in \mathcal{T}'$. Let now $(\pi(u_n))$ be a Cauchy sequence in E/M. From this Cauchy sequence we can extract a "C-sequence" $(\pi(u_{n_k}))$, *i.e.* a subsequence such that

$$\|\pi(u_{n_k}) - \pi(u_{n_{k+1}})\| < 2^{-k}, \tag{3.6}$$

for all k. This is equivalent to the existence of a sequence $(w_{k+1/2})$ of elements of M such that

$$\|u_{n_k} - u_{n_{k+1}} - w_{k+1/2}\| < 2^{-k}, \tag{3.7}$$

for all k. Define $v_0 = 0$, $v_1 = w_{1/2}, \ldots, v_k = v_{k-1} + w_{k-1/2}, \ldots$ Then

$$\|\hat{u}_k - \hat{u}_{k+1}\| < 2^{-k}, \quad \hat{u}_k := u_{n_k} - v_k, \tag{3.8}$$

so, assuming that E is complete, the sequence $(\hat{u}_k)$ is converging to an element $\hat{u} \in E$. Since π is continuous, $\pi(\hat{u}_k) = \pi(u_{n_k})$ is converging to $\pi(\hat{u})$, but then the whole Cauchy sequence $(\pi(u_n))$ is converging to $\pi(\hat{u})$: we can conclude that E/M is complete. □

3.1.3 Hyperplanes

Proposition 3.4 (Hyperplanes)**.** *Let E be a normed vector space and let $H \subset E$. The following assertions are equivalent.*

1. *H is the kernel of a non-trivial linear map $\varphi\colon E \to \mathbb{R}$,*
2. *the space E/H has dimension 1,*

3. $H \neq E$, and for every space M such that $H \subset M \subset E$, either $M = H$, or $M = E$.

If any of this assertion is realized, then $E = H \oplus \langle v \rangle$ for a certain $v \in E$, and we say that H is an hyperplane *of E.*

Proof of Proposition 3.4. We have $1 \Rightarrow 2$ by the Isomorphism Theorem (Theorem 3.2). The property 2 implies $E = H \oplus \langle v \rangle$ for a $v \in E$. Indeed, $E/H = \langle \pi(v) \rangle$ for a $v \in E$, which is equivalent to $E = H \oplus \langle v \rangle$. Setting $\varphi(u + \lambda v) = \lambda$ for $u \in H$ then shows that $2 \Rightarrow 1$. If 2 is realized and $H \subset M \subset E$, then M/H is a subspace of E/H of dimension 0 or 1. If the dimension is 0, then $M = H$. If the dimension is 1, then $M/H = E/H$, *i.e.* every element v in E is of the form $u + w$ with $u \in H$ and $w \in M$. Since $H \subset M$, we obtain $E = M$. We can generalize the reasoning used above: if E/H has dimension n and $(\pi(v_i))_{1,n}$ is a basis of E/H, then

$$E = H \oplus \langle v_1, \dots, v_n \rangle. \tag{3.9}$$

This shows that non-2 implies non-3 (consider $M = H \oplus \langle v_1 \rangle$ if $n > 1$) and concludes the proof. □

Theorem 3.5 (Closed hyperplane)**.** *Let E be a real normed vector space. Then $H \subset E$ is a* closed *hyperplane of E if, and only if, it is the kernel of a* continuous *linear form on E. An hyperplane which is not closed is dense.*

Proof of Theorem 3.5. We use the characterization 1 of Proposition 3.4 to write H as the kernel of a linear form $\varphi \colon E \to \mathbb{R}$. If φ is actually continuous, then H is closed. If H is closed, then we can consider the induced linear map

$$\varphi^\sharp \colon E/H \to \mathbb{R}, \quad u + H \mapsto \varphi(u). \tag{3.10}$$

Since E/H has finite dimension, $\varphi^\sharp$ is continuous. We endow E/H with the quotient norm (3.5) and use the fact that the continuity of $\varphi^\sharp$ is equivalent to the fact that it is bounded: there exists $C \geq 0$ such that

$$|\varphi^\sharp(\pi(u))| \leq C\|\pi(u)\|, \tag{3.11}$$

for all $u \in E$. Then, for all $u \in E$,

$$|\varphi(u)| = |\varphi^\sharp(\pi(u))| \leq C\|\pi(u)\| \leq C\|u\|, \tag{3.12}$$

since $\|\pi\| = 1$, and φ is continuous. If H is an hyperplane which is not dense, then there is an open ball $B(u_0, r)$ which doest not intersect H. Assume that $H = \{\varphi = 0\}$, where φ is a linear functional and assume, without loss of generality, that $\varphi(x_0) > 0$. Then $\varphi \geq 0$ on $B(u_0, r)$. Indeed, suppose by contradiction that there is a $u_1 \in B(u_0, r)$ such that $\varphi(u_1) < 0$. Then φ will have a zero on the segment $[u_0, u_1]$: more precisely,

$$\varphi(tu_0 + (1-t)u_1) = 0 \text{ for } t = \frac{-\varphi(u_1)}{\varphi(u_0) - \varphi(u_1)} \in (0, 1), \tag{3.13}$$

and this contradicts $H \cap B(u_0, r) = \emptyset$. That $\varphi \geq 0$ on $B(u_0, r)$ gives

$$0 < \varphi(u_0) + r\varphi(v), \quad \forall v \in B(0, 1). \tag{3.14}$$

Since $B(0,1)$ is symmetric $-B(0,1) = B(0,1)$, it follows that φ is bounded, with $\|\varphi\| \leq r^{-1}\varphi(u_0)$. We deduce that φ is continuous and H closed. Note that, if $H = \mathrm{Ker}(\varphi)$ is closed, then the complementary set is open so, again, we can conclude that there is an open ball that does not intersect H and then use the arguments above to show that φ is continuous, and give an alternative proof of the first assertion of the theorem. □

3.2 Hahn-Banach theorem

3.2.1 Analytic form of the Hahn-Banach theorem: extension of linear functionals

Definition 3.1. Let E be a real vector space over $\mathbb{R}$. An application $p\colon E \to \mathbb{R}$ is said to be

- *positively homogeneous* if $p(tu) = tp(u)$ for all $t > 0$,
- *subadditive* if $p(u+v) \leq p(u) + p(v)$, for all $u, v \in E$.

For example, a norm is a positively homogeneous and subadditive function. If $(p_\alpha)_{\alpha\in A}$ is a collection of semi-norms on a real vector space E, then a sum $\sum_{\alpha\in A_0} p_\alpha$, where A_0 is finite, is a positively homogeneous and subadditive function.

Theorem 3.6 (Hahn-Banach - Analytic version)**.** *Let E be a real vector space, let p be a positively homogeneous and subadditive function on E. Let M be a linear subspace of E and $\varphi\colon M \to \mathbb{R}$ a linear functional which is dominated by p:*

$$\varphi(u) \leq p(u), \quad \forall u \in M. \tag{3.15}$$

Then φ can be extended to a linear functional $E \to \mathbb{R}$ which remains dominated by p.

Proof of Theorem 3.6. We can assume $M \neq E$. Let us first show that we can extend φ to $M' = M \oplus \langle w\rangle$ if $w \notin M$. Denote by φ' such an extension, assuming that it exists. By decomposing each $v \in M'$ as $u + \lambda w$ where $u \in M$, we should have

$$\varphi(u) + \lambda\varphi'(w) \leq p(u + \lambda w). \tag{3.16}$$

If $\lambda = 0$, then (3.16) is satisfied by hypothesis, whatever the choice of the value $\varphi'(w)$. If $\lambda \neq 0$, we can as well replace u by λu in (3.16). Then, using the positive homogeneity of p, it is sufficient to satisfy the cases $\lambda = \pm 1$, *i.e.*

$$\varphi'(w) \leq p(u+w) - \varphi(u), \quad -\varphi(u) - p(-u-w) \leq \varphi'(w) \tag{3.17}$$

We can choose a convenient value $\varphi'(w)$, provided

$$-\varphi(u) - p(-u-w) \leq p(v+w) - \varphi(v), \quad \forall u, v \in M, \tag{3.18}$$

but (3.18) is satisfied since

$$\varphi(v) - \varphi(u) = \varphi(v-u) \leq p(v-u) \tag{3.19}$$

by (3.18), and $p(v-u) \leq p(v+w) + p(-u-w)$ by subadditivy. This gives a satisfactory extension to $M \oplus \langle w\rangle$. To justify that we can go on this process and choose new values of the functional until we exhaust the whole space E, we need the Axiom of Choice. We use the equivalent form of Zorn's Lemma. Consider the following order on extensions of (φ, M) being dominated by p: we say that $(\varphi', M') \leq (\varphi'', M'')$ if we have inclusion of the graphs: $M' \subset M''$ and φ'' in restriction to M' is equal to φ'. If $\mathcal{F} = \{(\varphi_\alpha, M_\alpha); \alpha \in A\}$ is a totally ordered family, then $(\bar\varphi, \bar M)$ defined by $\bar M = \cup_{\alpha\in A} M_\alpha$, $\bar\varphi = \varphi_\alpha$ on M_α is an upper bound for $\mathcal{F}$. By Zorn's Lemma, the set of extensions of (φ, M) being dominated by p admits a maximal element (φ^*, M^*). Then $M^* = E$, otherwise we can repeat the construction done in the beginning of the proof to produce a contradiction. □

Theorem 3.7 (Hahn-Banach - Analytic version - Complex case)**.** *Let E be a vector space over $\mathbb{C}$, let p be a semi-norm on E. Let M be a linear subspace of E and $\varphi\colon M \to \mathbb{C}$ a complex linear functional which is dominated by p in the following sense*

$$|\varphi(u)| \leq p(u), \forall u \in M. \tag{3.20}$$

Then φ can be extended to a complex linear functional $E \to \mathbb{C}$ which remains dominated by p.

Proof of Theorem 3.7. Sketch of the proof: note that a complex linear functional ψ on a subspace N of E is of the form

$$\psi(u) = \theta(u) - i\theta(iu), \tag{3.21}$$

where $\theta\colon N \to \mathbb{R}$ can be seen as a real linear functional and that, conversely, if $\theta\colon N \to \mathbb{R}$ is a real linear functional, then ψ defined by (3.21) is a complex linear functional on N. Use Theorem 3.6 to conclude. □

We now apply the Hahn-Banach theorem in a normed vector space, to deduce a series of fundamental results. We denote by E^* the set of continuous linear functionals on E.

Theorem 3.8 (Corollary of the Hahn-Banach Theorem)**.** *Let E be a normed vector space, real or complex. We have the following statements*

1. *if M is a closed subspace of E and $v \in E \setminus M$, then there exists $\varphi \in E^*$ such that $\varphi|_M = 0$, $\|\varphi\|_{E^*} = 1$, $\varphi(v) = d(v, M) > 0$.*
2. *If $u \neq 0 \in E$, then there exists $\varphi \in E^*$ such that $\|\varphi\|_{E^*} = 1$ and $\varphi(u) = \|u\|_E$.*
3. *The space E^* separates points on E.*
4. *Define the map $J\colon E \to E^{**}$ (where E^{**} denotes the topological dual of E^*) by $Ju(\varphi) = \varphi(u)$. Then J is a linear isometry of E into E^{**}.*

Proof of Theorem 3.8. To prove *1.*, we define φ on $M \oplus \langle v\rangle$ by $\varphi(u + \lambda v) = \lambda\delta$, $\delta := d(v, M)$. Then, assuming $\lambda \neq 0$, we have

$$|\varphi(u + \lambda v)| = |\lambda|\delta \leq |\lambda| \|\lambda^{-1}u + v\|_E = \|u + \lambda v\|_E, \tag{3.22}$$

for all $u \in M$. The bound $|\varphi(u + \lambda v)| \leq \|u + \lambda v\|_E$ remains true when $\lambda = 0$ so we can apply the Hahn-Banach theorem with $p(u) = \|u\|_E$ to conclude that there exists $\varphi \in E^*$ such that $\|\varphi\|_{E^*} \leq 1$ and $\varphi(v) = \delta$. Let $\varepsilon > 0$ and let $u \in M$ be such that $\|u - v\|_E < \delta + \varepsilon$. Then

$$|\varphi(u - v)| = \delta \geq \|u - v\|_E - \varepsilon. \tag{3.23}$$

Since $\|u - v\|_E \geq \delta$, (3.23) gives

$$\|\varphi\|_{E^*} \geq 1 - \frac{\varepsilon}{\|u - v\|_E} \geq 1 - \frac{\varepsilon}{\delta},$$

from which we deduce that $\|\varphi\|_{E^*} = 1$. The statement *2.* follows from *1.* with $M = \{0\}$ and we deduce *3.* from *2.* by considering the vector $u - v$ if $u \neq v$. The map J defined in *4.* is clearly linear. For all $\varphi \in E^*$, we have $|Ju(\varphi)| = |\varphi(u)| \leq \|u\|_E \|\varphi\|_{E^*}$ so $\|Ju\|_{E^{**}} \leq \|u\|_E$. Using *2.* we obtain $\|Ju\|_{E^{**}} = \|u\|_E$. □

Exercise 3.2. Use *2.* in Theorem 3.8, instead of (2.63), to simplify slightly the proof of the lemma 2.10.

In practice, we can very often avoid the use of the general version of the Hahn-Banach theorem. The following statement is sufficient in many cases. It is a countable feature that makes the use of the axiom of choice irrelevant. We give a version on real vector spaces, but, of course, the complex version of Theorem 3.7 can be adapted as well.

Theorem 3.9 (Hahn-Banach Theorem without the axiom of choice)**.** *Let E be a real vector space, let p be a positively homogeneous and subadditive function on E. Let M be a linear subspace of E and $\varphi\colon M \to \mathbb{R}$ a linear functional which is dominated by p. Assume*

- *either E has a countable basis,*
- *or E is a separable Banach space and p is a norm.*

Then φ can be extended to a linear functional $E \to \mathbb{R}$ which remains dominated by p.

Proof of Theorem 3.9. In exercises class. □

3.2.2 Geometric form of the Hahn-Banach theorem: separation

The Minkowski gauge (see Proposition-Definition 3.11 below) is the central tool to deduce the geometric form of the Hahn-Banach theorem from the analytic one. Minkowski's gauges are also relevant to the study of Fréchet spaces (done in Section 4), which we also prepare here.

Definition 3.3 (Topological vector space)**.** A vector space X on $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ is said to be a *topological vector space* if it is endowed with a topology such that:

1. the linear operations
$$X \times X \to X, \quad (u, v) \mapsto u + v \quad \text{and} \quad \mathbb{K} \times X \to X, \quad (\lambda, u) \mapsto \lambda u \tag{3.24}$$
are continuous,
2. all sets $\{u\}$ reduced to a point $u \in X$ are closed.

Remark 3.1 (Second condition of Definition 3.3). Using the two conditions *1.* and *2.*, one can show that a topological vector space is Hausdorff, [Rud73, p.10].
If X is a vector space and $A, B \subset X$, $\lambda \in \mathbb{K}$, we use the notations
$$A + B = \{u + v; u \in A, v \in B\}, \quad \lambda A = \{\lambda u; u \in A\}. \tag{3.25}$$

Definition 3.4. A subset A of a vector space X is said to be

1. *symmetric* if $-A = A$,
2. *convex* if $tA + (1-t)A \subset A$ for all $t \in [0, 1]$.

A topological vector space X is said to be *locally convex* if each point has a neighbourhood base consisting of convex neighbourhoods.

Proposition 3.10 (Base of convex symmetric neighbourhoods)**.** *Let X be a locally convex topological vector space. Then the origin has a neighbourhood base consisting of convex symmetric neighbourhoods.*

Proof of Proposition 3.10. Let V be a convex neighbourhood of 0. Then $-V$ is also convex. A finite intersection of open sets is open, idem for convex sets, so $U = V \cap (-V)$ is a symmetric convex neighbourhood of the origin contained in V. This construction gives a neighbourhood base of the origin consisting of convex symmetric neighbourhoods. □

Proposition 3.10 will be used later in the proof of Theorem 4.1, but the interest in conciliating the properties of convexity and symmetry already appears in the following result.

Proposition-Definition 3.11 (Minkowski's gauge of a convex set)**.** *Let C be a convex neighbourhood of 0 in a real topological vector space X. The (Minkowski) gauge p of C is defined as follows:*

$$p_C(u) = \inf\left\{t > 0; \frac{u}{t} \in C\right\}, \tag{3.26}$$

for all u in X. Then p_C is positively homogeneous, subadditive, and we have

$$u \in C \Rightarrow p_C(u) \leq 1, \quad p_C(u) < 1 \Rightarrow u \in C \tag{3.27}$$

for all $u \in X$. If C is open, then

$$p_C(u) < 1 \iff u \in C. \tag{3.28}$$

We also have the following properties:

1. *if, additionally, C is symmetric, then p_C is a continuous semi-norm.*
2. *if X is a normed vector space, then there exists $K \geq 0$ such that $p_A(x) \leq K\|x\|_X$ for all $x \in X$.*

Proof of Proposition-Definition 3.11. First note that $\lambda \mapsto \lambda u$ is continuous from $[0, 1]$ to X. Since C is a neighbourhood of 0, we have $\lambda u \in C$ for $\lambda \in [0, \lambda_u]$ where $\lambda_u > 0$. Then $p_C(u) \leq \lambda_u^{-1} < +\infty$. If $r > 0$, and $u \in X$, then

$$p_C(ru) = \inf\left\{t > 0; \frac{ru}{t} \in C\right\} = r \inf\left\{t > 0; \frac{u}{t} \in C\right\} = rp_C(u).$$

If $u, v \in X$ and $t_u, t_v > 0$ are such that $u \in t_uC$, $v \in t_vC$, then, setting $\theta = \frac{t_u}{t_u+t_v}$, we have

$$u + v \in t_uC + t_vC = (t_u + t_v)\left[\theta C + (1 - \theta)C\right] \subset (t_u + t_v)C,$$

so $p_C(u + v) \leq t_u + t_v$. This implies $p_C(u + v) \leq p_C(u) + p_C(v)$. Obviously, $u \in C$ implies $p_C(u) \leq 1$. If $\lambda u \in C$ for a $\lambda \geq 1$ then, since $0 \in C$ as well, we have $u \in C$. So $p_C(u) < 1$ implies $u \in C$. If C is open and $u \in C$, then $\theta u \in C$ for a certain $\theta > 1$ (we use the continuity of $\theta \mapsto \theta u$ at $\theta = 1$). Then $p_C(\theta u) = \theta p_C(u) \leq 1$ by (3.27), so $p_C(u) < 1$. Let us come back to the case where C is a neighbourhood of 0 and assume now that it is also symmetric. Then p_C is symmetric as well, so it is a semi-norm. Let $u \in X$ and $\varepsilon > 0$. We want to find a neighbourhood $V_\varepsilon(u)$ of u such that $|p_C(u) - p_C(v)| \leq \varepsilon$ for all $v \in V_\varepsilon(u)$. By subadditivity and symmetry, we have

$$|p_C(u) - p_C(v)| \leq p_C(u - v)$$

for all $v \in X$, so it is sufficient to prove the continuity at $u = 0$, and then take $V_\varepsilon(u) = u + V_\varepsilon(0)$. At $u = 0$, we have $p_C(u) = 0$ and $p_C(v) \leq \varepsilon$ if $v \in \varepsilon C$, so considering $V_\varepsilon(0) = \varepsilon C$, which is a neighbourhood of 0, we obtain the continuity of p_C. It is a general fact that $p_C \leq p_{C'}$ if $C' \subset C$, where C' is a convex neighbourhood of 0. If X is a normed space and $\varepsilon > 0$ is such that $C' := B(0, \varepsilon) \subset C$, then

$$p_C(x) \leq p_{C'}(x) = \frac{1}{\varepsilon}\|x\|_X,$$

for all $x \in X$. □

Notation: let X be a real topological vector space. If $H = \mathrm{Ker}(\varphi)$ is an hyperplane, we denote by H^+ and H^- the half-spaces

$$H^+ = \{\varphi \geq 0\}, \quad H^- = \{\varphi \leq 0\}.$$

We also denote by H^{++} and H^{--} the strict half-spaces

$$H^{++} = H^+ \setminus H = \{\varphi > 0\}, \quad H^{--} = \{\varphi < 0\}.$$

An affine hyperplane is the translate $w + H$ of a (vectorial) hyperplane by a vector $w \in E$. The associated half-spaces are

$$H_w^+ = w + H^+, \quad H_w^- = w + H^-,$$

and similarly for H_w^{++}, H_w^{--}.

Definition 3.5 (Separation by affine hyperplanes)**.** Given $A, B \subset X$ and an affine hyperplane $H_w = w + H$ of E, we say that

1. H_w separates A and B if $A \subset H_w^{--}$, $B \subset H_w^+$,
2. H_w strictly separates A and B if there exists $w' \in H^{++}$ such that $A \subset H_w^-$, $B \subset w' + H_w^+$.

Remark 3.2 (Symmetric separation). If H_w separates A and B, then we have in particular

$$A \subset H_w^-, \quad B \subset H_w^+, \quad (A - B) \cap H = \emptyset. \tag{3.29}$$

Let $\psi \colon X \to \mathbb{R}$ be such that $H = \mathrm{Ker}(\psi)$. Note that, up to the substitution of $-\psi$ to ψ, the statement (3.29) is symmetric in A and B, so we can as well conclude that

$$A \subset H_w^+, \quad B \subset H_w^-, \quad (A - B) \cap H = \emptyset, \tag{3.30}$$

where H_w is an affine hyperplane.

Remark 3.3 (Open linear functional). Let X be a topological vector space. *A continuous non-trivial linear functional* $\varphi \colon X \to \mathbb{R}$ *is open.* Indeed, let $u \in X$, and let V be a neighbourhood of the origin. There is some $\varepsilon > 0$ and some open neighbourhood W of the origin such that $tv \in u + V$ for all $v \in u + W$ and all t with $|t - 1| < \varepsilon$ (we use the continuity of $(t, v) \mapsto tv$ at $(1, u)$). If $\lambda = \varphi(u) > 0$, this will give

$$(\lambda(1 - \varepsilon), \lambda(1 + \varepsilon)) \subset \varphi(u + V).$$

If $\lambda < 0$, we deduce similarly that there is a neighbourhood I of λ in $\mathbb{R}$ such that $I \subset \varphi(u+V)$. If $\lambda = \varphi(u) = 0$, we consider a $w \in X$ such that $\varphi(w) \neq 0$. We know that there is a neighbourhood I of $\varphi(w)$ in $\mathbb{R}$ such that $I \subset \varphi(w + V)$. Then $J := I - \varphi(w)$ is a neighbourhood of 0 in $\mathbb{R}$ such that $J \subset \varphi(u + V)$.

Theorem 3.12 (Separation by hyperplanes)**.** *Let A, B be two non-empty convex subsets of a real normed vector space E such that $A \cap B = \emptyset$. If*

1. *A is open and $B = \{v\}$ is reduced to a point, then there is a closed affine hyperplane that separates A and B,*
2. *A is open, then there exists a closed affine hyperplane that separates A and B,*
3. *A is closed and $B = \{v\}$ is reduced to a point, then there exists a closed affine hyperplane that strictly separates A and B,*

4. *A is closed and B is compact, then there exists a closed affine hyperplane that strictly separates A and B.*

Remark 3.4 (Generalization). Theorem 3.12 is true in the more general context of locally convex, real or complex, topological vector spaces, [Rud73, Theorem 3.4, p.58].

Proof of Theorem 3.12. It is clear that *3.* is a particular case of *4.* and that *1.* is a particular case of *2.*, but it is worth emphasizing these cases. Assume that A is open, convex and non-empty and $v \notin A$. Up to a translation, we can assume that $0 \in A$. Define the linear form ψ on $\langle v \rangle$ by $\psi(\lambda v) = \lambda$ and consider the Minkowski gauge p_A of A. If $\lambda < 0$, then $\psi(\lambda v) < 0 \leq p_A(\lambda v)$. If $\lambda \geq 0$, we use the fact that $p_A(v) \geq 1$ since $v \notin A$ (by (3.27)), hence

$$\psi(\lambda v) = \lambda \leq \lambda p_A(v) = p_A(\lambda v).$$

We use the Hahn Banach Theorem to extend ψ to a linear functional φ defined on E and dominated by p_A. The point *2.* of Proposition-Definition 3.11 ensures that φ is continuous. Since A is open, we have $p_A(u) < 1$ if $u \in A$, so $\varphi(u) \leq p_A(u) < 1 = \varphi(v)$ and *1.* follows. Let us prove *2.* now. Let

$$A' = A - B = \bigcup_{v \in B} A - \{v\}.$$

Then A' is open, convex, non-empty and $v' := 0 \notin A'$. By *1.* applied to A' and $v' = 0$, we have $A - B \subset H_w^{--}$, $0 \in H_w^+$, where $H = \mathrm{Ker}(\varphi)$, $\varphi \in X^*$. This implies $\varphi(u) < \varphi(v)$ for all $u \in A$, $v \in V$. So $\varphi(A)$ and $\varphi(B)$ are two disjoints intervals (since convex) of $\mathbb{R}$. Since $\varphi(A)$ is open by Remark 3.3, the infimum λ of $\varphi(B)$ satisfies

$$\varphi(u) < \lambda \leq \varphi(v), \quad \forall u \in A, v \in B. \tag{3.31}$$

There exists $w \in X$ such that $\varphi(w) = \lambda$ (φ is surjective since non-trivial). We have then $A \subset H_w^{--}$ and $B \subset H_w^+$. Let us prove *3.* We have $v \in E \setminus A$, which is open. Let $r > 0$ such that $B(v,r) \subset E \setminus A$. By *2.*, there exists a closed hyperplane $H_w = H + w$ which separates A and $B(r,v)$. Using Remark 3.2, we have

$$A \subset H_w^-, \quad B(v,r) \subset H_w^+, \quad (A - B(v,r)) \cap H = \emptyset.$$

Let u_0 be a point in A and let $w_0' = v - u_0$. Then $w_0' \in H^{++}$ and $w' := \frac{r}{2}\frac{w_0'}{\|w_0'\|}$ is also in H^{++}. We have $v - w' \in B(v,r) \subset H_w^+$, so $v \in w' + H_w^+$. This is the desired result. The proof of *4.* is left as an exercise. □

3.2.3 Applications of the Hahn-Banach theorem in finite dimension

All this section is taken from this post on Terence Tao's blog.

Theorem 3.13 (Farkas' Lemma)**.** *Let $P_1, \dots, P_d$ be some affine functional on $\mathbb{R}^d$. Then only one of the following statements is satisfied:*

1. *there exists $x \in \mathbb{R}^d$ such that $P_1(x) \geq 0, \dots, P_n(x) \geq 0$,*
2. *there exists $q_1 \geq 0, \dots, q_n \geq 0$ such that $q_1 P_1(x) + \dots + q_n P_n(x) = -1$ for all $x \in \mathbb{R}^d$.*

Proof of Theorem 3.13. It is clear that *1.* implies non-*2.* Suppose now that *1.* is not satisfied. Each affine functional P_i can be written $P_i(x) = z_i \cdot x - b_i$, where $z_i \in \mathbb{R}^d$, $b_i \in \mathbb{R}$. Let

$M \in \mathcal{M}_{d,n}(\mathbb{R})$ be the matrix with $z_1, \ldots, z_n$ on its lines and let b be the vector in $\mathbb{R}^n$ with components b_i. Then *1.* means that $b \in A$, where A is the closed convex set

$$A = \left\{ y \in \mathbb{R}^n ; \exists x \in \mathbb{R}^d, y \leq Mx \right\},$$

where $y \leq z$ means $y_i \leq z_i$ for all i if $y, z \in \mathbb{R}^n$. The statement *2.* means that there is a $q \in \mathbb{R}^n$ with non-negative components such that $q \cdot (Mx - b) = -1$ for all $x \in \mathbb{R}^d$. This is equivalent to

$$M^* q = 0, \quad q \cdot b = 1,$$

so up to a step of rescaling, *2.* is equivalent to the existence of a vector $q \in \mathbb{R}^n$ such that

$$\forall i,\ q_i \geq 0, \quad M^* q = 0, \quad q \cdot b > 0. \tag{3.32}$$

If $b \notin A$, then, by Theorem 3.12, *3.*, there is an affine hyperplane that separates strictly A and b: there exists $q \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ such that

$$q \cdot b > \lambda, \quad q \cdot y \leq \lambda, \quad \forall y \in A. \tag{3.33}$$

Let us prove that q satisfies (3.32). Since $0 \in A$, $\lambda \geq 0$, so $q \cdot b > 0$. Since $Mx \in A$, we have $M^* q \cdot x = q \cdot Mx \leq \lambda$ for all $x \in \mathbb{R}^d$, which is equivalent (as one can check) to $M^* q = 0$. Since $Mx - z \in A$ if z has non negative components, we have, for all such z,

$$q \cdot (Mx - z) = -q \cdot z \leq \lambda,$$

and this is equivalent to the fact that q has non-negative components. □

We give an application of Farkas' lemma to the proof of the minimax theorem for zero-sum games. The framework is the following one: player one (Alice) has the choice between n moves, player two (Bob) has the choice between m moves. To a conjoint choice (i, j) of moves is associated a cost $c_{i,j}$: Alice gets $c_{i,j}$ (so Alice wins $c_{i,j}$ if $c_{i,j} \geq 0$ and loses an amount of $|c_{i,j}|$ if $c_{i,j} < 0$). Bob gets $-c_{i,j}$. *Example:* the paper-rock-scissors games, where $n = m = 3$ and the expected table of costs is

	paper	rock	scissors
paper	$(0,0)$	$(1,-1)$	$(-1,1)$
rock	$(-1,1)$	$(0,0)$	$(1,-1)$
scissors	$(1,-1)$	$(-1,1)$	$(0,0)$

$$C = (c_{i,j}) = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}.$$

Here is a modelling postulate: a *strategy* is a random variable I on the set of moves, or, equivalently, a probability law on the set of moves. So a strategy p for Alice is a vector $p \in \mathbb{R}^n$ with non-negative entries such that the components p_i add up to 1. The corresponding random variable is I with law $\mathbb{P}(I = i) = p_i$. We have similar considerations for Bob, with notations q and J. The *expected pay-off* for Alice is then

$$F(p, q) = \mathbb{E}\left[c_{I,J}\right] = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{i,j} p_i q_j = Cp \cdot q. \tag{3.34}$$

Theorem 3.14 (Minimax theorem)**.** *There exists some strategies p_* and q_* such that*

1. *p_* is optimal for Alice, in the sense that Alice can expect to win at least $F(p_*, q_*)$: for all strategy q of Bob, $F(p_*, q_*) \leq F(p_*, q)$*

2. q_* *is optimal for Bob, in the sense Bob can expect to lose at most* $F(p_*, q_*)$*: for all strategy* p *of Alice,* $F(p_*, q_*) \geq F(p, q_*)$.

We can also say that (p_*, q_*) is a Nash equilibrium in Alice and Bob's zero-sum game: if Bob chooses the strategy q_*, Alice cannot do better than choosing p_* as a strategy (we use 2.) and, if Alice chooses the strategy p_*, then Bob cannot do better than choosing q_* as a strategy.

Proof of Theorem 3.14. Our aim is to prove that

$$\max_p \min_q F(p, q) = \min_q \max_p F(p, q). \tag{3.35}$$

Then the argument (p_*, q_*) of the common value $F(p_*, q_*)$ in (3.35) will give us the desired strategies. Typically, this "minimax" result (3.35) is valid in the saddle configuration where $F(p, \cdot)$ is convex for every p and $F(\cdot, q)$ is concave for every q. We have $F(p, q) \leq \max_p F(p, q)$ for all p, q, taking the min with respect to q and then the max with respect to p shows that

$$\max_p \min_q F(p, q) \leq \min_q \max_p F(p, q). \tag{3.36}$$

Assume by contradiction that

$$\max_p \min_q F(p, q) < \gamma < \min_q \max_p F(p, q), \tag{3.37}$$

for a given $\gamma \in \mathbb{R}$. Then the system

$$p_1, \dots, p_n \geq 0, \quad p_1 + \dots + p_n = 1, \quad F(p, q) \geq \gamma, \ \forall q, \tag{3.38}$$

has no solution. By convexity of $q \mapsto F(p, q)$, this is equivalent to the fact that the system

$$p_1, \dots, p_n \geq 0, \quad p_1 + \dots + p_n = 1, \quad (Cp)_j = \sum_{i=1}^{n} c_{i,j} p_i \geq \gamma, \ \forall j \in \{1, \dots, m\}, \tag{3.39}$$

has no solution. By linearity of $p \mapsto Cp$, we can replace the equation in (3.39) by an inequality and claim that the system

$$p_1, \dots, p_n \geq 0, \quad p_1 + \dots + p_n \leq 1, \quad (Cp)_j = \sum_{i=1}^{n} c_{i,j} p_i \geq \gamma, \ \forall j \in \{1, \dots, m\}, \tag{3.40}$$

has no solution. Indeed, if (3.40) has a solution p, then this solution p is non trivial. Otherwise $\gamma \leq 0$, but (using the first inequality in (3.37)) we have $\gamma > \min_q F(0, q) = 0$, and obtain a contradiction. So $\theta := p_1 + \dots + p_n \in (0, 1]$. Replacing p by $\theta^{-1} p$ gives a probability distribution solution to (3.39). Considering (3.40) now, we can apply Farkas' Lemma: there are some vectors $\zeta \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$, $\xi \in \mathbb{R}^m$ with non-negative components such that

$$-1 = p \cdot \zeta + \alpha(1 - p \cdot \mathbf{1}) + (Cp - \gamma \mathbf{1}) \cdot \xi, \quad \forall p \in \mathbb{R}^n, \tag{3.41}$$

where $\mathbf{1}$ is the vector of $\mathbb{R}^n$ or $\mathbb{R}^m$ with all components equal to 1. If p is a probability distribution, then the first two terms in the right-hand side of (3.41) are non-negative, so

$$F(p, \xi) = Cp \cdot \xi \leq -1 + \gamma \mathbf{1} \cdot \xi < \gamma \mathbf{1} \cdot \xi. \tag{3.42}$$

In particular $\xi \not\equiv 0$ and, dividing ξ by $\mathbf{1} \cdot \xi$, we obtain a probability distribution q_ξ such that

$$\min_q \max_p F(p, q) \leq \max_p F(p, q_\xi) < \gamma.$$

This contradicts the second inequality in (3.37). □

3.2.4 Uniform convexity, projection on convex sets

Definition 3.6 (Uniform convexity)**.** A normed vector space E is said to be *uniformly convex* if for all $\varepsilon > 0$, there exists $\delta > 0$ such that,

$$(\|u\| = \|v\| = 1, \quad \|u - v\| \geq \varepsilon) \Longrightarrow \left\| \frac{u+v}{2} \right\| < 1 - \delta, \tag{3.43}$$

for all $u, v \in E$.

To cite [Cla36], the space E is uniformly convex if "the mid-point of a variable chord of the unit sphere of the space cannot approach the surface of the sphere unless the length of the chord goes to zero". By the parallelogram identity

$$\|u + v\|_H^2 + \|u - v\|_H^2 = 2(\|u\|_H^2 + \|v\|_H^2). \tag{3.44}$$

a Hilbert space is uniformly convex. Typical examples of non uniformly convex spaces are given by Lebesgue spaces L^1 or L^∞, or, in finite dimension, $\mathbb{R}^d$ with norm

$$\|x\|_1 = \sum_{i=1}^{d} |x_i| \quad \text{or} \quad \|x\|_\infty = \max_{1 \leq i \leq d} |x_i|.$$

(Draw a picture of the unit balls). For instance, in $\mathbb{R}^2$ with the norm $\|\cdot\|_\infty$, the points $(1, 1)$ and $(1, -1)$ do not satisfy (3.43). We prove in Corollary 3.17 below that the spaces L^p are uniformly convex when $1 < p < +\infty$. Let us come back to the case of Hilbert spaces: it turns out that the existence of an operator "projection on a convex", well-known in Hilbert spaces, is also true in uniformly convex Banach spaces.

Theorem 3.15 (Projection on a convex)**.** *Let K be a non-empty closed convex set in a uniformly convex Banach space E. Then, for all $u \in E$, the distance $d(u, K)$ is reached at a unique point $\pi_K(u)$. The map $u \mapsto \pi_K(u)$ defined in this way is continuous.*

Proof of Theorem 3.15. If $u \in K$, then $\pi_K(u) = u$. Assume $u \notin K$ and let $\eta = d(u, K)$. Then $\eta > 0$. Let v_n be a sequence in K such that $\eta_n := \|u - v_n\| \to \eta$. We can assume $\eta_n \neq 0$ and consider the unit vectors $w_n = \eta_n^{-1}(u - v_n)$. We will show that (w_n) is Cauchy. Since E is complete, we will deduce that $w_n \to w$ for a certain $w \in E$, and thus $v_n \to v := u - \eta w$. This point v will realize the distance $d(u, K)$. Such a point is necessarily unique: if $v' \in K$ is an other point that realizes the distance, then $w' = \eta^{-1}(u - v)$ is a unit vector with $\|w - w'\| = \eta^{-1}\|v - v'\|$ and then

$$\|w - w'\| = \varepsilon > 0 \Rightarrow \left\| \frac{w + w'}{2} \right\| < 1 - \delta \Rightarrow \left\| u - \frac{v + v'}{2} \right\| < (1 - \delta)\eta, \tag{3.45}$$

a contradiction (in (3.45), δ is given by (3.43)). To prove that the sequence (w_n) is Cauchy, we modify slightly the arguments used to show the uniqueness of $\pi_K(u)$: let $\varepsilon > 0$ and $\delta > 0$ given by (3.43). If $\|w_n - w_m\| \geq \varepsilon$, then $\|w_n + w_m\| < 2(1 - \delta)$ and

$$\frac{1}{2} \left\| \frac{u - v_n}{\eta_n} + \frac{u - v_m}{\eta_m} \right\| = \frac{1}{2} \left\| \frac{\eta_n + \eta_m}{\eta_n \eta_m} u - \frac{1}{\eta_n} v_n - \frac{1}{\eta_m} v_m \right\| < (1 - \delta). \tag{3.46}$$

The point

$$v^* = \frac{\eta_n \eta_m}{\eta_n + \eta_m} \left(\frac{1}{\eta_n} v_n + \frac{1}{\eta_m} v_m \right)$$

belongs to K, as a convex combination of v_n and v_m. Consequently, $\eta \leq d(u, v^*)$ and by (3.46) we obtain

$$\eta < (1-\delta)\frac{2\eta_n\eta_m}{\eta_n + \eta_m}. \tag{3.47}$$

The quotient $\frac{2\eta_n\eta_m}{\eta_n+\eta_m}$ approaches 1 when $n, m \to +\infty$, so (3.47) cannot be satisfied for $n, m \geq n_0$ if n_0 is large enough, which means that (w_n) satisfies the Cauchy condition $\|w_n - w_m\| < \varepsilon$ for $n, m \geq n_0$ and complete the proof of existence of $\pi_K(u)$. Let us now study the continuity of the map π_K. To that purpose, note that we have proved above a little more than the mere existence of a point realizing the distance. We have established the following fact: *if $u \in E \setminus K$ and (v_n) is a minimizing sequence, in the sense that $d(u, v_n) \to d(u, K)$, then $v_n \to \pi_K(u)$.* If $u \in K$ and $v \in B(u, r)$, then

$$\|\pi_K(v) - \pi_K(u)\| = \|\pi_K(v) - u\| \leq r + \|\pi_K(v) - v\| = r + d(v, K) \leq 2r, \tag{3.48}$$

so π_K is continuous on K. Our aim is to show that it is continuous on $E \setminus K$: let (u_n) be a sequence converging to $u \notin K$. The distance function $v \mapsto d(v, K)$ is 1-Lipschitz continuous, so $\eta_n := d(u_n, K)$ converges to $\eta := d(u, K)$. Set $v_n = \pi_K(u_n)$. We have

$$\eta \leq d(u, v_n) \leq d(u, u_n) + d(u_n, v_n) = d(u, u_n) + \eta_n. \tag{3.49}$$

The right-hand side in (3.49) converges to η so $d(u, v_n) \to \eta$. Therefore (v_n) is a minimizing sequence, and $v_n \to \pi_K(u)$. □

Proposition 3.16 (Hanner's inequality)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $1 \leq p \leq 2$. For all $u, v \in L^p(X)$, we have*

$$\left|\|u\|_{L^p(X)} + \|v\|_{L^p(X)}\right|^p + \left|\|u\|_{L^p(X)} - \|v\|_{L^p(X)}\right|^p \leq \|u+v\|^p_{L^p(X)} + \|u-v\|^p_{L^p(X)}, \tag{3.50}$$

and (by application of (3.50) to $u+v$ and $u-v$):

$$\left|\|u+v\|_{L^p(X)} + \|u-v\|_{L^p(X)}\right|^p + \left|\|u+v\|_{L^p(X)} - \|u-v\|_{L^p(X)}\right|^p \leq 2^p\left(\|u\|^p_{L^p(X)} + \|v\|^p_{L^p(X)}\right). \tag{3.51}$$

If $2 \leq p < +\infty$, then the inequalities are reversed.

Corollary 3.17 (Uniform convexity of Lebesgue spaces)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space. If $1 < p < +\infty$, then $L^p(X)$ is uniformly convex.*

Proof of Theorem 3.16. We assume $1 \leq p < 2$. We will establish the following general inequality: there exists some functions

$$\alpha, \beta \colon [0,1] \to \mathbb{R}_+, \tag{3.52}$$

such that, for all $0 \leq A, B$, for all $r \in [0,1]$,

$$\alpha(r)A^p + \beta(r)B^p \leq |A+B|^p + |A-B|^p, \tag{3.53}$$

and equality holds in (3.53) if $r = \min(B/A, A/B)$. To deduce (3.50) from this result, we apply (3.53) to $A = |u(x)|$, $B = |v(x)|$ and integrate with respect to $x \in X$ to obtain

$$\alpha(r)\|u\|^p_{L^p(X)} + \beta(r)\|v\|^p_{L^p(X)} \leq \|u+v\|^p_{L^p(X)} + \|u-v\|^p_{L^p(X)}. \tag{3.54}$$

Without loss of generality, we can assume $0 < \|u\|_{L^p(X)} \leq \|v\|_{L^p(X)}$. Then we choose

$$r = \frac{\|u\|_{L^p(X)}}{\|v\|_{L^p(X)}}$$

in (3.54) to conclude. To guess the form of α and β, first note that we can assume $B \leq A$ in (3.53) and, dividing by A, consider the inequality

$$\alpha(r) + \beta(r)\sigma^p \leq (1+\sigma)^p + (1-\sigma)^p, \quad r, \sigma \in [0,1], \tag{3.55}$$

with equality for $r = \sigma$. Set

$$\varphi(\sigma) = (1+\sigma)^p + (1-\sigma)^p - (\alpha(r) + \beta(r)\sigma^p).$$

Then

$$\varphi'(\sigma) = p\left[(1+\sigma)^{p-1} - (1-\sigma)^{p-1} - \beta(r)\sigma^{p-1}\right].$$

We would like $\varphi'(r) = 0$ since φ should reach a minimum at $\sigma = r$, so we set

$$\beta(r) = \left[(1+r)^{p-1} - (1-r)^{p-1}\right] r^{1-p}. \tag{3.56}$$

The equality in (3.55) will be realized for $r = \sigma$ if

$$\alpha(r) = (1+r)^{p-1} + (1-r)^{p-1}. \tag{3.57}$$

Conversely, one can check that (3.55) is satisfied with β and α given by (3.56)-(3.57). This concludes the proof. □

Proof of Corollary 3.17. If $u, v \in L^p(X)$ satisfy

$$\|u\|_{L^p(X)} = \|v\|_{L^p(X)} = 1, \quad \|u - v\|_{L^p(X)} = 2\varepsilon > 0, \tag{3.58}$$

then (3.51) gives

$$(\theta+\varepsilon)^p + (\theta-\varepsilon)^p \leq 2, \quad \theta := \left\|\frac{u+v}{2}\right\|. \tag{3.59}$$

From (3.53), we deduce that

$$\alpha(r)\theta^p + \beta(r)\varepsilon^p \leq 2, \tag{3.60}$$

where $r \in [0,1]$ is arbitrary. Let us take $r = \varepsilon$. We have

$$\alpha(\varepsilon) = 2 + p(p-1)\varepsilon^2 + \mathcal{O}(\varepsilon^3), \quad \beta(\varepsilon) = 2(p-1)\varepsilon^{2-p} + \mathcal{O}(\varepsilon^{3-p}).$$

By (3.60), we obtain

$$\theta^p \leq 1 - (p-1)^2\varepsilon^2 + \mathcal{O}(\varepsilon^3) \Rightarrow \theta \leq 1 - \frac{(p-1)^2}{p}\varepsilon^2 + \mathcal{O}(\varepsilon^3). \tag{3.61}$$

This shows that we can associate a δ to ε such that (3.58) implies $\left\|\frac{u+v}{2}\right\| < 1 - \delta$. □

To complete this section on uniformly convex spaces, we give the following result, which will be used later in the proof that uniformly Banach spaces are reflexive spaces (Section 5.4.3).

Lemma 3.18 (Uniform convexity, extended criterion)**.** *If E is uniformly convex, then, given $\varepsilon > 0$, there exists $\delta' > 0$ such that*

$$(\|u\| \leq 1, \|v\| \leq 1, \quad \|u - v\| \geq \varepsilon) \Longrightarrow \left\|\frac{u+v}{2}\right\| < 1 - \delta', \tag{3.62}$$

for all $u, v \in E$.

Proof of Lemma 3.18. Let $\alpha > 0$. If $\|u\|$ or $\|v\| < 1 - 2\alpha$, then, using the triangular inequality, the bound

$$\left\| \frac{u+v}{2} \right\| \leq \frac{1}{2}(\|u\| + \|v\|) < 1 - \alpha \tag{3.63}$$

is satisfied. Assume $\|u\|$ and $\|v\| \geq 1 - 2\alpha$ (we will choose α depending on ε and the delta associated to $\varepsilon/2$ by (3.43)). Let $U = \frac{u}{\|u\|}$ and $V = \frac{v}{\|v\|}$. We have then

$$\|U - u\| = 1 - \|u\| < 2\alpha,$$

so

$$\|U - V\| \geq \varepsilon - 4\alpha, \quad \|u + v\| \leq \|U + V\| + 4\alpha.$$

Let $\delta = \delta(\varepsilon) > 0$ be associated to condition (3.43) with $\varepsilon/2$ instead of ε. Assume α smaller than $\varepsilon/8$. Let us apply (3.43) to (U, V): we obtain

$$\|u + v\| \leq \|U + V\| + 4\alpha < 1 + 4\alpha - \delta.$$

It is sufficient to take α smaller than $\varepsilon/8$ and $\delta/8$ to obtain (3.62) with $\delta' = \delta/2$. □

4 Fréchet spaces

It is said that Banach spaces, as we know them, were called that way by Maurice Fréchet. In return, Stephan Banach would have given the name of "Fréchet spaces" to topological vector spaces which admit a complete metric, invariant by translation, compatible with the original topology. Depending on the authors, such spaces are sometimes called F-spaces, while Fréchet spaces have the additional property to be locally convex. In practice, Fréchet spaces very often appear as spaces endowed with a countable family of semi-norms. We have already encountered the Schwartz space in Definition 2.3. Some other standard functional spaces are Fréchet spaces: let U be an open set of $\mathbb{R}^d$ (non-necessarily bounded, it may be the whole space $\mathbb{R}^d$) and let (K_n) be an exhaustive sequence of compacts of U: $K_n \subset K_{n+1}$ is compact and $U = \cup_n K_n$. We can consider

1. the spaces $L^p_{\text{loc}}(U)$, for $p \in [1, +\infty]$, with the semi norms $p_n(u) = \|u\|_{L^p(K_n)}$,
2. the spaces $C^k(U)$, for $k \in \mathbb{N} \cup \{\infty\}$, of continuous functions with continuous differential up to order k, with the semi-norms

$$p_n(u) = \sup_{|\alpha| \leq k \wedge n} \sup_{x \in K_n} |\partial_x^\alpha u(x)|. \tag{4.1}$$

In (4.1), α is a multi-index in $\mathbb{N}^d$, $|\alpha| = \alpha_1 + \cdots + \alpha_d$ and $k \wedge n = \min(k, n)$.

Theorem 4.1 (On the way to Fréchets spaces). *Let X be a topological vector space. The following assertions are equivalent*

1. *there exists a countable family of continuous semi-norms which is separating,*
2. *X is locally convex and admits a metric, invariant by translation, compatible with the original topology.*

If realized, and $\{p_n; n \in \mathbb{N}\}$ is a countable family of continuous semi-norms which is separating, then a neighbourhood base of the origin consists in the finite intersections of the sets

$$V_{k,n} = \left\{ u \in X; p_n(u) < \frac{1}{k} \right\}. \tag{4.2}$$

In the proof of Theorem 4.1 we associate to a separating family of continuous semi-norms a metric invariant by translation, compatible with the original topology (see (4.4)). If we can show additionally that d is complete (this is relatively easy in practical situations), then X turns out to be a Fréchet space.

Remark 4.1 (Increasing sequence of semi-norms). In practice, the countable family $\{p_n; n \in \mathbb{N}\}$ of semi-norms is often increasing: $p_n(u) \leq p_{n+1}(u)$ for all n. If this is not the case, then considering

$$\bar{p}_n = \sum_{j \leq n} p_j$$

will do a transposition to this "monotone" situation. If $\{p_n; n \in \mathbb{N}\}$ is increasing, then $\{V_{k,n}\}$ in (4.2) is a neighbourhood base of the origin (no need to make finite intersections). In any case, a linear map $T\colon X \to Y$ between Fréchet spaces with respective countable families of semi-norms $\{p_n; n \in \mathbb{N}\}$ and $\{q_m; m \in \mathbb{N}\}$ is continuous if, and only if, for each m, there exists $C_m \geq 0$ and $N_m \in \mathbb{N}$ such that

$$q_m(T(u)) \leq C_m \max_{1 \leq n \leq N_m} p_n(u), \tag{4.3}$$

for all $u \in X$. See Section 5.2 on bounded sets in Fréchet spaces for more details on that point.

Proof of Theorem 4.1. Let $\{p_n; n \in \mathbb{N}\}$ be a countable family of continuous semi-norms which is separating. Define

$$d(u, v) = \sum_{n \in \mathbb{N}} 2^{-n} \Phi(p_n(u - v)), \quad \Phi(p) := \frac{p}{1 + p}. \tag{4.4}$$

Then d is a metric on X. Indeed, it is symmetric since each p_n is. We have $d(u, v) = 0$ if, and only if $u = v$, because $\{p_n\}$ is separating. The triangular inequality is satisfied because p_n is subadditive and Φ also: by algebraic manipulations, one can see that $\Phi(p + q) \leq \Phi(p) + \Phi(q)$ is equivalent to

$$\frac{q}{1 + p} + \frac{p}{1 + q} \leq q + p,$$

which is obviously true. By definition, the metric d is invariant by translation. Let $\mathcal{T}$ denote the original topology, and let $\mathcal{T}_d$ be the topology generated by the open balls associated to d. Let $r > 0$ and $u \in B(0, r)$. We will show that there exists a $\mathcal{T}$-neighbourhood of u included in $B(0, r)$, which means that $B(0, r)$ is open for $\mathcal{T}$ and proves the inclusion $\mathcal{T}_d \subset \mathcal{T}$. Actually, the neighbourhood that we exhibit will be a finite intersections of sets $u + V_{k,n}$ as in (4.2). Since each p_n is continuous, it will be a $\mathcal{T}$-neighbourhood indeed. Let $\varepsilon = r - d(u, 0)$ and let $N \geq 0$ be such that $\sum_{n > N} 2^{-n} < \varepsilon/2$. By invariance by translation of d, we have

$$d(v, 0) \leq d(u, 0) + d(u - v, 0) = r - \varepsilon + d(u - v, 0).$$

By our choice of index N, we have then

$$d(v, 0) \leq r - \varepsilon/2 + N \max_{1 \leq n \leq N} p_n(u - v). \tag{4.5}$$

We choose k such that $N < k\varepsilon/2$. By (4.5), the intersection of the sets $u + V_{k,n}$ over $n \in \{1, \ldots, N\}$ is a subset of $B(0, r)$. To prove the converse inclusion $\mathcal{T} \subset \mathcal{T}_d$, we consider an open set U of $\mathcal{T}$ and a point $u \in U$. Our aim is to find $r > 0$ such that $B(u,r) \subset U$. By continuity of the semi-norm p_0, there exists $k > 1$ such that $u + V_{k,0} \subset U$. If $v \in B(u,r)$, and $0 < r < 1$, then

$$\Phi(p_0(u-v)) \leq d(u,v) < r \Rightarrow p_0(u-v) < \frac{r}{1-r}. \tag{4.6}$$

Choosing r small enough, we obtain $B(u,r) \subset u + V_{k,0} \subset U$. The sets $V_{k,n}$ are convex, so X is locally convex. This establishes *2*. Conversely, assume that *2* is satisfied. Since X is locally convex, there is, by Proposition 3.10, for all $n \geq 1$, a convex symmetric set V_n such that $V_n \subset B(0, 1/n)$. Let p_n be the gauge of V_n. By Proposition-Definition 3.11, p_n is a continuous semi-norm. The family p_n is separating since $p_n(u) = 0$ implies $u \in B(0, 1/n)$ for all n. □

5 Weak topologies

5.1 Definition of the weak topology, weakly open, closed, bounded sets

Definition 5.1 (Weak topology)**.** Let X be a topological vector space and let X^* be the set of continuous linear functionals on X. The weak topology on X is the coarsest topology that makes all the maps $\varphi \in X^*$ continuous.

Proposition 5.1 (Weak topology)**.** *Let X be a topological vector space. Denote by X_w the space X endowed with the weak topology. Then X_w is a topological vector space. A neighbourhood base of the origin is given by the finite intersections of sets of the form*

$$V_{\varphi,\varepsilon} = \{u \in X; |\varphi(u)| < \varepsilon\}. \tag{5.1}$$

In particular, X_w is locally convex. If X has finite dimension, then $X_w = X$. In all generality, the topological dual of X_w is X^*, i.e. $X_w^* = X^*$.

Proof of Proposition 5.1. By definition, the topology of X_w is generated by the class

$$\left\{\varphi^{-1}(B); B \text{ open in } \mathbb{R}\right\}. \tag{5.2}$$

The continuity of the sum and the multiplication by a scalar in X_w are left as an exercise. That each set reduced to a single element $\{u\}$ is closed is a direct consequence of the separation property *3.* in Theorem 3.8. Since the topology of $\mathbb{R}$ is generated by open "balls" (=segments) it is sufficient to restrict to such B in (5.2). When a topology is generated by a class $\mathcal{E}$ (containing the whole space and the empty set), a neighbourhood base is given by the unions of the finite intersections of elements of $\mathcal{E}$ (use Proposition 4.2 and Proposition 4.4 in [Fol99, p.115] for instance). In our case, a neighbourhood base for the weak topology is given by the finite intersections of the sets

$$V_{\varphi,\varepsilon,\alpha} = \{u \in X; |\varphi(u) - \alpha| < \varepsilon\}. \tag{5.3}$$

In particular, the finite intersections of sets as in (5.1) form a neighbourhood base of the origin. Assume that X has finite dimension and let us show that $X = X_w$. We will need the following result: a space of finite dimension in a topological vector space is closed, see [Rud73, Theorem 1.32]. Let $(e_i)_{1,d}$ be a basis of X. Let e_i^* denote the linear functional $X \to \mathbb{R}$ defined by $e_i^*(e_j) = \delta_{ij}$. Then $\mathrm{Ker}(e_i^*)$ has dimension $d-1$ so it is an hyperplane, and it is closed. By Theorem 3.5, e_i^* is continuous. In particular, $\Phi\colon u \mapsto (e_i^*(u))_{1,d}$ is an homeomorphism of X onto $\mathbb{R}^d$, with inverse $(x_i)_{1,d} \mapsto \sum_{i=1}^d x_i e_i$ (note that the inverse Φ^{-1} is continuous because X is a topological space) and the images by Φ^{-1} of the balls for the ℓ^∞-norm of $\mathbb{R}^d$ give a neighbourhood base of X. If $r > 0$, $x \in \mathbb{R}^d$, then

$$\Phi^{-1}(B_{\ell^\infty}(x,r)) = \bigcap_{i=1}^{d} \{v \in X; |e_i^*(v) - x_i| < r\} = \bigcap_{i=1}^{d} V_{e_i^*,x_i,r}$$

is weakly open, so open sets are weakly open and $X = X_w$. In all generality now, consider a (non-trivial) continuous linear functional $\psi\colon X_w \to \mathbb{R}$. Then $H = \mathrm{Ker}(\psi)$ is closed for the weak topology, and $X \setminus H$ is open for the weak topology. Let $v \in X \setminus H$. There exists $\varphi_1, \dots, \varphi_n \in X^*$ and $\varepsilon_1, \dots, \varepsilon_n > 0$ such that W, the intersection of the sets $v + V_{\varphi_i,\varepsilon_i}$, satisfies $W \subset X \setminus H$. If $n = 1$, then $V_{\varphi_1,\varepsilon_1}$ contains $\mathrm{Ker}(\varphi_1)$, hence $X \setminus H$ contains the affine hyperplane $v + \mathrm{Ker}(\varphi_1)$. Clearly (drawing...), this is possible only if $H = \mathrm{Ker}(\varphi_1)$, *i.e.* $\psi = \lambda\varphi_1$ for a given $\lambda \in \mathbb{R} \setminus \{0\}$. More generally, when n is any positive integer, we have

$$\bigcap_{i=1}^{n} \mathrm{Ker}(\varphi_i) \subset \mathrm{Ker}(\psi). \tag{5.4}$$

Indeed, if (5.4) is not satisfied then there is a $w \in X$ such that tw is in the intersection of the sets $V_{\varphi_i,\varepsilon_i}$ for all $t \in \mathbb{R}$ and $\psi(w) \neq 0$. But then we obtain an element in $H \cap W$ by considering $v + tw$ for $t = -\psi(w)^{-1}\psi(v)$. This is a contradiction, and thus (5.4) is true. It is well-known (see Lemma 5.2 below) that (5.4) is equivalent to the existence of $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ such that

$$\psi = \lambda_1\varphi_1 + \cdots + \lambda_n\varphi_n. \tag{5.5}$$

We conclude that $\psi \in X^*$. □

Remark 5.1 (Weak topology with less functionals). Let Y be a subspace of X^* that separates points in X. We can consider the Y-weak topology on X (the coarsest topology that makes all the element of Y continuous). As above (same proof), this defines a locally convex topological vector space, whose topological dual is Y.

Lemma 5.2. *Let X be a vector space, and $\psi\colon X \to \mathbb{R}$, $\varphi_i\colon X \to \mathbb{R}$, $i = 1, \dots, n$ be some linear functional. Then there is equivalence between the following statements:*

1. *the inclusion* (5.4) *is satisfied,*
2. *the relation* (5.5) *is satisfied for some $\lambda_1, \dots, \lambda_n \in \mathbb{R}$,*
3. *there exists $M \geq 0$ such that*

$$|\psi(u)| \leq M \max_{1\leq i\leq n} |\varphi_i(u),| \tag{5.6}$$

for all $u \in X$.

Proof of Lemma 5.2. It is clear that

2. ⇒ *3.* ⇒ *1.*

Assume that *1.* is satisfied. Let $\Phi\colon X \to \mathbb{R}^{n+1}$ the linear map with n first components $\varphi_1, \dots, \varphi_n$ and last component $-\psi$. Let v be the vector with all n first components equal to 0 and last component equal to 1. Then $v \notin \mathrm{Im}(\Phi)$. By *1.* in Theorem 3.8, there exists a continuous linear form φ on $\mathbb{R}^{n+1}$ such that $\varphi \equiv 0$ on $\mathrm{Im}(\Phi)$ and $\varphi(v) > 0$ (note that we use the Hahn-Banach theorem in finite dimension here). There exists $\lambda \in \mathbb{R}^{n+1}$ such that $\varphi(u) = \lambda \cdot u$ (canonical scalar product in $\mathbb{R}^{n+1}$) for all $u \in \mathbb{R}^{n+1}$. We have then

$$\lambda_{n+1} = \varphi(v) > 0, \quad \lambda \cdot \Phi(x) = 0, \; \forall x \in X. \tag{5.7}$$

Dividing by λ_{n+1} if necessary, we can assume $\lambda_{n+1} = 1$. Then the second identity in (5.7) gives *2*. □

Remark 5.2 (Weak closure of the unit sphere). In the proof of Proposition 5.1, we have used the fact that, in infinite dimension, a finite intersection of sets as in (5.1) contains a whole vector space. It is clear, from that respect, that, in infinite dimension, a point in the unit ball $B(0,1)$ of a normed vector space E, cannot be an interior point for the weak topology. The points in $E \setminus \bar{B}(0,1)$ are interior points for the weak topology on the contrary. It follows that the weak closure of the sphere $S(0,1)$ is the unit ball $\bar{B}(0,1)$. See exercises class for the details.

Terminology: we have used the term "weak closure" in Remark 5.2. This means "closure in the weak topology". Similarly we will employ the terms "weakly open", "weakly closed", "weakly bounded". Only this last term requires an explanation.

Definition 5.2 (Bounded set in a topological vector space)**.** In a topological vector space, a set B is said to be bounded if, for all neighbourhood V of the origin, there exists $t > 0$ such that $B \subset tV$.

One checks then that a set B in X is weakly bounded if, and only if, for all $\varphi \in X^*$, the set $\varphi(B)$ is bounded in $\mathbb{R}$.

Theorem 5.3 (Weakly closed sets)**.** *Let X be a locally convex real topological vector space. If $A \subset X$ is weakly closed, then it is closed. If A is convex and closed, then A is weakly closed.*

Proof of Theorem 5.3. We do the proof in the case where X is a normed space (in the general case, the proof is the same, simply use the generalization of Theorem 3.12 mentioned in Remark 3.4). First, by definition, weak open sets are open, so weak closed sets are closed. Assume that A is convex and closed and let $v \in X \setminus A$. Then v can be separated strictly from A by a closed affine hyperplane (Theorem 3.12). This implies $v \in V_{\varphi,\varepsilon,\alpha}$ for some $\varphi \in X^*$, $\varepsilon > 0$, $\alpha \in \mathbb{R}$ (with the notation in (5.3)). So v is interior to $E \setminus A$ for the weak topology. □

Remark 5.3 (Weak closure of a set). Let X be a topological vector space. For a given set $A \subset X$ denote by $\mathrm{co}(A)$ the set of convex combinations of elements of A and by $\overline{\mathrm{co}}(A)$ the closure (for the original topology) of this set. One may wonder if the weak closure of A would not precisely be $\overline{\mathrm{co}}(A)$. Denote by $w\mathrm{cl}(A)$ the weak closure of A. Since $A \subset \overline{\mathrm{co}}(A)$, and since $\overline{\mathrm{co}}(A)$ is weakly closed by Theorem 5.3, we have $w\mathrm{cl}(A) \subset \overline{\mathrm{co}}(A)$. The inclusion can be strict as shown by the following example: $A = \{u, v\}$, where $u \neq v \in X$. The closure of the convex envelop of A is the segment $[u, v]$, but A is weakly closed since $A = \{u\} \cup \{v\}$ is the union of two weakly closed sets. The inclusion $w\mathrm{cl}(A) \subset \overline{\mathrm{co}}(A)$ is used to proved Mazur's theorem: *if (u_n) is a sequence that converges weakly to an $u \in X$, then there is a sequence (v_m) converging strongly to u such that each v_m is a finite convex combination of the elements u_n.* Just consider $A = \{u_n; n \in \mathbb{N}\}$. The context of Mazur's theorem gives an hint to an other example with $w\mathrm{cl}(A) \subsetneq \overline{\mathrm{co}}(A)$. Let H be a separable Hilbert space with orthonormal basis (e_n). Then (e_n) is converging weakly to 0 (see Exercises class on the weak topology in Hilbert spaces). Let $A = \{e_n; n \in \mathbb{N}\} \cup \{0\}$. Then A is weakly compact (if $(U_i)_{i \in I}$ is a cover of A by weakly open sets and $0 \in U_j$, then U_j contains all the e_n, for $n \geq n_0$ and a certain n_0, so we can find a finite subcover of A). Since X_w

is Hausdorff, A is weakly closed. On the other hand, using the notation $\hat{u}(n) = \langle u, e_n \rangle$, we have

$$\overline{\text{co}}(A) = \left\{ u \in H; \hat{u}(n) \geq 0, \forall n, \sum_{n \in \mathbb{N}} \hat{u}(n) \leq 1 \right\}, \tag{5.8}$$

which is much bigger than A. It is quite clear that we have (5.8), but here is a complete proof. Call Δ_1 the right-hand side of (5.8). It is convex and closed and contains A, so $\overline{\text{co}}(A) \subset \Delta_1$. If v is an element in $\Delta_1 \setminus \overline{\text{co}}(A)$, then we can separate strictly v from $\overline{\text{co}}(A)$: there exists $w \in H$, $\alpha \in \mathbb{R}$, $\varepsilon > 0$ such that $\langle w, v \rangle_H > \alpha + \varepsilon$ and $\langle w, u \rangle_H \leq \alpha$ for all $u \in \overline{\text{co}}(A)$. Taking $u = 0$ gives $\alpha \geq 0$ and taking $u = e_n$ gives $\hat{w}(n) \leq \alpha$ for all n. Since

$$\alpha + \varepsilon \leq \langle w, v \rangle_H = \sum_{n \in \mathbb{N}} \hat{w}(n)\hat{v}(n) \leq \alpha \sum_{n \in \mathbb{N}} \hat{v}(n),$$

we contradict $v \in \Delta_1$.

Corollary 5.4 (Weakly lower semi-continuous functions)**.** *Let X be a locally convex real topological vector space. A map $f\colon X \to \mathbb{R}$ which is convex and continuous is weakly-lower semi-continuous.*

Proof of Corollary 5.4. Recall that a map $g\colon Z \to \mathbb{R}$, where Z is a topological vector space, is lower semi-continuous if, for all $a \in \mathbb{R}$, the set $\{g \leq a\}$ is closed. Typical examples of lower semi-continuous functions are given by:

1. the supremum $g = \sup_{\alpha \in A} g_\alpha$ of a family of continuous functions g_α. Indeed,

$$\{g \leq a\} = \bigcap_{\alpha \in A} \{g_\alpha \leq a\}$$

   is closed.

2. the characteristic function $g = \mathbf{1}_U$ of an open set U. Indeed, depending on the value of a, the set $\{g \leq a\}$ is either the empty set, Z, or $Z \setminus U$.

The corollary follows from the fact that, for all $a \in \mathbb{R}$, the set $\{f \leq a\}$ is closed and convex, hence weakly closed. □

Combined with the Banach-Alaoglu theorem, the result of Corollary 5.4 is used in particular to establish the existence of minimizers in the Calculus of variations, see Section 5.4.4.

Theorem 5.5 (Weakly bounded sets)**.** *Let X be a locally convex real topological vector space. A subset $A \subset X$ is weakly bounded if, and only if, it is bounded.*

Proof of Theorem 5.5. The proof has been done in exercises class (Exercise 1-3 of TD3) when X is a normed space and uses the Banach-Steinhaus theorem and the fact that X^* is complete. We give here a more constructive proof, taken from [LL01, Theorem 2.12, p.52], also valid when X is a normed vector space. For a proof in the general case of locally convex topological vector space, see [Rud73, Theorem 3.18] (beware that the proof then uses the Banach-Alaoglu theorem, which has not been stated yet, see Theorem 5.12).

Step 0. Assume by contradiction that there is a sequence (u_n) in A such that, although the sequence $(\varphi(u_n))$ is bounded for very $\varphi \in X^*$, (u_n) is not bounded in X.

Step 1. Reduction to the case $\|u_n\|_X = 4^n$. Up to extraction of a subsequence, we can assume that $\|u_n\|_X \to +\infty$. By extracting a further subsequence if necessary, we can assume that $\|u_n\|_X \geq 4^n$ for all n. Consider then $\tilde{u}_n = \frac{4^n}{\|u_n\|_X} u_n$. We have $\|\tilde{u}_n\|_X = 4^n$ and

$$|\varphi(\tilde{u}_n)| \leq |\varphi(u_n)|, \tag{5.9}$$

so $(\varphi(\tilde{u}_n))$ is bounded for every $\varphi \in X^*$ as well.

Step 2. Contradiction by construction of a functional $\bar{\varphi}$. By *1* in Theorem 3.8, there exists $\varphi_n \in X^*$ with $\|\varphi_n\|_{X^*} = 1$ and $\varphi_n(u_n) = 4^n$. Let (ε_n) be a sequence of elements of $\{-1, 1\}$ constructed by recursion as follows: we set $\varepsilon_0 = 1$ and if $n > 0$, ε_n is chosen such that

$$\varepsilon_n \quad \text{and} \quad \sum_{j=0}^{n-1} 3^{-j} \varepsilon_j \varphi_j(u_n) \tag{5.10}$$

have the same sign. Since $\varphi_n(u_n) > 0$, this has the consequence that

$$\left| \sum_{j=0}^{n} 3^{-j} \psi_j(u_n) \right| \geq |3^{-n} \varphi_n(u_n)| = \frac{4^n}{3^n}, \tag{5.11}$$

where we have set $\psi_j = \varepsilon_j \varphi_j$. The linear functional

$$\bar{\varphi} = \sum_{j=0}^{\infty} 3^{-j} \psi_j \tag{5.12}$$

is an element of X^*, since the series is absolutely convergent in X^* ($\|\psi_j\|_{X^*} \leq 1$ for all j) and X^* is complete. We have

$$\left| \sum_{j=n+1}^{\infty} 3^{-j} \psi_j(u_n) \right| \leq 4^n \sum_{j \geq n+1} 3^{-j} = \frac{1}{2} \frac{4^n}{3^n}. \tag{5.13}$$

From (5.11)-(5.12)-(5.13), we obtain $|\bar{\varphi}(u_n)| \geq \frac{4^n}{3^n} - \frac{1}{2}\frac{4^n}{3^n} = \frac{1}{2}\frac{4^n}{3^n}$, which contradicts the fact that $(\bar{\varphi}(u_n))$ is bounded. □

5.2 Bounded sets in Fréchet spaces

To complete Section 4, and since the notion of bounded set in a topological vector space was given in Definition 5.2 without much development, we will discuss this notion with more details. This is also a way to be more explicit than in Remark 4.1. First, let us give some examples of bounded sets in a general topological vector space X. A set B reduced to a single point $\{u\}$ is a bounded set. Indeed, by continuity of $\lambda \mapsto \lambda u$, there is for all neighbourhood V of 0, a $\lambda > 0$ such that $\lambda u \in V$. This gives $B \subset tV$ with $t = \lambda^{-1}$. It follows that a finite set is bounded. If (u_n) is a sequence which tends to 0 in X, then $B = \{u_n; n \in \mathbb{N}\}$ is bounded. Indeed, for all neighbourhood V of 0, there is an N such that $u_n \in V$ for all $n \geq N$. Since $\{u_n; n < N\}$ is finite, hence bounded, the result follows. If, more generally, (u_n) is a sequence which tends to a given u in X, then $B = \{u_n; n \in \mathbb{N}\}$ is bounded. To prove this, we set $v_n = u_n - u$, and consider the bounded sets $\{v_n; n \in \mathbb{N}\}$ and $\{u\}$. To conclude we must show that the sum of two bounded sets is a bounded set. This is an easy consequence of the following result: *given a neighbourhood V of 0, there exists a neighbourhood W of 0 such that $W + W \subset V$.* To prove this we use the continuity of $(u, v) \mapsto u + v$: there exists W_1, W_2 neighbourhoods of 0 such that $W_1 + W_2 \subset V$. Then we set $W = W_1 \cap W_2$. Similar arguments show that, if (u_n) is a Cauchy sequence, then $\{u_n; n \in \mathbb{N}\}$ is a bounded set. Let us give a last example: a compact set is bounded. Indeed, if K is compact, then each $u \in K$ is in $t_u V$ for a $t_u > 0$. Thus K can be covered by the sets $t_u V$ for $u \in L$, where $L \subset K$ is finite. We obtain $K \subset tV$ with $t = \max_{u \in L} t_u$.

Proposition 5.6 (Bounded sets)**.**
1. *If E is a normed space, then a set B is bounded if, and only if it is bounded in the customary sense: the norm $\| \cdot \|_E$ is bounded on B (this is a straightforward application of the definition with $V = \bar{B}(0, \varepsilon)$, $\varepsilon > 0$, as neighbourhood base of 0).*

2. *Let X be a Fréchet space with a countable family $\{p_n; n \in \mathbb{N}\}$ of increasing, continuous semi-norms. A set B in X is bounded if, and only if, every p_n is bounded on B: there exists a family of numbers $M_n \geq 0$ such that*

$$B \subset \bigcap_{n \in \mathbb{N}} \{u \in X; p_n(u) \leq M_n\}. \tag{5.14}$$

3. *A linear continuous map between two topological vector spaces X and Y sends bounded sets on bounded sets. The reciprocal statement is true if X and Y are Fréchet spaces.*

4. *Let X (*resp. *Y) be a Fréchet space with a countable family $\{p_n; n \in \mathbb{N}\}$ (*resp. *$\{q_m; m \in \mathbb{N}\}$) of increasing, continuous semi-norms. A linear map $\Lambda \colon X \to Y$ is continuous if, and only if, for all m, there exists n and $C_{n,m} \geq 0$ such that $q_m(\Lambda(u)) \leq C_{n,m} p_n(u)$ for all $u \in X$.*

To give the proof of *3.* in Proposition 5.6, we will need the following result.

Lemma 5.7 (Improving convergence in metric spaces)**.** *Let (u_n) be a sequence converging to 0 in a topological vector space X which admits a compatible metric. Then there is a sequence $(\gamma_n) \uparrow +\infty$, such that $(\gamma_n u_n)$ is converging to 0.*

Proof of Lemma 5.7. Let d be the compatible metric: we have $\delta_n := d(u_n, 0) \to 0$. There is an increasing sequence (n_k) such that $\delta_n < 3^{-k}$ if $n \geq n_k$. Set $n_0 = 0$ and $\gamma_n = 2^k$ if $n_k \leq n < n_{k+1}$. Then $(\gamma_n) \uparrow +\infty$ and $(\gamma_n \delta_n)$ is converging to 0, so $(\gamma_n u_n)$ is converging to 0. □

Proof of Proposition 5.6. Let us prove *2.* A neighbourhood base of the origin is given by the sets $V_{n,\varepsilon} = \{u \in X, p_n(u) < \varepsilon\}$ (Theorem 4.1). If B is bounded then there is for all n a $t_n \geq 0$ such that $B \subset t_n V_{n,1}$. This gives (5.14) with $M_n = t_n$. Assume now that $p_n(u) \leq M_n$ for all $u \in B$. Then $B \subset t_n V_{n,\varepsilon}$, $t_n := M_n/\varepsilon$ for all n, ε, which shows that B is bounded. Let us now consider a linear map $\Lambda \colon X \to Y$ between two topological vector spaces. If Λ is continuous, V a neighbourhood of 0 in Y and B a bounded set in X, then $\Lambda^{-1}(V)$ is a neighbourhood of 0 in X. Let $t \geq 0$ be such that $B \subset t\Lambda^{-1}(V)$. Then $\Lambda(B) \subset tV$, so $\Lambda(B)$ is bounded. Assume that X and Y are Fréchet spaces and that Λ sends bounded sets on bounded sets. We want to show that Λ is continuous. Since a neighbourhood base of the origin in Y is given by the sets $U_{m,\varepsilon} = \{v \in Y; q_m(v) < \varepsilon\}$, and by homogeneity, it is sufficient to prove that, given $m \in \mathbb{N}$, $\Lambda^{-1}(U_{m,1})$ is open. If we reach this conclusion, then there is a n and $\varepsilon > 0$ such that $V_{n,\varepsilon} \subset \Lambda^{-1}(U_{m,1})$, which implies $q_m(\Lambda(u)) \leq C_{n,m} p_n(u)$ for all $u \in X$, with $C_{n,m} = \varepsilon^{-1}$. To prove that $\Lambda^{-1}(U_{m,1})$ is open, we show that the complementary set is closed. Since X admits a compatible metric, we can use a sequential characterization, so let us consider a sequence (u_n) converging to $u \in X$ such that that $q_m(\Lambda u_n) \geq 1$ for all n. We apply Lemma 5.7 to $u_n - u$.

Since $B = \{\gamma_n(u_n - u)\}$ is bounded, and since Λ sends bounded sets into bounded sets, there exists $t \geq 0$ such that $\Lambda(B) \subset tU_{m,1}$. This implies $q_m(\Lambda(u_n - u)) \leq t\gamma_n^{-1}$. We deduce that

$$1 \leq q_m(\Lambda(u_n)) \leq q_m(\Lambda(u_n - u)) + q_m(\Lambda(u)) \leq t\gamma_n^{-1} + q_m(\Lambda(u))$$

and obtain the desired inequality $1 \leq q_m(\Lambda(u))$ at the limit $[n \to +\infty]$. □

5.3 Topological dual spaces of some standard Banach and Fréchet spaces

You may skip in this part of the course the first two sections 5.3.1 and 5.3.2, and only consider the list of results given in Section 5.3.3.

5.3.1 Additive set functions

Definition 5.3 (Bounded additive set functions)**.** Let $\mathcal{A}$ be an algebra of sets. A set function $\mu\colon \mathcal{A} \to \mathbb{R}$ is said to be *additive*, or *finitely additive*, if $\mu(\emptyset) = 0$ and

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i), \tag{5.15}$$

for every disjoint sets $A_1, \ldots, A_n \in \mathcal{A}$. The set function μ is said to be *bounded* if there exists $M \geq 0$ such that $|\mu(A)| \leq M$ for all $A \in \mathcal{A}$.

The variation $|\mu|$ of a finitely additive measure is defined as the following set function:

$$|\mu|(A) = \sup\left\{\sum_{i=1}^{n} |\mu(A_i)|\right\}, \tag{5.16}$$

where the supremum is taken over all finite partitions of a set $A \in \mathcal{A}$ by elements $A_1, \ldots, A_n$ of $\mathcal{A}$. Then $|\mu|$ is also an additive function and, if μ is bounded (say by M), then $|\mu|$ is bounded (at least by $2M$) and, we have, for all $A \in \mathcal{A}$,

$$|\mu(A)| \leq |\mu|(A)| \leq 2|\mu(A)|. \tag{5.17}$$

(To prove (5.17), one can use the decomposition (2.3) for instance.)

Definition 5.4 (Regular bounded additive measure)**.** Let X be a topological space and $\mathcal{A}$ a sub-algebra of the Borel σ-algebra. An additive set function μ on $\mathcal{A}$ is said to be *regular* if, for all $A \in \mathcal{A}$ and $\varepsilon > 0$, there is a closed set $F \in \mathcal{A}$ and an open set $G \in \mathcal{A}$ such that $F \subset A \subset G$ and $|\mu|(G \setminus F) < \varepsilon$.

Theorem 5.8 (Alexandrov)**.** *A bounded and regular finitely additive function on the Borel σ-algebra of a compact Hausdorff space X is countably additive. More precisely, if A_n, $n = 1, 2, \cdots$ are some disjoint Borel sets and A denote their union, then*

$$\sum_{n=1}^{\infty} |\mu(A_n)| < +\infty, \tag{5.18}$$

and $\mu(A)$ is equal to the sum of the $\mu(A_n)$ over $n \in \mathbb{N} \setminus \{0\}$.

Proof of Theorem 5.8. For all finite N, we have

$$\sum_{n=1}^{N} |\mu|(A_n) = |\mu| \left(\bigcup_{n=1}^{N} A_n \right) \leq |\mu|(A),$$

so

$$\sum_{n=1}^{\infty} |\mu|(A_n) \leq |\mu(A)| \leq |\mu|(A) < +\infty, \tag{5.19}$$

which gives (5.18) by (5.17). We will show first that $|\mu|$ is countably additive. Since (5.19) gives one of the desired inequalities, it is sufficient to prove that

$$\sum_{n=1}^{\infty} |\mu|(A_n) \geq |\mu|(A). \tag{5.20}$$

Let $\varepsilon > 0$. There is a closed set $F \subset A$ and some open sets $G_n \supset A_n$ such that $|\mu|(A \setminus F) < \varepsilon$ and $|\mu|(G_n \setminus A_n) < 2^{-n}\varepsilon$ for all n. Since F is compact, there is a finite N such that

$$F \subset \bigcup_{n=1}^{N} G_n \quad \Rightarrow \quad |\mu|(F) \leq \sum_{n=1}^{N} |\mu|(G_n). \tag{5.21}$$

We have then

$$|\mu|(A) \leq |\mu|(F) + \varepsilon \leq \sum_{n=1}^{N} |\mu|(G_n) + \varepsilon \leq \sum_{n=1}^{\infty} |\mu|(A_n) + 2\varepsilon,$$

which gives (5.20) by letting $\varepsilon \to 0$. Using the decomposition

$$A = \left(\bigcup_{n=1}^{N} A_n \right) \bigcup B_N, \quad B_N := \bigcup_{n>N} A_n,$$

we have, by finite additivity of μ and countable additivity of $|\mu|$,

$$\left| \mu(A) - \sum_{n=1}^{N} A_n \right| = |\mu(B_N)| \leq |\mu|(B_N) = \sum_{n>N} |\mu|(A_n). \tag{5.22}$$

The right-hand side of (5.22) tends to 0 when $N \to +\infty$ by (5.19), so μ is countably additive. □

Notations: we will use the following notations.

- $\mathrm{ba}(X, \mathcal{A})$ is the set of bounded finitely additive set functions on a given algebra $\mathcal{A}$ of subsets of a set X,
- $\mathrm{rba}(X)$ is the set of regular bounded finitely additive set functions on the Borel σ-algebra $\mathcal{A}$ of a topological space X,
- $\mathrm{ca}(X, \mathcal{A})$ is the set of bounded countably additive set functions on a given σ-algebra $\mathcal{A}$ on a set X,
- $\mathrm{rca}(X)$ is the set of regular bounded countably additive set functions on the Borel σ-algebra $\mathcal{A}$ of a topological space X.

With the terminology of Section 2.1, $\mathrm{rca}(X)$ is the set of regular signed Borel measures on the Borel sets of X. Alexandrov's theorem (Theorem 5.8) shows that the inclusion $\mathrm{rca}(X) \subset \mathrm{rba}(X)$ is an equality when X is compact Hausdorff. The set $\mathrm{ba}(X, \mathcal{A})$ is endowed with the norm

$$\|\mu\|_{\mathrm{ba}(X,\mathcal{A})} = |\mu|(X). \tag{5.23}$$

If (μ_n) is a Cauchy sequence in $\mathrm{ba}(X, \mathcal{A})$, then, for each $A \in \mathcal{A}$, the sequence $(\mu_n(A))$ is Cauchy. It is therefore convergent in $\mathbb{R}$, and if we call $\mu(A)$ the limit, we can check that this defines a set function $\mu \in \mathrm{ba}(X, \mathcal{A})$, and that $\|\mu_n - \mu\|_{\mathrm{ba}} \to 0$. Therefore $\mathrm{ba}(X, \mathcal{A})$ is a Banach Space. The space $\mathrm{ca}(X, \mathcal{A})$ is closed in $\mathrm{ba}(X, \mathcal{A})$, so it is a Banach Space. When X is a topological space, the sets $\mathrm{rba}(X)$, $\mathrm{rca}(X)$ are closed in $\mathrm{ba}(X)$ and are Banach spaces also.
We will also use the following additional notations.

- $\mathrm{B}(X)$ is the set of bounded real-valued functions on a space X, with norm

$$\|u\|_{\mathrm{B}(X)} = \sup_{x\in X} |u(x)|. \tag{5.24}$$

- If $\mathcal{A}$ is an algebra of sets on a space X, then $\mathrm{B}(X, \mathcal{A})$ is the subset of $\mathrm{B}(X)$, defined by taking the closure of the set of $\mathcal{A}$-simple functions (finite linear combinations of characteristic functions of elements in $\mathcal{A}$) for the sup-norm (5.24). When $\mathcal{A}$ is a σ-algebra, this set is also denoted $\mathrm{BM}(X, \mathcal{A})$, where the "M" is for "measurable"). The notation is consistent since $\mathrm{B}(X, \mathcal{A})$ is precisely[4] the set of *bounded $\mathcal{A}$-measurable* functions on X. When X is a topological space, we also use the notation $\mathrm{BM}(X)$ for $\mathrm{B}(X, \mathcal{A})$ where $\mathcal{A}$ is the Borel σ-algebra.

- If X is a topological space, $\mathrm{BC}(X)$ is the set of bounded and continuous functions on X, with the sup norm (5.24).

If X is a topological space, then $\mathrm{BC}(X)$ and $\mathrm{BM}(X)$ are closed subspace of $\mathrm{B}(X)$. Since $\mathrm{B}(X)$ is complete (direct consequence of the fact that $\mathbb{R}$ is complete), all three spaces are Banach spaces.

5.3.2 A list of topological dual spaces

1. Let $\mathcal{A}$ be an algebra of subsets of a set X. The topological dual of $\mathrm{B}(X, \mathcal{A})$ is $\mathrm{ba}(X, \mathcal{A})$. The map

$$\Phi \colon \mathrm{ba}(\mathcal{A}) \to \mathrm{B}(X, \mathcal{A})^*, \quad \mu \mapsto \left(f \mapsto \int_{\mathbb{R}^d} f d\mu \right) \tag{5.25}$$

is an isomorphism with the isometry property $\|\Phi(\mu)\|_{\mathrm{B}(X,\mathcal{A})^*} = \|\mu\|_{\mathrm{ba}(X,\mathcal{A})}$. See [DS58, IV.5.1] (*note:* the integral against an element of $\mathrm{ba}(X, \mathcal{A})$ is defined in [DS58, III.2]). Note also that the special case $\mathcal{A} = \mathcal{P}(X)$ gives a representation of the dual of $\mathrm{B}(X)$ by bounded finitely additive set functions defined on all subsets of X.

2. Let X be a topological space. Assume that X is normal (which means that if F_1 and F_2 are two disjoint closed sets of X, then there are disjoint open sets G_1, G_2 such that $F_i \subset G_i$). The topological dual of $\mathrm{BC}(X)$ is $\mathrm{rba}(X)$ and we have an isomorphism as in (5.25), see [DS58, IV.6.2]. Since $\mathrm{BC}(X) \subset \mathrm{BM}(X)$, we have $\mathrm{BM}(X)^* \subset \mathrm{BC}(X)^*$, so the fact that $\mathrm{rba}(X)$ is a subset of $\mathrm{ba}(X)$ can be surprising at first. The essential point in the proof of the identity $\mathrm{BC}(X)^* = \mathrm{rba}(X)$ is to establish that, when X is normal, every $\lambda \in \mathrm{ba}(X)$ can be represented by a $\mu \in \mathrm{rba}(X)$, in the sense that the integrals of bounded continuous functions against the two measures always coincide.

[4] to justify this, simply observe that a bounded measurable function is limit for the sup-norm (5.24) of a sequence of simple functions - take a look at [Rud87, Theorem 1.17] for instance.

3. Let X be a topological space. Assume that X is compact Hausdorff. The topological dual of $\mathrm{C}(X)$ is $\mathrm{rca}(X)$ and we have an isomorphism as in (5.25). See [DS58, IV.6.3]. The result is also a consequence of: - the previous result 2., - the fact that a compact Hausdorff space is normal [Fol99, Proposition 4.25], - Alexandrov's theorem (Theorem 5.8).

4. Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space and let $1 \leq p < +\infty$. Then the topological dual of $L^p(X, \mu)$ is $L^{p'}(X, \mu)$, where $p' = \frac{p}{p-1}$ is the conjugate exponent to p. More precisely, the map

$$\Phi \colon L^{p'}(X, \mu) \to (L^p(X, \mu))^*, \quad v \mapsto \left(u \mapsto \int_X uv d\mu\right) \tag{5.26}$$

is an isomorphism with the isometry property $\|\Phi(v)\|_{(L^p(X,\mu))^*} = \|v\|_{L^{p'}(X,\mu)}$. The classical proof uses the Radon-Nykodim theorem, see [Rud87, Theorem 6.16].

5. Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space. The dual of $L^\infty(X, \mu)$ is $\mathrm{ba}(X, \mathcal{A}, \mu)$, which is the subset of $\mathrm{ba}(X, \mathcal{A})$ formed by the bounded finitely additive set functions ν with the property that

$$\mu(A) = 0 \Rightarrow \nu(A) = 0, \tag{5.27}$$

for all $A \in \mathcal{A}$. See [DS58, IV.8.16].

5.3.3 A list of topological dual spaces: spaces built on $\mathbb{R}^d$

Theorem 5.9 (Some dual spaces)**.** *Let U be an open set in $\mathbb{R}^d$, H a compact set in $\mathbb{R}^d$ and E a Borel set in $\mathbb{R}^d$.*

1. (Fundamental) *The topological dual of $C_0(\mathbb{R}^d)$, the space of continuous functions on $\mathbb{R}^d$ which tend to 0 at infinity is $\mathcal{M}_b(\mathbb{R}^d)$, the set of (bounded by definition) signed Borel measures on $\mathbb{R}^d$, with norm $\|\mu\|_{\mathcal{M}_b(\mathbb{R}^d)} = |\mu|(\mathbb{R}^d)$ (see Section 2.2). The map*

$$\Phi \colon \mathcal{M}_b(\mathbb{R}^d) \to \left(C_0(\mathbb{R}^d)\right)^*, \quad \mu \mapsto \left(f \mapsto \int_{\mathbb{R}^d} f d\mu\right) \tag{5.28}$$

is an isometry.

2. (Fundamental) *The topological dual of $C(H)$ is $\mathcal{M}(H)$, the set of signed Borel measure on H.*

3. (Complement) *The topological dual of $C(U)$ is $\mathcal{M}_c(U)$, the set of signed Borel measures on $\mathbb{R}^d$ supported in a compact of U.*

4. (Complement) *Let $C_c(U)$ denote the set of continuous functions on U with compact support. There is a topology (the inductive topology, see Section 5.3.4) on $C_c(U)$ such that the topological dual of $C_c(U)$ is $\mathcal{M}_{\mathrm{loc}}(U)$, the set of "measures" on U whose restriction to K are in $\mathcal{M}(K)$ for all compact subset K of U.*

5. (Fundamental) *For $1 \leq p < +\infty$, the topological dual of $L^p(E)$ is $L^{p'}(E)$, where p' is the conjugate exponent to p.*

6. (Complement) *For $1 \leq p < +\infty$, the topological dual of $L^p_{\mathrm{loc}}(U)$ is $L^{p'}_c(U)$, where p' is the conjugate exponent to p, and $L^{p'}_c(U)$ denote the set of functions in $L^{p'}(U)$ which vanish a.e. outside a compact of U.*

7. (Complement) *The topological dual of* $L^\infty(E)$ *is* $\mathrm{ba}(E, \mathcal{A}, \mu)$*, where* $\mathcal{A}$ *is the trace of the Borel* σ*-algebra on* E *and* μ *is the restriction of the Lebesgue measure to* E*.*

8. (Complement) *For* $1 \leq p < +\infty$*, let* $W^{1,p}(U)$ *be the Sobolev space (see Section 7)*

$$W^{1,p}(U) = \left\{u \in L^p(U); \nabla u \in L^p(U;\mathbb{R}^d)\right\}, \quad \|u\|_{W^{1,p}(U)} = \|u\|_{L^p(U)} + \sum_{i=1}^d \|\partial_{x_i} u\|_{L^p(U)}. \tag{5.29}$$

Any φ *in the topological dual of the Sobolev space* $W^{1,p}(U)$ *is represented by some elements* $v_0, v_1, \ldots, v_d \in L^{p'}(U)$*, in the sense that*

$$\varphi(u) = \int_U u(x)v_0(x)dx + \sum_{i=1}^d \int_U \partial_{x_i} u(x) v_i(x) dx, \tag{5.30}$$

for all $u \in W^{1,p}(U)$*. We have then*

$$\|\varphi\|_{(W^{1,p}(U))^*} = \inf\left\{\max_{0\leq i\leq d} \|v_i\|_{L^{p'}(U)}\right\}, \tag{5.31}$$

where the infimum is taken over all $v_0, v_1, \ldots, v_d \in L^{p'}(U)$ *satisfying* (5.30)*, and this infimum is attained.*

9. *the dual of the Fréchet space* $C^\infty(\mathbb{R}^d)$ *is the set of distributions with compact support (see Section 6.2).*

10. *the dual of the Schwartz space* $\mathscr{S}(\mathbb{R}^d)$ *(see Definition 2.3) is a subset of the space of distributions on* $\mathbb{R}^d$*, called the space of* tempered *distribution (see Section 6.2).*

Remark 5.4 (Dual of the Sobolev space). The result *8.* is not as obvious as one may believe, look at the proof first to be convinced of this fact. Consider also the case where U is bounded and of class C^2. Then the trace operator $\gamma\colon W^{1,p}(U) \to L^p(\partial U)$ is a continuous operator (see Section 7.3.6 or [Eva10, Section 5.5], where U of class C^1 is sufficient), so, given $g \in L^{p'}(\partial U)$, the map

$$W^{1,p}(U) \to \mathbb{R}, \quad u \mapsto \int_{\partial U} \gamma u(x) g(x) d\sigma(x) \tag{5.32}$$

is a continuous linear functional on $W^{1,p}(U)$. One can also use the injections of Sobolev space to build linear functional on $W^{1,p}(U)$ which cannot be cast under the form (5.30) so easily. Assume that U is bounded and of class C^2. If $p < d$, we have an injection $W^{1,p}(U) \hookrightarrow L^{p^*}(U)$, where $\frac{1}{p^*} = \frac{1}{p} - \frac{1}{d}$ (see Section 7.3.5 or [Eva10, p.279]), so

$$W^{1,p}(U) \to \mathbb{R}, \quad u \mapsto \int_U u(x)w(x)d(x) \tag{5.33}$$

is a continuous linear functional on $W^{1,p}(\mathbb{R}^d)$ if $w \in L^q(U)$, $\frac{1}{q} = \frac{1}{p'} + \frac{1}{d}$. If $p > d$, we have an injection $W^{1,p}(U) \hookrightarrow C(\bar{U})$ (see Section 7.3.5 or [Eva10, p.283]), so a Dirac mass δ_{x_0}, where $x_0 \in U$, will be a continuous linear functional on $W^{1,p}(U)$.

Note that we use both the notations introduced in Section 5.3.1 and some other notations that appear frequently. So $\mathcal{M}_b(\mathbb{R}^d) = \mathrm{rca}(\mathbb{R}^d)$ and $\mathcal{M}(K) = \mathrm{rca}(K)$. Only the statement *7.* requires to read section 5.3.1 and 5.3.2 in detail.

Proof of Theorem 5.9. **Proof of *1*.:** this is a theorem of representation of F. Riesz: Theorem 2.5.

Proof of *2*.: see 3. in Section 5.3.2.

Proof of *3*.: let $\varphi \in C(U)^*$. Recall that the topology of Fréchet space of $C(U)$ is described in Section 4. This topology is independent on the choice of the exhaustive sequence (K_n). Without loss of generality, we will assume that, for all n large enough, K_n is a subset of the interior of K_{n+1}. This property is satisfied by the sequence

$$K_n = \bar{B}(0,n) \cap \left\{x \in U; d(x, \partial U) \geq n^{-1}\right\}. \tag{5.34}$$

for instance. By Remark 4.1, there should exists n and a constant $M \geq 0$ such that

$$|\varphi(u)| \leq M\|u\|_{C(K_n)}, \quad \forall u \in C(U). \tag{5.35}$$

Let χ be a continuous function such that $K_n \prec \chi \prec K_{n+1}$: $\chi \equiv 1$ on K_n and χ is supported in K_{n+1}. Such a χ exists since K_n is a subset of the interior of K_{n+1}. Let $\psi\colon C(K_{n+1}) \to \mathbb{R}$, $\psi\colon u \mapsto \varphi(u\chi)$. Then ψ is a linear functional, which is continuous since

$$|\psi(u)| \leq M\|u\chi\|_{C(K_n)} = M\|u\|_{C(K_n)} \leq M\|u\|_{C(K_{n+1})}$$

by (5.35). If $\mu \in \mathcal{M}(K_{n+1})$ is representing ψ, then

$$\left|\int_{K_{n+1}} u d\mu\right| = |\psi(u)| \leq M\|u\|_{C(K_n)}, \tag{5.36}$$

for all $u \in C(K_{n+1})$. As a consequence of (5.36), μ is supported in K_n.

Proof of *4*.: we will not prove this point now. Indeed, the topology on $C_c(U)$ has not been specified. Instead, let us do the following important remark: the topology considered on $C_c(U)$ is such that

- a sequence (u_n) is converging to u in $C_c(U)$ if there exists a fixed compact $K \subset U$ such that all the functions u_n, u are supported in K and $u_n \to u$ in $C(K)$,
- a linear functional $\varphi\colon C_c(U) \to \mathbb{R}$ is continuous if, and only if, for all compact K of U, there exists a constant $M_K \geq 0$ such that, for all $u \in C_c(U)$ supported in K, one has $|\varphi(u)| \leq M_K\|u\|_{C(K)}$.

Note also that, this topology is not a topology of Fréchet space. More details are given in Section 5.3.4 and then in Section 6.1.1 (and the proof is actually fully given in Proposition 5.10).

Proof of *5*.: this is a direct consequence of 4. in Section 5.3.2.

Proof of *6*.: let φ be a continuous linear functional on $L^p_{\mathrm{loc}}(U)$. Recall that the topology of Fréchet space of $L^p_{\mathrm{loc}}(U)$ is described in Section 4. By Remark 4.1, there should exists n and a constant $C \geq 0$ such that

$$|\varphi(u)| \leq C\|u\|_{L^p(K_n)}, \quad \forall u \in L^p_{\mathrm{loc}}(U). \tag{5.37}$$

Let $\mathrm{J}_n\colon L^p(K_n) \to L^p_{\mathrm{loc}}(U)$ be the map which associates to $u \in L^p(K_n)$ the extension of u by 0 in $U \setminus K_n$. Then (5.37) gives $\varphi \circ \mathrm{J}_n$ as an element of the topological dual of $L^p(K_n)$. By 5., there is a representative $v \in L^{p'}(K_n)$ for this linear functional. We denote by V the extension of v by 0 in $U \setminus K_n$. We have, for all $u \in L^p_{\mathrm{loc}}(U)$,

$$\int_U V u dx = \int_{K_n} v u|_{K_n} dx = \varphi \circ \mathrm{J}_n(u|_{K_n}).$$

To conclude, we just need to show that $\varphi \circ \mathrm{J}_n(u|_{K_n}) = \varphi(u)$. Since

$$w := \mathrm{J}_n(u|_{K_n}) - u = -u\mathbf{1}_{U\setminus K_n},$$

this follows from the linearity of φ and (5.37) since $\|w\|_{L^p(K_n)} = 0$.

Proof of 7.: see 5. in Section 5.3.2.

Proof of 8.: (see [Ada75, Theorem 3.8]) on $L^p(U;\mathbb{R}^{d+1}) \simeq (L^p(U))^{d+1}$, we consider the norm

$$\|F\|_{L^p(U;\mathbb{R}^{d+1})} = \sum_{i=1}^{d+1} \|F_i\|_{L^p(U)}. \tag{5.38}$$

Let $\Psi\colon W^{1,p}(U) \to L^p(U;\mathbb{R}^{d+1})$ denote the injection $u \mapsto \begin{pmatrix} u \\ \nabla u \end{pmatrix}$. Let M denote the range of Ψ (endowed with the norm (5.38)) and let Φ denote the inverse operator $M \to W^{1,p}(U)$. Since Ψ is an isometry, M is complete, so Φ is continuous. If φ is a continuous linear functional on $W^{1,p}(U)$, then $\Lambda := \varphi \circ \Phi$ is a continuous linear functional on M. By the Hahn-Banach theorem, it can be extended as a continuous linear functional $\tilde{\Lambda}$ on $L^p(U;\mathbb{R}^{d+1})$ with norm $\|\tilde{\Lambda}\| = \|\Lambda\|$. Using 5., we can represent $\tilde{\Lambda}$ by an element $V \in L^{p'}(U;\mathbb{R}^{d+1})$ with components $v_0, \dots, v_d$, the norm of V being given by

$$\|V\| = \max_{0\le i\le d} \|v_i\|_{L^{p'}(U)}.$$

We have then

$$\varphi(u) = \Lambda(\Psi(u)) = \int_U \sum_{i=1}^{d+1} \Psi(u)_i v_{i-1} dx = \int_U u(x)v_0(x)dx + \sum_{i=1}^{d} \int_U \partial_{x_i} u(x) v_i(x) dx,$$

which gives (5.30). We have also

$$\|\Lambda\| = \|\tilde{\Lambda}\| = \max_{0\le i\le d} \|v_i\|_{L^{p'}(U)}, \tag{5.39}$$

The norm of φ is bounded from above by the inf in (5.31), (5.39) shows that the inf is attained with equality. □

5.3.4 The bounded-open topology and the final/inductive topology

Let X, Y be two topological spaces. The product topology on the set Y^X is generated by the sets

$$\mathcal{U}(x, U) = \{f\colon X \to Y; f(x) \in U\},$$

where x is a point in X and U an open set in Y. The *compact-open* topology on Y^X is the topology generated by the sets

$$\mathcal{U}(K, U) = \{f\colon X \to Y; f(K) \subset U\},$$

where K is a compact in X and U an open set in Y.

Definition 5.5 (Bounded-open topology)**.** Let X, Y be two topological vector spaces. The *bounded-open* topology on Y^X is the topology generated by the sets

$$\mathcal{U}(B, U) = \{f\colon X \to Y; f(B) \subset U\},$$

where B is a bounded set in X and U an open set in Y.

Recall that a set B in X is bounded if, for all neighbourhood V of the origin, there exists $t \geq 0$ such that $B \subset tV$. We are particularly interested in the case where X and Y are Fréchet spaces. We assume that there in an increasing sequence of continuous semi-norms (p_n) (*resp.* (q_m)) on X (*resp.* Y). Recall (*cf.* Remark 4.1) that a linear map $T\colon X \to Y$ is continuous if, and only if, for all $m \in \mathbb{N}$, there exists $n \in \mathbb{N}$ and $C_{m,n} \geq 0$ such that $q_m(T(u)) \leq C_{m,n} p_n(u)$ for all $u \in X$. We can prove the following assertions (left as an exercise):

1. a set B is bounded in X if, and only if, for all $n \in \mathbb{N}$, there exists $M_n \geq 0$ such that $p_n(u) \leq M_n$ for all $u \in B$ (*Hint:* consider the neighbourhoods $\{u \in X; p_n(u) < 1\}$, or see Section 5.2),

2. let $L(X, Y)$ denote the set of linear maps $X \to Y$. The sets

$$\mathcal{U}_{n,m,\varepsilon} = \{T \in L(X, Y); p_n(u) < 1 \Rightarrow q_m(T(u)) < \varepsilon\}$$

form a basis of the origin of the bounded-open topology on $L(X, Y)$.

3. In the case where $Y = \mathbb{R}$, with norm given by the absolute value, the convergence of sequences of continuous linear functional for the open-bounded topology on X^* corresponds to the uniform convergence on bounded sets of X.

To conclude the discussion, the open-bounded topology on X^* coincides with the topology of the norm $\|\cdot\|_{X^*}$ when X is a Banach space and constitutes a natural generalization of this topology in the context of Fréchet spaces.

Definition 5.6 (Final topology)**.** Let A be a set of indices and let X, Z_α, $\alpha \in A$ be some topological space. Let also some maps $f_\alpha\colon Z_\alpha \to X$ be given. The *final topology* on X is the finest topology that makes all the maps f_α continuous. A set U is open for this topology if, and only if, the sets $f_\alpha^{-1}(U)$ are open in Z_α for each $\alpha \in A$. If Y is a topological space, a function $g\colon X \to Y$ is continuous if, and only if, all the functions $g \circ f_\alpha\colon Z_\alpha \to Y$ are continuous.

The quotient topology discussed in the proof of Proposition 3.3 is an example of final topology, when the single map $\pi\colon X \to X/M$ is considered. Below we will consider the case where the maps f_α are some natural injections. One also speaks of *inductive topology* in that case. Let U be an open subset of $\mathbb{R}^d$. Let (K_n) be an exhaustive sequence of compacts in U. On $C_c(U)$, we can consider (as in (4.1) when $k = 0$) the topology generated by the semi norms

$$p_{K_n}(u) = \sup_{x \in K_n} |u(x)|.$$

The resulting space is not complete, since a sequence of functions of $C_c(U)$ may converge in this topology to a function which is not compactly supported: consider the functions $u_k(x) = \min(1, [kd(x, \partial U) - 1]^+)$. Indeed, $p_{K_n}(u_k - u_l) = 0$ for all k, l large enough, but $u_k \to 1$ on U. Let us now consider an inductive topology on $C_c(U)$. If K is a compact set with $K \subset U$, we denote by $C_K(U)$ the set of functions $u \in C_c(U)$ supported in K. Let $\mathcal{T}$ denote the inductive topology on $C_c(U)$ associated to the injections $i_K\colon C_K(U) \to C_c(U)$. This topology has in particular the following properties:

1. the space $(C_c(U), \mathcal{T})$ is a locally convex topological vector space,

2. if B is a bounded set then there is a fixed K such that $B \subset C_K(U)$ and a $M \geq 0$ such that $p_K(u) \leq M$ for all $u \in B$,

3. if (u_n) is a sequence converging to an element $u \in C_c(U)$, then there is a compact K such that all the functions u_n, and u, are supported in K and $p_K(u_n - u) \to 0$,

4. the space $(C_c(U), \mathcal{T})$ is complete.

Proof. We admit the first point. Let $B \subset C_c(U)$ and assume that there is no compact K such that $B \subset C_K(U)$. Then there is a sequence (x_n) in U with no limit point in U and some functions $u_n \in B$ such that $\theta_n := |u_n(x_n)| > 0$. Let

$$V = \bigcap_{n\geq 1} \left\{ u \in C_c(U); |u(x_n)| < \frac{\theta_n}{n} \right\}.$$

If K is a compact, $K \subset U$, then K contains only a finite number of the points x_n so

$$i_K^{-1}(V) = V \cap C_K(U) = \bigcap_{n\in R} \left\{ u \in C_K(U); |u(x_n)| < \frac{\theta_n}{n} \right\},$$

where R is finite, is a finite intersection of open sets and is open in $C_K(U)$. So V is a neighbourhood of 0 and there is no $t \geq 0$ such that $B \subset tV$, since otherwise $\theta_n < t\frac{\theta_n}{n}$ for all n. The point *3.* follows from *2.* since the set

$$B = \{u_n; n \in \mathbb{N}\} \cup \{u\}$$

is bounded (See Section 5.2). Consider now a Cauchy sequence (u_n) in $C_c(U)$. It also defines a bounded set (See Section 5.2 also), so there is a compact K such that all the functions u_n are supported in K. Since $C(K)$ is complete, the sequence is convergent. □

Proposition 5.10 (Locally signed measures)**.** *The dual of $C_c(U)$ (endowed with the inductive topology) consists precisely in the space $\mathcal{M}_{\text{loc}}(U)$, which consists in differences of non-negative Radon measures: α is in the dual of $C_c(U)$ if, and only if, there exists two non-negative measures ρ and ν which are finite on compact subsets of U such that*

$$\alpha(u) = \int_U u d\rho - \int_U u d\nu, \quad \forall u \in C_c(U). \tag{5.40}$$

If realized, then we can assume that the measures ρ and ν in (5.40) are mutually singular, and that they are given by

$$\rho(A) = \mu(A \cap \{\sigma = +1\}), \quad \nu(A) = \mu(A \cap \{\sigma = -1\}), \quad A \in \mathcal{B}(U), \tag{5.41}$$

where μ is a non-negative Borel measure μ on U which is finite on the compact subsets of U and a Borel map $\sigma \colon U \to \{-1, +1\}$. In that case, we also have

$$\alpha(u) = \int_U u(x)\sigma(x) d\mu(x), \tag{5.42}$$

for all $u \in C_c(U)$.

Proof of Proposition 5.10. By definition of the inductive topology, α is in the topological dual of $C_c(U)$ if, and only if, $\alpha \circ i_K$ is a linear continuous functional on $C_K(U)$ for each compact $K \subset U$. This means that there exists $N_K \geq 0$ such that

$$|\alpha(u)| \leq N_K \max_{x\in K} |u(x)|, \quad u \in C_K(U). \tag{5.43}$$

It is clear that (5.43) is realized (with $N_K = \rho(K) + \nu(K)$) if (5.40) is satisfied. Conversely, assume that we have (5.43) for all compact $K \subset U$. We will admit that the version of the Riesz' theorem given in Theorem 2.3 remains valid when the space $\mathbb{R}^d$ is replaced by U (apply [Sim83, Theorem 4.1] with $X = U$, $H = \mathbb{R}$). We obtain that there is a non-negative Borel measure μ on U which is finite on the compact subsets of U and a Borel map $\sigma \colon U \to \{-1. +1\}$ such that (5.42) is satisfied for all $u \in C_c(U)$. We then define ρ and ν by (5.41) to obtain (5.40). □

Consider now the space $L_c^{p'}(U)$ discussed in *6.* of Theorem 5.9. We can associate at least two topologies to $L_c^{p'}(U)$:

- the "bounded-open" topology, since $L_c^{p'}(U)$ is a dual space,
- the inductive topology, induced by the injections $i_K\colon L_K^{p'}(U) \to L_c^{p'}(U)$, where, given a compact $K \subset U$, $L_K^{p'}(U)$ is the set of functions $v \in L^{p'}(U)$ such that $\|v\|_{L^{p'}(U\setminus K)} = 0$.

As an exercise, you can check that the two topologies coincide.

5.3.5 Separable spaces

Different spaces are considered in the previous section 5.3.3. Some are separable, some are not. We give the following results without proof.

Separable spaces: if K is compact, then $C(K)$ is separable (Use the Stone-Weierstrass theorem, or some other method...). This has the consequence that $C_0(\mathbb{R}^d)$ is separable, that $C(U)$ is separable, that $C^\infty(U)$ is separable (add a step of convolution), that the Schwartz space $\mathscr{S}(\mathbb{R}^d)$ is separable, that $L^p(U)$ is separable if $1 \leq p < +\infty$ (by truncature and regularization, any $u \in L^p(U)$ can be approximated in $L^p(U)$ by some continuous compactly supported functions), that $W^{1,p}(U)$ is separable if $1 \leq p < +\infty$.

Non-separable spaces: sets of measures $\mathcal{M}(X)$ endowed with the total variation norm are not separable in general since

$$\|\delta_x - \delta_y\| = |\delta_x - \delta_y|(X) = 1 \tag{5.44}$$

as soon as $x \neq y$. If A is a set which is dense in $\mathcal{M}(X)$, there is for each $x \in X$ an element $\mu \in A$ such that $\|\delta_x - \mu\| < 1/2$. By (5.44), this defines an injection of X in A, which cannot be countable if X is not. Similarly, we have $\|\mathbf{1}_A - \mathbf{1}_B\|_{L^\infty(X,\mu)} = 1$ if the symmetric difference $(A \setminus B) \cup (B \setminus A)$ has a positive measure, so $L^\infty(X,\mu)$ is in general not separable. In particular, $L^\infty(U)$ is not separable (consider a cube $Q(x_0) := x_0 + (0,\varepsilon)^d$ in U, and all the subcubes $Q(x) \cap Q(x_0)$ for $x \in Q(x_0)$ for instance).

Theorem 5.11 (Separable dual)**.** *Let X be a normed vector space. If X^* is separable, then X also.*

This theorem is sometimes used to establish that some spaces are not reflexive, *i.e.* $X \subsetneq X^{**}$ (see Definition 5.10 below). For instance, $L^1(U)$ is not reflexive since $L^\infty(U)$ is not separable, while $L^1(U)$ is. Similarly, if K is a compact subset of $\mathbb{R}^d$, then $C(K)$ is not reflexive since $\mathcal{M}(K)$ is not separable while $C(K)$ is.

Proof of Theorem 5.11. Let $\{\varphi_n \in \mathbb{N}\}$ be a countable dense family in X^*. For each n and $\varepsilon > 0$, there exists u_n in the closed unit ball of X such that $\|\varphi_n\|_{X_*} \leq |\varphi_n(u_n)| + \varepsilon$. We take $\varepsilon = \frac{1}{2}\|\varphi_n\|_{X_*}$, so that $\|\varphi_n\|_{X_*} \leq 2|\varphi_n(u_n)|$. Let M_0 be the $\mathbb{Q}$-vector space generated by the family $\{u_n; n \in \mathbb{N}\}$: it is the set of linear combinations $\sum_{n\in J} \lambda_n u_n$, where J is finite, $\lambda_n \in \mathbb{Q}$. Then M_0 is countable and dense in the closure of M, the $\mathbb{R}$-vector space generated by the family $\{u_n; n \in \mathbb{N}\}$. To conclude, it is sufficient to prove that $\bar{M}$, the closure of M is the whole space X. Assume not, and let $v \in E \setminus \bar{M}$. By the Hahn-Banach theorem (we use *1.* in Theorem 3.8), there is a $\varphi \in X^*$ such that $\varphi \equiv 0$ on $\bar{M}$ and $\varphi(v) \neq 0$. Let $\varepsilon > 0$. There exists $n \in \mathbb{N}$ such that $\|\varphi - \varphi_n\|_{X^*} < \varepsilon$. Since $\varphi(u_n) = 0$, we have then

$$\|\varphi\|_{X^*} < \varepsilon + \|\varphi_n\|_{X^*} \leq \varepsilon + 2|\varphi_n(u_n)| = \varepsilon + 2|(\varphi - \varphi_n)(u_n)| \leq \varepsilon + 2\|\varphi - \varphi_n\|_{X^*} < 3\varepsilon,$$

and we conclude that $\varphi = 0$: a contradiction. □

5.4 Weak-star topology

Definition 5.7 (weak-$*$ topology)**.** Let X be a topological vector space and let X^* be the set of continuous linear functionals on X. The weak-star topology on X^* (written weak-$*$) is the coarsest topology on X^* that makes all the evaluation maps $\pi_u\colon \varphi \mapsto \varphi(u)$ for $u \in X$ continuous.

We identify X with a subspace of X^{**} via the map $J\colon X \to X^{**}$ defined by $Ju = \pi_u$ (this an injective map, we have also seen in Theorem 3.8 that it is an isometry if X is a normed vector space):

$$Ju(\varphi) = \pi_u(\varphi) = \varphi(u), \quad u \in X, \varphi \in X^{**}. \tag{5.45}$$

Then X is separating points in X^*. By Remark 5.1, the weak-$*$ topology on X^* is the X-weak topology on X^*. Endowed with this topology, X^* is a locally convex topological vector space. A neighbourhood base of the origin is given by the finite intersections of the sets

$$V_{u,\varepsilon} = \{\varphi \in X^*; |\varphi(u)| < \varepsilon\}. \tag{5.46}$$

A continuous linear function on weak-$*$ X^* is of the form π_u for a given $u \in X$. If $J(X)$ is a strict subset of X^{**} (*i.e.* X is not reflexive, *cf.* Definition 5.10), then consider a $T \in X^{**} \setminus J(X)$ and the space

$$H = \{\varphi \in X^*; T(\varphi) = 0\}.$$

Then H is a closed hyperplane of X^*. In particular, it is a closed convex subset of X^*. It is not weak-$*$ closed however, otherwise Theorem 3.5 would imply that there exists a weak-$*$ continuous linear form S on X^* such that $H = \operatorname{Ker}(S)$. But then T is proportional to S, and $S = Ju$ for a given $u \in X$, so $T \in J(X)$: contradiction.

Theorem 5.12 (Banach-Alaoglu)**.** *Let X be a topological space and V a neighbourhood of the origin in X. Then*

$$K = \{\varphi \in X^*; \forall u \in V, |\varphi(u)| \leq 1\} \tag{5.47}$$

is compact for the weak-$*$ topology.

When X is a normed vector space, the theorem is often applied to $V = \bar{B}(0,1)$. It says then that *the unit ball of X^* is compact for the weak-$*$ topology.*

Proof of Theorem 5.12. We begin with a first observation on the weak-$*$ topology on X^*: it is generated by the sets

$$\{\varphi \in X^*; \pi_u(\varphi) \in U\} = X^* \cap \{\varphi\colon X \to \mathbb{R}; \pi_u(\varphi) \in U\},$$

where U is open in $\mathbb{R}$. The weak-$*$ topology on X^* is therefore the trace on X^* of the product topology on $\mathbb{R}^X$. The linearity conditions

$$\varphi(u+v) = \varphi(u) + \varphi(v), \quad \varphi(\lambda u) = \lambda\varphi(u)$$

can be written as

$$(\pi_{u+v} - \pi_u - \pi_v)(\varphi) = 0, \quad (\pi_{\lambda u} - \lambda\pi_u)(\varphi) = 0.$$

They are closed conditions in $\mathbb{R}^X$, so the set $L(X,\mathbb{R})$ of linear maps $X \to \mathbb{R}$ is closed for the product topology in $\mathbb{R}^X$. Similarly, the set

$$F = \{\varphi \in \mathbb{R}^X; \forall u \in V, |\varphi(u)| \leq 1\} = \bigcap_{u \in V} \pi_v^{-1}([-1,1]) \tag{5.48}$$

is closed in $\mathbb{R}^X$. Moreover, an element in $L(X,\mathbb{R})\cap F$ is weakly-$*$ continuous, so in X^*. Indeed, if $\varphi \in F$ is linear and $\varepsilon > 0$, then $V \subset \varphi^{-1}([-1,1])$, so $(\varepsilon/2)V \subset \varphi^{-1}((-\varepsilon,\varepsilon))$. It follows that $K = L(X,\mathbb{R})\cap F$ and that K is closed for the product topology. If $K \subset K'$ where K' is compact in $\mathbb{R}^X$, then we conclude that K is compact for product topology, so for the weak-$*$ topology. To exhibit K', fix $u \in X$. The map $\lambda \mapsto \lambda u$ is continuous and takes the value 0 at 0, so there exists $\lambda_u > 0$ such that $\lambda u \in V$ for all $\lambda \in (-\lambda_u, \lambda_u)$. It follows that $|\varphi(u)| \le M_u := 2\lambda_u^{-1}$ if $\varphi \in K$. We obtain

$$K \subset K', \quad K' = \prod_{u\in X}[-M_u, M_u].$$

The set K' is compact for the product topology by Tychonov's theorem. This concludes the proof. □

Let us give more details in the case where X is separable. We will see that we do not need Tychonov's theorem (nor the Axiom of Choice thus) to establish Theorem 5.12. First, we give the following important result.

Theorem 5.13 (Metrizability of polar sets for the weak-star topology)**.** *Let X be a topological space and V a neighbourhood of the origin in X. Assume that X is separable. Then*

$$K = \{\varphi \in X^*; \forall u \in V, |\varphi(u)| \le 1\} \tag{5.49}$$

is metrizable for the weak-$*$ topology.

Proof of Theorem 5.13. We want to prove, more precisely, that the trace of the weak-$*$ topology $\mathcal{T}_{w*}$ on K is metrizable. Let $\{u_n; n \in \mathbb{N}\}$ be a dense countable set in X and let d be the metric defined by (compare with (4.4))

$$d(\varphi,\psi) = \sum_{n\in\mathbb{N}} 2^{-n}\Phi(p_n(\varphi-\psi)), \quad \Phi(p) := \frac{p}{1+p} \quad p_n(\varphi) := |\varphi(u_n)|, \tag{5.50}$$

where $\varphi,\psi \in X^*$. Since $\{u_n; n\in\mathbb{N}\}$ is dense, the semi-norms p_n are separating, so d is a metric on K. We employed the same notations as in the proof of Theorem 4.1 on purpose. Indeed, we can deduce from (the proof of) Theorem 4.1 that the topology $\mathcal{T}_d$ associated to the metric (5.50) is the topology generated by the sets $\mathcal{V}_{n,\varepsilon} = \{\varphi \in X^*; p_n(\varphi) < \varepsilon\}$ and their translates. Since each p_n is weakly-$*$ continuous, each set $\mathcal{V}_{n,\varepsilon}$ is weakly-$*$ open, *i.e.* $\mathcal{T}_d \subset \mathcal{T}_{w*}$. Now, let $\mathcal{U}$ be open for the weak-$*$ topology. Let $\psi \in \mathcal{U}$. There is a $\varepsilon > 0$ and some elements $v_1,\cdots,v_n \in X$ such that

$$\psi + \bigcap_{j=1}^{n} \mathcal{V}_{v_j,\varepsilon} \subset \mathcal{U}, \quad \mathcal{V}_{v,\varepsilon} = \{\varphi \in X^*, |\varphi(v)| < \varepsilon\}.$$

For each $j \in \{1,\dots,n\}$, there exists N_j such that

$$v_j - u_{N-j} \in (\varepsilon/3)V, \quad |\psi(v) - \psi(u_{N_j})| < \varepsilon/3. \tag{5.51}$$

Let $\varphi \in K \cap (\psi + \mathcal{V}_{N_j,\varepsilon/3}$: $\varphi = \psi + \theta$, $|\theta(u_{N_j})| < \varepsilon/3$. Then the decomposition

$$\theta(v_j) = \varphi(v_j) - \psi(v_j) = \theta(u_{N_j}) + (\varphi(v_j) - \varphi(u_{N_j})) - (\psi(v_j) - \psi(u_{N_j})),$$

and the fact that $|\varphi(v_j - u_{N_j})| \le \varepsilon/3$ since $\varphi \in K$ show that $\theta \in \mathcal{V}_{v_j,\varepsilon}$. It follows that

$$K \cap \mathcal{W} \subset K \cap \mathcal{U}, \quad \mathcal{W} := \psi + \bigcap_{j=1}^{n} \mathcal{V}_{u_{N_j},\varepsilon/3}.$$

Since $\mathcal{W} \in \mathcal{T}_d$, we deduce that the traces on K of $\mathcal{T}_d$ and $\mathcal{T}_{w*}$ coincide. This completes the proof. □

Proof of Theorem 5.12 in the case where X is separable. By Theorem 5.13 (the weak-$*$ topology on K is metrizable), it is sufficient to consider a sequence (φ_m) in K and to show that it admits a subsequence which is convergent for the metric d defined in (5.50). For all n, there exists $t_n \geq 0$ such that $u_n \in t_n V$ (same reasoning as in the proof of Theorem 5.12). Then $|\varphi_m(u_n)| \leq t_n$ since $\varphi_m \in K$. By a diagonal argument, there is a subsequence still denoted (φ_m) such that $(\varphi_m(u_n))$ is convergent (and therefore Cauchy), for all n. If $u \in X$, then for all $\varepsilon > 0$, there exists N such that $u - u_N \in \varepsilon V$ and m_0 such that $|\varphi_p(u_N) - \varphi_q(u_N)| < \varepsilon$ for $p, q \geq m_0$. Then $|\varphi_p(u) - \varphi_q(u)| < 2\varepsilon$ for $p, q \geq m_0$, so $(\varphi_m(u))$ is Cauchy. We denote by $\varphi(u)$ the limit of $\varphi_m(u)$. Then φ is the limit of (φ_m) for the product topology and, as in the proof of Theorem 5.12, we can show that φ is linear and $\varphi \in F$ (defined in (5.48)), so $\varphi \in K$. □

5.4.1 Some applications of the Banach-Alaogu theorem

Weak convergence of measures. Let K be a compact subset of $\mathbb{R}^d$. The space $X = C(K)$ is separable. Let (μ_n) be a sequence of signed measures which is bounded for the total variation norm: there exists $M \geq 0$ such that $|\mu_n|(K) \leq M$ for all $n \in \mathbb{N}$. Then (μ_n) is a sequence in

$$K = \{\mu \in C(K)^*; \forall u \in V, |\mu(u)| \leq 1\}, \quad V := \{u \in C(K); \|u\|_{C(K)} < M^{-1}\}.$$

By the Banach-Alaoglu theorem and the fact that the weak-star topology on K is metrizable, we deduce that there is a subsequence of (μ_n) still denoted (μ_n) and a signed measure μ on K such that

$$\forall u \in C(K), \int_K u d\mu_n \to \int_K u d\mu, \tag{5.52}$$

when $n \to +\infty$.

Weak convergence in $L^q(U)$, $1 < q \leq +\infty$. Let U be an open set in $\mathbb{R}^d$. Let $q \in (1, +\infty]$. Then $L^q(U)$ is the dual of $L^p(U)$, p being the conjugate exponent to q. If (u_n) is a sequence bounded in $L^q(U)$, then there exists a subsequence of (u_n) still denoted (u_n) and a $u \in L^q(U)$ such that

$$\forall v \in L^p(U), \int_U u_n v dx \to \int_U u v dx, \tag{5.53}$$

when $n \to +\infty$.

Weak convergence in $W^{1,q}(U)$, $1 < q \leq +\infty$. The arguments will be given in details later, when we study Sobolev spaces, but we can already mention the following result. Let U be an open set in $\mathbb{R}^d$. Let $q \in (1, +\infty]$. If (u_n) is a sequence bounded in $W^{1,q}(U)$, then there exists a subsequence of (u_n) still denoted (u_n) and a $u \in W^{1,q}(U)$ such that we have the strong convergence in of the functions:

$$u_n \to u \text{ in } L^q(U), \tag{5.54}$$

and the weak convergence of the derivatives:

$$\forall v \in L^p(U), \int_U u_n v dx \to \int_U u v dx, \quad \int_U \partial_{x_i} u_n v dx \to \int_U \partial_{x_i} u v dx, \tag{5.55}$$

when $n \to +\infty$.

The case $q = 1$. Let U be an open bounded subset of $\mathbb{R}^d$. To a sequence (u_n) bounded in $L^1(U)$, we associate the sequence of signed measures μ_n on $K = \bar{U}$ given by

$$\mu_n(A) = \int_{A \cap U} u_n dx, \quad \text{i.e.} \int_K v d\mu_N = \int_U v u_n dx,$$

for all $v \in C(K)$. Then (μ_n) is bounded in $\mathcal{M}(K)$: we deduce that there exists a signed measure μ on K and a subsequence of (u_n) still denoted (u_n) such that

$$\forall v \in C(K), \int_U u_n v dx \to \int_K v d\mu, \tag{5.56}$$

when $n \to +\infty$. For instance (5.56) is the mode of convergence $\rho_n \to \delta_0$, when (ρ_n) is an approximation of the unit. If U is of class C^1, considering a sequence u_n which concentrates on the boundary ∂U can be used to defined a "surface" measure on $\partial U \subset \bar{U}$ (see the proof of the Green Formula in the course "Analyse et EDP" last year, or (A.4) and the paragraph below).

Weak convergence of measures - 2. (To be skipped at first reading) In probability theory, different modes of convergence of random variables are considered. Let E be a metric space, and (X_n) a sequence of random variables on E: there is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that all the functions $X_n \colon \Omega \to E$ are measurable, where the σ-algebra one considers on E is the Borel σ-algebra. One says that the sequence (X_n) *converges in law* (or *in distribution*) to a random variable X on E if

$$\mathbb{E}\left[u(X_n)\right] \to \mathbb{E}\left[u(X)\right], \tag{5.57}$$

for all continuous and bounded functions $u \colon E \to \mathbb{R}$ ($u \in \mathrm{BC}(E)$ with the notations of Section 5.3.2). Let μ_n and μ denote the law of X_n and X respectively. Then (5.57) means that

$$\langle u, \mu_n \rangle \to \langle u, \mu \rangle, \tag{5.58}$$

for all $u \in \mathrm{BC}(E)$, where

$$\langle u, \mu \rangle := \int_E u d\mu.$$

The situation is the following one: we are given a sequence of Borel probability measures on E and we are wondering if the weak-$*$ convergence[5] (5.58) occurs. There is a fundamental criterion related to this question.

Definition 5.8 (Tight sequence of probability measures)**.** Let E be a metric space, and let (μ_n) be a sequence of Borel probability measures on E. The sequence (μ_n) is said to be *tight* if, for all $\varepsilon > 0$, there exists a compact $K \subset E$ such that $\mu_n(K) \geq 1 - \varepsilon$ for all n.

Proposition 5.14 (Tight probability measure)**.** *Suppose that the metric space E is separable and complete. Then a single probability measure on E is tight.*

Proof of Proposition 5.14. Let μ be a Borel probability measure on E. Let A be a dense countable subset of E and let (r_k) be a sequence decreasing to 0. For each k, the space E is covered by the balls $B(x, r_k)$, $x \in A$. There is a finite set $A_k \subset A$ such that

$$\mu(G_k) > 1 - \frac{\epsilon}{2^k}, \quad G_k := \bigcup_{x \in A_k} B(x, r_k).$$

Let K denote the closure of $G = \bigcap_{k \in \mathbb{N}} G_k$. Then K is closed and totally bounded in E complete, so is compact, and we have $\mu(K) \geq \mu(G) > 1 - \varepsilon$. □

Theorem 5.15 (Prokhorov)**.** *Let E be a metric space, and let (μ_n) be a sequence of Borel probability measures on E. If (μ_n) is tight, then there is a subsequence still denoted (μ_n), which is convergent in the sense of (5.58). If the metric space E is separable and complete, and (μ_n) satisfies (5.58), then (μ_n) is tight.*

[5] this is indeed a weak-$*$ convergence, because *1.* every probability measure on a metric space is regular, [Bil99, Theorem 1.1], *2.* a metric space is a normal space, *3.* the topological dual of $\mathrm{BC}(E)$ is $rca(E)$, *cf.* Section 5.3.2

See [Bil99], Theorem 5.1 and Theorem 5.2.

Exercise 5.9. The aim of this exercise is to prove the useful implication of Prokhorov's theorem in the case $E = \mathbb{R}^d$. So we consider a sequence of Borel probability measures on $\mathbb{R}^d$ which is tight.

1. Show that it is sufficient to establish (5.58) for $u \in C_0(\mathbb{R}^d)$.

   If (5.58) is satisfied for every $u \in C_0(\mathbb{R}^d)$ and $u \in \mathrm{BC}(\mathbb{R}^d)$, we can show that every member in (5.58) can be approached with an arbitrary precision by similar terms with a test-function $\tilde{u} \in C_0(\mathbb{R}^d)$. For $\varepsilon > 0$, consider a compact K such that all the μ_n are supported in K, up to ε and (using Proposition 5.14) a compact K' such that $\mu(K') > 1 - \varepsilon$. Replacing K by K' (and ε by 2ε) if necessary, we can assume that $K = K'$. Then we set $\tilde{u} = u\tilde{\chi}$, where $\tilde{\chi}$ is a truncate function $\tilde{\chi} \in C_c(\mathbb{R}^d)$ such that $\tilde{\chi} \equiv 1$ on K.

2. Prove the result.

   We have $|\mu_n|(\mathbb{R}^d) = \mu_n(\mathbb{R}^d) = 1$ for all n. By the duality between $C_0(\mathbb{R}^d)$ and the space of signed Borel measures on $\mathbb{R}^d$, and Banach-Alaoglu's theorem, we deduce that there exists a signed measure μ and a subsequence still denoted (μ_n) such that (5.58) is satisfied for every $u \in C_0(\mathbb{R}^d)$. By 1., it is sufficient to show that μ is a probability measure to conclude. Let us first prove that μ is a non-negative measure. By Proposition 2.2, there is a Borel set A_- such that μ^- is concentrated on A_-. Assume by contradiction that $\mu(A_-) < 0$. By regularity of μ (cf. Remark 2.1), there is, for a given $\varepsilon > 0$, a compact set K and open set U such that $K \subset A_- \subset U$ and $|\mu|(U \setminus K) < \varepsilon$. Let u be a continuous function such that $K \prec u \prec U$. By (5.58) and the fact that $u \geq 0$, $\mu_n \geq 0$, we have $\langle u, \mu \rangle \geq 0$. At the same time,

$$\langle u, \mu \rangle = \int_{A_+} u d\mu^+ - \int_{A_-} u d\mu^- \leq \mu_+(U) - \mu_-(K) \leq \mu(A) + 2\varepsilon$$

   is strictly negative if ε is small enough, and this is a contradiction. So $\mu \geq 0$. Let now $m = \mu(\mathbb{R}^d)$. We can write $\mu = m\tilde{\mu}$, where $\tilde{\mu}$ is a probability measure. An easy generalization of the point 1. shows then that (5.58) holds true when u is a bounded and continuous function. Taking $u \equiv 1$, we obtain $m = 1$.

5.4.2 Reflexive spaces

Definition 5.10 (Reflexive space)**.** A topological vector space X is said to be reflexive if the map J defined in (5.45) is surjective: $J(X) = X^{**}$.

Let E be a reflexive Banach space. Then we can "transfer" the result of the Banach-Alaoglu theorem, applied to E^{**} endowed with the weak-$*$ topology defined by E^*, to E (also endowed with the weak-$*$ topology defined by E^*) to deduce that the closed unit ball $\bar{B}_E(0,1)$ is compact for the weak topology (the Lesbesgue spaces $L^p(U)$ for $1 < p < +\infty$ are reflexive and we can already observe the transfer argument in the case of the spaces $L^q(U)$ detailed above, *cf.* (5.53)). More precisely, let $J\colon E \to E^{**}$ be the isometry defined in (5.45). The inclusion $J(\bar{B}_E(0,1)) \subset \bar{B}_{E^{**}}(0,1)$ becomes an equality when E is reflexive. By Banach-Alaoglu's Theorem, the closed ball $\bar{B}_{E^{**}}(0,1)$ is compact for the weak-$*$ topology $\sigma(E^{**}, E^*)$, so we want to prove that $J^{-1}\colon E^{**} \to E$ is continuous, when E^{**} has the topology $\sigma(E^{**}, E^*)$ and E has the weak topology $\sigma(E, E^*)$. By definition of the topology $\sigma(E, E^*)$, we must check that each map $\varphi \circ J^{-1}$ is continuous, where $\varphi \in E^*$. For such a φ, for $\xi = J(u) \in E^{**}$, we have

$$\varphi \circ J^{-1}(\xi) = \varphi(u) = \xi(\varphi),$$

so $\varphi \circ J^{-1} = \pi_\varphi$ has the desired continuity property. Actually, we can state the following result.

Theorem 5.16 (Kakutani)**.** *Let E be Banach space. Then E is reflexive if, and only if, the closed unit ball $\bar{B}_E(0,1)$ is weakly compact.*

We will not prove Kakutani's theorem (see [Bre11, Theorem 3.17]). We will not prove the following result either.

Theorem 5.17 (Eberlein-Šmulian)**.** *Let E be Banach space. Regarding the weak topology, it is equivalent for a set $A \subset E$ to be compact or sequentially compact.*

By sequentially compact, we mean that every sequence in A admits a (weakly) convergent subsequence. See exercises class for the proof of the implication (compact) $\Rightarrow$ (sequentially compact), which is the most useful one for us. If needed, a reference for the complete proof is [DS58, V.6.1]. Using Eberlein-Šmulian's Theorem in particular, we can prove the following result.

Theorem 5.18 (Weakly compact sets)**.** *Let E be a reflexive Banach space. Let K be a bounded, closed and convex subset of E. Then K is compact and sequentially compact for the weak topology.*

Proof of Theorem 5.18. By Theorem 5.3, the set K is weakly closed. By Banach-Alaoglu's Theorem and the discussion that precedes Kakutani's Theorem, the ball $\bar{B}_E(0,r)$ $(r>0)$ is weakly compact. Since $K \subset \bar{B}_E(0,r)$ for a $r>0$, we deduce that K is weakly compact. By Eberlein-Šmulian's Theorem (the "easy implication"), K is also weakly sequentially compact. □

We insist on this fact: note well that we do not need the space E to be separable to obtain the sequential compactness in Theorem 5.18 (but separability is involved for good in the proof, by the fact that the closure of the vector space generated by a countable family of vectors is separable, see Exercises class).

5.4.3 Milman-Pettis Theorem

Our aim in this section is to prove the following result

Theorem 5.19 (Milman-Pettis)**.** *A uniformly convex Banach space is reflexive.*

We need some preliminary results first.

Lemma 5.20 (Helly)**.** *Let E be a Banach space, let d be a non-negative integer, and let $\Psi\colon E \to \mathbb{R}^d$ a continuous linear map and $x \in \mathbb{R}^d$. There is equivalence between*

1. *x is in the (strong) closure of $\Psi(\bar{B}_E(0,1))$,*
2. *for all $y \in \mathbb{R}^d$, we have*
$$|x \cdot y| \leq \|\Psi \cdot y\|_{E^*}, \tag{5.59}$$
where $x \cdot y = \sum_{i=1}^d x_i y_i$ is the canonical scalar product on $\mathbb{R}^d$.

Proof of Lemma 5.20. It is clear that *1.* implies *2.* if, additionally, $x \in \Psi(\bar{B}_E(0,1))$. By continuity, we obtain the general case. Next, denote by A the closure of $\Psi(\bar{B}_E(0,1))$ and suppose that $x \notin A$. Since A is closed, convex, non-empty, we can separate strictly A and x by an affine hyperplane: there exists $y \in \mathbb{R}^d$, $\gamma \in \mathbb{R}$, $\varepsilon > 0$, such that

$$a \cdot y < \gamma < \gamma + \varepsilon < x \cdot y, \tag{5.60}$$

for all $a \in A$. Since $0 = \Psi(0) \in A$, we have $\gamma > 0$. Taking $a = \Psi(u)$, $u \in \bar{B}_E(0,1)$, we deduce from (5.60) that $\|\Psi \cdot y\|_{E^*} < |x \cdot y|$. □

Lemma 5.21 (Goldstine)**.** *Let E be a Banach space, and let $J\colon E \to E^{**}$ be the isometry defined in* (5.45)*. Then $J(\bar{B}_E(0,1))$ is dense in $\bar{B}_{E^{**}}(0,1)$ for the weak-$*$ topology on E^{**}.*

Proof of Lemma 5.21. Let $\xi \in \bar{B}_{E^{**}}(0,1)$ and let V be a neighbourhood of 0 for the weak-$*$ topology on E^{**}. We want to show that $\xi + V$ intersects $J(\bar{B}_E(0,1))$. We can assume that V has the form

$$V = \bigcap_{i=1}^{d} \{\zeta \in E^{**}, |\zeta(\varphi_i)| < \varepsilon\},$$

where $\varepsilon > 0$ and $\varphi_1, \dots, \varphi_d \in E^*$. Let $\Psi(u) = (\varphi_1(u))_{1,d}$ and $x = (\xi(\varphi_i))_{1,d}$. Then $\Psi\colon E \to \mathbb{R}^d$ is linear and continuous. By linearity of ξ, and since $\|\xi\|_{E^{**}} \leq 1$, we have, for all $y \in \mathbb{R}^d$,

$$|x \cdot y| = |\xi(\Psi \cdot y)| \leq \|\xi\|_{E^{**}} \|\Psi \cdot y\|_{E^*} \leq \|\Psi \cdot y\|_{E^*}.$$

By Helly's Lemma, we deduce that x is in the closure of $\Psi(\bar{B}_E(0,1))$. So, ε being positive, there exists $u \in \bar{B}_E(0,1)$ such that $|x_i - \varphi_i(u)| < \varepsilon$ for all i. This means precisely that $Ju \in \xi + V$, so $\xi + V$ intersects $J(\bar{B}_E(0,1))$. □

Proof of Theorem 5.19. Recall that uniform convexity has been defined in Section 3.2.4. We want to prove that the inclusion $J(E) \subset E^{**}$ is an equality or, equivalently, that the inclusion $J(\bar{B}_E(0,1)) \subset \bar{B}_{E^{**}}(0,1)$ is an equality. Since J is an isometry and $\bar{B}_E(0,1)$ is complete, $J(\bar{B}_E(0,1))$ is complete and therefore closed in E^{**}. It is sufficient to prove that, given $\xi \in \bar{B}_{E^{**}}(0,1)$ and $\varepsilon > 0$, there exists $u \in \bar{B}_E(0,1)$ such that $\|Ju - \xi\|_{E^{**}} \leq \varepsilon$. If $\xi = 0$, then there is nothing to prove so, up to a rescaling procedure, we can assume that $\|\xi\|_{E^{**}} = 1$. To the modulus $\varepsilon > 0$, we can associate a $\delta > 0$ such that the criterion of uniform convexity (3.62) is satisfied. Assume by contradiction that, for all $u \in \bar{B}_E(0,1)$, we have

$$\xi \in W_u := E^{**} \setminus \bar{B}_{E^{**}}(Ju, \varepsilon).$$

The set W_u is open in E^{**} endowed with the weak-$*$ topology because $\bar{B}_{E^{**}}(Ju, \varepsilon) = Ju + \varepsilon \bar{B}_{E^{**}}(0,1)$ and $\bar{B}_{E^{**}}(0,1)$ is closed for the weak-$*$ topology. By Goldstine's lemma, there exists $v \in \bar{B}_E(0,1)$ such that $Jv \in W_u$. This means that

$$\|u - v\|_E = \|Ju - Jv\|_{E^{**}} > \varepsilon. \tag{5.61}$$

To get a contradiction, we need u and v to have the additional property

$$\|u + v\|_E \geq 2(1 - \delta). \tag{5.62}$$

We exploit the duality: we know that $\|u + v\|_E \geq \varphi(u + v)$ if $\varphi \in \bar{B}_{E^*}(0,1)$. So (5.62) will be realized if $\varphi(u) \geq 1 - \delta$ and $\varphi(v) \geq 1 - \delta$ for a given $\varphi \in \bar{B}_{E^*}(0,1)$. We proceed as follows: since $1 = \|\xi\|_{E^{**}}$, there exists $\varphi \in \bar{B}_{E^*}(0,1)$ such that $\xi(\varphi) \geq 1 - \delta/2$. Consider the neighbourhood of the origin (for the weak-$*$ topology)

$$V = \{\zeta \in E^{**}; |\zeta(\varphi)| < \delta/2\}.$$

By Goldstine's lemma, there exists $u \in \bar{B}_E(0,1)$ such that $Ju \in \xi + V$. In particular, $\varphi(u) \geq 1 - \delta$. We then correct the choice of v made above by taking $v \in \bar{B}_E(0,1)$ such that, not only $Jv \in W_u$ but also $Jv \in \xi + V$. Such a v exists by Goldstine's lemma since $\xi \in W_u$, so $W_u \cap (\xi + V)$ is a neighbourhood of ξ for the weak-$*$ topology. With this choice of v, we have simultaneously (5.61) and (5.62), which is a contradiction. □

5.4.4 Minimization of functionals

Let E be a Banach space, $I\colon E \to \mathbb{R}$ a given functional and A a subset of E. We consider the problem of minimizing I on A. Here are few standards examples considered in the calculus of variation (there are more examples related to the minimization of the Dirichlet functional, there are also a lot of different examples furnished by the study of functional inequalities, or being related to some questions of "shape optimization").

1. Minimization of the Dirichlet functional

$$I[u] = \frac{1}{2}\int_U |\nabla u|^2 dx - \int_U wudx, \quad A = E = W_0^{1,2}(U). \tag{5.63}$$

   Here U is a bounded open subset of $\mathbb{R}^d$, $w \in L^2(U)$ and $W_0^{1,2}(U)$ is the Sobolev space of functions that vanish on ∂U.

2. Minimization of the area functional

$$I[u] = \int_U \sqrt{1+|\nabla u|^2} dx - \int_U wudx, \quad A = E = W_0^{1,1}(U). \tag{5.64}$$

   Here U is a bounded open subset of $\mathbb{R}^d$, $w \in L^\infty(U)$ and $W_0^{1,1}(U)$ is the Sobolev space of functions that vanish on ∂U.

3. Non-linear eigenvalues: minimization of the Dirichlet functional

$$I[u] = \frac{1}{2}\int_U |\nabla u|^2 dx - \int_U wudx, \quad E = W_0^{1,2}(U), \quad A = \{u \in E; J(u) = 0\}, \tag{5.65}$$

   where $J\colon L^2(U) \to \mathbb{R}$ is continuous and convex.

Consider now the following hypotheses.

1. The functional I is continuous, convex and coercive on E: there exists $\alpha > 0$ and $M \geq 0$ such that

$$I[u] \geq \alpha\|u\|_E - M, \tag{5.66}$$

   for all $u \in E$.

2. The set A is non-empty, closed and convex.

3. The set E is reflexive.

Proposition 5.22 (Minimization of functional). *Assume that I, E, A satisfy 1.-2.-3. Then the functional I admits a minimum on A.*

Proof of Proposition 5.22. Since $I[u]$ is bounded from below by (5.66), we have

$$I_* := \inf_{u \in A} I[u] > -\infty.$$

Let (u_n) be a minimizing sequence: $u_n \in A$, $I[u_n] \to I_*$. Given $v \in A$, we have $I[u_n] \leq I[v] + 1$ for n large enough, so, by the coercivity condition (5.66),

$$\|u_n\|_E \leq C := \alpha^{-1}(I[v] + M + 1).$$

We conclude that the sequence (u_n) is *bounded*. By Kakutani's theorem (the easy implication), there is a subsequence still denoted (u_n) which converges weakly to a given $u_* \in E$. As A is closed and convex, it is weakly closed, so $u_* \in A$. As I is continuous (so lower semi-continuous) and convex, it is weakly lower semi-continuous. Consequently

$$I[u_*] \leq \liminf_{n\to+\infty} I[u_n] = \lim_{n\to+\infty} I[u_n] = \inf_{u\in A} I[u].$$

This shows that the infimum is attained. □

Proposition 5.22 can be applied to the examples (5.63) and (5.65), but not to (5.64), since $W_0^{1,1}(U)$ is not a reflexive space. To minimize the area functional by such a direct approach, one has to work in the space of functions of bounded variation, [Giu84]. See the discussion on the case $q = 1$ in Section 5.4.1.

6 Distribution theory

Let U be an open subset of $\mathbb{R}^d$. The theory of "distributions" elaborated by Laurent Schwartz gives a suitable framework to work on "generalized" functions defined on U and to extend some standard operations, like differentiation, composition by diffeomorphism, multiplication by smooth functions, translation... from a well-known class of functions to the class of distributions. Here are some examples of distributions (the justification that these are indeed distributions will be given later, once distributions have been defined, or in exercises class).

1. A function in $C^k(U)$ $(k \geq 0)$.
2. A function in $L^1_{\mathrm{loc}}(U)$.
3. A measure μ on U which is finite on compact subsets of U.
4. The "principal value" p.v. $\left(\frac{1}{x}\right)$ $(d = 1, U = \mathbb{R})$.

If α is one of the three distributions of examples *1.*, *2.*, *3.* above and u a smooth, compactly supported test-function on U, then we know how to give a meaning to define an action $\langle \alpha, u\rangle$ using the theory of integration. In the fourth case *4.* we also exploit the averaging effect of integration to give a meaning to $\langle \alpha, u\rangle$. More precisely, on can check that, for $u \in C_c^\infty(\mathbb{R})$, the three following quantities

$$\lim_{\varepsilon\to 0}\int_{|x|>\varepsilon} \frac{u(x)}{x}dx, \quad \int_{\mathbb{R}} \frac{u(x)-u(0)}{x}\chi(x)dx, \quad \int_{\mathbb{R}}\int_0^1 u'(tx)\chi(x)dtdx \tag{6.1}$$

are equal. The function χ in (6.1) is a "bump function": a function $\chi \in C_c^\infty(\mathbb{R})$ such that $0 \leq \chi \leq 1$ and such that the sets $\{\chi = 1\}$ and $\{\chi = 0\}$ have convenient properties. Here we assume that $\{\chi = 1\}$ contains the support of u and we also assume assume that χ is an even function. Any of the quantities in (6.1) can be used thus to define the action $\langle \alpha, u\rangle$ for $\alpha :=$ p.v. $\left(\frac{1}{x}\right)$.

We have mentioned some operations that can be defined in a convenient way for distributions. There are some other operations that are of great interest and defined without too much difficulties: localization, definition of a support. Some important operations require more care to be defined: convolution, Fourier transform, tensor product for instance. The definition of the product of two distributions is a very delicate question.

6.1 Distribution: definition, elementary operations

6.1.1 Space of distributions

Recall first that the support of a function, as we consider it, is the *closed* support: if U is an open set of $\mathbb{R}^d$ and $v \in C(U)$, the support $\operatorname{supp}(v)$ of a continuous function v is defined as the *closure* (in U) of the set $\{x \in U; v(x) \neq 0\}$. If we set $\mathcal{Z}_v = U \setminus \operatorname{supp}(v)$, then $\mathcal{Z}_v$ is the interior of $\{x \in U; v(x) = 0\}$: the largest open set on which $v = 0$. The space $C_c^\infty(U)$, sometimes also denoted $\mathcal{D}(U)$, is the set of infinitely differentiable, compactly supported functions on U. If K is a compact subset of U, the set $C_K^\infty(U)$ (or $\mathcal{D}_K(U)$) is the subset of $\mathcal{D}(U)$ consisting in the functions compactly supported in K. Each space $\mathcal{D}_K(U)$ is a Fréchet space with family of semi norms

$$p_{K,j}(u) = \sup_{|m|\leq j} \sup_{x\in K} |\partial_x^m u(x)|,$$

where $m \in \mathbb{N}^d$, $|m| = m_1 + \cdots + m_d$, $\partial_x^m = \partial_{x_1}^{m_1} \cdots \partial_{x_d}^{m_d}$. On $\mathcal{D}(U)$, we consider the inductive topology associated to the injections $i_K : \mathcal{D}_K(U) \to \mathcal{D}(U)$. As in Section 5.3.4, where we considered the inductive topology on $C_c(U)$, it is possible to prove the following facts: *a set B is bounded in $\mathcal{D}(U)$ if*

1. *there exists a compact $K \subset U$ such that all the functions $u \in B$ are supported in K,*
2. *for all $j \geq 0$, there exists $M_j \geq 0$, such that $p_{K,j}(u) \leq M_j$ for all $u \in B$.*

We also have the following fact: *a sequence (u_n) in $\mathcal{D}(U)$ converges to an element $u \in \mathcal{D}(U)$ if*

1. *there exists a compact $K \subset U$ such that all the functions u_n and u are supported in K,*
2. *u_n converges uniformly to u on K and all the derivatives $\partial_x^m u_n$, $m \in \mathbb{N}^d$ converge uniformly to $\partial_x^m u$ on K.*

The space of distributions $\mathcal{D}'(U)$ is by definition the dual space of $\mathcal{D}(U)$. By definition of the inductive topology, a linear functional $\alpha : \mathcal{D}(U) \to \mathbb{R}$ is continuous if, and only if, all the linear functionals $i_K \circ \alpha : \mathcal{D}_K(U) \to \mathbb{R}$ are continuous. This means that *α is a distribution on U if, for all compact $K \subset U$, there exists $j \in \mathbb{N}$ and a constant $C_{K,j}$ such that*

$$|\langle \alpha, u \rangle| \leq C_{K,j} \sup_{|m|\leq j} \sup_{x\in K} |\partial_x^m u(x)|, \tag{6.2}$$

for all $u \in \mathcal{D}_K(U)$. In (6.2), we have used the notation $\langle \alpha, u \rangle$ for $\alpha(u)$, traditional notation for the duality $\mathcal{D}'(U) - \mathcal{D}(U)$ that we will keep all along.

Let us come back to the examples considered previously. To a function $u \in L^1_{\mathrm{loc}}(U)$, we associate α_u defined by

$$\langle \alpha_u, w \rangle = \int_U u(x) w(x) dx. \tag{6.3}$$

Then α_u satisfies (6.2) with $j = 0$:

$$|\langle \alpha_u, w \rangle| \leq C_K \sup_{x\in K} |w(x)|, \quad \forall w \in \mathcal{D}_K(U),$$

with $C_K = \|u\|_{L^1(K)}$, so α_u is a distribution. The map $L^1_{\mathrm{loc}}(U) \to \mathcal{D}'(U)$ given by $u \mapsto \alpha_u$ is an *injection* of $L^1_{\mathrm{loc}}(U)$ in $\mathcal{D}'(U)$. In particular, since $L^p(U)$ injects in $L^1_{\mathrm{loc}}(U)$ for $p \in [1, +\infty]$, we also have an injection of $L^p(U)$ in $\mathcal{D}'(U)$. The injective character of $u \mapsto \alpha_u$ will be proved in exercises class. We frequently simply denote by u the distribution α_u.

A generalization is furnished by $\mathcal{M}_{\mathrm{loc}}(U)$, the space described in Proposition 5.10. This space $\mathcal{M}_{\mathrm{loc}}(U)$ was introduced in Section 5.3.3 as the dual to $C_c(U)$, when $C_c(U)$ is endowed with the inductive topology associated to the injections $C_K(U) \hookrightarrow C_c(U)$. With that definition, $\alpha \in \mathcal{M}_{\mathrm{loc}}(U)$ if, and only if, for all compact $K \subset U$, there exists $C_K \geq 0$ such that

$$|\langle \alpha_u, w\rangle| \leq C_K \sup_{x\in K} |w(x)|, \quad \forall w \in \mathcal{D}_K(U). \tag{6.4}$$

So it is clear that such a α is a distribution. We have also proved in Proposition 5.10 that each such α can be written (see (5.42)) as

$$\alpha(u) = \int_U u(x)\sigma(x)d\mu(x), \quad \forall u \in C_c(U), \tag{6.5}$$

where μ is a non-negative Borel measure μ on U which is finite on the compact subsets of U and a Borel map $\sigma\colon U \to \{-1, +1\}$. If we introduce the two mutually singular measures μ_+ and μ_- given by

$$\mu_\pm(A) = \mu(A \cap \{\sigma = \pm 1\}), \quad A \in \mathcal{B}(U), \tag{6.6}$$

then μ_+ and μ_- are some non-negative measures finite on compact subsets of U and

$$\alpha(u) = \int_U u(x)d\mu_+(x) - \int_U u(x)d\mu_-(x), \quad \forall u \in C_c(U). \tag{6.7}$$

If $v \in L^1_{\mathrm{loc}}(U)$, we set $\sigma(x) = \mathrm{sgn}(v(x))$ and $d\mu(x) = |v(x)|dx$. Then α_v is given by (6.5).

On the space $\mathcal{D}'(U)$, we will consider the weak-$*$ topology, so *a sequence of distributions* (α_n) *converges to a distribution* $\alpha \in \mathcal{D}'(U)$ *if, for all* $u \in \mathcal{D}(U)$, $\langle \alpha_n, u\rangle \to \langle \alpha, u\rangle$. In particular one can check the followings fact: *convergence in* $L^1_{\mathrm{loc}}(U)$ *(so in particular convergence in* $L^p(U)$*) implies convergence in the sense of distributions.*

Proposition 6.1 (Weak-strong convergence for distributions)**.** *If* (u_n) *is converging to* u *in* $\mathcal{D}(U)$ *and* (α_n) *is converging to* α *in* $\mathcal{D}'(U)$*, then* $\langle \alpha_n, u_n\rangle \to \langle \alpha, u\rangle$.

Proof of Proposition 6.1. Let K be a compact subset of U such that $u_n \to u$ in $\mathcal{D}_K(U)$. The restriction $\beta_n := \alpha_n \circ i_K$ is a continuous linear functional on $\mathcal{D}_K(U)$. By hypothesis, for each $u \in \mathcal{D}_K(U)$, $(\beta_n(u))$ is convergent, and thus bounded. We admit the fact that $\mathcal{D}_K(U)$ is a Fréchet space and that the Banach-Steinhaus theorem (or *uniform boundedness principle*) is valid in the framework of Fréchet spaces, giving the conclusion that (β_n) is uniformly bounded on bounded sets. We expand the product $\langle \alpha_n, u_n\rangle = \langle \beta_n, u_n\rangle$ as

$$\langle \beta_n, u_n\rangle = \langle \beta_n, u_n - u\rangle + \langle \beta_n - \beta, u\rangle + \langle \beta, u\rangle. \tag{6.8}$$

The set $B = \{u_n - u; n \in \mathbb{N}\}$ is bounded (see Section 5.2), so there exists $M \geq 0$ such that $|\langle \beta_n, v\rangle| \leq M$ for all $v \in B$. Let $\varepsilon > 0$ and let V be a neighbourhood of 0 in $\mathcal{D}_K(U)$. Let $t > 0$ be such that $B \subset tV$. For n large enough, we have $u_n - u \in \varepsilon V$ and $|\langle \beta_n - \beta, u\rangle| < \varepsilon$, so

$$|\langle \beta_n, u_n - u\rangle + \langle \beta_n - \beta, u\rangle| < (1 + tM)\varepsilon.$$

We deduce from (6.8) that $\langle \alpha_n, u_n\rangle \to \langle \alpha, u\rangle$. ☐

6.1.2 Bump functions, partition of unity, support

To work with distributions, we need the existence of *bump functions*: if $K \subset U$ in $\mathbb{R}^d$ with K compact and U open, then there exists a function χ of class C^∞ such that $K \prec \chi \prec U$. The easiest way to prove this is to consider the convolution (with a kernel having a sufficiently small support) of the characteristic function of a set W such that W is open, $\bar{W}$ compact, and $K \subset W$, $\bar{W} \subset U$. More precisely, we can take

$$W = \{x \in \mathbb{R}^d; d(x, K) < \varepsilon\}, \quad \chi = \rho_\varepsilon * \mathbf{1}_W,$$

for ε small enough, where $\rho_\varepsilon(x) = \varepsilon^{-d}\rho_1(\varepsilon x)$, with $\rho_1 \in C^\infty_{\bar{B}(0,1)}(\mathbb{R}^d)$, $0 \le \rho_1$, and ρ_1 of integral 1. We are left with the proof of the existence of ρ_1. We can consider

$$\rho(x) = \theta(|x|^2)\theta(1 - |x|^2), \quad \theta(r) = \begin{cases} \exp\left(-\frac{1}{1-r}\right) & r < 1, \\ 0 & r \ge 1, \end{cases}$$

and renormalize ρ to satisfy the condition of integral 1. To work with distributions, we also need the existence of C^∞ *partitions of unity*. Recall that a family of functions $v_i \colon \mathbb{R}^d \to \mathbb{R}_+$, $i \in I$, is said to be a (locally finite) partition of unity on $A \subset \mathbb{R}^d$ if every $x \in X$ has a neighbourhood on which all the v_i but a finite number of them are identically 0 and if $\sum_{i \in I} v_i = 1$ everywhere on A. If Γ is an open cover of A, a partition of unity on A is said to be subordinate to Γ if each v_i is supported in a certain open set U_i member of Γ.

Proposition 6.2 (Partition of unity)**.** *Let K be a compact of $\mathbb{R}^d$, U open in $\mathbb{R}^d$. Let $\Gamma = \{U_1, \dots, U_n\}$ be a finite open cover of K. Then there exists a finite partition of unity $v_1, \dots, v_m$ on K, subordinate to Γ, such that each $v_i \in \mathcal{D}(\mathbb{R}^d)$. If $K \subset U$ with U open, then we can ensure that each $v_i \in \mathcal{D}(U)$.*

Proof of Proposition 6.2. The last assertion of the theorem is obtained by multiplication of each v_i by χ, a C^∞ bump function such that $K \prec \chi \prec U$, so it is sufficient to consider the case $U = \mathbb{R}^d$. If $n = 1$, the existence of a C^∞ bump function χ_1 such that $K \prec \chi \prec U_1$ gives the result. Consider the case $n = 2$ (the general case $n \ge 2$ will be treated similarly). If we have shown that $K \subset K_1 \cup K_2$ with K_i compact, $K_i \subset U_i$, then we can exhibit some C^∞ bump functions χ_i such that $K_i \prec \chi_i \prec U_i$. Let $w = \chi_1 + \chi_2$ and $\bar{w} = w + (1 - \chi_3)$, where χ_3 is a C^∞ bump functions such that $K \prec \chi_3 \prec U_3$, $U_3 := \{w > 0\}$. We have $K \subset U_3$ since $w \ge 1$ on K and $\bar{w} > 0$ everywhere in $\mathbb{R}^d$ by construction, with $\bar{w} = w$ on K. We set $v_i = \chi_i / \bar{w}$ to obtain the desired partition of unity. There remains to prove the existence of K_1 and K_2. For each $x \in K \cap U_i$, there is an open ball $B(x, r_x)$ with $B(x, r_x) \subset U_i$. We can find a cover of K by the balls $B(x, r_x/2)$ for $x \in L$ with L finite subset of K. We then take $K_i = \bigcup_{x \in L \cap U_i} \bar{B}(x, r_x/2)$. □

Let $\alpha, \beta \in \mathcal{D}(U)$ and let V be an open subset of U. We say that $\alpha = \beta$ on V if $\langle \alpha, u \rangle = \langle \beta, u \rangle$ for all $u \in \mathcal{D}(V)$. We then have the following consistence result: if $\alpha = \beta$ on V_i, $i \in I$, where each V_i is an open subset of U, then $\alpha = \beta$ on $W = \bigcup_{i \in I} V_i$. Indeed, let $u \in \mathcal{D}(W)$ and let K be a compact subset of W such that u is supported in K. There is a finite cover of K by $\Gamma = \{V_i; i \in I_K\}$, where I_K is finite. Let $\{v_1, \dots, v_m\}$ be a C^∞ partition of unity on K subordinate to Γ. Then $uv_j \in \mathcal{D}(V_i)$ for a certain $i \in I_K$, so $\langle \alpha, uv_j \rangle = \langle \beta, uv_j \rangle$. It follows that

$$\langle \alpha, u \rangle = \sum_{j=1}^m \langle \alpha, uv_j \rangle = \sum_{j=1}^m \langle \beta, uv_j \rangle = \langle \beta, u \rangle,$$

and $\alpha = \beta$ on W. Take now $\beta = 0$. The consistency result shows that the class $\mathscr{Z}$ of open sets V such that $\alpha = 0$ on V, ordered by inclusion, has a maximal element, given by $\mathcal{Z}_\alpha := \bigcup_{V \in \mathscr{Z}} V$.

Definition 6.1 (Support of a distribution)**.** The support $\mathrm{supp}(\alpha)$ of a distribution $\alpha \in \mathcal{D}(U)$ is the complementary in U of the largest open set $\mathcal{Z}_\alpha$ on which $\alpha = 0$.

According to this definition, a point x is not in the support of α if, and only if, $\alpha = 0$ on an open neighbourhood V of x. So a point x is in the support of α if, for all open neighbourhood V of x, there is a $u \in \mathcal{D}(V)$ with $|\langle \alpha, u\rangle| > 0$. Recall that the (closed) support $\mathrm{supp}(v)$ of a continuous function v is defined as the closure (in U) of the set $\{x \in U; v(x) \neq 0\}$. If we set $\mathcal{Z}_v = U \setminus \mathrm{supp}(v)$ (this is the interior of $\{x \in U; v(x) = 0\}$, the largest open set on which $v = 0$), then we have (by definition, and then by taking the complementary)

$$\langle \alpha, u\rangle = 0 \Longleftarrow \mathrm{supp}(u) \subset \mathcal{Z}_\alpha \iff \mathrm{supp}(\alpha) \subset \mathcal{Z}_u. \tag{6.9}$$

The last two conditions in (6.9) are also equivalent to the condition

$$\mathrm{supp}(u) \cap \mathrm{supp}(\alpha) = \emptyset \iff d(\mathrm{supp}(u), \mathrm{supp}(\alpha)) > 0, \tag{6.10}$$

with the definitions

$$d(K, F) := \min_{x\in K} d(x, F), \quad d(x, F) := \inf_{y\in F} |x - y|,$$

for K compact and F closed in $\mathbb{R}^d$. Note well that, in (6.9), it is *not true* that $\langle \alpha, u\rangle = 0$ implies (6.10) (why?).

Proposition 6.3 (Support of a measure)**.** *Let U be an open set in $\mathbb{R}^d$ and let α be a locally signed measure on U. Then the support of α as a measure and the support of α as a distribution are the same.*

Recall that α is given by (6.5) and (6.7). The support $S(\alpha)$ of α seen as a measure is defined as follows:

$$S(\alpha) = \{x \in U; \mu(V) > 0 \text{ for all open neighbourhood } V \text{ of } x\}.$$

Proof of Proposition 6.3. Let Σ denote the support of α seen as a distribution. If $x \in \Sigma$ and V is an open neighbourhood of x, then there is a $u \in \mathcal{D}(V)$ such that $|\langle \alpha, u\rangle| > 0$. Since

$$\langle \alpha, u\rangle = \int_U u d\mu_+ - \int_U u d\mu_-, \tag{6.11}$$

at least one of the integrals in (6.11), say the first one, is non trivial. Replacing u by $-u$ if necessary, we can assume that $\langle \mu_+, u\rangle > 0$. In particular, $\langle \mu_+, u_+\rangle > 0$. By regularity of the measure

$$\theta \colon A \mapsto \int_A u_+ d\mu_+, \quad A \in \mathcal{B}(V),$$

there is a compact $K \subset V$ such that $\theta(K) > 0$. Since u_+ is continuous, it is bounded by a given M on K, and we see that $M\mu_+(K) \geq \theta(K) > 0$ and therefore $\mu_+(K) > 0$. We have then

$$\mu(V) \geq \mu_+(V) \geq \mu_+(K) > 0.$$

This being true for every neighbourhood V of x, we have $x \in S(\alpha)$. If, conversely, $x \in S(\alpha)$ and V is an open neighbourhood of x, then $\mu(V) > 0$. Assume for instance $\mu_+(V) > 0$. By regularity of μ, there is a compact $K \subset V$ such that $\mu_+(K) > 0$. Let χ be a C^∞ bump function with $K \prec \chi \prec V$. We have $\chi \in \mathcal{D}(V)$ and $\langle \mu_+, \chi\rangle \geq \mu_+(K) > 0$. We need to adjust χ to be certain that

$$\int_U \chi d\mu_+ \neq \int_U \chi d\mu_-.$$

We will set $A_\pm = \{\sigma = \pm 1\}$ (*cf.* (6.6)) so that $\mu_\pm(B) = \mu(A_\pm \cap B)$. Let $\varepsilon > 0$ and let H_ε, W_ε be some compact and open sets in U such that $H_\varepsilon \subset A_+ \subset W_\varepsilon$ and $\mu(W_\varepsilon \setminus H_\varepsilon) < \varepsilon$. Let θ_ε be a C^∞ bump function with $H_\varepsilon \prec \theta_\varepsilon \prec W_\varepsilon$. We have

$$\int_U \theta_\varepsilon \chi d\mu_- \le \int_U \theta_\varepsilon d\mu_- \le \mu(A_- \cap W_\varepsilon) = \mu(W_\varepsilon \setminus A_+) < \varepsilon,$$

and

$$\int_U \theta_\varepsilon \chi d\mu_+ \ge \mu(H_\varepsilon \cap K) \ge \mu(W_\varepsilon \cap K) - \varepsilon \ge \mu(A_+ \cap K) - \varepsilon = \mu_+(K) - \varepsilon,$$

so

$$\langle \alpha, \chi\theta_\varepsilon \rangle \ge \mu_+(K) - 2\varepsilon > 0,$$

if ε is chosen sufficiently small. Since $\chi\theta_\varepsilon \in \mathcal{D}(V)$ and since V is an arbitrary neighbourhood of x, we obtain $x \in \Sigma$. □

Exercise 6.2 (Injection of locally integrable functions)**.** Explain how to deduce from Proposition 6.3 that the map $u \mapsto \alpha_u$ defined in (6.3) is injective.

Exercise 6.3 (Convergence and support)**.** Prove that, if (α_n) is a sequence of distributions on an open set $U \subset \mathbb{R}^d$ which converges to a distribution $\alpha \in \mathcal{D}'(U)$, then $\operatorname{supp}(\alpha) \subset \liminf \operatorname{supp}(\alpha_n)$.

6.1.3 Order of a distribution, derivation of a distribution

Definition 6.4 (Order of a distribution)**.** Let U be an open set in $\mathbb{R}^d$ and let α be a distribution on U. Let K be a compact subset of U. The minimal integer j such that (6.2) is realized is called the order of the distribution α on K. If one can find an integer j such that α is of order j on every compact subset of U, then α is said to be of order j.

Exercise 6.5. Show that, if α is of order j, then the product $\langle \alpha, u \rangle$ can be extended to functions $u \in C^j_c(U)$. Show that the distribution p.v. $\left(\frac{1}{x}\right)$, defined by (6.1), is of order 1.

Definition 6.6 (Derivation of a distribution)**.** Let U be an open set in $\mathbb{R}^d$ and let α be a distribution on U. Let $m \in \mathbb{N}^d$ be a multi-index. The derivation ∂_x^m of α is defined by the duality formula

$$\langle \partial_x^m \alpha, u \rangle = (-1)^{|m|} \langle \alpha, \partial_x^m u \rangle, \forall u \in \mathcal{D}(U). \tag{6.12}$$

If α is of order j on K, then $\partial_x^m \alpha$ is a distribution of order at most $j + |m|$ on K.

We have to justify the last assertion of Definition 6.6. If (6.2) is satisfied, then clearly

$$|\langle \alpha, \partial_x^m u \rangle| \le C_{K,j} \sup_{|m'| \le j + |m|} \sup_{x \in K} |\partial_x^{m'} u(x)|, \tag{6.13}$$

for all $u \in \mathcal{D}_K(U)$. This shows that $\partial_x^m \alpha$ is a distribution, of order at most $j + |m|$ on K (note well that differentiation of distribution is not always increasing the order: if $u \in \mathcal{D}(\mathbb{R}^d)$, it defines a distribution of order 0 with all derivatives of order 0). The following properties are elementary consequences of the equivalent property for smooth functions and/or the duality formula (6.12):

1. the map $\partial_x^m \colon \alpha \mapsto \partial_x^{m'} \alpha$ is continuous on $\mathcal{D}'(\mathbb{R}^d)$. Note well, in particular, that convergence in $L^1_{\text{loc}}(U)$ (and so convergence in $L^p(U)$) implies convergence of the derivatives in the sense of distributions.
2. For all $m, m' \in \mathbb{N}^d$, $\partial_x^m(\partial_x^{m'} \alpha) = \partial_x^{m+m'} \alpha$.

3. We have the inclusion $\mathrm{supp}(\partial_x^m \alpha) \subset \mathrm{supp}(\alpha)$.

Note also, to come back to the example of the distribution p.v. $\left(\frac{1}{x}\right)$, defined by (6.1), that we have

$$\text{p.v.}\left(\frac{1}{x}\right) = \partial_x(\log(|x|)), \tag{6.14}$$

with a slight abuse of notation (we indicate the variable x), with log being the inverse of exp.

Exercise 6.7. Prove (6.14) (check that you recover all the expressions given in (6.1)).

Proposition 6.4 (Order of distributions)**.** *We have the following results.*

1. *a distribution with compact support has a finite order,*
2. *distributions of order* 0 *are locally signed measures.*

Proof of Proposition 6.4. Let α be a distribution with compact support $K_0 \subset U$. Let V be an open set with compact closure such that $K_0 \subset V \subset \bar{V} \subset U$. Let χ be a bump function such that $\bar{V} \prec \chi \prec U$. Let $u \in \mathcal{D}(U)$. With the notations of (6.9), we have

$$\mathrm{supp}(\alpha) = K_0 \subset V \subset \mathcal{Z}_{(1-\chi)u},$$

so $\langle \alpha, u \rangle = \langle \alpha, \chi u \rangle$. Let K be such that $\chi \in \mathcal{D}_K(U)$ and let j be the order of α on K. For all $u \in \mathcal{D}(U)$, we have

$$|\langle \alpha, u \rangle| = |\langle \alpha, \chi u \rangle| \leq C_{K,j} p_{K,j}(\chi u). \tag{6.15}$$

We use the bound

$$p_{K,j}(\chi u) \leq C(j) p_{K,j}(\chi) p_{K,j}(u), \tag{6.16}$$

which follows from the Leibniz formula

$$\partial_x^m(\chi u) = \sum_{p+q=m} \binom{m}{p} \partial_x^p \chi \partial_x^q u, \quad \binom{m}{p} = \frac{m!}{p!q!} \text{ (with } p+q=m), \quad m! = m_1! \cdots m_d!. \tag{6.17}$$

We obtain a bound uniform with respect to the support of u, which shows that α is of order j. Let now α be a distribution of order 0. By Proposition 5.10, α is locally a signed measure. ☐

Proposition 6.5 (Punctual support)**.** *A distribution supported on a singleton* $\{z\}$ *is a finite linear combination of the Dirac mass at* z *and of its derivatives.*

Proof of Proposition 6.5. Let α be a distribution supported on a singleton $\{z\}$. Since α has compact support, there is a bump function χ_1 such that $\langle \alpha, u \rangle = \langle \alpha, u\chi_1 \rangle$ for all $u \in \mathcal{D}(U)$ (*cf.* the proof of Proposition 6.4 above). The bump function χ_1 is constantly equal to 1 in a neighbourhood of z. Actually, we can assume $\bar{B}(z,\delta) \prec \chi_1 \prec B(z, 2\delta)$ for a given $\delta > 0$. If

$$\alpha = \sum_{|m| \leq j} c_m \partial_x^m \delta_0, \tag{6.18}$$

then

$$(\forall |m| \leq j,\ \partial_x^m u(z) = 0) \Rightarrow \langle \alpha, u \rangle = 0. \tag{6.19}$$

The converse implication is also true (compare with Lemma 5.2). Indeed, assuming that (6.19) is satisfied and that 0 is in the support of $u \in \mathcal{D}(U)$, we can write the Taylor expansion

$$u(x) = \sum_{|m| \leq j} \frac{x^m}{m!} \partial_x^m u(z) + v(x). \tag{6.20}$$

Since $\langle \alpha, u\rangle = \langle \alpha, u\chi_1\rangle$ and since $\langle \alpha, \chi_1 v\rangle = 0$ by (6.19), we see that α satisfies (6.18), with

$$c_m := \langle \alpha, v_m\rangle, \quad v_m(x) = \chi_1(x)\frac{x^m}{m!}.$$

Let us prove (6.19). Let j be the order of α. Let $\chi_n(x) = \chi_1(z + n(x - z))$. Since $\chi_n \equiv 1$ in a neighbourhood of z, we have $\langle \alpha, u\rangle = \langle \alpha, u\chi_n\rangle$, so $|\langle \alpha, u\rangle| \leq C p_{K,j}(u\chi_n)$. If $u(z) = 0$, then $|u(x)| \leq p_{\bar{B}(z,2\delta),1}(u)|x - z|$ for all $x \in \bar{B}(z, 2\delta)$, so

$$p_{K,0}(u\chi_n) \leq p_{\bar{B}(z,2\delta),1}(u)\frac{2\delta}{n},$$

since $\chi_n(x) \neq 0$ implies $|x - z| < 2\delta/n$. If $u(z) = 0$ and $\nabla u(z) = 0$, then

$$|u(x)| \leq C_2 p_{\bar{B}(z,2\delta),2}(u)|x - z|^2, \tag{6.21}$$

for all $x \in \bar{B}(z, 2\delta)$ (use Taylor's formula with integral remainder to justify (6.21)), so

$$|\partial_{x_j}(u\chi_n)(x)| \leq |\partial_{x_j}u(x)| + n|u(x)|\|\chi_1'\|_{L^\infty} \leq C(\delta)n^{-1}.$$

Similar arguments for higher derivatives show that, if $\partial_x^m u(z) = 0$ for all multi-index of length $|m| \leq j$, then $p_{K,j}(u\chi_n) \to 0$ when $n \to +\infty$. It follows that $\langle \alpha, u\rangle = 0$, and (6.19) is satisfied. □

6.1.4 Some operations on distributions

We have already defined the differentiation of distributions, based on a duality formula. With a similar approach, we can define the following operations.

1. *Multiplication by a smooth function.* If $\alpha \in \mathcal{D}'(U)$, $w \in C^\infty(U)$, we define the distribution $w\alpha$ by
$$\langle w\alpha, u\rangle = \langle \alpha, wu\rangle, \tag{6.22}$$
for all $u \in \mathcal{D}(U)$. Multiplication of a distribution by a bump function is a standard tool to localize distributions. This operation was used in the proof of *1.* of Proposition 6.4 for instance. It is an exercise to check on the duality formula (6.12) that the Leibniz' Formula (6.17) remains valid for the product of a distribution by a smooth function. We also have the following property:
$$\mathrm{supp}(w\alpha) \subset \mathrm{supp}(w) \cap \mathrm{supp}(\alpha). \tag{6.23}$$
The inclusion can be strict, as shown by the example $d = 1$, $w(x) = x$, $\alpha = \delta_0$, for which $w\alpha = 0$. Let us also use this result to emphasize the difficulty to define properly the product of two distributions. Indeed, assume that a product with "good" properties (associativity, commutativity if one of the factor is a smooth function) is defined on $\mathcal{D}'(\mathbb{R}^d)$. We have $x\delta_0 = \delta_0 x = 0$, so $(\delta_0 x)\mathrm{p.v.}\left(\frac{1}{x}\right) = 0$. By associativity,
$$(\delta_0 x)\mathrm{p.v.}\left(\frac{1}{x}\right) = \delta_0\left(x\mathrm{p.v.}\left(\frac{1}{x}\right)\right),$$
and since $x\mathrm{p.v.}\left(\frac{1}{x}\right) = 1$ (use (6.1)), we obtain $0 = \delta_0$, which is absurd.

Exercise 6.8. Let H denote the Heavyside function $H(x) = \mathbf{1}_{x>0}$. Show that $H\delta_0$ can not be well defined in $\mathcal{D}'(\mathbb{R})$, but that $H(x)\delta_{y=0}$ is well-defined in $\mathcal{D}'(\mathbb{R}^2)$.

2. *Translation.* If $\alpha \in \mathcal{D}'(U)$, $z \in \mathbb{R}^d$, we define the distribution $\tau_z\alpha$ on $V := z + U$ by

$$\langle \tau_z\alpha, u\rangle = \langle \alpha, \tau_{-z}u\rangle, \tag{6.24}$$

for all $u \in \mathcal{D}(V)$.

3. *Composition with a diffeomorphism.* If $\alpha \in \mathcal{D}'(U)$, and $\psi\colon U \to V$ is a C^∞-diffeomorphisms with Jacobian determinant $J\psi$, we define the distribution $\alpha \circ \psi$ on V by

$$\langle \alpha \circ \psi, u\rangle = \langle \alpha, (J\psi)^{-1} u \circ \psi^{-1}\rangle, \tag{6.25}$$

for all $u \in \mathcal{D}(V)$. In particular, for $\psi\colon \mathbb{R}^d \to \mathbb{R}^d$ given by $\psi(x) = -x$, denoting $\check{u} = u \circ \psi$, we define, for $\alpha \in \mathcal{D}'(\mathbb{R}^d)$, the distribution $\check{\alpha}$ by

$$\langle \check{\alpha}, u\rangle = \langle \alpha, \check{u}\rangle, \tag{6.26}$$

for all $u \in \mathcal{D}(\mathbb{R}^d)$.

4. *Convolution with a function.* Let $\alpha \in \mathcal{D}'(\mathbb{R}^d)$, $w \in \mathcal{D}(\mathbb{R}^d)$. Based on the following formula (valid when $\alpha \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ for instance)

$$\alpha * w(x) = \int_{\mathbb{R}^d} \alpha(y) w(x-y) dy,$$

we define the *real number* $\alpha * w(x)$ for $x \in \mathbb{R}^d$ by

$$\alpha * w(x) = \langle \alpha, \tau_x \check{w}\rangle. \tag{6.27}$$

It is left as an exercise to check that the formula (6.22), (6.24), (6.25) define distributions (*i.e.* that (6.2) is satisfied). Regarding convolution, we have the following result.

Proposition 6.6 (Convolution of a distribution by a function)**.** *Let $\alpha \in \mathcal{D}'(\mathbb{R}^d)$, $w \in \mathcal{D}(\mathbb{R}^d)$. Then we have $\alpha * w \in C^\infty(\mathbb{R}^d)$ with*

$$\partial_x^m(\alpha * w) = (\partial_x^m \alpha) * w = \alpha * \partial_x^m w, \tag{6.28}$$

for every multi-index $m \in \mathbb{N}^d$. We also have

$$\langle \alpha * w, u\rangle = \langle \alpha, u * \check{w}\rangle, \tag{6.29}$$

for all $u \in \mathcal{D}(\mathbb{R}^d)$ and the inclusion

$$\mathrm{supp}(\alpha * w) \subset \mathrm{supp}(\alpha) + \mathrm{supp}(w), \tag{6.30}$$

which shows that $\alpha * w$ has a compact support if α has a compact support. If α has compact support L and order $k \in \mathbb{N}$, say $|\langle \alpha, u\rangle| \leq C p_{L,k}(u)$, then

$$p_{K,j}(\alpha * w) \leq C p_{K-L,k+j}(w). \tag{6.31}$$

Remark 6.1 (Unit for the convolution product). The Dirac mass δ_0 is the unit for the convolution product: $\delta_0 * u = u$ for all $u \in \mathcal{D}(\mathbb{R}^d)$. Indeed, using (6.27), we have

$$\delta_0 * u(x) = \langle \delta_0, \tau_x \check{u}\rangle = u(x). \tag{6.32}$$

Proof of Proposition 6.6. we prove the existence of the differentials of $\alpha * w$ by recursion on the order of differentiation. We will simply treat the case of order 1, *i.e.* ∂_{x_j} for a $j \in \{1, \ldots, d\}$, since higher derivatives are treated similarly. For $t \neq 0$, using (6.27) and the formula $\tau_{x+z} = \tau_x \tau_z$, we have

$$\frac{\alpha * w(x + te_j) - \alpha * w(x)}{t} = \langle \alpha, \tau_x D_{j,t} \check{w} \rangle, \quad D_{j,t} u(x) := \frac{u(x + te_j) - u(x)}{t},$$

whereas, for all fixed $u \in \mathcal{D}(\mathbb{R}^d)$, $D_{j,t_n} u \to \partial_{x_j} u$ in $\mathcal{D}(\mathbb{R}^d)$ for all sequence (t_n) in $\mathbb{R}^*$ which tends to 0. Indeed, the expansion

$$(D_{j,t} u - \partial_{x_j} u)(x) = \int_0^1 [\partial_{x_j} u(x + t\theta e_j) - \partial_{x_j} u(x)] d\theta = \int_0^1 \int_0^1 \partial_{x_j}^2 u(x + tr\theta e_j) t\theta dr d\theta,$$

gives, for K compact and $l \in \mathbb{N}$, the estimate $p_{K,l}(D_{j,t} u - \partial_{x_j} u) \leq p_{L,l+2}(u)|t|$ for $|t| < 1$, where L is a compact which contains all the points at distance at most 1 of K. We have established the existence of $\partial_{x_j}(\alpha * w)$ therefore, and altogether the formula $\partial_{x_j}(\alpha * w) = \alpha * \partial_{x_j} w$. By recursion, we will obtain $\partial_x^m(\alpha * w) = \alpha * \partial_x^m w$. The second identity in (6.28) will follow from (6.29) and the standard property of distributivity of the differentiation for the product of convolution of functions. Since

$$\langle \alpha * w, u \rangle = \int_{\mathbb{R}^d} \langle \alpha, \tau_x \check{w} \rangle u(x) dx, \quad \int_{\mathbb{R}^d} \tau_x \check{w}(y) u(x) dx = u * \check{w}(y),$$

the formula (6.29) will be established if we can justify the exchange of α with the integral over x. This is done by approximation of the integral by Riemann sums. Without loss of generality, we can assume that u is supported in the unit cube $Q = (0,1)^d$. By regularity of the integrand, we have

$$\int_Q \langle \alpha, \tau_x \check{w} \rangle u(x) dx = \sum_{Q' \subset Q} \int_{Q'} \langle \alpha, \tau_x \check{w} \rangle u(x) dx = N^{-d} \sum_{Q' \subset Q} \langle \alpha, \tau_{x'} \check{w} \rangle u(x') + o(1), \tag{6.33}$$

when $N \to +\infty$, where the sum over $Q' \subset Q$ is done on the cubes $Q' = x' + N^{-d} Q$ of size N^{-d}, with corner at $x' \in N^{-d} \mathbb{Z}^d \cap [0,1)^d$. We have

$$N^{-d} \sum_{Q' \subset Q} \langle \alpha, \tau_{x'} \check{w} \rangle u(x') = \langle \alpha, v_N \rangle, \quad v_N(y) = N^{-d} \sum_{Q' \subset Q} \tau_{x'} \check{w}(y) u(x').$$

To prove that $v_N \to u * \check{w}$ in $\mathcal{D}(\mathbb{R}^d)$, it will be sufficient to prove that all the sequence v_N is supported in a single compact set K and that $v_N \to u * \check{w}$ uniformly on K. The result can be transferred to derivatives then, by replacing w with one of its derivative. We can take $K = \operatorname{supp}(w) + \bar{Q}$. The fact that $v_N \to u * \check{w}$ uniformly on K is deduced from the C^1-regularity of u and w:

$$v_N(y) - u * \check{w}(y) = \sum_{Q' \subset Q} \int_{Q'} [\varphi_y(x') - \varphi_y(x)] dx, \quad \varphi_y(x) := w(x - y) u(x),$$

so

$$v_N(y) - u * \check{w}(y) = \sum_{Q' \subset Q} \int_{Q'} \int_0^1 (\nabla_x \varphi_y)(\theta x' + (1 - \theta) x) \cdot (x' - x) d\theta dx$$

and

$$p_{K,0}(v_N - u * \check{w}) \leq 2N^{-1} p_{K,1}(w) p_{K,1}(u),$$

which gives $v_N \to u * \check{w}$ uniformly on K as desired. Let us establish the inclusion of the supports (6.30) now. We can use directly the definition (6.27), or (6.29), and this is this the first option we choose. If $\alpha * w(x) \not 0 =$, then $\mathrm{supp}(\alpha) \cap \mathrm{supp}(\tau_x \check{w}) \neq \emptyset$, so there exists $z \in \mathrm{supp}(\alpha)$ such that $z \in \mathrm{supp}(\tau_x \check{w})$. We check that $\mathrm{supp}(\tau_x \check{w}) = x - \mathrm{supp}(w)$, so that $x \in \mathrm{supp}(\alpha) + \mathrm{supp}(w)$. It follows that $\{\alpha * w \neq 0\} \subset \mathrm{supp}(\alpha) + \mathrm{supp}(w)$. If F is closed and K compact then $F + K$ is closed (use a sequential characterization of the closure for instance), so we obtain (6.30). Let $w \in \mathcal{D}(\mathbb{R}^d)$. We have the point-wise bound

$$|\alpha * w(x)| = |\langle \alpha, \tau_x \check{w} \rangle| \leq C p_{L,k}(\tau_x \check{w}) \leq C p_{K-L,k}(w),$$

for all $x \in K$, so $p_{K,0}(\alpha * w) \leq C p_{M-L,k}(w)$, which is (6.31) for $j = 0$. Replacing u by $\partial_x^m u$ (and using (6.28)) gives the general case. □

Corollary 6.7 (Density of smooth functions)**.** *Let $\alpha \in \mathcal{D}'(\mathbb{R}^d)$. Let (ρ_n) be an approximation of the unit on $\mathbb{R}^d$:*

$$\rho_n(x) = n^d \rho_1(nx), \quad \rho_1 \in \mathcal{D}(B(0,1)), \quad \rho_1 \geq 0, \quad \int_{\mathbb{R}^d} \rho_1(x) dx = 1. \tag{6.34}$$

Then $(\alpha * \rho_n)$ is a sequence of C^∞-functions on $\mathbb{R}^d$ which converges to α in $\mathcal{D}'(\mathbb{R}^d)$. Let χ_1 be a C^∞-bump function such that $\bar{B}(0,1) \prec \chi_1 \prec B(0,2)$ and set $\chi_n(x) = \chi_1(n^{-1}x)$. Then $\chi_n(\alpha * \rho_n)$ is a sequence of functions of $\mathcal{D}(\mathbb{R}^d)$ which converges to α in $\mathcal{D}'(\mathbb{R}^d)$.

Proof of Corollary 6.7. We use the formula (6.29): if $u \in \mathcal{D}(\mathbb{R}^d)$, then

$$\langle \chi_n(\alpha * \rho_n), u \rangle = \langle \alpha, (\chi_n u) * \check{\rho}_n \rangle.$$

For n large enough, we have $\chi_n u = u$, and since $u * \check{\rho}_n \to u$ in $\mathcal{D}(\mathbb{R}^d)$ (proof left as exercise), the result follows. □

Proposition 6.8 (Convolution of distributions)**.** *Let $\alpha, \beta \in \mathcal{D}'(\mathbb{R}^d)$. If one of the two distributions, say β, is compactly supported, then we can define the convolution product $\alpha * \beta$ by*

$$\langle \alpha * \beta, u \rangle = \langle \alpha, \check{\beta} * u \rangle, \tag{6.35}$$

for all $u \in \mathcal{D}(\mathbb{R}^d)$. We then have the following properties:

$$\mathrm{supp}(\alpha * \beta) \subset \mathrm{supp}(\alpha) + \mathrm{supp}(\beta), \tag{6.36}$$

and, for all multi-index $m \in \mathbb{N}^d$,

$$\partial_x^m(\alpha * \beta) = (\partial_x^m \alpha) * \beta = \alpha * (\partial_x^m \beta). \tag{6.37}$$

If $\gamma \in \mathcal{D}'(\mathbb{R}^d)$ is also compactly supported, then

$$\beta * \gamma = \gamma * \beta, \quad (\alpha * \beta) * \gamma = \alpha * (\beta * \gamma). \tag{6.38}$$

Remark 6.2 (Unit for the convolution product). The Dirac mass δ_0 is the unit for the convolution product: $\alpha * \delta_0 = \alpha$. Indeed, using Remark 6.1 and (6.35), we have

$$\langle \alpha * \delta_0, u \rangle = \langle \alpha, \check{\delta}_0 * u \rangle = \langle \alpha, \delta_0 * u \rangle = \langle \alpha, u \rangle.$$

Proof of Proposition 6.8. Let us first check that (6.35) defines a distribution. If $u \in \mathcal{D}_{K_0}(\mathbb{R}^d)$, we set $K = -\operatorname{supp}(\beta) + K_0$. Then K is compact and $\operatorname{supp}(\check{\beta} * u) \subset K$ by (6.30). There exists $j \in \mathbb{N}$, $C_1 \geq 0$ such that α satisfies the condition (6.2): $|\langle \alpha, v\rangle| \leq C_1 p_{K,j}(v)$. With $v = \check{\beta} * u$, this gives $|\langle \alpha * \beta, u\rangle| \leq C_1 p_{K,j}(\check{\beta} * u)$. We apply (6.31) to deduce that

$$p_{K,j}(\check{\beta} * u) \leq C_2 p_{K+L,j+k}(u),$$

and the estimate $|\langle \alpha * \beta, u\rangle| \leq C_1 C_2 p_{K+L,j+k}(u)$, where $L = \operatorname{supp}(\beta)$ and k is the order of β. Let now $x \in \operatorname{supp}(\alpha * \beta)$: for all neighbourhood V of x, there is a $u \in \mathcal{D}(V)$ such that $\langle \alpha * \beta, u\rangle = \langle \alpha, \check{\beta} * u\rangle \neq 0$. We have then $\operatorname{supp}(\alpha) \cap \operatorname{supp}(\check{\beta} * u) \neq \emptyset$. Let z be an element in the intersection. By (6.30), we have

$$\operatorname{supp}(\check{\beta} * u) \subset \operatorname{supp}(\check{\beta}) + \operatorname{supp}(u) = -\operatorname{supp}(\beta) + \operatorname{supp}(u),$$

so there is $z' \in \operatorname{supp}(\beta)$ and $y \in \operatorname{supp}(u)$ such that $z = -z' + y$, which means that

$$\operatorname{supp}(u) \subset \operatorname{supp}(\alpha) + \operatorname{supp}(\beta).$$

Consequently $x \in \operatorname{supp}(\alpha) + \operatorname{supp}(\beta)$ is a necessary condition, and we obtain (6.36). The formula (6.37) is deduced from (6.28) and (6.35). For instance, we have

$$\langle \partial_x^m(\alpha * \beta), u\rangle = (-1)^{|m|}\langle \alpha, \check{\beta} * (\partial_x^m u)\rangle = \langle \partial_x^m \alpha, \check{\beta} * u\rangle = \langle (\partial_x^m \alpha) * \beta, u\rangle,$$

for all $u \in \mathcal{D}(\mathbb{R}^d)$. To justify the commutation formula and associativity property in (6.36), we observe that the formula are true if α, β, γ are all elements of $\mathcal{D}(\mathbb{R}^d)$, as a consequence of the Fubini Theorem. The result in the general case will follow by approximation. Let us give the details of this step for the commutation formula $\beta * \gamma = \gamma * \beta$, that is to say

$$\langle \beta, \check{\gamma} * u\rangle = \langle \gamma, \check{\beta} * u\rangle, \tag{6.39}$$

where $u \in \mathcal{D}(\mathbb{R}^d)$. Let $\gamma = \gamma_n$ in (6.39) with $\gamma_n \in \mathcal{D}(\mathbb{R}^d)$ and $\gamma_n \to \gamma$ in $\mathcal{D}'(\mathbb{R}^d)$. We can assume (construct γ_n by convolution) that each γ_n and γ are supported in the same compact K and that we have

$$|\langle \gamma_n, u\rangle| \leq C p_{K,j}(u), \tag{6.40}$$

for a uniform constant C and a uniform order j. We can pass to the limit in the right-hand side of (6.39) of course. To justify the limit in the left-hand side of (6.39), we simply note that

$$\check{\gamma}_n * u \to \check{\gamma} * u \text{ in } \mathcal{D}(\mathbb{R}^d).$$

Indeed all the functions are supported in $L := K + \operatorname{supp}(u)$ by (6.30) and we see that, up to the substitution of u by $\partial_x^m u$, it is sufficient to prove that $\check{\gamma}_n * u \to \check{\gamma} * u$ uniformly on L. We have

$$\check{\gamma}_n * u(x) = \langle \check{\gamma}_n, \tau_x \check{u}\rangle = \langle \gamma_n, \tau_{-x} u\rangle \to \langle \gamma, \tau_{-x} u\rangle = \check{\gamma} * u(x),$$

for all $x \in L$. This is only a punctual convergence, but the functions $\check{\gamma}_n * u$ are equi-continuous, since by (6.40) and (6.31), we have the Lipschitz bound $p_{L,1}(\check{\gamma}_n * u) \leq C p_{K+L,j+1}(u)$. So the convergence is uniform. We obtain (6.39) when $\beta \in \mathcal{D}(\mathbb{R}^d)$ and $\gamma \in \mathcal{D}'(\mathbb{R}^d)$ is compactly supported. The same reasoning then shows that the results holds true when $\beta \in \mathcal{D}'(\mathbb{R}^d)$ is compactly supported. The proof of the associativity property in (6.36) is left as an exercise. $\square$

6.1.5 Fundamental solutions of PDEs - I

If $P(X) = \sum_{|m|\leq j} a_m X^m$ is a complex polynomial in the d variables $X_1, \dots, X_d$, we introduce the following notations

$$D^m = (2\pi i)^{-|m|}\partial_x^m, \quad P(D) = \sum_{|m|\leq j} a_m D^m. \tag{6.41}$$

We say that $P(D)$ is a linear differential operator with constant coefficients. The choice of the convention for the derivative D comes from the formula

$$\mathcal{F}(P(D)u)(\xi) = P(\xi)(\mathcal{F}u)(\xi), \tag{6.42}$$

where $\mathcal{F}$ is the Fourier Transform defined in (2.151):

$$\mathcal{F}v(\xi) = \hat{v}(\xi) = \int_{\mathbb{R}^d} v(x)e^{-2\pi i x\cdot\xi}dx. \tag{6.43}$$

We deduce (6.42) from the formulas (2.162).

Definition 6.9 (Fundamental solution)**.** Let $P(D)$ be a linear differential operator with constant coefficients as in (6.41). A *fundamental solution* for the Poisson equation associated to $P(D)$ is a distribution Φ on $\mathbb{R}^d$ such that $P(D)\Phi = \delta_0$.

The Poisson equation associated to $P(D)$ is the equation

$$P(D)u = f, \tag{6.44}$$

where f is the data and u the unknown. Since the Dirac mass δ_0 is the unit for the convolution product, and in virtue of (6.28), one way to solve (6.44) is to set $u = \Phi * f$, where Φ is fundamental solution. This has even a meaning if f is a distribution with compact support. When

$$P(\xi) = 4\pi^2\left(\xi^1 + \cdots + \xi_d^2\right) = 4\pi^2|\xi|^2,$$

which corresponds to $P(D) = -\Delta$, (6.44) is simply called the Poisson equation.

Proposition 6.9 (Fundamental solution of the Poisson Equation)**.** *Let $\omega_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}$ denote the volume of the unit sphere in $\mathbb{R}^d$. Let $\Phi(x)$ be defined by*

$$\Phi(x) = \begin{cases} -\dfrac{1}{2\pi}\log(|x|) & \text{if } d = 2, \\ \dfrac{1}{(d-2)\omega_d}\dfrac{1}{|x|^{d-2}} & \text{if } d \geq 3. \end{cases} \tag{6.45}$$

Then Φ is a fundamental solution of the Poisson Equation.

Proof of Proposition 6.9. Note first that Φ is locally integrable, because, using polar coordinates, and writing $\Phi(x) = \bar{\Phi}(r)$, $r = |x|$, we have

$$\int_{B(0,1)} |\Phi(x)|dx = \omega_d \int_0^1 |\bar{\Phi}(r)|r^{d-1}dr < +\infty.$$

By similar computations, we see that $\nabla\Phi$ is locally integrable also, but that, on the contrary, $\partial^2_{x_i x_j}\Phi$ has a singularity at 0 which may be not integrable. We will have to work therefore to prove that Φ is a fundamental solution. First, let us specify the computations in question. With the notation $r = |x| = \sqrt{|x|^2}$, we have

$$\nabla \bar{\Phi}(r) = \bar{\Phi}'(r) \nabla r = \bar{\Phi}'(r) \frac{x}{r}. \tag{6.46}$$

We then use the formulas

$$\Delta w = \operatorname{div}(\nabla w), \quad \operatorname{div}(w\Psi) = w \operatorname{div}(\Psi) + \nabla w \cdot \Psi, \tag{6.47}$$

where $w\colon \mathbb{R}^d \to \mathbb{R}$, $\Psi\colon \mathbb{R}^d \to \mathbb{R}^d$, to obtain, since $\operatorname{div}(x) = d$,

$$\Delta \bar{\Phi}(r) = d \frac{\bar{\Phi}'(r)}{r} + \frac{\partial}{\partial r} \left(\frac{\bar{\Phi}'(r)}{r} \right) \frac{x \cdot x}{r}$$

and thus

$$\Delta \bar{\Phi}(r) = \bar{\Phi}''(r) + (d-1) \frac{\bar{\Phi}'(r)}{r}. \tag{6.48}$$

Let Φ_ε be the approximation of Φ defined by

$$\Phi_\varepsilon(x) = \begin{cases} -\dfrac{1}{4\pi} \log(|x|^2 + \varepsilon^2) & \text{if } d = 2, \\ \dfrac{1}{(d-2)\omega_d} \dfrac{1}{(|x|^2 + \varepsilon^2)^{d/2-1}} & \text{if } d \geq 3. \end{cases} \tag{6.49}$$

By (6.48), we have $-\Delta \Phi_\varepsilon(x) = g_\varepsilon(x) := \varepsilon^{-d} g(\varepsilon^{-1} x)$, where

$$g(x) = \begin{cases} \dfrac{1}{\pi} \dfrac{1}{(|x|^2 + 1)^2} & \text{if } d = 2, \\ \dfrac{d}{\omega_d} \dfrac{1}{(|x|^2 + 1)^{d/2+1}} & \text{if } d \geq 3. \end{cases} \tag{6.50}$$

Let $I(d)$ denote the integral of g over $\mathbb{R}^d$. Using polar coordinates, we compute $I(2) = 1$ and, with an additional integration by parts, $I(d) = I(d-2)$ for $d \geq 3$, with

$$I(1) := \int_0^\infty \frac{1}{(r^2 + 1)^{3/2}} dr.$$

With the change of variable $s = 1/r$, we compute

$$I(1) = \int_0^\infty \frac{s}{(s^2 + 1)^{3/2}} ds = 1.$$

So all the integrals $I(d)$ are equal to 1. Consequently $g_\varepsilon \to \delta_0$ in $\mathcal{D}'(\mathbb{R}^d)$. Since (Φ_ε) converges to Φ in $L^1_{\text{loc}}(\mathbb{R}^d)$, it also converges to Φ in $\mathcal{D}'(\mathbb{R}^d)$. We deduce that $-\Delta \Phi = \delta_0$. □

6.2 Tempered distributions

6.2.1 Definition of tempered distributions

We denote by $\hat{f}$ or $\mathcal{F} f$, depending on the context, the Fourier transform of a function f, *cf.* (6.43). For $f, g \in L^1(\mathbb{R}^d)$, Fubini's Theorem gives the symmetry formula

$$\int_{\mathbb{R}^d} \hat{f}(y) g(y) dy = \iint_{\mathbb{R}^d \times \mathbb{R}^d} f(x) g(y) e^{-2\pi i x \cdot y} = \int_{\mathbb{R}^d} f(x) \hat{g}(x) dx, \tag{6.51}$$

Based on (6.51), we may define the Fourier transform $\hat{\alpha}$ of a distribution $\alpha \in \mathcal{D}'(\mathbb{R}^d)$ by the duality formula

$$\langle \hat{\alpha}, u \rangle = \langle \alpha, \hat{u} \rangle. \tag{6.52}$$

This is not possible however, the reason being that $\mathcal{D}(\mathbb{R}^d)$ is not stable by $u \mapsto \hat{u}$. Indeed, if we examine the formula

$$\hat{u}(\xi) = \int_{\mathbb{R}^d} u(x) e^{-2\pi i x \cdot \xi} dx,$$

there is no reason for $\hat{u}$ to be compactly supported. The fact is, that we have the following result (in the same spirit as Proposition 2.22).

Proposition 6.10 (Localization in space and frequency are not compatible)**.** *Let $u \in \mathscr{S}(\mathbb{R}^d)$. If both u and $\hat{u}$ are compactly supported, then $u = 0$.*

Proof of Proposition 6.10. If u is compactly supported in $B(0, R)$, then

$$\hat{u}(\xi) = \int_{B(0,R)} e^{-2\pi i x \cdot \xi} u(x) dx$$

can be extended as an holomorphic function $\mathbb{C}^d \to \mathbb{C}$ given by

$$z \mapsto \int_{B(0,R)} e^{-2\pi i x \cdot z} u(x) dx,$$

so $\hat{u} = 0$ if it vanishes on a subset of $\mathbb{R}^d$ having an accumulation point. □

Exercise 6.10. Show that, if $u \in \mathscr{S}(\mathbb{R}^d)$ and $\hat{u}$ is compactly supported, then there exists $\psi \in \mathscr{S}(\mathbb{R}^d)$ such that $u = u * \psi$:

$$u(x) = \int_{\mathbb{R}^d} u(x - y)\psi(y) dy, \quad \forall x \in \mathbb{R}^d.$$

Hint: consider a bump function χ such that $\hat{u} = \hat{u}\chi$.

A first idea is to restrict (6.52) to the case where α is compactly supported, but we will see that it is possible to consider the largest class of tempered distributions. Recall that the Schwartz space was defined in Definition 2.3: we denote by $\mathscr{S}(\mathbb{R}^d)$ the space of infinitely differentiable functions whose derivatives decay faster at infinity than any polynomial: $v \in \mathscr{S}(\mathbb{R}^d)$ if v is of class C^∞ and all the semi-norms

$$q_{N,k}(v) = \sup_{x \in \mathbb{R}^d, |\alpha| \leq k} (1 + |x|^2)^{N/2} |\partial_x^\alpha v(x)| \tag{6.53}$$

are finite. We have seen in (2.160) that $\mathscr{S}(\mathbb{R}^d) \subset L^p(\mathbb{R}^d)$ for all $p \in [1, +\infty]$. The formulas (2.162), or Proposition 2.21, show that the class $\mathscr{S}(\mathbb{R}^d)$ is stable by the Fourier Transform $\mathcal{F}$. Since $\mathscr{S}(\mathbb{R}^d) \subset L^1(\mathbb{R}^d)$, we can apply the Fourier inversion formula to deduce that *the Fourier Transform $\mathcal{F}$ is an isomorphism on the Schwartz space $\mathscr{S}(\mathbb{R}^d)$ of inverse $\check{\mathcal{F}}$* (remember the notation $\check{\theta}(x) = \theta(-x)$).

Definition 6.11 (Tempered distributions)**.** The dual of the Schwartz space $\mathscr{S}(\mathbb{R}^d)$ is called the set of *tempered distributions* and denoted $\mathscr{S}'(\mathbb{R}^d)$.

A tempered distribution is a distribution, since we have an injection $\mathcal{D}(\mathbb{R}^d) \hookrightarrow \mathscr{S}(\mathbb{R}^d)$, which automatically implies $\mathscr{S}'(\mathbb{R}^d) \subset \mathcal{D}'(\mathbb{R}^d)$. In more details, if $\alpha \in \mathscr{S}'(\mathbb{R}^d)$, then there exists $C \geq 0$, $N, k \geq 0$, such that

$$|\langle \alpha, u\rangle| \leq C q_{N,k}(u). \tag{6.54}$$

If K is a compact of $\mathbb{R}^d$, we have then $|\langle \alpha, u\rangle| \leq C(K) p_{K,k}(u)$, where

$$C(K) = C \sup_{x\in K} (1+|x|^2)^N.$$

We will now prove that $\mathcal{D}(\mathbb{R}^d)$ is dense in $\mathscr{S}(\mathbb{R}^d)$, with the consequence that a distribution satisfying (6.54) for some N, k, C and for all $u \in \mathcal{D}(\mathbb{R}^d)$ can be extended in a unique way as a tempered distribution.

Proposition 6.11 (Dense subset of the Schwartz space)**.** *The set $\mathcal{D}(\mathbb{R}^d)$ is dense in $\mathscr{S}(\mathbb{R}^d)$.*

The Schwartz space $\mathscr{S}(\mathbb{R}^d)$ is a Fréchet space. A sequence (u_n) converges to an element u in $\mathscr{S}(\mathbb{R}^d)$ if, for all neighbourhood V of the origin, $u_n \in u+V$ for n large enough. A base of neighbourhoods of the origin is given by the sets $\{q_{N,K} < \varepsilon\}$ so (u_n) converges to an element u in $\mathscr{S}(\mathbb{R}^d)$ if, and only if, for all N, k, ε, we have $q_{N,k}(u_n - u) < \varepsilon$ for n large enough, *i.e.* $q_{N,k}(u_n - u) \to 0$ when $n \to +\infty$. This can also be seen by considering the following metric (*cf.*(4.4))

$$d(u,v) = \sum_{j\in\mathbb{N}} 2^{-j} \Phi(q_{N_j,k_j}(u-v)), \quad \Phi(q) := \frac{q}{1+q}, \tag{6.55}$$

where $(N_j) \uparrow +\infty$, $(k_j) \uparrow +\infty$. Since the topology is metrizable, we can use a sequential criterion to characterize dense sets in $\mathscr{S}(\mathbb{R}^d)$.

Proof of Proposition 6.11. Let χ_1 be a bump function with $\bar{B}(0,1) \prec \chi_1 \prec B(0,2)$ and $\chi_n(x) = \chi_1(n^{-1}x)$. Let $u \in \mathscr{S}(\mathbb{R}^d)$ and let $u_n = u\chi_n$. Each $u_n \in \mathcal{D}(\mathbb{R}^d)$ and we are going to show that $u_n \to u$ in $\mathscr{S}(\mathbb{R}^d)$. Let $N, k \geq 0$. To prove $q_{N,k}(u - u_n) \to 0$, we first consider the case $k = 0$. We have then

$$\sup_{|x|>n} (1+|x|^2)^N |(1-\chi_n)u(x)| \leq \sup_{|x|>n} (1+|x|^2)^N |u(x)| \leq n^{-2} q_{N+1,0}(u),$$

so $q_{N,0}(u_n - u) \to 0$ when $n \to +\infty$. If $k > 0$ and $|m| = k$, then

$$\partial_x^m (u - u_n) = (1-\chi_n)\partial_x^m u + \eta_n,$$

where η_n contains some derivatives of u and some derivatives of χ_n or order at least 1. More precisely, the Leibniz formula gives

$$\eta_n = -\sum_{p+q=m, |p|\geq 1} \binom{m}{p} \partial_x^q u \partial_x^p \chi_n.$$

Since

$$|\nabla\chi_n(x)| = n^{-1} |(\nabla\chi_1)(n^{-1}x)| \leq \|\nabla\chi_1\|_{L^\infty(\mathbb{R}^d)} n^{-1},$$

and, similarly, $|\partial^m \chi_n(x)| \leq C_m(\chi_1) n^{-|m|}$, we have

$$q_{N,0}(\eta_n) \leq \sum_{p+q=m, |p|\geq 1} C(m,p,\chi_1) q_{N,k}(u) n^{-|p|} \leq C(m,\chi_1,u) n^{-1},$$

so $q_{N,0}(\eta_n) \to 0$ when $n \to 0$. The case $k = 0$ gives us

$$\lim_{n\to+\infty} q_{N,0}((1-\chi_n)\partial_x^m u) = 0.$$

Since $q_{N,0}$ is subadditive, we obtain

$$\lim_{n\to+\infty} q_{N,0}(\partial_x^m (u-u_n)) = 0.$$

Finally, the convergence $q_{N,k}(u-u_n) \to 0$ follows since

$$q_{N,k}(v) = \sup_{|m|\leq k} q_{N,0}(\partial_x^m v).$$

□

In parallel to this discussion on distributions and tempered distributions, we can discuss distributions and compactly supported distributions. The space of compactly supported distributions (on $\mathbb{R}^d$) is usually denoted by $\mathcal{E}'(\mathbb{R}^d)$. We have the following characterization of $\mathcal{E}'(\mathbb{R}^d)$. To state it, let us recall that $C^\infty(\mathbb{R}^d)$, the space of infinitely differentiable functions on $\mathbb{R}^d$ is a Fréchet space, when endowed with the family of semi-norms

$$p_{K,j}(u) = \sup_{x\in K} \sup_{|m|\leq j} |\partial^m u(x)|,$$

associated to the topology of uniform convergence of a function and its derivatives over compact sets. In particular, a linear map $C^\infty(\mathbb{R}^d) \to \mathbb{R}$ is continuous if, and only if, there exists a compact K, an index $j \in \mathbb{N}$ and $A \geq 0$ such that

$$|\alpha(u)| \leq A p_{K,j}(u), \tag{6.56}$$

for all $u \in C^\infty(\mathbb{R}^d)$.

Proposition 6.12 (Compactly supported distribution)**.** *The space $\mathcal{E}'(\mathbb{R}^d)$ of compactly supported distributions is the dual space to $C^\infty(\mathbb{R}^d)$. More precisely:*

1. *any element in the topological dual of $C^\infty(\mathbb{R}^d)$ gives rise (by restriction to $\mathcal{D}(\mathbb{R}^d)$) to a compactly supported distribution,*
2. *any compactly supported distribution can be extended in a unique way as a continuous linear functional on $C^\infty(\mathbb{R}^d)$.*

Proof of Proposition 6.12. If α is an element of the topological dual of $C^\infty(\mathbb{R}^d)$, we define $\langle \alpha, u\rangle = \alpha(u)$ for $u \in \mathcal{D}(\mathbb{R}^d)$. By hypothesis, α satisfies (6.56). If L is a compact subset of $\mathbb{R}^d$, and $u \in C_L^\infty(\mathbb{R}^d)$, then

$$p_{K,j}(u) \leq \sup_{x\in\mathbb{R}^d} \sup_{|m|\leq j} |\partial_x^m u(x)| = p_{K,j}(u),$$

so $|\langle \alpha, u\rangle| \leq A p_{L,j}(u)$ and α is a distribution (of finite order j). If u is supported in $\mathbb{R}^d \setminus K$, then $\langle \alpha, u\rangle = 0$ by (6.56), so α is supported in K. This proves the first assertion *1*. The second point *2.* uses the fact that $\mathcal{D}(\mathbb{R}^d)$ is dense in $C^\infty(\mathbb{R}^d)$. The proof of this result is easier than the proof of Proposition 6.11: we consider the same cut-off function χ_n and also set $u_n = \chi_n u$, the function u being given in $C^\infty(\mathbb{R}^d)$. If L is a compact subset of $\mathbb{R}^d$ and $k \in \mathbb{N}$, then $p_{L,k}(u_n - u) = 0$ as soon as $L \subset \bar{B}(0,n)$, so clearly $p_{L,k}(u_n - u) \to 0$ when $n \to +\infty$. Consider now a compactly supported distribution α. Our aim is to prove (6.56) for all $u \in \mathcal{D}(\mathbb{R}^d)$, then we can extend α by density as a continuous linear functional on $C^\infty(\mathbb{R}^d)$. Let χ be a bump function such that $\chi = 1$ on $\mathrm{supp}(\alpha)$. Let $K = \mathrm{supp}(\chi)$. There exists $j \in \mathbb{N}$ and $A \geq 0$ such that (6.56) is satisfies for all $u \in \mathcal{D}(\mathbb{R}^d)$ supported in K. If u is a general element of $\mathcal{D}(\mathbb{R}^d)$, then $\langle \alpha, u\rangle = \langle \alpha, \chi u\rangle$, so

$$|\langle \alpha, u\rangle| \leq A p_{K,j}(\chi u) \leq A' p_{K,j}(u) \tag{6.57}$$

for a given constant A' (we use the Leibniz formula to get the second inequality in (6.57), the procedure is standard now). This concludes the proof. □

Note well, to conclude this paragraph, that we have $\mathcal{D}(\mathbb{R}^d) \hookrightarrow \mathscr{S}(\mathbb{R}^d) \hookrightarrow C^\infty(\mathbb{R}^d)$, so any compactly supported distribution is a tempered distribution, both being distributions.

6.2.2 Operations on tempered distributions

In this section, we will give the details of some admissible operations on tempered distributions (Fourier transform in particular). Let us first list some examples of tempered distributions.

1. a compactly supported distribution is a tempered distribution,
2. if $u \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, then u is a tempered distribution, provided

$$\int_{\mathbb{R}^d} \frac{|u(x)|}{(1+|x|^2)^N} dx < +\infty,$$

   for some $N \geq 0$,
3. $u(x) = e^x$ is not tempered, but $u(x) = e^x \cos(e^x)$ is tempered (use integration by parts).

Exercise 6.12. Give a function $w \in C^\infty(\mathbb{R})$ such that

1. there is no polynomial function P such that $|w(x)| \leq |P(x)|$ for all $x \in \mathbb{R}$,
2. w defines a tempered distribution.

Multiplication by moderately growing function. A function $w \in C^\infty(\mathbb{R}^d)$ is said to be moderately growing (or slowly increasing sometimes) if w and all its derivatives have at most polynomial growth at infinity: for all $m \in \mathbb{N}^d$, there exists $N = N(m) \geq 0$ and $C = C(m) \geq 0$ such that

$$|\partial_x^m w(x)| \leq C(1+|x|^2)^N, \tag{6.58}$$

for all $x \in \mathbb{R}^d$. If $u \in \mathscr{S}(\mathbb{R}^d)$ and w is moderately growing, then $wu \in \mathscr{S}(\mathbb{R}^d)$ by the Leibniz formula, so we can define the product of w by a tempered distribution w as in (6.22): $\langle w\alpha, u\rangle := \langle \alpha, wu\rangle$.

Differentiation. If $m \in \mathbb{N}^d$ is a multi-index, then $\partial_x^m : \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^d)$ is continuous since $q_{N,k}(\partial_x^m u) \leq q_{N,k+|m|}(u)$. We can define the differentiation of a tempered distribution $\alpha \in \mathscr{S}'(\mathbb{R}^d)$ by

$$\langle \partial_x^m \alpha, u\rangle = (-1)^{|m|}\langle \alpha, u\rangle. \tag{6.59}$$

Fourier Transform. If $\alpha \in \mathscr{S}'(\mathbb{R}^d)$, Proposition 2.21 shows that we can define the tempered distribution $\hat{\alpha}$ by the duality formula (6.52). By duality, the Fourier Transform $\mathcal{F}$ is an isomorphism on $\mathscr{S}'(\mathbb{R}^d)$, of inverse $\check{\mathcal{F}}$. The duality (6.52) also transfers the properties (2.162) to the Fourier transform on $\mathscr{S}'(\mathbb{R}^d)$. Expressed in terms of $D_x^m = (2\pi i)^{-|m|}\partial_x^m$, we have therefore

$$\mathcal{F}(D_x^m \alpha) = \xi^m \mathcal{F}\alpha, \quad D_\xi^m \mathcal{F}\alpha = (-1)^{|m|}\mathcal{F}(x^m \alpha), \tag{6.60}$$

for all $\alpha \in \mathscr{S}'(\mathbb{R}^d)$. Note that each term in (6.60) is well defined: either we multiply a tempered distribution by a moderately growing function, or we differentiate a tempered distribution (or apply the Fourier transform), this is why it can be simply justified by duality. Since

$$\mathcal{F}u(0) = \int_{\mathbb{R}^d} u(x)dx = \langle 1, u \rangle, \quad u \in \mathscr{S}(\mathbb{R}^d),$$

(6.52) gives us the following expression for the Fourier transform of δ_0: $\mathcal{F}(\delta_0) = 1$. The Fourier transform maps the unit for the convolution to the unit for the punctual product, which is consistent with the homomorphism property (6.63) proved below.

Exercise 6.13. We use the notations in (6.41). Prove that

$$\mathcal{F}(P(D)\delta_0)(\xi) = P(\xi), \tag{6.61}$$

for all $\xi \in \mathbb{R}^d$. The formula (6.61) shows that, for $\alpha := P(D)\delta_0$, $\mathcal{F}(\alpha)$ is a function, and more precisely a moderately growing function. Generalize this to the case where α is a distribution with compact support, and prove that $\mathcal{F}(\alpha)(\xi) = \langle \alpha, e_\xi \rangle$ where $e_\xi \in C^\infty(\mathbb{R}^d)$ is given by $e_\xi(x) := \exp(-2\pi i x \cdot \xi)$.

Solution: by (6.60), we have $\mathcal{F}(P(D)\delta_0)(\xi) = P(\xi)\mathcal{F}(\delta_0) = P(\xi)$. We have seen in Proposition 6.12 that, if α is a distribution with compact support, then it can be extended in a unique manner as a continuous linear form on $C^\infty(\mathbb{R}^d)$ (we still denote by α this extension). Then $\langle \alpha, e_\xi \rangle$ is well defined. By linearity and continuity, we have

$$\int_{\mathbb{R}^d} u(\xi)\langle \alpha, e_\xi \rangle d\xi = \langle \alpha, \mathcal{F}u \rangle, \tag{6.62}$$

for all $u \in \mathcal{D}(\mathbb{R}^d)$. To prove (6.62), we approach the integral by a Riemann sum, as in the proof of Proposition 6.6 (we will not give the details). Since $\langle \alpha, \mathcal{F}u \rangle = \langle \mathcal{F}\alpha, u \rangle$, (6.62) is the identity $\langle \mathcal{F}\alpha, u \rangle = \langle \theta, u \rangle$, where $\theta(\xi) := \langle \alpha, e_\xi \rangle$. Admit for the moment that θ is a moderately growing function. It is a locally integrable function then, and the identity $\mathcal{F}\alpha = \theta$ follows from the injection of $L^1_{\text{loc}}(\mathbb{R}^d)$ in $\mathcal{D}'(\mathbb{R}^d)$. Since α is compactly supported, (6.56) is realized for some K, j, A. We have

$$|\partial_x^m e_\xi(x)| = |(-2\pi i \xi)^m| \le (2\pi)^{|m|}(1 + |\xi|^2)^{|m|/2},$$

so $p_{K,j}(e_\xi) \le (2\pi)^j(1 + |\xi|^2)^{j/2}$ and θ is therefore moderately growing.

Convolution. If $\alpha \in \mathscr{S}'(\mathbb{R}^d)$, $u \in \mathscr{S}(\mathbb{R}^d)$, we define the number $\alpha * u(x) = \langle \alpha, \tau_x \check{u} \rangle$ as in (6.27). We have the following results then.

Proposition 6.13 (Convolution $\mathscr{S}'$-$\mathscr{S}$)**.** *Let $\alpha \in \mathscr{S}'(\mathbb{R}^d)$, $w \in \mathscr{S}(\mathbb{R}^d)$, $m \in \mathbb{N}^d$. Then $\alpha * w$ is a moderately growing function and*

$$\mathcal{F}(\alpha * w) = \mathcal{F}\alpha \cdot \mathcal{F}w. \tag{6.63}$$

Moreover, the distributivity of the derivation holds true:

$$\partial_x^m(\alpha * w) = (\partial_x^m \alpha) * w = \alpha * \partial_x^m w. \tag{6.64}$$

We also have

$$\langle \alpha * w, u\rangle = \langle \alpha, u * \check{w}\rangle, \tag{6.65}$$

for all $u \in \mathscr{S}(\mathbb{R}^d)$.

Proof of Proposition 6.13. Let $A \geq 0$, $N, k \geq 0$ such that $|\langle \alpha, u\rangle| \leq A q_{N,k}(u)$ for all $u \in \mathscr{S}(\mathbb{R}^d)$. For $m \in \mathbb{N}^d$ with $|m| \leq k$ and $w \in \mathscr{S}(\mathbb{R}^d)$, we have

$$(1+|y|^2)^N |\partial_y^m \tau_x w(y)| = (1+|y|^2)^N |\partial_y^m w(y-x)| \leq C(N)(1+|x|^2)^N q_{N,k}(w), \tag{6.66}$$

for all $x, y \in \mathbb{R}^d$, where

$$C(N) = \left[\sup_{y\in\mathbb{R}^d} \frac{(1+|y-x|^2)}{(1+|x|^2)(1+|y|^2)}\right]^N.$$

The constant $C(N)$ is finite, with $C(N) \leq 2^N$, since

$$\langle x-y\rangle \leq \sqrt{2}\langle x\rangle\langle y\rangle, \quad \langle x\rangle := (1+|x|^2)^{1/2}. \tag{6.67}$$

Indeed, we have $|x-y|^2 \leq 2(|x|^2+|y|^2)$ and so $\langle x-y\rangle^2 \leq 2\langle x\rangle^2\langle y\rangle^2$. We deduce from (6.66) that

$$|\alpha * w(x)| \leq AC(N) q_{N,k}(w)(1+|x|^2)^N, \tag{6.68}$$

for all $x \in \mathbb{R}^d$. The estimate (6.68) shows that $\alpha * w$ is a locally bounded function, hence a distribution. The estimate (6.68) also shows that, given $u \in \mathcal{D}(\mathbb{R}^d)$, the map $w \mapsto \langle \alpha * w, u\rangle$ is continuous on $\mathscr{S}(\mathbb{R}^d)$. Therefore, since $\mathcal{D}(\mathbb{R}^d)$ is dense in $\mathscr{S}(\mathbb{R}^d)$, (6.65) is a consequence of the identity already proved in the case $w, u \in \mathcal{D}(\mathbb{R}^d)$ (see (6.29)). The distributivity of the derivation for the convolution on $\mathscr{S}(\mathbb{R}^d)$ and (6.65) imply now (6.64). If $n \in \mathbb{N}^d$, the result (6.68) applied to $\partial_x^n w$ instead of w gives the bound

$$|\partial_x^n(\alpha * w)(x)| \leq AC(N) q_{N,k+|n|}(w)(1+|x|^2)^N,$$

for all $x \in \mathbb{R}^d$. We deduce that $\alpha * w$ is a moderately growing function. To establish the formula $\mathcal{F}(\alpha * w) = \mathcal{F}\alpha \cdot \mathcal{F}w$, we use (6.52) and (6.65): for $u \in \mathscr{S}(\mathbb{R}^d)$,

$$\begin{aligned}\langle \mathcal{F}(\alpha * w), u\rangle = \langle \alpha * w, \mathcal{F}u\rangle = \langle \alpha, (\mathcal{F}u) * \check{w}\rangle = \langle \check{\mathcal{F}}\mathcal{F}\alpha, (\mathcal{F}u) * \check{w}\rangle = \langle \mathcal{F}\alpha, \check{\mathcal{F}}((\mathcal{F}u) * \check{w})\rangle \\ = \langle \mathcal{F}\alpha, u\mathcal{F}w\rangle = \langle \mathcal{F}w\mathcal{F}\alpha, u\rangle. \end{aligned} \tag{6.69}$$

In the sequence of equalities (6.69), the crucial identity is

$$\check{\mathcal{F}}((\mathcal{F}u) * \check{w}) = u\mathcal{F}w,$$

which is simply a consequence (6.63) in the case $\alpha \in \mathscr{S}(\mathbb{R}^d)$. This particular case follows from Fubini's theorem since

$$\int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^d} |\alpha(x-y)w(y)|dy\right]dx = \|\alpha\|_{L^1(\mathbb{R}^d)}\|w\|_{L^1(\mathbb{R}^d)} < +\infty,$$

and thus

$$\begin{aligned}\mathcal{F}(\alpha * w)(\xi) &= \int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^d} \alpha(x-y)w(y)dy\right]e^{-2\pi i x\cdot\xi}dx \\ &= \int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^d} \alpha(x-y)e^{-2\pi i x\cdot\xi}dx\right]w(y)dy = \int_{\mathbb{R}^d}(\mathcal{F}\alpha)(\xi)e^{-2\pi i y\cdot\xi}w(y)dy = (\mathcal{F}\alpha)(\xi)(\mathcal{F}w)(\xi).\end{aligned}$$

□

In Proposition 6.8, we have seen that the convolution product $\mathcal{D}' * \mathcal{D}$ can be extended to $\mathcal{D}' * \mathcal{E}'$. We can therefore define the convolution product of a tempered distribution with a compactly supported distribution. We will show that it is a tempered distribution.

Proposition 6.14 (Convolution of a tempered distribution with a compactly supported distribution)**.** *Let α be a tempered distribution and β be a compactly supported distribution. Then $\alpha * \beta$ is a tempered distribution.*

Proof of Proposition 6.14. By Exercise 6.13, $\mathcal{F}\beta$ is a moderately growing function. The product $\mathcal{F}\alpha\mathcal{F}\beta$ is a tempered distribution therefore, and the distribution $\check{\mathcal{F}}(\mathcal{F}\alpha\mathcal{F}\beta)$ is tempered. If $u \in \mathcal{D}(\mathbb{R}^d)$, we have

$$\langle \check{\mathcal{F}}(\mathcal{F}\alpha\mathcal{F}\beta), u\rangle = \langle \mathcal{F}\alpha\mathcal{F}\beta, \check{\mathcal{F}}u\rangle = \langle \mathcal{F}\alpha, \mathcal{F}\beta\check{\mathcal{F}}u\rangle = \langle \alpha, \mathcal{F}(\mathcal{F}\beta\check{\mathcal{F}}u)\rangle.$$

Since $\check{\mathcal{F}}u = \mathcal{F}\check{u}$ and $\mathcal{F}(\beta * \check{u}) = \mathcal{F}\beta\mathcal{F}\check{u}$ by (6.63), it turns out that

$$\check{\mathcal{F}}(\mathcal{F}\beta\check{\mathcal{F}}u) = \beta * \check{u}.$$

It follows that

$$\langle \check{\mathcal{F}}(\mathcal{F}\alpha\mathcal{F}\beta), u\rangle = \langle \alpha, \check{\beta} * u\rangle,$$

and as $\langle \alpha * \beta, u\rangle = \langle \alpha, \check{\beta} * u\rangle$ by definition, we see that $\alpha * \beta$ coincide on $\mathcal{D}(\mathbb{R}^d)$ with a tempered distribution, so is (after extension by density) a tempered distribution as well. □

6.2.3 Fundamental solutions of PDEs - II

Fundamental solution. To find fundamental solutions to evolution equations of the form $\partial_t u + P(D_x)u = 0$, where $P(D_x)$ is a linear differential operator with constant coefficients in the variable $x \in \mathbb{R}^d$ and $t \in \mathbb{R}$ or $t \geq 0$, we will use a partial Fourier transform in the variable x only. For $u \in \mathscr{S}(\mathbb{R} \times \mathbb{R}^d)$, we set (sp=space)

$$(\mathcal{F}_{\text{sp}}u)(t,\xi) = \int_{\mathbb{R}^d} u(t,x)e^{-2\pi i x\cdot\xi}dx, \quad \xi \in \mathbb{R}^d.$$

Then $\mathcal{F}_{\text{sp}} \colon \mathscr{S}(\mathbb{R} \times \mathbb{R}^d) \to \mathscr{S}(\mathbb{R} \times \mathbb{R}^d)$ is bijective of inverse

$$(\check{\mathcal{F}}_{\text{sp}}u)(t,x) = \int_{\mathbb{R}^d} u(t,x)e^{2\pi i x\cdot\xi}d\xi, \quad x \in \mathbb{R}^d.$$

By the duality formula $\langle \mathcal{F}_{\text{sp}}\alpha, u\rangle = \langle \alpha, \mathcal{F}_{\text{sp}}u\rangle$ we can extend $\mathcal{F}_{\text{sp}}$ to $\mathscr{S}'(\mathbb{R} \times \mathbb{R}^d)$. We have then the following identities in $\mathscr{S}'(\mathbb{R} \times \mathbb{R}^d)$

$$D_x^m \mathcal{F}_{\text{sp}}\alpha = \xi^m \mathcal{F}_{\text{sp}}\alpha, \quad \mathcal{F}_{\text{sp}}(X^m\alpha) = (-D_\xi)^m \mathcal{F}_{\text{sp}}\alpha, \tag{6.70}$$

where (recall (6.41)) $D_x^m = (2\pi i)^{-|m|}\partial_x^m$, and

$$\mathcal{F}_{\text{sp}}(\partial_t^n \alpha) = \partial_t^n \mathcal{F}_{\text{sp}}\alpha, \tag{6.71}$$

for $n \in \mathbb{N}, m \in \mathbb{N}^d$. A tempered distribution α is then a solution to the equation

$$\partial_t\alpha + P(D_x)\alpha = \delta_{(0,0)}, \tag{6.72}$$

(where $\delta_{(0,0)}$ denotes the Dirac mass at $(0,0)$) if, and only if

$$\partial_t \mathcal{F}_{\text{sp}}\alpha + P\mathcal{F}_{\text{sp}}\alpha = \mathcal{F}_{\text{sp}}\delta_{(0,0)}. \tag{6.73}$$

We compute

$$\langle \mathcal{F}_{\mathrm{sp}}\delta_{(0,0)}, u\rangle = (\mathcal{F}_{\mathrm{sp}}u)(0,0) = \int_{\mathbb{R}^d} u(0,x)dx = \langle \delta_{t=0}\otimes \mathbf{1}_{\mathrm{freq}}, u\rangle,$$

where (freq=frequency) $\mathbf{1}_{\mathrm{freq}}(\xi) = 1$ for all ξ. So a α is a fundamental solution for $\partial_t + P(D)$ if, and only if, $\mathcal{F}_{\mathrm{sp}}\alpha$ is solution to

$$\partial_t \mathcal{F}_{\mathrm{sp}}\alpha + P\mathcal{F}_{\mathrm{sp}}\alpha = \delta_{t=0}\otimes \mathbf{1}_{\mathrm{freq}}. \tag{6.74}$$

Weak solution to the Cauchy Problem. Suppose now that we have found a fundamental solution α for $\partial_t + P(D)$. We want to solve the Cauchy Problem

$$\begin{cases} \partial_t v + P(D_x)v = f & \text{in } (0,+\infty)\times\mathbb{R}^d \\ v(0,\cdot) = u_0 & \text{in } \mathbb{R}^d, \end{cases} \tag{6.75}$$

where the data are $u_0 \in \mathcal{D}(\mathbb{R}^d)$, $f \in \mathcal{D}(\mathbb{R}\times\mathbb{R}^d)$. We first merge both equations into the single equation

$$\partial_t v + P(D_x)v = f + \delta_{t=0}\otimes u_0, \tag{6.76}$$

and solve (6.76) by setting

$$v = \alpha * (f + \delta_{t=0}\otimes u_0). \tag{6.77}$$

This defines a tempered distribution since f is in the Schwartz class and $\delta_{t=0}\otimes u_0$, defined by

$$\langle \delta_{t=0}\otimes u_0, w\rangle = \int_{\mathbb{R}^d} u_0(x)w(0,x)dx, \quad w\in\mathscr{S}(\mathbb{R}\times\mathbb{R}^d),$$

is a compactly supported distribution (apply Proposition 6.14). The distributivity of the derivation (6.64) ensures that

$$\partial_t v + P(D_x)v = (\partial_t\alpha + P(D_x)\alpha) * (f+\delta_{t=0}\otimes u_0) = \delta_{(0,0)} * (f+\delta_{t=0}\otimes u_0) = f + \delta_{t=0}\otimes u_0.$$

Using duality, it appears that (6.76) is equivalent to the fact that

$$\langle v, \partial_t w + P(D_x)^* w\rangle + \int_{\mathbb{R}}\int_{\mathbb{R}^d} f(t,x)w(t,x)dxdt + \int_{\mathbb{R}^d} u_0(x)w(0,x)dx = 0, \tag{6.78}$$

for all $w\in\mathscr{S}(\mathbb{R}\times\mathbb{R}^d)$, where, given $P(\xi) = \sum_{|m|\le k} a_m\xi^m$, we set

$$P(D_x)^* = \sum_{|m|\le k} a_m(-1)^{|m|}D_x^m.$$

Assume now that

$$v \text{ is represented by a locally integrable function } u \text{ in } \mathscr{S}'(\mathbb{R}\times\mathbb{R}^d). \tag{6.79}$$

Then we can rewrite (6.78) as

$$\iint_{\mathbb{R}\times\mathbb{R}^d} u(t,x)(\partial_t w(t,x) + P(D_x)^* w(t,x))dxdt \\ + \iint_{\mathbb{R}\times\mathbb{R}^d} f(t,x)w(t,x)dxdt + \int_{\mathbb{R}^d} u_0(x)w(0,x)dx = 0, \tag{6.80}$$

for all $w\in\mathscr{S}(\mathbb{R}\times\mathbb{R}^d)$. Equation (6.80) is called the weak formulation of the Cauchy Problem (6.75).

Fundamental solution of the Heat Equation. The Heat Equation is associated to the operator $\partial_t - \Delta_x$, *i.e.* $P(\xi) = 4\pi^2|\xi|^2$. The equation (6.74) is then ($\beta = \mathcal{F}_{\rm sp}\alpha$)

$$\partial_t\beta(\cdot,\xi) + 4\pi^2|\xi|^2\beta(\cdot,\xi) = \delta_{t=0},$$

which we solve by setting

$$\beta(t,\xi) = H(t)e^{-4\pi^2|\xi|^2 t},$$

where $H(t) = \mathbf{1}_{t>0}$ is the Heavyside function. Since the Gaussian $G_a\colon x \mapsto e^{-\pi^2 a|x|^2}$ satisfies $\hat{G_a}(\xi) = a^{-d/2}G_{a^{-1}}(x)$ for $a > 0$ ([Fol99, Prop. 8.24, p.251]), we obtain the fundamental solution

$$\alpha(t,x) = H(t)K_t(x), \quad K_t(x) := \frac{1}{(4\pi t)^{d/2}}e^{-\frac{|x|^2}{4t}}, \quad (t,x) \in \mathbb{R}\times\mathbb{R}^d. \tag{6.81}$$

The set of functions $\{K_t; t > 0\}$ is called the Heat Kernel. Since α is a moderately growing function, the discussion above (and Fubini's theorem to justify the expression of $\alpha * f$ in (6.82) below) shows that

$$u(t) = H(t)K_t *_x u_0 + \int_0^t K_{t-s} *_x f(s)ds \tag{6.82}$$

satisfies

$$\iint_{\mathbb{R}\times\mathbb{R}^d} u(t,x)(\partial_t w(t,x) - \Delta w(t,x))dxdt \\ + \iint_{\mathbb{R}\times\mathbb{R}^d} f(t,x)w(t,x)dxdt + \int_{\mathbb{R}^d} u_0(x)w(0,x)dx = 0, \tag{6.83}$$

for all $w \in \mathscr{S}(\mathbb{R}\times\mathbb{R}^d)$. One can extend (6.82)-(6.83) for less regular/supported data u_0 and f in various ways, using the following property: for all $r \geq p$ and $t > 0$,

$$\|K_t * u\|_{L^r(\mathbb{R}^d)} \leq C(p,r)\frac{1}{t^{\frac{d}{2}(\frac{1}{p}-\frac{1}{r})}}\|u\|_{L^p(\mathbb{R}^d)}. \tag{6.84}$$

Exercise 6.14. Prove (6.84) and show that, given $f \in C(\mathbb{R}_+; L^p(\mathbb{R}^d))$, $u_0 \in L^p(\mathbb{R}^d)$, the Cauchy Problem

$$\begin{cases} \partial_t u - \Delta u = f & \text{in } (0,+\infty)\times\mathbb{R}^d \\ u(0,\cdot) = u_0 & \text{in } \mathbb{R}^d, \end{cases} \tag{6.85}$$

admits a weak solution $u \in C(\mathbb{R}_+; L^p(\mathbb{R}^d))$. *Hint:* use the Young inequality for convolution

$$\|K_t * u\|_{L^r(\mathbb{R}^d)} \leq \|K_t\|_{L^q(\mathbb{R}^d)}\|u\|_{L^p(\mathbb{R}^d)}, \quad \frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r},$$

and use the homogeneity property $K_t(x) = t^{-d/2}K_1(t^{-1/2}x)$ to prove (6.84).

Fundamental solution of the Wave Equation. The Cauchy Problem for the Wave Equation is

$$\begin{cases} \partial_t^2 u - \Delta u = f, & \text{in } (0,+\infty)\times\mathbb{R}^d \\ u(0,\cdot) = u_0,\ \partial_t u(0,\cdot) = v_0 & \text{in } \mathbb{R}^d. \end{cases} \tag{6.86}$$

It can be put in the form of a first order equation for unknowns taking values in $\mathbb{R}^2$:

$$\begin{cases} \partial_t U + P(D_x)U = F & \text{in } (0,+\infty)\times\mathbb{R}^d \\ U(0,\cdot) = U_0 & \text{in } \mathbb{R}^d, \end{cases} \tag{6.87}$$

where

$$U = \begin{pmatrix} u \\ v \end{pmatrix}, \; F = \begin{pmatrix} 0 \\ f \end{pmatrix}, \; P(D_x) = \begin{pmatrix} 0 & -1 \\ -\Delta & 0 \end{pmatrix}.$$

The Fourier Transform in x gives the equation

$$\partial_t B(t,\xi) + M(\xi)B(t,\xi) = \begin{pmatrix} 0 \\ \delta_{t=0} \otimes \mathbf{1}_{\text{freq}} \end{pmatrix}, \; M(\xi) = \begin{pmatrix} 0 & -1 \\ |\xi|^2 & 0 \end{pmatrix} \tag{6.88}$$

for the partial Fourier transform $B = \mathcal{F}_{\text{sp}} A$ of the fundamental solution A, with

$$\partial_t A + P(D_x)A = \begin{pmatrix} 0 \\ \delta_{(0,0)} \end{pmatrix}.$$

We solve (6.88) by taking $B(t,\xi) = H(t)E(t,\xi)$, where $E(0,\xi) = 1$ and

$$\partial_t E(t,\xi) + M(\xi)E(t,\xi) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Longrightarrow E(t,\xi) = \exp(-tM(\xi)) \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{6.89}$$

We compute $M(\xi)^2 = -|\xi|^2 \mathrm{I}_2$, so on the diagonal of $\exp(-tM(\xi))$ we will have

$$\sum_{n=0}^{\infty} (-1)^n \frac{(t|\xi|)^{2n}}{(2n)!} = \cos(t|\xi|).$$

Since $\partial_t \exp(-tM(\xi)) = -M(\xi)\exp(-tM(\xi))$ and $M(\xi)^{-1} = -|\xi|^{-2} M(\xi)$, we can deduce the terms on the anti-diagonal and obtain

$$\exp(-tM(\xi)) = \begin{pmatrix} \cos(t|\xi|) & \sin(t|\xi|)/|\xi| \\ -\sin(t|\xi|) & \cos(t|\xi|) \end{pmatrix}, \quad E(t,\xi) = \begin{pmatrix} \sin(t|\xi|)/|\xi| \\ \cos(t|\xi|) \end{pmatrix}.$$

We conclude that

$$B(t,\xi) = H(t) \begin{pmatrix} \sin(t|\xi|)/|\xi| \\ \cos(t|\xi|) \end{pmatrix}$$

gives an expression of the Fourier Transform in x of the fundamental solution of the wave equation. One then has to compute the inverse Fourier Transform. See [Zui02, p.122] for instance.

Fundamental solution of the Schrödinger Equation. The Schrödinger Equation is associated to the operator $\partial_t - i\Delta_x$, where $i^2 = -1$. The fundamental solution for the Schrödinger Equation is

$$\alpha(t,x) = \frac{H(t)}{t^{d/2}} K^S \left(\frac{x}{t^{1/2}} \right), \quad K^S(x) = \frac{e^{-id\pi/2}}{(4\pi)^{d/2}} e^{i\frac{|x|^2}{4t}}.$$

See [Zui02, p.123] on that topic.

7 Sobolev spaces

We follow for some parts L.C. Evans, [Eva10], and C. Zuily, [Zui02]. When the Sobolev space is defined on an open bounded subset U, the geometry of the boundary ∂U is involved in a number of results (extension, trace, approximation by functions smooth up to the boundary...). A number of authors work locally, using a transport by diffeomorphism onto a flat boundary. We use a global approach, by means of a parametrization of a tubular neighbourhood of the boundary based on the natural splitting into tangential and normal coordinates. The results we give are not optimal however. Those optimal statements can be found in the classical book by R.A. Adams, [Ada75]. We also refer to the on-line preprint by J. Droniou, [Dro01].

7.1 Definition and first properties

Definition 7.1 (Sobolev space)**.** Let U be an open subset. Let $k \in \mathbb{N}$, $p \in [1, +\infty]$. The space $W^{k,p}(U)$ is the set of functions in $L^p(U)$ such that $\partial_x^m u \in L^p(U)$ for all $m \in \mathbb{N}^d$ with $|m| \leq k$. The norm on $W^{k,p}(U)$ is

$$\|u\|_{W^{k,p}(U)} = \left[\sum_{|m|\leq k} \|\partial_x^m u\|_{L^p(U)}^p\right]^{1/p}, \quad 1 \leq p < +\infty, \tag{7.1}$$

and

$$\|u\|_{W^{k,\infty}(U)} = \max\left\{\|\partial_x^m u\|_{L^\infty(U)}; |m| \leq k\right\},$$

if $p = +\infty$.

Equivalent norms when $k = 1$. We use the following convention of notation

$$\|\nabla u\|_{L^p(U)} = \|v\|_{L^p(U)}, \quad v := |\nabla u|. \tag{7.2}$$

We have then

$$\|u\|_{W^{1,2}(U)} = \left[\|u\|_{L^2(U)}^2 + \|\nabla u\|_{L^2(U)}^2\right]^{1/2}.$$

When $p \in [1, +\infty)$ is different from 2 and $k = 1$, we will often use the norm

$$\|u\|_{W^{1,p}(U)} = \left[\|u\|_{L^p(U)}^p + \|\nabla u\|_{L^p(U)}^p\right]^{1/p}. \tag{7.3}$$

This is an abuse of notation, but the norm in (7.3) is equivalent to the norm in (7.1) when $k = 1$ by equivalence of the norms ℓ_2 and ℓ_p on $\mathbb{R}^d$.

Bessel spaces. When $p = 2$ and $U = \mathbb{R}^d$, the formula $\mathcal{F}(D_x^m \alpha) = \xi^m \mathcal{F}\alpha$ for α a tempered distribution (*cf.* (6.60)), shows that

$$W^{k,2}(\mathbb{R}^d) = \left\{u \in L^2(\mathbb{R}^d); \int_{\mathbb{R}^d} (1 + 4\pi^2|\xi|^2)^k |\mathcal{F}u(\xi)|^2 d\xi < +\infty\right\},$$

and that the norm $\|\cdot\|_{W^{k,2}}(\mathbb{R}^d)$ is equivalent to the norm

$$u \mapsto \left[\int_{\mathbb{R}^d} (1 + 4\pi^2|\xi|^2)^k |\mathcal{F}u(\xi)|^2 d\xi\right]^{1/2}.$$

For $s \in \mathbb{R}$, the Bessel space $H^{s,p}(\mathbb{R}^d)$ (also denoted $L^{s,p}(\mathbb{R}^d)$ is defined as the set of functions $u \in L^p(\mathbb{R}^d)$ such that the tempered distribution $(\mathrm{Id} - \Delta)^{s/2} u$ is (represented by) an element of $L^p(\mathbb{R}^d)$. The associated norm is

$$\|u\|_{H^{s,p}(\mathbb{R}^d)} = \|(\mathrm{Id} - \Delta)^{s/2} u\|_{L^p(\mathbb{R}^d)}$$

The operator $(\mathrm{Id} - \Delta)^{s/2}$ is defined in $\mathscr{S}'(\mathbb{R}^d)$ by the formula

$$\mathcal{F}\left[(\mathrm{Id} - \Delta)^{s/2} u\right] = (1 + 4\pi^2|\xi|^2)^{s/2} \mathcal{F}u.$$

It is clear then that $H^{k,2}(\mathbb{R}^d) = W^{k,2}(\mathbb{R}^d)$ when $k \in \mathbb{N}$ (the equality here means that they coincide as sets and that the norms are equivalent). That $H^{k,p}(\mathbb{R}^d) = W^{k,p}(\mathbb{R}^d)$ is satisfied more generally for $p \in (1, +\infty)$ is true, but the proof requires the theory of singular integrals

Weak estimates. In Definition 7.1, $\partial_x^m u$ is understood in the sense of distribution. This can be formalized by considering the injective map $L^p(U) \to \mathcal{D}'(U)$ which to $u \in L^p(U)$ associates the distribution obtained by integration of the test-functions against u. Nevertheless it is not strictly necessary to use the theory of distribution to assert that $\partial_x^m u \in L^p(U)$ in Definition 7.1. The notion of weak derivative in L^p is sufficient for that purpose, [Eva10, p.256]. In relation with the definition of Sobolev spaces by use of weak notions of derivatives, we give the following proposition.

Proposition 7.1 (Sobolev and weak inequalities)**.** *Let U be a non-empty open set in $\mathbb{R}^d$. Let $p \in [1, +\infty]$ and let $u \in L^p(U)$ satisfy the following estimate: there exists a constant $A \geq 0$ such that*

$$\left| \int_U u(x)\partial_{x_i} v(x)dx \right| \leq A\|v\|_{L^{p'}(U)}, \quad \forall v \in \mathcal{D}(U), \tag{7.4}$$

where $p' = \frac{p}{p-1}$ is the conjugate exponent to p. Then

1. *If $1 < p \leq +\infty$, then (7.4) is a characterization of the fact that $u \in W^{1,p}(U)$,*
2. *if $p = 1$, then (7.4) characterizes the fact that $u \in \mathrm{BV}(U)$, the space of functions "of bounded variation". This is the space of functions $u \in L^1(U)$ such that the derivatives $\partial_{x_i} u$ are represented by signed measures on U (recall that, with our convention, signed measures have a finite total variation precisely). This space is strictly bigger than $W^{1,1}(U)$.*

Proposition 7.1 prevents us (at least when $p = 1$) from the possibility of working on weak inequalities when manipulating Sobolev functions. However, weak inequalities are not sufficient in themselves. For instance, assume $1 < r \leq p, q$ and $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$. We expect the following statement to be true:

$$u \in W^{1,p}(U),\ v \in W^{1,q}(U) \Rightarrow uv \in W^{1,r}(U) \text{ and } \nabla(uv) = u\nabla v + v\nabla u \text{ a.e.} \tag{7.5}$$

How can we prove (7.5)? An answer will be furnished by Proposition 7.2 below (see Exercises class).

Proof of Proposition 7.1. Assume first $1 < p \leq +\infty$. Then $\mathcal{D}(U)$ is dense in $L^{p'}(U)$ (we have $p' < +\infty$) so $v \mapsto \langle \partial_{x_i} u, v\rangle$ can be extended as a continuous linear form on $L^{p'}(U)$. The dual space of $L^{p'}(U)$ is $L^p(U)$, so there exists $w_i \in L^p(U)$ such that $\langle \partial_{x_i} u, v\rangle = \langle w_i, v\rangle$ for all $v \in \mathcal{D}(\mathbb{R}^d)$. This means precisely that the distribution $\partial_{x_i} u$ is represented by an element of $L^p(U)$. If $p = 1$ now, the closure of $\mathcal{D}(U)$ for the L^∞ norm is the space $\mathrm{BC}(U)$ of bounded continuous functions on U. Actually, (7.4) for $p = 1$ tells us that the distribution $\partial_{x_i} u$ is of order 0, and we have seen that distributions of order 0 are signed measures. It is clear that $\mathrm{BV}(U)$ is strictly bigger than $W^{1,1}(U)$. The example with $d = 1$, $U = \mathbb{R}^d$, $u(x) = H(x)$ (Heavyside function), for which $u' = \delta_0$, can easily be generalized to higher dimension and general open non-empty set U. $\square$

Some examples. If $U = B(0,1)$ and $u(x) = |x|^s$, $s \in \mathbb{R}$, then $u \in W^{k,p}(U)$ if, and only if the integral

$$\int_{B(0,1)} |x|^{p(s-k)} dx = |S(0,1)| \int_0^1 r^{p(s-k)+d-1} dt$$

is finite, *i.e.* $s > \frac{k}{p} - d$. If $U = \mathbb{R}^d$ and $u = \mathbf{1}_{B(0,1)}$, we have $u \in L^p(\mathbb{R}^d)$ for all $p \in [1, +\infty]$. We have

$$\langle \nabla u, v\rangle = -\int_{S(0,1)} xv(x)d\sigma(x), \tag{7.6}$$

where σ is the surface measure on $S(0,1)$ (see Proposition 7.6 below for more details on that point, or simply consider polar coordinates). For any $p \in [1,+\infty]$, we cannot have $\nabla u \in L^p(\mathbb{R}^d;\mathbb{R}^d)$, otherwise σ would be absolutely continuous with respect to the Lebesgue measure. So u is not belonging to any Sobolev space $W^{k,p}(\mathbb{R}^d)$, $k \geq 1$.

Local approximation by smooth functions. We give now a result of approximation by smooth functions (restricted to the case $k=1$ for facility), the approximation being valid only locally in U. Better approximation results can be given, as in Theorem 7.7 below, but this first statement will be useful in many places.

Proposition 7.2 (Local approximation by smooth functions)**.** *Let U be an open subset of $\mathbb{R}^d$. Let $1 \leq p < +\infty$ and let $u \in W^{1,p}(U)$. There is a sequence of functions $u_n \in C^\infty(U) \cap W^{1,p}(U)$ such that $\|u - u_n\|_{W^{1,p}(V)} \to 0$ when $n \to +\infty$, for all open set V with compact closure $\bar{V} \subset U$.*

Proof of Proposition 7.2. Let (ρ_n) be a standard approximation of the unit, as in (2.139). Let $\tilde{u}$ denote the extension[6] of u by 0 outside U. Set $v_n = \rho_n * \tilde{u}$ and define u_n as the restriction of v_n to U. We will see that (u_n) has the desired property. Let V be an open set with compact closure $\bar{V} \subset U$. There exists $\eta > 0$ and W an open set with compact closure $\bar{W} \subset U$ which contains the η-neighbourhood $V + B(0,\eta)$. Let χ be a bump function such that $\bar{W} \prec \chi \prec U$. Then $\tilde{u} = \chi u$ in W, so

$$v_n = w_n := \rho_n * (\chi u), \quad u = w := \chi u$$

in V if $n > \eta^{-1}$ since ρ_n is supported in $B(0,n^{-1})$. It will be sufficient now to prove that $w_n \to w$ in $W^{1,p}(\mathbb{R}^d)$. Standard results on the convolution give $w_n \to w$ in $L^p(\mathbb{R}^d)$. By Proposition 6.6, we know that $\partial_{x_i} w_n = \rho_n * (\partial_{x_i} w)$ (identity between C^∞ functions). Again, standard results on the convolution with an approximation of the unit give $\partial_{x_i} w_n \to \partial_{x_i} w$ in $L^p(\mathbb{R}^d)$. This concludes the proof. $\square$

Calculus of variations. The Sobolev norms appear in a great number of classical problems in the theory of partial differential equations (PDEs), especially when one looks for some a priori estimates on the solution. Actually, some standard equations are derived as Euler-Lagrange equations for a functional that involves some Sobolev norms. The most famous example is the Dirichlet energy (where $f \in L^2(U)$ for instance)

$$J(u) = \int_U \frac{|\nabla u|^2}{2} dx - \int_U f u dx. \tag{7.7}$$

Suppose that J admits a minimum over a set $A \subset W^{1,2}(U)$. Suppose also that u is *interior* to A. Then u is a critical point: for all $v \in W^{1,2}(U)$, $DJ(u) \cdot v = 0$. Note that the functional J is the sum of a quadratic term with a linear term, so we compute

$$J(u+v) = J(u) + \int_U (\nabla u \cdot \nabla v - f v) dx + \int_U \frac{|\nabla v|^2}{2} dx. \tag{7.8}$$

The last term in (7.8) is smaller than $\|v\|^2_{W^{1,2}(U)}$, so

$$DJ(u) \cdot v = \int_U (\nabla u \cdot \nabla v - f v) dx. \tag{7.9}$$

[6] beware that we may have $\tilde{u} \notin W^{1,p}(\mathbb{R}^d)$, see the example (7.6). See also Theorem 7.8 for the construction of accurate extensions.

By integration by parts (we will discuss later to what extent this is licit), we also obtain the expression

$$\forall v \in W_0^{1,2}(U),\ DJ(u)\cdot v = \int_U (-\Delta u - f) v dx, \tag{7.10}$$

where the subscript 0 in $v \in W_0^{1,2}(U)$ indicates that $v \in W^{1,2}(U)$ and $v = 0$ on ∂U (this will also been explained in greater details). Finally (modulo a discussion on the boundary conditions that should be done), we obtain the Poisson Equation $-\Delta u = f$ as the Euler-Lagrange equation (*i.e.* the equation for critical points) of the Dirichlet Energy.

7.2 Sobolev spaces defined on the whole space

Here is a list of properties of the Sobolev spaces $W^{k,p}(\mathbb{R}^d)$. We employ the following notation: if $u \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, we denote by u^* the precise representative of u defined by

$$u^*(x) = \begin{cases} \lim\limits_{r\to 0+} \dfrac{1}{|B(x,r)|}\displaystyle\int_{B(x,r)} u(y)dy & \text{if the limit exists,} \\ 0 & \text{otherwise.} \end{cases} \tag{7.11}$$

By Lebesgue's differentiation theorem, Theorem 2.27, we know that $u^* = u$ a.e., so u^* is in the class of u for the equivalence relation of "equality a.e." By definition also, we have $u^* = v^*$ everywhere if $u = v$ a.e. We use this precise representative when some continuity properties of u are put forward. Indeed, if v is continuous, then $v = v^*$ everywhere, so if u admits a representative that is continuous, then this representative is necessarily u^*.

1. For all $p \in [1, +\infty]$, the Sobolev spaces $W^{k,p}(\mathbb{R}^d)$ are complete.
2. For $1 \leq p < +\infty$, the Sobolev spaces $W^{k,p}(\mathbb{R}^d)$ are separable (false for $p = +\infty$, *cf.* the case $k = 0$).
3. Approximation by smooth functions: $\mathcal{D}(\mathbb{R}^d)$ is dense in $W^{k,p}(\mathbb{R}^d)$ for $1 \leq p < +\infty$ (false for $p = +\infty$, *cf.* the case $k = 0$)
4. A function $u \in W^{1,\infty}(\mathbb{R}^d)$ if, and only if, u^* is a globally Lipschitz function.
5. Operation on Sobolev spaces: if $F \in C^1(\mathbb{R})$ is has bounded derivatives and $F(0) = 0$, then $u \mapsto F \circ u$ defines a linear continuous map $W^{1,p}(\mathbb{R}^d) \to W^{1,p}(\mathbb{R}^d)$ and we have
$$\partial_{x_i} F(u) = F'(u)\partial_{x_i} u \tag{7.12}$$
for all $i \in \{1, \dots, d\}$. *Note: the result can be extended to the case where F is a globally Lipschitz continuous function, see Section 7.4.*
6. Gagliardo-Nirenberg-Sobolev Inequality: if $1 \leq p < d$, then
$$\|u\|_{L^{p^*}(\mathbb{R}^d)} \leq \frac{p(d-1)}{d-p}\|\nabla u\|_{L^p(\mathbb{R}^d)}, \quad \frac{1}{p^*} := \frac{1}{p} - \frac{1}{d}, \tag{7.13}$$
for all $u \in W^{1,p}(\mathbb{R}^d)$, so we have a continuous injection $W^{1,p}(\mathbb{R}^d) \hookrightarrow L^{p^*}(\mathbb{R}^d)$.
7. Morrey's inequality: if $d < p$, then there exists a constant $C(p,d) \geq 0$ such that
$$\|u^*\|_{\mathrm{BC}^{0,\mu}(\mathbb{R}^d)} \leq C(p,d)\|u\|_{W^{1,p}(\mathbb{R}^d)}, \quad \mu := 1 - \frac{d}{p}, \tag{7.14}$$
for all $u \in W^{1,p}(\mathbb{R}^d)$, so we have a continuous injection $W^{1,p}(\mathbb{R}^d) \hookrightarrow \mathrm{BC}^{0,\mu}(\mathbb{R}^d)$ (modulo the map $u \mapsto u^*$).

In (7.14), we have introduced the space $\mathrm{BC}^{0,\mu}(\mathbb{R}^d)$ of μ-Hölder bounded continuous functions u (here $0 \le \mu \le 1$), which consists of functions $u \in \mathrm{BC}(\mathbb{R}^d)$ (bounded and continuous functions) with finite norm

$$\|u\|_{\mathrm{BC}^{0,\mu}(\mathbb{R}^d)} = \|u\|_{\mathrm{BC}(\mathbb{R}^d)} + [u]_{C^{0,\mu}(\mathbb{R}^d)},$$

where

$$\|u\|_{\mathrm{BC}(\mathbb{R}^d)} = \sup_{x\in\mathbb{R}^d} |u(x)|, \quad [u]_{C^{0,\mu}(\mathbb{R}^d)} := \sup_{x\neq y\in\mathbb{R}^d} \frac{|u(x)-u(y)|}{|x-y|^\mu}.$$

As a closed subspace of the Banach space of bounded functions $\mathrm{B}(\mathbb{R}^d)$ (endowed with the sup-norm), $\mathrm{BC}^{0,\mu}(\mathbb{R}^d)$ is a Banach space.

The points *1.*-*2.*-*3.*-*5.* will be proved in exercises class. For *4.*, we refer to Theorem 7.16 below.

Remark 7.1 (Index of regularity). Define the following indexes of regularity:

$$\mathrm{ind}(W^{k,p}(\mathbb{R}^d)) = \frac{k}{d} - \frac{1}{p}, \quad \mathrm{ind}(\mathrm{BC}^{k,\mu}(\mathbb{R}^d)) = \frac{k+\mu}{d}.$$

Then note that the index of regularity is preserved in the "Sobolev's injections" (7.13) and (7.14). An other way to find a posteriori the values of p^* and μ is to consider the behaviour of the various norms at stake under the rescaling $u \mapsto \theta_\lambda u$, where $\theta_\lambda u(x) = u(\lambda x)$ (see Exercises class). It works for (7.14) if you replace it with the "homogeneous" estimate

$$[u]_{C^{0,\mu}(\mathbb{R}^d)} \le C(p,d)\|\nabla u\|_{L^p(\mathbb{R}^d)}, \quad \mu := 1 - \frac{d}{p}, \tag{7.15}$$

which is true (see the proof of Morrey's inequality below, see (7.22) in particular).

7.2.1 Proof of Morrey's inequality

Let $A \subset \mathbb{R}^d$ be Borel set with positive Lebesgue measure $|A|$. We introduce the notation

$$⨍_A f(x)dx = \frac{1}{|A|}\int_A f(x)dx,$$

where f is a non-negative measurable function. By C_i, $i = 1, 2, \ldots$ we denote various depending on the dimension d and on the exponent p only. We will use some results established in Section 2.5.4, in particular the estimate

$$|u(x) - u(y)| \le C_1|x-y|(M_{2r}[|\nabla u|](x) + M_{2r}[|\nabla u|](y)), \quad r = |x-y|, \tag{7.16}$$

where

$$M_r[|\nabla u|](x) = \sup_{0<t<r} ⨍_{B(x,t)} |\nabla u(z)|dz. \tag{7.17}$$

In (7.16), the function u is assumed to be smooth and compactly supported: $u \in \mathcal{D}(\mathbb{R}^d)$. The estimate (7.16) is deduced from (2.263), which was established for a vector field $a\colon \mathbb{R}^d \to \mathbb{R}^d$. Considering $a(x) = u(x)e_1$ where e_1 is the first basis vector of $\mathbb{R}^d$ will indeed give (7.16). The estimate (7.16) as such is not completely satisfactory because, if we want to use the L^p-bound (2.249) on the local maximal function, we will need to integrate somehow with respect to x and y. We will first prove the following inequality: for all $\rho \le 1$,

$$\left| ⨍_{B(x,\rho)} u - ⨍_{B(x,\rho/2)} u \right| \le C_2\rho^\mu \|\nabla u\|_{L^p(\mathbb{R}^d)}. \tag{7.18}$$

By iteration and triangular inequality, (7.18) gives

$$\left| \fint_{B(x,\rho)} u - \fint_{B(x,2^{-n}\rho)} u \right| \leq C_2 \sum_{k=0}^{n-1} 2^{-k\mu} \rho^{\mu} \|\nabla u\|_{L^p(\mathbb{R}^d)} \leq C_3 \rho^{\mu} \|\nabla u\|_{L^p(\mathbb{R}^d)}, \tag{7.19}$$

where

$$C_3 = C_2 \sum_{k=0}^{\infty} 2^{-k\mu} = \frac{C_2}{1-2^{-\mu}}.$$

Taking the limit $[n \to +\infty]$, (7.19) gives

$$\left| \fint_{B(x,\rho)} u - u(x) \right| \leq C_3 \rho^{\mu} \|\nabla u\|_{L^p(\mathbb{R}^d)}. \tag{7.20}$$

To conclude then, we may apply (7.20) with $\rho = r = |x-y|$ to obtain, by triangular inequality,

$$|u(x) - u(y)| \leq 2C_3 r^{\alpha} \|\nabla u\|_{L^p(\mathbb{R}^d)} + \left| \fint_{B(x,r)} u - \fint_{B(y,r)} u \right|. \tag{7.21}$$

The bound

$$[u]_{C^{0,\mu}(\mathbb{R}^d)} \leq C_4 \|\nabla u\|_{L^p(\mathbb{R}^d)}, \tag{7.22}$$

will follow from the inequality

$$\left| \fint_{B(x,r)} u - \fint_{B(y,r)} u \right| \leq C_5 r^{\alpha} \|\nabla u\|_{L^p(\mathbb{R}^d)}, \quad r = |x-y|. \tag{7.23}$$

Both (7.18) and (7.23) can be deduced from the single estimate

$$\left| \fint_{B(x,\rho_1)} u - \fint_{B(y,\rho_2)} u \right| \leq C_6 (|x-y| + \rho_1 + \rho_2)(\rho_1^{-d/p} + \rho_2^{-d/p}) \|\nabla u\|_{L^p(\mathbb{R}^d)}. \tag{7.24}$$

Let us establish (7.24). We have

$$\left| \fint_{B(x,\rho_1)} u - \fint_{B(y,\rho_2)} u \right| \leq \fint_{B(x,\rho_1)} \fint_{B(y,\rho_2)} |u(x') - u(y')| dx' dy'.$$

If $x' \in B(x,\rho_1)$ and $y' \in B(y,\rho_2)$, then $|x'-y'| \leq R := |x-y| + \rho_1 + \rho_2$ so

$$|u(x') - u(y')| \leq C_1 (|x-y| + \rho_1 + \rho_2)(M_{2R}[|\nabla u|](x') + M_{2R}[|\nabla u|](y'))$$

by (7.16). By integration with respect to x' and y', we obtain

$$\begin{aligned} &\left| \fint_{B(x,\rho_1)} u - \fint_{B(y,\rho_2)} u \right| \\ &\qquad \leq \frac{C_1(|x-y| + \rho_1 + \rho_2)}{|B(x,\rho_1)||B(y,\rho_2)|} \left[\int_{B(x,\rho_1)} \int_{B(y,\rho_2)} (M_{2R}[|\nabla u|](x') + M_{2R}[|\nabla u|](y')) dx' dy' \right], \end{aligned}$$

that is

$$\left| \fint_{B(x,\rho_1)} u - \fint_{B(y,\rho_2)} u \right|$$
$$\leq C_1(|x-y| + \rho_1 + \rho_2) \left[\fint_{B(x,\rho_1)} M_{2R}[|\nabla u|] + \fint_{B(y,\rho_2)} M_{2R}[|\nabla u|](y') \right]. \quad (7.25)$$

By Hölder's inequality and (2.249), we have

$$\fint_{B(x,\rho_1)} M_{2r}[|\nabla u|] \leq \frac{C_6}{\rho_1^d} |B(x,\rho_1)|^{1/p'} \|M_{2R}[|\nabla u|]\|_{L^p(B(x,\rho_1)} \leq C_7 \rho_1^{-d/p} \|\nabla u\|_{L^p(\mathbb{R}^d)}.$$

Together with the estimate (7.25), this gives us (7.24). To complete the estimate (7.22), we must also give a bound

$$\|u\|_{\mathrm{BC}(\mathbb{R}^d)} \leq C_8 \|u\|_{W^{1,p}(\mathbb{R}^d)}. \quad (7.26)$$

We start from (7.20) with $\rho = 1$ to write

$$|u(x)| \leq \left| \fint_{B(x,1)} u \right| + \left| \fint_{B(x,1)} u - u(x) \right| \leq C_9 \int_{B(x,1)} |u(x)| dx + C_3 \|\nabla u\|_{L^p(\mathbb{R}^d)}. \quad (7.27)$$

By Hölder's inequality, we have

$$\int_{B(x,1)} |u(x)| dx \leq C_{10} \int_{B(x,1)} |u(x)|^p dx \leq C_{10} \|u\|_{L^p(\mathbb{R}^d)},$$

hence

$$|u(x)| \leq C_{10} \|u\|_{L^p(\mathbb{R}^d)} + C_3 \|\nabla u\|_{L^p(\mathbb{R}^d)}.$$

Taking the sup over $x \in \mathbb{R}^d$, we obtain (7.26). At this point, we have proved (7.14) when $u \in \mathcal{D}(\mathbb{R}^d)$. If u is an arbitrary function of $W^{1,p}(\mathbb{R}^d)$, we consider a sequence (u_n) of $\mathcal{D}(\mathbb{R}^d)$ which converges to u in $W^{1,p}(\mathbb{R}^d)$. The Morrey estimate applied to $u_p - u_q$ shows that (u_n) is Cauchy in the Banach space $\mathrm{BC}^{0,\mu}(\mathbb{R}^d)$: it admits a limit v in $\mathrm{BC}^{0,\mu}(\mathbb{R}^d)$. By convergence of the norms, we have

$$\|v\|_{\mathrm{BC}^{0,\mu}(\mathbb{R}^d)} \leq C(p,d) \|u\|_{W^{1,p}(\mathbb{R}^d)}, \quad (7.28)$$

Since both convergences in $W^{1,p}(\mathbb{R}^d)$ and $\mathrm{BC}^{0,\mu}(\mathbb{R}^d)$ imply convergence in the sense of distributions, we have $u = v$ in $\mathcal{D}'(\mathbb{R}^d)$, and therefore $u = v$ as $L^1_{\mathrm{loc}}(\mathbb{R}^d)$ functions, which means that $u = v$ a.e. It follows that $u^* = v^*$. Since v is continuous, $v = v^* = u^*$ and (7.28) is the desired Morrey estimate. (7.14).

Remark 7.2 (Limit case $p = d$). When $p = d$, the space dimension, we have $W^{1,d}(\mathbb{R}^d) \hookrightarrow \mathrm{BMO}$, where BMO is the space of functions with *bounded mean oscillations*. The mean oscillation of a locally integrable function u on a set A of positive measure being defined as

$$\fint_A |u - u_A| dx, \quad u_A := \fint_A u dx,$$

BMO is the space of locally integrable functions u such that the semi-norm

$$[u]_{\mathrm{BMO}} = \sup \left\{ \fint_Q |u - u_Q| dx;\ Q \text{ cube in } \mathbb{R}^d \right\}$$

is finite. This semi-norm is a norm on the quotient space BMO/$\mathbb{R}$, where $\mathbb{R}$ represents the set of constant functions on $\mathbb{R}^d$, and the resulting space is a Banach space. The function $x \mapsto \ln|x|$ is an example of functions in BMO which is not bounded. If $\chi \in \mathcal{D}(\mathbb{R}^d)$ and $\chi(0) \neq 0$, then $x \mapsto \chi(x)\ln(|\ln|x||)$ is a function in $W^{1,d}(\mathbb{R}^d)$ which is not in $L^\infty(\mathbb{R}^d)$.

Let us prove that there exists a constant $A(d)$ such that

$$[u]_{\mathrm{BMO}} \leq A(d)\|\nabla u\|_{L^d(\mathbb{R}^d)}, \tag{7.29}$$

for all $u \in W^{1,d}(\mathbb{R}^d)$. Assume first $u \in \mathcal{D}(\mathbb{R}^d)$. Let $x, y \in Q$, where $Q \subset \mathbb{R}^d$ is a cube of side length ℓ and volume ℓ^d. We have $|x-y| \leq r := C_{11}\ell$ so, by integration in (7.16), we obtain

$$⨍_Q |u - u_Q|dx \leq ⨍_Q ⨍_Q |u(x') - u(y')|dx'dy' \leq C_{12}\ell ⨍_Q M_{2r}[|\nabla u|]dx. \tag{7.30}$$

By the Hölder inequality,

$$⨍_Q M_{2r}[|\nabla u|]dx = \frac{1}{|Q|}\int_Q M_{2r}[|\nabla u|]dx \leq |Q|^{-1+\frac{d-1}{d}}\left[\int_Q |M_{2r}[|\nabla u|]|^d dx\right]^{1/d},$$

and thus

$$⨍_Q M_{2r}[|\nabla u|]dx \leq \ell^{-1}\|\nabla u\|_{L^d(\mathbb{R}^d)}, \tag{7.31}$$

by (7.17). Returning to (7.30), we see that (7.31) gives (7.29). In the general case where $u \in W^{1,d}(\mathbb{R}^d)$, we approximate u in $W^{1,d}(\mathbb{R}^d)$ by a sequence (u_n) of elements of $\mathcal{D}(\mathbb{R}^d)$. The estimate (7.29) applied to $u_p - u_q$ shows that (u_n) is Cauchy in BMO/$\mathbb{R}$, so $u_n \to u + c$ for the BMO semi-norm, where c is a constant. Passing to the limit $[n \to +\infty]$ in the estimate

$$[u_n]_{\mathrm{BMO}} \leq A(d)\|\nabla u_n\|_{L^d(\mathbb{R}^d)}$$

yields (7.29).

7.2.2 Proof of the Gagliardo-Nirenberg-Sobolev Inequality

We will use the following result.

Lemma 7.3 (Gagliardo)**.** *Let $d \geq 2$, let $w_1, \ldots, w_d$ be some positive functions in $L^{d-1}(\mathbb{R}^{d-1})$ and let v be a non-negative measurable function $\mathbb{R}^d \to \mathbb{R}$. Assume*

$$v(x) \leq \prod_{i=1}^{d} w_i(\hat{x}_i), \tag{7.32}$$

for all $x \in \mathbb{R}^d$, where $\hat{x}_i$ is the vector of $\mathbb{R}^{d-1}$ obtained by removing the component x_i from the components of x: $\hat{x}_i = (x_j)_{j \neq i}$. Then $v \in L^1(\mathbb{R}^d)$ and

$$\|v\|_{L^1(\mathbb{R}^d)} \leq \prod_{i=1}^{d-1} \|w_i\|_{L^{d-1}(\mathbb{R}^{d-1})}. \tag{7.33}$$

Proof of Lemma 7.3. Let us consider the case $d = 2$ first: we have $v(x, y) \leq w_1(y)w_2(x)$ so $\|v\|_{L^1(\mathbb{R}^2)} \leq \|w_1\|_{L^1(\mathbb{R})}\|w_2\|_{L^1(\mathbb{R})}$ by Fubini's theorem. When $d = 3$ now, we have

$$v(x_1, x_2, x_3) \leq w_1(x_2, x_3)w_2(x_1, x_3)w_3(x_1, x_2).$$

By integration with respect to x_1 and the Cauchy-Schwarz inequality,

$$\int_{\mathbb{R}} v(x_1,x_2,x_3)dx_1 \le w_1(x_2,x_3)\left[\int_{\mathbb{R}} |w_2(x_1,x_3)|^2 dx_1\right]^{1/2}\left[\int_{\mathbb{R}} |w_3(x_1,x_2)|^2 dx_1\right]^{1/2}.$$

We integrate with respect to x_2 next and use the Cauchy-Schwarz inequality again to obtain

$$\int_{\mathbb{R}^2} v(x_1,x_2,x_3)dx_1 dx_2 \le \left[\int_{\mathbb{R}} |w_1(x_2,x_3)|^2 dx_2\right]^{1/2}\left[\int_{\mathbb{R}} |w_2(x_1,x_3)|^2 dx_1\right]^{1/2} \|w_3\|_{L^2(\mathbb{R}^2)}.$$

At last, we integrate with respect to x_3 and use the Cauchy-Schwarz inequality one last time to get (7.33). In greater dimensions $d \ge 4$, the proof follows the same procedure, except that we need to replace the Cauchy-Schwarz inequality by the generalized Hölder inequality

$$\int_{\mathbb{R}} |z_1(t)\dots z_k(t)|dt \le \prod_{i=1}^{k} \|z_i\|_{L^k(\mathbb{R})}, \tag{7.34}$$

which can be deduced from the normalized case $\|z_i\|_{L^k(\mathbb{R})} = 1$ and from the convexity inequality

$$z_1 \dots z_k \le \frac{1}{k}\left(z_1^k + \dots + z_k^k\right), \tag{7.35}$$

for $z_1, \dots, z_k \ge 0$. The inequality (7.35) is a convexity inequality when written for $z_i = e^{-t_i}$. □

We can now give the proof of (7.13). We begin with the case $p = 1$. We have $1^* = \frac{d}{d-1}$ then. By an argument of approximation similar to the one used in Section 7.2.1, it is sufficient to consider the case $u \in \mathcal{D}(\mathbb{R}^d)$. We can write

$$u(x) = \int_{-\infty}^{x_1} \partial_{x_1} u(y_1, x_2, \dots, x_d)dy_1,$$

and obtain

$$|u(x)| \le [w_1(\hat{x}_1)]^{d-1}, \quad w_1(\hat{x}_1) := \left[\int_{\mathbb{R}} |\partial_{x_1} u(y_1, x_2, \dots, x_d)|dy_1\right]^{\frac{1}{d-1}}.$$

A similar reasoning on the $(d-1)$ other variables gives

$$|u(x)| \le [w_i(\hat{x}_i)]^{d-1}, \quad w_i(\hat{x}_i) := \left[\int_{\mathbb{R}} |\partial_{x_i} u(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_d)|dy_i\right]^{\frac{1}{d-1}}.$$

It follows that

$$v(x) := |u(x)|^{1^*} = |u(x)|^{\frac{d}{d-1}} \le \prod_{i=1}^{d} w_i(\hat{x}_i).$$

Each w_i is in $L^{d-1}(\mathbb{R}^{d-1})$ with $\|w_i\|_{L^{d-1}(\mathbb{R}^{d-1})} \le \|\nabla u\|_{L^1(\mathbb{R}^d)}^{\frac{1}{d-1}}$. By Gagliardo's lemma 7.3, $u \in L^{\frac{d}{d-1}}(\mathbb{R}^d)$ and

$$\|u\|_{L^{\frac{d}{d-1}}(\mathbb{R}^d)} = \|v\|_{L^1(\mathbb{R}^d)}^{\frac{d-1}{d}} \le \left[\prod_{i=1}^{d} \|\nabla u\|_{L^1(\mathbb{R}^d)}^{\frac{1}{d-1}}\right]^{\frac{d-1}{d}} = \|\nabla u\|_{L^1(\mathbb{R}^d)}.$$

When $1 < p < d$, we consider $v = |u|^r$ (and still $u \in \mathcal{D}(\mathbb{R}^d)$) where $r > 0$ will be chosen later. By approximation of $s \mapsto |s|^r$ by smooth functions, we can justify the identity $|\nabla v| = r|u|^{r-1}|\nabla u|$. By (7.13) in the case $p = 1$, we can estimate

$$\left[\int_{\mathbb{R}^d} |u|^{r\frac{d}{d-1}} dx\right]^{\frac{d-1}{d}} \leq r \int_{\mathbb{R}^d} |u|^{r-1}|\nabla u| dx.$$

By Hölder's inequality,

$$\left[\int_{\mathbb{R}^d} |u|^{r\frac{d}{d-1}} dx\right]^{\frac{d-1}{d}} \leq r \left[\int_{\mathbb{R}^d} |u|^{p'(r-1)} dx\right]^{1/p'} \|\nabla u\|_{L^p(\mathbb{R}^d)}. \tag{7.36}$$

We choose r such that $r\frac{d}{d-1} = p'(r-1)$. We expect this common value to be p^*, *i.e.*

$$r\frac{d}{d-1} = \frac{dp}{d-p} = p'(r-1),$$

which is realized indeed for $r = \frac{p(d-1)}{d-p}$ since

$$p'(r-1) = \frac{p}{p-1}\frac{p(d-1)-(d-p)}{d-p} = \frac{dp}{d-p}.$$

With this choice of the parameter r, (7.36) reads

$$\|u\|^r_{L^{p^*}(\mathbb{R}^d)} \leq r\|u\|^{r-1}_{L^{p^*}(\mathbb{R}^d)}\|\nabla u\|_{L^p(\mathbb{R}^d)},$$

which gives (7.13).

7.2.3 Sobolev's embeddings

In the following result, we use the following notion: a map $T\colon E \to F$ between two Banach spaces E and F is said to be *compact* if it maps bounded sets on relatively compact sets.

Theorem 7.4 (Sobolev's embeddings)**.** 1. *If $1 \leq p < d$, then we have a continuous injection $W^{1,p}(\mathbb{R}^d) \hookrightarrow L^r(\mathbb{R}^d)$ for all $r \in [p, p^*]$. If $r \in [p, p^*)$, we obtain a compact operator by restriction to a bounded domain. More precisely, if A is a bounded set in $W^{1,p}(\mathbb{R}^d)$ and $R > 0$, then*

$$A_R := \{u|_{B(0,R)}; u \in A\} \tag{7.37}$$

is relatively compact in $L^r(B(0,R))$.

2. *For $p = d$, we have an injection $W^{1,d}(\mathbb{R}^d) \hookrightarrow L^r(\mathbb{R}^d)$ for all $r \in [d, +\infty)$. If A is a bounded set in $W^{1,d}(\mathbb{R}^d)$ and $R > 0$, then A_R defined in (7.37) is relatively compact in $L^r(B(0,R))$ for all $r \geq d$.*

3. *If $d < p$, we have (modulo the composition with $u \mapsto u^*$) a continuous injection $W^{1,d}(\mathbb{R}^d) \hookrightarrow \mathrm{BC}^{0,\nu}(\mathbb{R}^d)$, for all $0 < \nu \leq \mu := 1 - d/p$. If $\nu < \mu$, we obtain a compact operator by restriction to a compact set: if A is bounded in $W^{1,p}(\mathbb{R}^d)$ and $R > 0$, then the set*

$$A_R := \left\{u^*|_{\bar{B}(0,R)}; u \in A\right\} \tag{7.38}$$

is relatively compact in $\mathrm{BC}^{0,\nu}(\bar{B}(0,R))$.

Remark 7.3 (Sliding bump function). The Sobolev norm on $\mathbb{R}^d$ is invariant by translation. If $\psi \in \mathcal{D}(\mathbb{R}^d)$ and (z_n) is a sequence of points in $\mathbb{R}^d$, then $(\tau_{z_n}\psi)$ is therefore a bounded sequence in $W^{1,p}(\mathbb{R}^d)$. By considering a sequence (z_n) such that $|z_n| \to +\infty$, we see the necessity to restrict to bounded or compact sets in Theorem 7.4 to obtain some results of compactness.

Proof of Theorem 7.4. **Step 1. Case** $1 \le p < d$**.** If $u \in L^p \cap L^q(\mathbb{R}^d)$ with $1 \le p < q < +\infty$ and $r \in [p,q]$, then

$$\|u\|_{L^r(\mathbb{R}^d)} \le \|u\|_{L^p(\mathbb{R}^d)}^{\theta}\|u\|_{L^q(\mathbb{R}^d)}^{1-\theta}, \quad \frac{1}{r} = \frac{\theta}{p} + \frac{1-\theta}{q}. \tag{7.39}$$

To prove (7.39), one can use Hölder's inequality: for $1 \le m \le +\infty$,

$$\|u\|_{L^r(\mathbb{R}^d)}^r = \int_{\mathbb{R}^d} |u|^{r\theta}|u|^{r(1-\theta)}dx \le \|u\|_{L^{mr\theta}}^{r\theta}\|u\|_{L^{m'r(1-\theta)}}^{r(1-\theta)}.$$

We choose m such that $mr\theta = p$, then

$$\frac{1}{m'} = 1 - \frac{1}{m} = 1 - \frac{r\theta}{p} = \frac{r(1-\theta)}{q},$$

so $m'r(1-\theta) = q$. It follows from (7.13) and (7.39) with $q = p^*$ that, for $1 \le p < d$, we have a continuous injection $W^{1,p}(\mathbb{R}^d) \hookrightarrow L^r(\mathbb{R}^d)$ for all $r \in [p,p^*]$. Assume $r \in [p,p^*)$ now and let A be a bounded set in $W^{1,p}(\mathbb{R}^d)$, say $\|u\|_{W^{1,p}(\mathbb{R}^d)} \le M$ for all $u \in A$. To prove that A_R is relatively compact in $L^r(B(0,R))$, we will use the Riesz-Fréchet-Kolmogorov criterion, Theorem 2.15. The set A_R is bounded in $L^r(B(0,R))$ since we have proved that $W^{1,p}(\mathbb{R}^d) \hookrightarrow L^r(\mathbb{R}^d)$. We now have to prove the uniform convergence of the translations (equi-continuity in L^r): we use (7.39) to obtain

$$\|\tau_z u - u\|_{L^r(\mathbb{R}^d)} \le \|\tau_z u - u\|_{L^p(\mathbb{R}^d)}^{\theta}\|\tau_z u - u\|_{L^{p^*}(\mathbb{R}^d)}^{1-\theta} \le 2\|u\|_{L^{p^*}(\mathbb{R}^d)}^{1-\theta}\|\tau_z u - u\|_{L^p(\mathbb{R}^d)}^{\theta}. \tag{7.40}$$

In (7.40), we have $\frac{1}{r} = \frac{\theta}{p} + \frac{1-\theta}{p^*}$, so $\theta > 0$ since $r < p^*$. We use the estimate (2.118) (which remains true for $u \in W^{1,p}(\mathbb{R}^d)$ by a density argument) and the bound $\|u\|_{L^{p^*}(\mathbb{R}^d)} \le C(p,d)M$ to obtain $\|\tau_z u - u\|_{L^r(\mathbb{R}^d)} \le C(p,d,M)\eta^{\theta}$. This gives the desired uniform control on the translations. We deduce that A_R is compact in $L^r(B(0,R)$.

Step 2. Case $p = d$**.** Once we have established the continuous injection $W^{1,d}(\mathbb{R}^d) \hookrightarrow L^r(\mathbb{R}^d)$ for all $r \in [d,+\infty)$, the proof of compactness is the same as in step 1, so we concentrate on the injection itself. Let $F \in C^1(\mathbb{R})$ with bounded derivatives, and such that $F(0) = 0$. This function F will be seen as a C^1 approximation of a truncate function $s \mapsto T_k(s)$ equal to s if $|s| \le k$ and to k if $|s| > k$. More precisely, we will assume $T_1 \le F \le T_2$. We have then $F(u) \in L^\infty(\mathbb{R}^d)$ with $\|F(u)\|_{L^\infty(\mathbb{R}^d)} \le 2$ and also $F(u) \in L^p(\mathbb{R}^d)$ if $u \in L^p(\mathbb{R}^d)$ since $|F(u)| \le |u|$, so

$$\|F(u)\|_{L^r(\mathbb{R}^d)} \le 2^{1-d/r}\|F(u)\|_{L^d(\mathbb{R}^d)}^{d/r} \le 2^{1-d/r}\|u\|_{L^d(\mathbb{R}^d)}^{d/r}, \tag{7.41}$$

if $u \in L^d(\mathbb{R}^d)$ and $r \ge d$. The function $H\colon u \mapsto u - F(u)$ is of class C^1, with bounded derivatives and satisfies $H(u) = 0$ for $|u| \le 1$. We have therefore

$$\nabla H(u) = H'(u)\nabla u = \mathbf{1}_{\{|u|\ge 1\}}H'(u)\nabla u.$$

The set $\{|u| \ge 1\}$ has finite measure, which can be bounded from above with the Markov inequality:

$$|\{|u| \ge 1\}| = |\{|u|^d \ge 1\}| \le \|u\|_{L^d(\mathbb{R}^d)}^d.$$

If $p \in [1,d)$, if follows by the Hölder inequality that

$$\|\nabla H(u)\|_{L^p(\mathbb{R}^d)} \le \|H'\|_{L^\infty} |\{|u|\ge 1\}|^{1/p^*} \|\nabla u\|_{L^d(\mathbb{R}^d)} \le \|H'\|_{L^\infty} \|u\|_{L^d(\mathbb{R}^d)}^{d/p^*} \|\nabla u\|_{L^d(\mathbb{R}^d)}.$$

By the Gagliardo-Nirenberg-Sobolev inequality (7.13),

$$\|H(u)\|_{L^{p^*}(\mathbb{R}^d)} \le \frac{p(d-1)}{d-p} \|H'\|_{L^\infty} \|u\|_{L^d(\mathbb{R}^d)}^{d/p^*} \|\nabla u\|_{L^d(\mathbb{R}^d)} \le \frac{p(d-1)}{d-p} \|H'\|_{L^\infty} \|u\|_{W^{1,d}(\mathbb{R}^d)}^{1+d/p^*}. \quad (7.42)$$

Any $r \ge d$ can be written $r = p^*$ where $p \in [1,d)$, so (7.41) and (7.42) give $u \in L^r(\mathbb{R}^d)$ with

$$\|u\|_{L^r(\mathbb{R}^d)} \le C(r,d) \left[\|u\|_{W^{1,d}(\mathbb{R}^d)}^{d/r} + \|u\|_{W^{1,d}(\mathbb{R}^d)}^{1+d/r} \right]. \quad (7.43)$$

To obtain the right power of $\|u\|_{W^{1,d}(\mathbb{R}^d)}$, we apply (7.43) to λu, $\lambda > 0$, to get

$$\|u\|_{L^r(\mathbb{R}^d)} \le C(r,d) \left[\lambda^{d/r-1} \|u\|_{W^{1,d}(\mathbb{R}^d)}^{d/r} + \lambda^{d/r} \|u\|_{W^{1,d}(\mathbb{R}^d)}^{1+d/r} \right]. \quad (7.44)$$

Taking $\lambda = \|u\|_{W^{1,d}(\mathbb{R}^d)}^{-1}$ gives $\|u\|_{L^r(\mathbb{R}^d)} \le 2C(r,d)\|u\|_{W^{1,d}(\mathbb{R}^d)}$ as expected.

Step 3. Case $d < p$. The Morrey inequality (7.14) gives the injection $W^{1,d}(\mathbb{R}^d) \hookrightarrow \mathrm{BC}^{0,\nu}(\mathbb{R}^d)$, as already seen. If $0 < \nu < \mu$, and $x, y \in \mathbb{R}^d$, then

$$\frac{|u(x)-u(y)|}{|x-y|^\nu} = \left[\frac{|u(x)-u(y)|}{|x-y|^\mu} \right]^{\nu/\mu} |u(x)-u(y)|^{1-\nu/\mu} \le 2^{1-\nu/\mu} \|u\|_{\mathrm{BC}(\mathbb{R}^d)}^{1-\nu/\mu} [u]_{\mathrm{BC}^{0,\mu}(\mathbb{R}^d)}^{\nu/\mu},$$

so

$$[u]_{\mathrm{BC}^{0,\nu}(\mathbb{R}^d)} \le 2^{1-\nu/\mu} \|u\|_{\mathrm{BC}(\mathbb{R}^d)}^{1-\nu/\mu} [u]_{\mathrm{BC}^{0,\mu}(\mathbb{R}^d)}^{\nu/\mu}. \quad (7.45)$$

By (7.45), there is an injection $\mathrm{BC}^{0,\mu}(\mathbb{R}^d) \hookrightarrow \mathrm{BC}^{0,\nu}(\mathbb{R}^d)$. Assume now that A is bounded in $W^{1,p}(\mathbb{R}^d)$ and let $K = \bar{B}(0,R)$. The set A_R defined in (7.38) is compact in $C(K)$ by Ascoli's Theorem. If (u_n) is a sequence in A, then there is a sequence still denoted (u_n) such that $v_n := u_n^*|_K$ is converging in $C(K)$. A proof similar to (7.45) gives

$$[v_p - v_q]_{\mathrm{BC}^{0,\nu}(K)} \le 2^{1-\nu/\mu} \|v_p - v_q\|_{C(K)}^{1-\nu/\mu} [v_p - v_q]_{\mathrm{BC}^{0,\mu}(K)}^{\nu/\mu}.$$

Since $[v_p - v_q]_{\mathrm{BC}^{0,\nu}(K)}$ is bounded from above uniformly in p, q, we see that (v_n) is Cauchy in $\mathrm{BC}^{0,\nu}(K)$, which is complete, so (v_n) is convergent in $\mathrm{BC}^{0,\nu}(K)$. □

7.3 Sobolev spaces defined on bounded domains

In this section, we study the Sobolev space $W^{1,p}(U)$, $p \in [1,+\infty]$, where U is a bounded open subset of $\mathbb{R}^d$. Various results (extension, trace, approximation, Green's formula) involve the geometry of the boundary ∂U of U. Section 7.3.1 gives the material to treat these questions of geometry.

7.3.1 Geometry of the boundary of bounded open sets

Definition 7.2 (Regular bounded open set)**.** Let U be an bounded open set of $\mathbb{R}^d$, $d \ge 2$. We say that U is of class C^k ($k \ge 1$) if there exists a function $\rho \colon \mathbb{R}^d \to \mathbb{R}$ of class C^k such that

$$U = \{x \in \mathbb{R}^d; \rho(x) < 0\}, \quad \nabla\rho(x) \ne 0, \ \forall x \in \partial U. \quad (7.46)$$

Remark 7.4 (Global equation). Let $\Gamma = \partial U$. If U is of class C^k, then, for x sufficiently close to a $z \in \Gamma$, $d_x\rho\colon v \mapsto \nabla\rho(x)\cdot v$ is a surjective map $\mathbb{R}^d \to \mathbb{R}$ since $\nabla\rho(x) \neq 0$. So ρ is a submersion, and Γ admits the global equation $\{\rho = 0\}$. In particular, Γ is a sub-manifold of dimension $d-1$ of $\mathbb{R}^d$ (an hypersurface of $\mathbb{R}^d$). The outward unit normal $\nu(z)$ to a point $z \in \Gamma$ is given by $\nu(z) = \frac{\nabla\rho(z)}{|\nabla\rho(z)|}$. It is a continuous function of z, so the manifold Γ is orientable and the normal bundle is trivial. It is known, conversely, that an hypersurface of $\mathbb{R}^d$ with trivial normal bundle admits a global equation.

Remark 7.5 (Jordan-Brouwer separation theorem). We say that U is C^k if the boundary ∂U is an hypersurface of class C^k and, in our description, ∂U is compact, orientable, closed as a manifold (*i.e.* without boundary). The Jordan-Brouwer theorem, [Lim88, McG16], states, conversely, that a compact, orientable, closed, connected hypersurface Γ of $\mathbb{R}^d$ separates $\mathbb{R}^d$ into two connected open components. One of this component, say U, is bounded since Γ is bounded, and we have then $\Gamma = \partial U$.

The outward unit normal $\nu(z)$ to a point $z \in \Gamma$ is given by $\nu(z) = \frac{\nabla\rho(z)}{|\nabla\rho(z)|}$. We will need to compute the differential of ν at a point x close to Γ. When U is of class C^2 at least, a careful calculation gives

$$d_x\nu(v)\cdot w = (A(x)v)\cdot w, \tag{7.47}$$

with

$$A(x) = \frac{1}{|\nabla\rho(x)|}\left[D^2\rho(x) - \frac{1}{|\nabla\rho(x)|^2}[\nabla\rho(x)\otimes\nabla\rho(x)]D^2\rho(x)\right]. \tag{7.48}$$

In (7.48), we have used the following notation: for $a, b \in \mathbb{R}^d$, the matrix $a\otimes b$ is the rank-1 matrix with (i,j)-element given by $a_i b_j$. It is clear that the matrix $A(x)$ is symmetric. By the spectral theorem, there is an orthogonal basis $(\varepsilon_i(x))_{1,d}$ in which $d_x\nu$ is diagonal. We can assume moreover that $\varepsilon_d(x) = \nu(x)$. Indeed, by differentiation of the identity $|\nu(x)|^2 = 1$, we have $d_x\nu(v) = 0$ for $v = \nu(x)$. Let $z \in \Gamma$. The tangent space $T_z\Gamma$ is the orthogonal to $\nu(z)$, so $(\varepsilon_i(z))_{1,d-1}$ is an orthogonal basis of $T_z\Gamma$. The eigenvalues $(\lambda_i(z))_{1,d-1}$ of $A(z)$ associated to $(\varepsilon_i(z))_{1,d-1}$ are called the *principal curvatures* of Γ at z.

7.3.2 Local parametrization in a tubular neighbourhood of ∂U

Consider the local parametrization of U near $\Gamma = \partial U$ given by

$$x = z - t\nu(z),\quad \nu(z) = \frac{\nabla\rho(z)}{|\nabla\rho(z)|},\quad z\in\Gamma,\ t>0. \tag{7.49}$$

We denote by $V_\varepsilon(\Gamma)$ the ε-neighbourhood

$$V_\varepsilon(\Gamma) = \{x\in\mathbb{R}^d; d(x,\Gamma) < \varepsilon\}.$$

We first prove the following result.

Proposition 7.5 (Parametrization of a tubular neighbourhood of Γ)**.** *Let U be an open bounded set of class C^k, according to Def. 7.2, with $k \geq 2$. Then there exists $\varepsilon > 0$ such that $V_\varepsilon(\Gamma)$ is diffeomorphic to $\Gamma\times(-\varepsilon,\varepsilon)$ via the C^{k-1}-map*

$$\Phi\colon \Gamma\times(-\varepsilon,\varepsilon)\to\mathbb{R}^d,\quad (z,t)\mapsto z - t\nu(z).$$

Moreover, each point $\Phi(z,t)$ has a unique closest point in Γ, namely z.

Proof of Proposition 7.5. Note that Φ can be seen as the restriction of a function $\tilde{\Phi}$ defined for η small enough on $V_\eta(\Gamma) \times (-\varepsilon, \varepsilon)$ and given by $(y,t) \mapsto y - t\nu(y)$. We have

$$d_{(x,t)}\tilde{\Phi}(y,s) = y - s\nu(x) - td_x\nu(y), \quad (y,s) \in \mathbb{R}^d \times \mathbb{R}. \tag{7.50}$$

We will deduce the expression of $d_{(z,0)}\Phi$ by restriction to $T_z\Gamma \times \mathbb{R}$: if $v \in T_z\Gamma$, with $v = \dot{\gamma}(0)$, $z = \gamma(0)$, where $\{\gamma(s); s \in (-\delta, \delta)\}$ is a curve in Γ, we have

$$\frac{d}{ds}\Phi(\gamma(s), t) = d_{(z,t)}\tilde{\Phi}(\dot{\gamma}(s), 0) = \dot{\gamma}(s) - td_z\nu(\dot{\gamma}(s)).$$

It follows that

$$d_{(z,t)}\Phi(v,s) = v - s\nu(z) - td_z\nu(v). \tag{7.51}$$

The differential $d_{(z,0)}\Phi\colon (v,s) \mapsto v - s\nu(z)$ is invertible: the equation $d_{(z,0)}\Phi(v,s) = y$ is solved by

$$y = v - s\nu(z), \quad s := -y \cdot \nu(z), \quad v := y + s\nu(z).$$

This corresponds to the orthogonal decomposition of y on $\langle\nu(z)\rangle^\perp \oplus \langle\nu(z)\rangle$. By the local inverse mapping theorem, Φ is a local diffeomorphism. To show that it is a global diffeomorphism on its image, we need to show that it is injective when ε is sufficiently small. Assume on the contrary that there are some sequences (z_n, t_n) and (z'_n, s_n) in $\Gamma \times (-1/n, 1/n)$ such that $\Phi(z_n, t_n) = \Phi(z'_n, s_n)$. Then $s_n, t_n \to 0$ and, by compactness, we can assume, by consideration of a subsequence if necessary, that $z_n \to z$, $z'_n \to z'$. Then $\Phi(z_n, t_n) \to z$, $\Phi(z'_n, s_n) \to z$, and so $z = z'$. Since Φ is a diffeomorphism in restriction to a neighbourhood of $V \times (-\delta, \delta)$ of $(z, 0)$, we will obtain a contradiction for n large enough. Let now $\Sigma_\varepsilon := \Phi(\Gamma \times (-\varepsilon, \varepsilon))$. We have $\Sigma_\varepsilon \subset V_\varepsilon(\Gamma)$. Indeed, if $x = z - t\nu(z) \in \Sigma_\varepsilon$, then $x \in V_\varepsilon(\Gamma)$ since

$$d(x, \Gamma) \leq |x - z| = t|\nu(z)| = t < \varepsilon.$$

Denote by $z_* \in \Gamma$ the point at which $d(x,\Gamma)$ is realized. The sphere of center x and radius $d(x,\Gamma)$ is tangent to Γ at z_*, so $x - z_*$ is orthogonal to $T_{z_*}\Gamma$. Indeed, let $(\gamma(s))_{|s|<\delta}$ be a curve in Γ such that $\gamma(0) = z_*$. Then $\varphi\colon s \mapsto |x - \gamma(s)|^2$ has a minimum at $s = 0$. By differentiation,

$$0 = \dot{\varphi}(0) = 2(x - z_*) \cdot \dot{\gamma}(0).$$

Since $T_{z_*}\Gamma^\perp$ is the straight line generated by $\nu(z_*)$, we have $x = z_* - \lambda\nu(z_*)$, where $\lambda = d(x,\Gamma)$ necessarily. As Φ is injective on $\Gamma \times (-\varepsilon, \varepsilon)$, we obtain $z = z_*$ and $t = \lambda = d(x,\Gamma)$. This concludes the proof. □

7.3.3 Measure on the boundary

Proposition 7.6 (Measure on the boundary)**.** *Assume that U is of class C^2. Then there is a finite measure σ on Γ endowed with the σ-algebra of Borel sets such that the Green formula*

$$\int_U \Psi(x) \cdot \nabla u(x)dx = -\int_U \mathrm{div}(\Psi)(x)u(x)dx + \int_{\partial U} \Psi(z) \cdot \nu(z)u(z)d\sigma(z) \tag{7.52}$$

is satisfied for all $\Psi \in C^1(\mathbb{R}^d; \mathbb{R}^d)$ and $u \in C^1(\mathbb{R}^d)$. We also have the following change of variable formula: let u be a continuous and bounded function on $V_\varepsilon(\Gamma)$. The integral of u on $V_\varepsilon(\Gamma)$ is given by

$$\int_{V_\varepsilon(\Gamma)} u(x)dx = \int_\Gamma \int_{-\varepsilon}^{\varepsilon} u \circ \Phi(z,t)\pi(z,t)d\sigma(z)dt, \tag{7.53}$$

with

$$\pi(z,t) := \prod_{i=1}^{d-1}(1 - t\lambda_i(z)), \tag{7.54}$$

where $(\lambda_i(z))_{1,d-1}$ *are the principal curvatures of* Γ *at* z*.*

The proof of Proposition 7.6 is given in Appendix A.

Remark 7.6 (Green's formula with $u \equiv 1$). It is sufficient to establish (7.52) in the case $u \equiv 1$:

$$\int_U \operatorname{div}(\Psi)(x)dx = \int_{\partial U} \Psi(z) \cdot \nu(z) d\sigma(z). \tag{7.55}$$

Apply (7.55) to Ψu then, and use the formula $\operatorname{div}(\Psi u) = \operatorname{div}(\Psi)u + \Psi \cdot \nabla u$ to recover the original Green formula (7.52).

Remark 7.7 (Expression in local coordinates). Assume that u is a continuous function supported in an open set W such that U and $\Gamma = \partial U$ admit a parametrization by local graph in W: there exists an open set V in $\mathbb{R}^{d-1}$ and a function $\psi \colon V \to \mathbb{R}$ of class C^k such that

$$U \cap W = \{(x', x_d) \in V \times \mathbb{R}; x_d > \psi(x')\} \cap W, \quad \Gamma \cap W = \{(x', x_d) \in V \times \mathbb{R}; x_d = \psi(x')\} \cap W. \tag{7.56}$$

Such a parametrization is always possible if W is chosen sufficiently small, and if we authorize ourselves a rotation of the axis (which means precisely that we should replace $U \cap W$ and $\Gamma \cap W$ in (7.56) by $R(U \cap W)$ and $R(\Gamma \cap W)$ respectively, where R is a rotation). Note that we obtain in particular a parametrization of $\Gamma \cap W$ by

$$g \colon V \to W, \quad x' \mapsto (x', \psi(x')).$$

We have then the following local expression of the measure σ:

$$\int_\Gamma u d\sigma = \int_{\mathbb{R}^{d-1}} u(g(x'))\sqrt{1 + |\nabla\psi(x')|^2}dx' = \int_{\mathbb{R}^{d-1}} u(x', \psi(x'))\sqrt{1 + |\nabla\psi(x')|^2}dx'. \tag{7.57}$$

Example 7.8 (Polar coordinates). Take $U = B(0, r)$. Using polar coordinates, we have, for $\varepsilon < r$,

$$\int_{V_\varepsilon(\Gamma)} u(x)dx = \int_{S(0,1)} \int_{r-\varepsilon}^{r+\varepsilon} u(sz)s^{d-1} ds d\sigma_{S(0,1)}(z).$$

Use the change of variable $z' = rz$ to write

$$\int_{V_\varepsilon(\Gamma)} u(x)dx = \int_\Gamma \int_{r-\varepsilon}^{r+\varepsilon} u(sz/r)(s/r)^{d-1} ds r^{d-1} d\sigma(z).$$

where σ is the measure on Γ. We have $\nu(z) = \frac{z}{r}$ for $z \in \Gamma$. Use next the change of variable $(s/r)z = z - t\nu(z)$, *i.e.* $s = r - t$. We get the formula

$$\int_{V_\varepsilon(\Gamma)} u(x)dx = \int_\Gamma \int_{-\varepsilon}^{\varepsilon} u(z - t\nu(z)) \left(1 - \frac{t}{r}\right)^{d-1} dt d\sigma(z).$$

This corresponds to (7.53) since all the principal curvatures are identically equal to r^{-1}.

7.3.4 Approximation by smooth functions

Theorem 7.7 (Approximation by smooth functions)**.** *Let U be a bounded open set of $\mathbb{R}^d$ of class C^2. Denote by $\mathrm{BC}^\infty(\bar{U})$ the set of restrictions to U of functions in $C^\infty(\mathbb{R}^d)$. If $1 \leq p < +\infty$, then $\mathrm{BC}^\infty(\bar{U})$ is dense in $W^{1,p}(U)$.*

Remark 7.9 (Approximation by functions smooth in the interior). Without any hypothesis on the regularity of ∂U, one can show that $C^\infty(U) \cap W^{1,p}(U)$ is dense in $W^{1,p}(U)$, [Eva10, p.265]. Functions in $C^\infty(U) \cap W^{1,p}(U)$ may be unbounded near the boundary of U, whereas functions in $\mathrm{BC}^\infty(\bar{U})$ are smooth "up to the boundary".

Proof of Theorem 7.7. We divide the proof into three steps.

Step 1. Definition of the approximation. Let (ρ_η) be a standard approximation of the unit:

$$\rho_\eta(x) = \eta^{-d}\rho_1(\eta^{-1}x), \quad \rho_1 \in C^\infty_{\bar{B}(0,1)}(\mathbb{R}^d), \quad \rho_1 \geq 0, \quad \int_{\mathbb{R}^d} \rho_1(x)dx = 1.$$

Let $\tilde{u}$ denote the extension of u by 0 outside U. We fix an $\varepsilon > 0$ such that the parametrization (7.49) of the tubular neighbourhood $V_{2\varepsilon}(\Gamma)$ of Γ is valid. Let $\eta \in (0, \varepsilon)$. If $x \in V_\eta(\Gamma)$, with $x = \Phi(z,t) = \Phi(z(x), t(x))$, we denote by

$$s_\eta(x) = \Phi(z(x), t(x) + \eta) = x - \eta\nu(z(x))$$

the point obtained by shifting x from a distance η in the interior normal direction $-\nu(z(x))$. Note that $x \mapsto z(x)$ is of class C^1 on $V_\eta(\Gamma)$ by Proposition 7.5. Let $(\chi_{\mathrm{b}}, \chi_{\mathrm{int}})$ be a C^∞ partition of unity subordinate to the open cover $(V_\varepsilon(\Gamma), U \setminus \bar{V}_{\varepsilon/2}(\Gamma))$ of $\bar{U}$. We define

$$v_\eta(x) = u(x)\chi_{\mathrm{int}}(x) + \tilde{u} \circ s_\eta(x)\chi_{\mathrm{b}}(x), \quad x \in \mathbb{R}^d,$$

and

$$w_\eta(x) = \rho_\eta * v_\eta(x), \quad x \in \mathbb{R}^d.$$

Clearly, $w_\eta \in C^\infty(\mathbb{R}^d)$ by the regularization properties of the convolution with smooth functions, so the restriction u_η of w_η to u is in $\mathrm{BC}^\infty(\bar{U})$.

Step 2. Convergence in $L^p(U)$. Using the expansion

$$\tilde{u}(x) = u(x)\chi_{\mathrm{int}}(x) + \tilde{u}(x)\chi_{\mathrm{b}}(x), \quad x \in \mathbb{R}^d, \tag{7.58}$$

we obtain $v_\eta - \rho_\eta * \tilde{u} = \rho_\eta * [(\tilde{u} - \tilde{u} \circ s_\eta)\chi_{\mathrm{b}}]$. We have

$$\|u_\eta - u\|_{L^p(U)} \leq \|v_\eta - \tilde{u}\|_{L^p(\mathbb{R}^d)} \leq \|v_\eta - \rho_\eta * \tilde{u}\|_{L^p(\mathbb{R}^d)} + \|\rho_\eta * \tilde{u} - \tilde{u}\|_{L^p(\mathbb{R}^d)} \tag{7.59}$$

and the last term in (7.59) converges to 0. We focus on the remaining term thus. The bump function χ_{b} being supported in $V_\varepsilon(\Gamma)$, we can use the bound

$$\|v_\eta - \rho_\eta * \tilde{u}\|_{L^p(\mathbb{R}^d)} \leq \|\rho_\eta\|_{L^1(\mathbb{R}^d)}\|\tilde{u} - \tilde{u} \circ s_\eta\|_{L^p(V_\varepsilon(\Gamma))} = \|\tilde{u} - \tilde{u} \circ s_\eta\|_{L^p(V_\varepsilon(\Gamma))}.$$

But $|s_\eta(x) - x| < \eta$ in $V_\varepsilon(\Gamma)$, so

$$\|v_\eta - \rho_\eta * \tilde{u}\|_{L^p(\mathbb{R}^d)} \leq \omega_{L^p}(\tilde{u}; \eta),$$

where the L^p-modulus of continuity is introduced in (2.95). Since $\omega_{L^p}(\tilde{u}; \eta) \to 0$ when $\eta \to 0$, we obtain $u_\eta \to u$ in $L^p(U)$.

Step 3. Convergence of the gradient in $L^p(U)$. When $x \in U$, only the values of v_η in $U + B(0,\eta)$ are involved in the computation of $w_\eta(x)$, so $\tilde{u} \circ s_\eta(x) = u \circ s_\eta(x)$: the extension by 0 is not manifest here, which is fortunate since $\tilde{u}$ may have a too severe discontinuity on ∂U and not belong to $W^{1,p}(\mathbb{R}^d)$ (see next section 7.3.5 for the construction of a more suitable extension). We used the extension by 0 of u only to define v_η globally on $\mathbb{R}^d$. Let now B be a ball with $\bar{B} \subset U + B(0,\eta)$: there exists $\eta' < \eta$ such that $B \subset U + B(0,\eta')$. If $x \in B$, then $s_\eta(x) \in V := U \setminus \bar{V}_{\eta-\eta'}(\Gamma))$. By Proposition 7.2, we can approximate u by smooth functions in V. We have then

$$\nabla u \circ s_\eta(x) = Ds_\eta(x)^*(\nabla u) \circ s_\eta(x), \tag{7.60}$$

Where $Ds_\eta(x)$ is the matrix in the canonical basis of $d_x s_\eta = \mathrm{Id} - \eta d_{z(x)}\nu \circ d_x z$. In particular, $Ds_\eta(x)$ is bounded in x, so the identity (7.60) remains true in our general case $u \in W^{1,p}(U)$. The expansion of Ds_η also justifies that

$$\nabla u_\eta(x) = \nabla v_\eta(x) = \nabla[u\chi_{\mathrm{int}}](x) + u \circ s_\eta(x)\nabla\chi_{\mathrm{b}}(x) + (\nabla u) \circ s_\eta(x)\chi_{\mathrm{b}}(x) + w_\eta,$$

for $x \in U$, where $|w_\eta| = \mathcal{O}(\eta)$ in $L^p(U)$. Using (7.58), which gives the expansion

$$\nabla u(x) = \nabla[u\chi_{\mathrm{int}}](x) + u(x)\nabla\chi_{\mathrm{b}}(x) + (\nabla u)(x)\chi_{\mathrm{b}}(x)$$

for $x \in U$, we see that the estimate of the difference $|\nabla u_\eta - \nabla u|$ in $L^p(U)$ will follow from the estimates of

$$\|\tilde{u} - \tilde{u} \circ s_\eta\|_{L^p(V_\varepsilon(\Gamma))}, \quad \|\widetilde{\nabla u} - \widetilde{\nabla u} \circ s_\eta\|_{L^p(V_\varepsilon(\Gamma))}. \tag{7.61}$$

But the terms in (7.61) can be controlled, as in Step 2., by

$$\omega_{L^p}(\tilde{u};\eta) + \sum_{i=1}^{d} \omega_{L^p}(\widetilde{\partial_{x_i} u};\eta),$$

which tends to 0 when $\eta \to 0$. This concludes the proof. □

7.3.5 Extension

Theorem 7.8 (Extension of Sobolev functions)**.** *Let $p \in [1,+\infty]$. Suppose that U is of class C^2 (according to Def. 7.2). Then, there is a linear continuous operator $E_p\colon W^{1,p}(U) \to W^{1,p}(\mathbb{R}^d)$ such that $E_p(u)|_U = u$ a.e. for all $u \in W^{1,p}(U)$.*

To rephrase Theorem 7.8, there exists a constant $C_p \geq 0$, and, for each $u \in W^{1,p}(U)$, an extension $E_p(u)$ of u in $W^{1,p}(\mathbb{R}^d)$, such that

$$\|E_p(u)\|_{W^{1,p}(\mathbb{R}^d)} \leq C_p \|u\|_{W^{1,p}(U)}. \tag{7.62}$$

We will see in the proof of Theorem 7.8 that the operator E_p is the same for all p. The examination of the proof of Theorem 7.8 also shows that we can assume that $E_p(u)$ is supported in W, where W is a given open set such that $\bar{U} \subset W$. By Theorem 7.4 applied to $E_p(u)$, we obtain the following result.

Corollary 7.9 (Sobolev's injection)**.** *Let $p \in [1,+\infty)$. Suppose that U is of class C^2.*

1. *If $1 \leq p < d$, then we have a continuous injection $W^{1,p}(U) \hookrightarrow L^r(U)$ for all $r \in [p,p^*]$. The injection is compact if $r \in [p,p^*)$.*
2. *For $p = d$, we have a compact injection $W^{1,d}(U) \hookrightarrow L^r(U)$ for all $r \in [d,+\infty)$.*

3. *If $d < p$, we have (modulo the composition with $u \mapsto u^*$) a continuous injection $W^{1,d}(\mathbb{R}^d) \hookrightarrow \mathrm{BC}^{0,\nu}(\bar{U})$, for all $0 < \nu \leq \mu := 1 - d/p$. This injection is compact if $\nu < \mu$.*

We have introduced the space $\mathrm{BC}^{0,\mu}(\bar{U})$, defined as the space of restrictions to U of functions in $\mathrm{BC}^{0,\mu}(\mathbb{R}^d)$ with norm

$$\mathrm{BC}^{0,\mu}(\bar{U}) = \inf \left\{ \|v\|_{\mathrm{BC}^{0,\mu}(\mathbb{R}^d)}; v \in \mathrm{BC}^{0,\mu}(\mathbb{R}^d), v = u \text{ in } U \right\}. \tag{7.63}$$

An alternative description is the following one: let $X = \mathrm{BC}^{0,\mu}(\mathbb{R}^d)$, and let $M = \{v \in X; v = 0 \text{ on } U\}$. Then $\mathrm{BC}^{0,\mu}(\bar{U})$ is the quotient space X/M with the quotient norm (3.5). Since M is closed, $\mathrm{BC}^{0,\mu}(\bar{U})$ is a Banach space (Proposition 3.3).

Proof of Theorem 7.8. We fix an $\varepsilon > 0$ such that the parametrization (7.49) of the tubular neighbourhood $V_\varepsilon(\Gamma)$ of Γ is valid. If $x \in V_\varepsilon(\Gamma)$ and $(z(x), t(x))$ is the notation for the couple (z, t) realizing (7.49), we set

$$S(x) = \Phi(z(x), |t(x)|) = z(x) - |t(x)|\nu(z(x))$$

From (7.53), it is clear that

$$\int_{V_\varepsilon(\Gamma)} \varphi(S(x))dx \leq 2 \int_{U \cap V_\varepsilon(\Gamma)} \varphi(x)dx, \tag{7.64}$$

for any non-negative integrable function φ on $U \cap V_\varepsilon(\Gamma)$. Let $(\chi_{\mathrm{b}}, \chi_{\mathrm{int}})$ be a C^∞ partition of unity subordinate to the open cover $(V_\varepsilon(\Gamma), U \setminus \bar{V}_{\varepsilon/2}(\Gamma))$ of $\bar{U}$. If $u \in L^p(\mathbb{R}^d)$, we set

$$Eu(x) = u(x)\chi_{\mathrm{int}}(x) + u(S(x))\chi_{\mathrm{b}}(x), \quad x \in \mathbb{R}^d. \tag{7.65}$$

By Theorem 7.7, we can assume that $u \in \mathrm{BC}^\infty(\bar{U})$ to establish (7.62). For u with this regularity, we can compute directly the derivative

$$\nabla(u\chi_{\mathrm{int}}) = \chi_{\mathrm{int}} \nabla u + u \nabla \chi_{\mathrm{int}},$$

so it is clear that $\|u\chi_{\mathrm{int}}\|_{W^{1,p}(\mathbb{R}^d)} \leq C(U, p, d)\|u\|_{W^{1,p}(U)}$. Set $v = Eu - u\chi_{\mathrm{int}}$. We apply (7.64) with $\varphi(x) = |u(x)|^p$ to obtain

$$\|v\|_{L^p(\mathbb{R}^d)} \leq 2\|u\|_{L^p(U)}.$$

By the chain-rule, we also have

$$\nabla v(x) = \chi_{\mathrm{b}}(x) DS(x)^*(\nabla u)(S(x)) + u(S(x))\nabla \chi_{\mathrm{b}}(x), \quad x \in V_\varepsilon(\Gamma) \setminus \Gamma,$$

where $DS(x)$ is the matrix in the canonical basis of $d_x S$. We have $d_x S = \mathrm{Id}$ if $x \in U$ and (differentiating the expression $S(x) = 2z(x) - x$), $d_x S = 2d_x z - \mathrm{Id}$ if $x \in V_\varepsilon(\Gamma) \setminus \bar{U}$. Therefore $DS(x)$ is uniformly bounded in $V_\varepsilon(\Gamma) \setminus \Gamma$. Using (7.64), we deduce that

$$\|\nabla v\|_{L^p(\mathbb{R}^d)} \leq C(U, p, d)\|\nabla u\|_{L^p(U)}.$$

This gives the desired result. □

Remark 7.10 (Extension of Lipschitz continuous functions). See Proposition 7.18 in Section 7.4.1.

7.3.6 Trace

Theorem 7.10 (Trace of Sobolev functions)**.** *Let $p \in [1, +\infty)$ and let U be a bounded open subset of $\mathbb{R}^d$ class C^2 (according to Def. 7.2). Then, there is a linear continuous operator $\gamma_p \colon W^{1,p}(U) \to L^p(\Gamma, \sigma)$ such that $\gamma_p(u) = u|_\Gamma$ for all $u \in \mathrm{BC}^\infty(\bar{U})$.*

Proof of Theorem 7.10. Define $\gamma_p(u) = u|_\Gamma$ for $u \in \mathrm{BC}^\infty(\bar{U})$. We will show that there exists a constant $A \geq 0$ such that

$$\|\gamma_p(u)\|_{L^p(\Gamma,\sigma)} \leq A\|u\|_{W^{1,p}(U)}, \quad u \in \mathrm{BC}^\infty(\bar{U}). \tag{7.66}$$

By Theorem 7.7, the space $\mathrm{BC}^\infty(\bar{U})$ is dense in $W^{1,p}(U)$, so (7.66) will allow us to extend γ_p as a linear continuous operator $\gamma_p \colon W^{1,p}(U) \to L^p(\Gamma, \sigma)$ which satisfies the bound (7.66) for all $u \in W^{1,p}(U)$. Fix $\varepsilon > 0$ such that the parametrization of Proposition 7.5 is valid in $V_\varepsilon(\Gamma)$. For $t \in (0, \varepsilon)$ and $z \in \Gamma$, we have

$$u(z) = u(\Phi(z,t)) + \int_0^t \frac{d}{ds} u(\Phi(z,s))ds = u(\Phi(z,t)) - \int_0^t (\nabla u)(\Phi(z,s)) \cdot \nu(z)ds. \tag{7.67}$$

By the triangular inequality, we obtain

$$|u(z)| \leq |u|(\Phi(z,t)) + \int_0^t |\nabla u|(\Phi(z,s))ds, \tag{7.68}$$

and

$$|u(z)|^p \leq 2^p|u|^p(\Phi(z,t)) + 2^p t^{p-1} \int_0^t |\nabla u|^p(\Phi(z,s))ds, \tag{7.69}$$

since $(a+b)^p \leq 2^p(a^p + b^p)$ for all $a, b \geq 0$. We integrate then (7.69) with respect to $z \in \Gamma$ and $t \in (0, \varepsilon)$. Introduce a constant $C \geq 0$ (depending on $D^2\rho$) such that $|\lambda_i(z)| \leq C$ for all $z \in \Gamma$. Suppose that $\varepsilon < (2C)^{-1}$. Then $1 - \varepsilon|\lambda_i(z)| > \frac{1}{2}$ for all $z \in \Gamma$, and $\pi(z,t)$ defined by the product (7.54) satisfies

$$(1/2)^{d-1} \leq \pi(z,t) \leq (3/2)^{d-1}. \tag{7.70}$$

The formula (7.53) (and Fubini's theorem for the last term) gives

$$\frac{\varepsilon}{2^{d-1}}\|u\|^p_{L^p(\Gamma,\sigma)} \leq 2^p\|u\|^p_{L^p(V_\varepsilon(\Gamma)\cap U)} + 2^p \int_\Gamma \int_0^\varepsilon \int_s^\varepsilon t^{p-1}\pi(z,t)dt|\nabla u|^p(\Phi(z,s))dsd\sigma(z). \tag{7.71}$$

We also have

$$\pi(z,t) \leq 3^{d-1}\pi(z,s) \tag{7.72}$$

for $0 \leq s \leq t \leq \varepsilon$ by (7.70), and so

$$\frac{\varepsilon}{2^{d-1}}\|u\|^p_{L^p(\Gamma,\sigma)} \leq 2^p\|u\|^p_{L^p(V_\varepsilon(\Gamma)\cap U)} + 3^{d-1}(2\varepsilon)^p\|\nabla u\|^p_{L^p(V_\varepsilon(\Gamma)\cap U)}. \tag{7.73}$$

It follows that

$$\|u\|_{L^p(\Gamma,\sigma)} \leq \frac{C(d,p)}{\varepsilon^{1/p}}\|u\|_{W^{1,p}(U)},$$

which is (7.66). □

Proposition-Definition 7.11 (Kernel of the trace map). *Let $p \in [1, +\infty)$ and let U be a bounded open subset of $\mathbb{R}^d$ class C^2. We denote by $W_0^{1,p}(U)$ the set $\{u \in W^{1,p}(U); \gamma_p(u) = 0\}$. Then $\mathcal{D}(U)$ is dense in $W_0^{1,p}(U)$ and, given $m < p$, we have*

$$\int_U \frac{|u(x)|^p}{d(x,U)^m} dx \leq C(p,m,U) \|u\|^p_{W^{1,p}(U)}, \quad \forall u \in W_0^{1,p}(U), \tag{7.74}$$

where $C(p,m,U)$ is an absolute constant depending on p, m, U only.

Remark 7.11 (Hardy-Sobolev Inequality). The inequality (7.74) holds true when $m = p$, but is more difficult to prove, [Haj99]:

$$\int_U \frac{|u(x)|^p}{d(x,U)^p} dx \leq C(p,d,U) \|u\|^p_{W^{1,p}(U)}, \quad \forall u \in W_0^{1,p}(U), \tag{7.75}$$

where $C(p,d,U)$ is an absolute constant depending on p, d, U only. The functional inequality (7.75) is sometimes called the Hardy-Sobolev Inequality.

Proof of Proposition-Definition 7.11. Fix $\varepsilon > 0$ such that the parametrization of Proposition 7.5 is valid in $V_\varepsilon(\Gamma)$. We divide the proof into three steps.

Step 1. Estimate on a tubular neighbourhood of the boundary. As a preliminary result, we will establish the estimate

$$\int_\Gamma \int_0^\varepsilon \frac{|v(\Phi(z,t)) - v(z)|^p}{t^m} \pi(z,t) d\sigma(z) dt \leq \varepsilon^{p-m} C(p,d,m) \int_{V_\varepsilon(\Gamma)\cap U} |\nabla v(x)|^p dx, \tag{7.76}$$

for all $v \in \mathrm{BC}^\infty(\bar{U})$, where $m \in [0,p)$. For such a v, we can use (7.67), (7.72) and the Hölder inequality, to obtain

$$|v(\Phi(z,t)) - v(z)|^p \pi(z,t) \leq 3^{d-1} t^{p-1} \int_0^t |\nabla v|^p (\Phi(z,s)) \pi(z,s) ds.$$

Then (7.53) will give (7.76) by integration, with a constant

$$C(p,d,m) = \frac{3^{d-1}}{\varepsilon^{p-m}} \int_0^\varepsilon \frac{dt}{t^{m-p+1}} = \frac{3^{d-1}}{p-m}.$$

Let now $u \in W^{1,p}(U)$. There is a sequence (v_n) of functions in $\mathrm{BC}^\infty(\bar{U})$ which converges to u in $W^{1,p}(U)$. By (7.66), $\gamma_p v_n \to \gamma u$ in $L^p(\Gamma, \sigma)$. We also have convergence of (v_n) to u in $L^p(U)$, so $(v_n \circ \Phi)$ is converging to $u \circ \Phi$ in $L^p(\Gamma \times (0,\varepsilon), \mu)$, where μ is the measure with density $\pi(z,t)$ with respect to $d\sigma(z)dt$. Up to subsequences, we can assume $\gamma_p v_n \to \gamma u$ σ-a.e. on Γ and $v_n \circ \Phi \to u \circ \Phi$ μ-a.e. on $\Gamma \times (0,\varepsilon)$. By Fatou's lemma in (7.76) written for v_n, we obtain

$$\int_\Gamma \int_0^\varepsilon \frac{|u(\Phi(z,t)) - \gamma_p u(z)|^p}{t^m} \pi(z,t) d\sigma(z) dt \leq \varepsilon^{p-m} C(p,d,m) \liminf_{n\to+\infty} \int_{V_\varepsilon(\Gamma)\cap U} |\nabla v_n(x)|^p dx. \tag{7.77}$$

As $\|\nabla v_n\|_{L^p(V_\varepsilon(\Gamma)\cap U)} \to \|\nabla u\|_{L^p(V_\varepsilon(\Gamma)\cap U)}$, we deduce from (7.77) the estimate

$$\int_\Gamma \int_0^\varepsilon \frac{|u(\Phi(z,t)) - \gamma_p u(z)|^p}{t^m} \pi(z,t) d\sigma(z) dt \leq \varepsilon^{p-m} C(p,d,m) \int_{V_\varepsilon(\Gamma)\cap U} |\nabla u(x)|^p dx, \tag{7.78}$$

for all $u \in W^{1,p}(U)$.

Step 2. Use of a cut-off function on a tubular neighbourhood of the boundary. Let $u \in W_0^{1,p}(U)$. Consider the cut-off function

$$\chi_r(x) = \theta(r^{-1}d(x,U)), \quad 0 < r < 1,$$

where $\theta \in C_c^\infty(0,+\infty)$ is non-decreasing, such that $0 \leq \theta \leq 1$ with $\theta \equiv 1$ on $[1,+\infty)$. For $r < \varepsilon$, we have $\chi_r(x) = \theta(t)$ with $x = \Phi(z,t)$ (consequence of Proposition 7.5). Since Φ, is of class C^1, $\chi_r \in C_c^1(U)$. Let $v_r = u\chi_r$. Then $v_r \in W^{1,p}(U)$ (*cf.* (7.5)) and v is compactly supported in U. Suppose that we manage to prove that $v_r \to u$ in $W^{1,p}(U)$ when $r \to 0$. Then an additional step of regularization by convolution (which we will not write in details, it is similar to the proof of Proposition 7.2, on the local approximation by smooth functions) ensures that $\mathcal{D}(U)$ is dense in $W^{1,p}(U)$. Let us establish the convergence $v_r \to u$ in $W^{1,p}(U)$ thus. We have, by (7.5),

$$\nabla v_r = \chi_r \nabla u + u\nabla\chi_r.$$

Let $\eta > 0$. By dominated convergence, there exists $r > 0$ such that

$$\left[\int_U |1-\chi_r|^p|u|^p dx + \sum_{i=1}^d \int_U |1-\chi_r|^p|\partial_{x_i}u|^p dx\right]^{1/p} < \eta.$$

This gives

$$\|v_r - u\|_{W^{1,p}(U)} < \eta + \|u\nabla\chi_r\|_{L^p(U)}. \tag{7.79}$$

We will use (7.78) to show that the remaining term in (7.79) can be made arbitrarily small. Since $x \mapsto d(x,U)$ is 1-Lipschitz continuous, we have $|\nabla_x d(x,U)| \leq 1$ (actually, $|\nabla_x d(x,U)| = 1$) for $x \in V_\varepsilon(\Gamma) \cap U$ so

$$|\nabla\chi_r(x)| \leq r^{-1}, \quad \operatorname{supp}(\nabla\chi_r) \subset V_r(\Gamma) \cap U. \tag{7.80}$$

Consequently $\|u\nabla\chi_r\|_{L^p(U)} \leq r^{-1}\|u\|_{L^p(V_r(\Gamma)\cap U)}$. We use (7.53) and (7.78) (where $\gamma_p u \equiv 0$) with $\varepsilon = r$ and $m = 0$ to obtain

$$\|u\nabla\chi_r\|_{L^p(U)} \leq C(p,d,m)^{1/p}\|\nabla u\|_{L^p(V_r(\Gamma)\cap U)}.$$

By dominated convergence, we have $C(p,d,m)^{1/p}\|\nabla u\|_{L^p(V_r(\Gamma)\cap U)} < \eta$ for r small enough, and thus $\|v_r - u\|_{W^{1,p}(U)} < 2\eta$. This is the desired result.

Step 3. Sub-optimal Hardy's inequality. Since $t = d(\Phi(z,t),U)$ in (7.78) (Proposition 7.5), we have

$$\int_{V_\varepsilon(\Gamma)\cap U} \frac{|u(x)|^p}{d(x,U)^m}dx \leq \varepsilon^{p-m}C(p,d,m)\int_U |\nabla u(x)|^p dx, \tag{7.81}$$

if $u \in W_0^{1,p}(U)$. We complete (7.81) with the estimate

$$\int_{U\setminus\bar{V}_\varepsilon(\Gamma)} \frac{|u(x)|^p}{d(x,U)^m}dx \leq \frac{1}{\varepsilon^m}\int_U |u(x)|^p dx,$$

to obtain (7.74). □

We have seen that the Green formula (7.52) is equivalent to (7.55). We have the following generalizations of both formulas for functions in Sobolev spaces.

Proposition 7.12 (Extended Green Formula)**.** *Let U be a bounded open subset of $\mathbb{R}^d$ class C^2. Let $\Psi\colon U \to \mathbb{R}^d$ such that $\Psi_i \in W^{1,1}(U)$ for all $i \in \{1, \dots, d\}$. Then*

$$\int_U \operatorname{div}(\Psi(x))dx = \int_{\partial U} \gamma_1 \Psi(x) \cdot \nu(x) d\sigma(x). \tag{7.82}$$

If $\Psi_i \in W^{1,p}(U)$ for all $i \in \{1, \dots, d\}$, where $1 \leq p < +\infty$, then

$$\int_U \operatorname{div}(\Psi(x))u(x)dx = -\int_U \Psi(x) \cdot \nabla u(x)dx + \int_{\partial U} \gamma_p \Psi(x) \cdot \nu(x)\gamma_{p'} u(x) d\sigma(x), \tag{7.83}$$

for all $u \in W^{1,p'}(U)$ (p' conjugate exponent to p).

Proof of Proposition 7.12. The second formula (7.83) follows from (7.82) applied to the product $u\Psi$, which is in $W^{1,1}(U)$ by (7.5). Formula (7.82) is deduced from (7.55) and approximation by functions in $\mathrm{BC}^\infty(\bar{U})$. □

7.3.7 Kernel of the trace operator

Let U be an open bounded set of $\mathbb{R}^d$. For $p \in [1, +\infty)$, let $H_0^{1,p}(U)$ denote the closure of $\mathcal{D}(U)$ in $W^{1,p}(U)$. We have seen that $H_0^{1,p}(U) = W_0^{1,p}(U)$ if U is sufficiently regular (Proposition-Definition 7.11). A fundamental property of functions in $H_0^{1,p}(U)$ is that they their extension by 0 provides function in $W^{1,p}(\mathbb{R}^d)$.

Proposition 7.13 (Extension by 0)**.** *The extension operator $E_0\colon \mathcal{D}(U) \to \mathcal{D}(\mathbb{R}^d)$ defined by*

$$E_0 u(x) = \begin{cases} u(x) & \text{if } x \in U, \\ 0 & \text{if } x \in \mathbb{R}^d \setminus U, \end{cases}$$

can be extended as linear continuous operator $H_0^{1,p}(U) \to W^{1,p}(\mathbb{R}^d)$ with norm $\|E_0\| \leq 1$. We also have

$$\|\nabla(E_0 u)\|_{W^{1,p}(U)} \leq \|\nabla u\|_{L^p(U)}, \tag{7.84}$$

for all $u \in H_0^{1,p}(U)$.

Proof of Proposition 7.13. This is an immediate consequence of the fact that 1) $\mathcal{D}(U)$ is dense in $H_0^{1,p}(U)$ for the Sobolev norm by definition, 2) $\|E_0 u\|_{W^{1,p}(\mathbb{R}^d)} \leq \|u\|_{W^{1,p}(U)}$ if $u \in \mathcal{D}(U)$. □

Theorem 7.14 (Poincaré's inequality)**.** *Let $p \in [1, +\infty)$. Let U be an open bounded set of $\mathbb{R}^d$. Then there is a constant $C = C(p, U) \geq 0$ such that*

$$\|u\|_{L^p(U)} \leq C \|\nabla u\|_{L^p(U)} \tag{7.85}$$

for all $u \in H_0^{1,p}(U)$. In particular, the norm $\|\cdot\|_{W^{1,p}(U)}$ is equivalent to the norm $u \mapsto \|\nabla u\|_{L^p(U)}$ on $H_0^{1,p}(U)$.

Proof of Theorem 7.14. Let E_0 be the operator "extension by 0". Assume first $p < d$ and let $p^* = \frac{pd}{d-p}$ as in (7.13). We have $\|u\|_{L^p(U)} \leq |U|^{1/}q\|u\|_{L^{p^*}(U)}$ by the Hölder inequality (q is the conjugate exponent to p^*/p), and thus, using the Gagliardo-Nirenberg-Sobolev inequality (7.13) and (7.84),

$$\|u\|_{L^p(U)} \leq |U|^{1/q} \|E_0 u\|_{L^{p^*}(\mathbb{R}^d)} \leq |U|^{1/q} \frac{p(d-1)}{d-p} \|\nabla(E_0 u)\|_{W^{1,p}(U)} \leq |U|^{1/q} \frac{p(d-1)}{d-p} \|\nabla u\|_{L^p(U)}.$$

If $d \leq p$, we fix $q < d$ such that $p \leq q^* = \frac{qd}{d-q}$. and use the following estimate $(1/m = 1 - p/q^*)$

$$\|u\|_{L^p(U)} \leq |U|^{1/m} \|u\|_{L^{q^*}(U)} \leq |U|^{1/m} \|E_0 u\|_{L^{q^*}(\mathbb{R}^d)}.$$

Since

$$\|E_0 u\|_{L^{q^*}(\mathbb{R}^d)} \leq \frac{q(d-1)}{d-q} \|\nabla(E_0 u)\|_{L^q(\mathbb{R}^d)}$$

and $\|\nabla(E_0 u)\|_{L^q(\mathbb{R}^d)} = \|\nabla u\|_{L^q(U)} \leq |U|^{1/l} \|\nabla u\|_{L^p(U)}$ $(1/l = 1 - q/p)$, we obtain the desired result. ☐

7.4 Operation of Lipschitz maps on Sobolev spaces

7.4.1 Lipschitz functions

Let U be an open set in $\mathbb{R}^d$. We define the space

$$\mathrm{Lip}(U) = \{u \in C(U); \mathrm{Lip}(u;U) < +\infty\}, \quad \mathrm{Lip}(u;U) = \sup_{x \neq y \in U} \frac{|u(x) - u(y)|}{|x-y|},$$

with norm $\|u\|_{\mathrm{Lip}(U)} = \sup_{x \in U} |u(x)| + \mathrm{Lip}(u;U)$. In this section, we prove the following two results.

Theorem 7.15 (Rademacher)**.** *Let U be an open set of $\mathbb{R}^d$. A Lipschitz function on U is differentiable a.e. on U.*

Theorem 7.16 ($W^{1,\infty}(U) = \mathrm{Lip}(U)$)**.** *Let U be an open set of $\mathbb{R}^d$. We have then*

1. *if $u \in \mathrm{Lip}(U)$, then $u \in W^{1,\infty}(U)$, with $\|\nabla u\|_{L^\infty(U)} \leq \mathrm{Lip}(u;U)$,*

2. *if C is a bounded convex set with $\bar{C} \subset U$ and $u \in W^{1,\infty}(U)$, then $u^* \in \mathrm{Lip}(C)$ and $\mathrm{Lip}(u^*;C) \leq \|\nabla u\|_{L^\infty(U)}$.*

To establish Rademacher's theorem 7.15, we will need this following version of the fundamental theorem of calculus.

Proposition 7.17 (Fundamental theorem of calculus for Lipschitz functions)**.** *Let $u \colon \mathbb{R} \to \mathbb{R}$ be a bounded Lipschitz continuous function. Then*

1. *$u \in W^{1,\infty}(\mathbb{R})$, with $\|u'\|_{L^\infty(\mathbb{R})} \leq \mathrm{Lip}(u)$,*

2. *u is differentiable a.e. and*

$$u(x) - u(0) = \int_0^x u'(y)dy, \tag{7.86}$$

for all $x \in \mathbb{R}$.

Remark 7.12. The fundamental theorem of calculus asserts that *2.* is satisfied in the more general case of an absolutely continuous functions, [Rud87, Theorem 7.18]. Note that item *1.* of Proposition 7.17 is *1.* of Theorem 7.16, when $d = 1$.

Proof of Proposition 7.17. To prove *2.*, we apply Proposition 7.1 with $p = +\infty$. It is sufficient to establish the bound

$$\left| \int_{\mathbb{R}} u(y) v'(y) dy \right| \leq \mathrm{Lip}(u) \|v\|_{L^1(\mathbb{R})}, \tag{7.87}$$

for all $v \in \mathcal{D}(\mathbb{R})$. To obtain (7.87), we use the fact that

$$u' = \lim_{\varepsilon\to 0} D_\varepsilon u \text{ in } \mathcal{D}'(\mathbb{R}), \quad D_\varepsilon u(x) := \frac{1}{\varepsilon}(u(x+\varepsilon) - u(x)),$$

see the proof of Proposition 6.6. Using the bound $|D_\varepsilon u(x)| \leq \mathrm{Lip}(u)$, (7.87) follows. The identity (7.86) is satisfied by $u * \rho_\varepsilon$, where (ρ_ε) is an approximation of the unit:

$$u * \rho_\varepsilon(x) - u * \rho_\varepsilon(0) = \int_0^x (u * \rho_\varepsilon)'(y)dy. \tag{7.88}$$

We have $u * \rho_\varepsilon(x) \to u(x)$ for all $x \in \mathbb{R}$ when $\varepsilon \to 0$, since u is continuous. We also have $(u*\rho_\varepsilon)' = u' * \rho_\varepsilon$ (apply (6.28)) and since $u' \in L^\infty(\mathbb{R})$, $(u*\rho_\varepsilon)' \to u'$ in $L^1_{\mathrm{loc}}(\mathbb{R})$. Therefore passing to the limit $\varepsilon \to 0$ in (7.88) gives (7.86). Once (7.86) is established, Lebesgue's differentiation theorem asserts that, for almost every $x \in \mathbb{R}$, the limit

$$\lim_{t\to 0^+} \frac{u(x+t) - u(x-t)}{2t} = \lim_{t\to 0^+} \frac{1}{2t}\int_{x-t}^{x+t} u'(y)dy \tag{7.89}$$

exists and is equal to $u'(x)$. A simple modification of the argument that we admit shows that, for almost every $x \in \mathbb{R}$, the limits

$$\lim_{t\to 0^+} \frac{u(x+t) - u(x)}{t} \quad \text{and} \quad \lim_{t\to 0^+} \frac{u(x) - u(x-t)}{t} \tag{7.90}$$

exist and are both equal to $u'(x)$, the function u being therefore differentiable a.e. □

Before we give the proof of Rademacher's theorem, let us also give some details on the extension of Lipschitz continuous functions.

Proposition 7.18 (Extension of Lipschitz continuous functions)**.** *Let U be an open bounded set of $\mathbb{R}^d$ (no regularity is assumed on ∂U) and let $u \in \mathrm{Lip}(U)$. Let*

$$Eu(x) = \inf_{z\in U} \{u(z) + \mathrm{Lip}(u;U)|x-z|\}.$$

Let $k > \sup_{x\in U}|u(x)|$ and define the operator (truncation at level k) $T_k(s) = \max(-k, \min(k,s))$. The function $T_k(Eu)$ is bounded and Lipschitz continuous on $\mathbb{R}^d$, $T_k(Eu) \in \mathrm{Lip}(\mathbb{R}^d)$, with $\mathrm{Lip}(T_k(Eu);\mathbb{R}^d) \leq \mathrm{Lip}(u;U)$ and $T_k(Eu)(x) = u(x)$ if $x \in U$.

A corollary of Proposition (7.18) is that the space $\mathrm{BC}^{0,1}(\bar{U})$ (defined in (7.63)) coincides (as normed vector space) with the space $\mathrm{Lip}(U)$.

Proof of Proposition 7.18. If $x \in U$, then $u(x) \leq u(z) + \mathrm{Lip}(u;U)|x-z|$ for all $z \in U$, so $Eu(x) = u(x)$. Let $x, y \in \mathbb{R}^d$. For all $z \in U$, we have

$$Eu(x) \leq u(z) + \mathrm{Lip}(u;U)|x-z| \leq u(z) + \mathrm{Lip}(u;U)|y-z| + \mathrm{Lip}(u;U)|x-y|.$$

Taking the inf over $z \in U$ gives $Eu(x) \leq Eu(y) + \mathrm{Lip}(u;U)|x-y|$, so Eu is $\mathrm{Lip}(u;U)$-globally Lipschitz continuous. We have $T_k(Eu(x)) = u(x)$ if $x \in U$ since k has been chosen greater than $\sup_{x\in U}|u(x)|$. To conclude there remains to check that $\mathrm{Lip}(T_k(F);\mathbb{R}^d) \leq \mathrm{Lip}(F;\mathbb{R}^d)$ for any Lipschitz function F on $\mathbb{R}^d$. This is left as an (easy) exercise. □

Proof of Theorem 7.15. The proof is not difficult, it uses two essential facts: 1) Proposition 7.17, 2) the fact that existence of directional derivatives, together with some consistency property, implies differentiability (an instance of this principle is the fact that existence of continuous partial derivatives implies differentiability, but this is not what we will use here). Note also that it is sufficient to consider the case $U = \mathbb{R}^d$, otherwise we can first apply the extension method described in Remark 7.10. So let u be a Lipschitz continuous function on $\mathbb{R}^d$. We begin with the existence of directional derivatives. Let $(e_i)_{1,n}$ be the canonical basis of $\mathbb{R}^d$ and let $x \in \mathbb{R}^d$. The function $\phi_x \colon t \mapsto u(x + te_1)$ is Lipschitz continuous. By Proposition 7.17, $t \mapsto \phi_x(t)$ is differentiable for a.e. $t \in \mathbb{R}$. Denote by E_1 the set

$$E_1 = \left\{ x \in \mathbb{R}^d ; \lim_{t \to 0} \frac{u(x + te_1) - u(x)}{t} \text{ does not exists} \right\}.$$

By continuity of u, we have

$$E_1^c = \left\{ x \in \mathbb{R}^d ; \lim_{t \to 0, t \in \mathbb{Q}} \frac{u(x + te_1) - u(x)}{t} \text{ exists} \right\}.$$

Using a characterization by a Cauchy condition of the convergence, we can then prove that E_c^1 is a Borel set, as well as E_1. By Fubini's theorem, we have,

$$|E_1| = \int_{\mathbb{R}^{d-1}} |E_1(y)| dy, \quad E_1(y) := E_1 \cap (y + \langle e_1 \rangle), \tag{7.91}$$

where $|E_1|$ is the d-dimensional Lebesgue measure of E_1 and $|E_1(y)|$ is the 1-dimensional Lebesgue measure of the section $E_1(y)$. More precisely, using the parametrization $s \mapsto y + se_1$ of $y + \langle e_1 \rangle$, we have

$$\begin{aligned} |E_1(y)| &= |\{ s \in \mathbb{R}; t \mapsto u(y + se_1 + te_1) \text{ is not differentiable at } t = 0 \}| \\ &= |\{ s \in \mathbb{R}; \phi_y \text{ is not differentiable at } s \}|, \end{aligned}$$

and thus $|E_1(y)| = 0$. This being true for all y, (7.91) shows that $|E_1| = 0$. Since e_1 has no privileged role here, we have $|E_a| = 0$ for all $a \in S(0,1)$, where

$$E_a = \left\{ x \in \mathbb{R}^d ; \lim_{t \to 0^+} \frac{u(x + ta) - u(x)}{t} \text{ does not exists} \right\}.$$

Define

$$D_a u(x) = \begin{cases} \lim_{t \to 0^+} \dfrac{u(x + ta) - u(x)}{t} & \text{if } x \in \mathbb{R}^d \setminus E_a, \\ 0 & \text{if } x \in E_a. \end{cases} \tag{7.92}$$

Then $D_a u$ is Borel-measurable, and thus $D_a u \in L^\infty(\mathbb{R}^d)$, with $\|D_a u\|_{L^\infty(\mathbb{R}^d)} \leq \text{Lip}(u)$. It is also clear that $D_a u$ is the directional derivatives of u in the sense of distributions:

$$\langle D_a u, v \rangle = -\langle u, D_a v \rangle, \quad v \in \mathcal{D}(\mathbb{R}^d), \quad D_a v(x) := \nabla v(x) \cdot a. \tag{7.93}$$

Define the gradient $\nabla u = (D_{e_i} u)_{1,n}$. Let $a \in S(0,1)$. We will show that

$$D_a u(x) = \nabla u(x) \cdot a, \tag{7.94}$$

for a.e. $x \in \mathbb{R}^d$. It is sufficient to prove (7.94) in the sense of distributions (both members are functions in $L^\infty(\mathbb{R}^d)$). Taking (7.93) into account, this amounts to establish that

$$\langle u, D_a v \rangle = \sum_{i=1}^d a_i \langle u, D_{e_i} v \rangle = \langle u, \nabla v \cdot a \rangle, \quad v \in \mathcal{D}(\mathbb{R}^d). \tag{7.95}$$

But (7.95) is clear, so (7.94) is true for a.e. $x \in \mathbb{R}^d$. Eventually, let us consider a countable set A of vectors $a \in S(0,1)$ such that A is dense in $S(0,1)$. We know that there is a negligible set E such that, if $x \in \mathbb{R}^d \setminus E$, then, for all $a \in A$, for all $i \in \{1, \dots, d\}$,

$$D_a u(x) = \lim_{t \to 0^+} \frac{u(x+ta) - u(x)}{t}, \quad D_{e_i} u(x) = \lim_{t \to 0^+} \frac{u(x+te_i) - u(x)}{t}, \quad D_a u(x) = \nabla u(x) \cdot a. \tag{7.96}$$

We will show that, if $x \in \mathbb{R}^d \setminus E$, then u is differentiable at x, *i.e.* for all $y \in \mathbb{R}^d$,

$$\lim_{t \to 0^+} \eta(x,y,t) = 0, \quad \eta(x,y,t) := \frac{u(x+ty) - u(x)}{t} - \nabla u(x) \cdot y. \tag{7.97}$$

This is trivial if $y = 0$ and, if $y \neq 0$, then

$$\eta(x,y,t) = |y|\eta(x,z,t|y|), \quad z = \frac{y}{|y|},$$

so it is sufficient to consider the case $y \in S(0,1)$: we assume that $|y| = 1$. If $z \in S(0,1)$, then

$$|\eta(x,y,t) - \eta(x,z,t)| \leq \left(\mathrm{Lip}(u) + \|\nabla u\|_{L^\infty(\mathbb{R}^d)}\right) |y - z|, \tag{7.98}$$

so $y \mapsto \eta(x,y,t)$ is Lipschitz continuous with respect to $y \in S(0,1)$, uniformly in t. The convergence (7.97) being satisfied for all y in the dense set A, it remains true when y is any element of the unit sphere. This concludes the proof. □

Proof of Theorem 7.16. If $u \in \mathrm{Lip}(U)$, then $u \in W^{1,\infty}(U)$ with $\|\nabla u\|_{L^\infty(U)} \leq \mathrm{Lip}(u;U)$. The proof is exactly the same as in dimension $d = 1$, and we refer to the proof of Proposition 7.17 for the details. Assume now that C is a bounded convex set with $\bar{C} \subset U$ and let $u \in W^{1,\infty}(U)$. If B is a ball with $B \subset U$, then $u \in W^{1,\infty}(B) \subset W^{1,p}(U)$ where p is any exponent strictly greater then d, so u^* is continuous on B by Corollary 7.9. Without loss of generality, we assume $u = u^*$. The set $\bar{C}$ is compact, included in U, so there exists $\varepsilon_0 > 0$ such that the ε_0-neighbourhood $V_{\varepsilon_0}(C)$ of C is included in U. Let (ρ_ε) be an approximation of the unit as in (2.139). Then, for $\varepsilon < \varepsilon_0$, and for $x, y \in C$,

$$\begin{aligned} |u * \rho_\varepsilon(x) - u * \rho_\varepsilon(y)| &= \left| \int_0^1 \nabla(u * \rho_\varepsilon)((1-t)x + ty) \cdot (y - x) dt \right| \\ &\leq \int_0^1 |\nabla(u * \rho_\varepsilon)((1-t)x + ty)| dt |x - y|. \end{aligned}$$

We have $\nabla(u * \rho_\varepsilon) = (\nabla u) * \rho_\varepsilon$ by (6.28) so

$$\|\nabla(u * \rho_\varepsilon)\|_{L^\infty(V_\varepsilon(C))} \leq \|\nabla u\|_{L^\infty(U)}.$$

It follows that

$$|u * \rho_\varepsilon(x) - u * \rho_\varepsilon(y)| \leq \|\nabla u\|_{L^\infty(U)} |x - y|.$$

We have also $u * \rho_\varepsilon(x) \to u(x)$ when $\varepsilon \to 0$ since u is continuous, so

$$|u(x) - u(y)| \leq \|\nabla u\|_{L^\infty(U)} |x - y|,$$

which is the desired result. □

7.4.2 Operation of Lipschitz functions on Sobolev spaces

Theorem 7.19 (Composition on the left with Lipschitz functions)**.** *Let U be an open subset of $\mathbb{R}^d$. Let $u \in W^{1,p}(U)$, $p \in [1, +\infty]$.*

1. *if A is a Borel subset of $\mathbb{R}$ of measure 0 then $\nabla u = 0$ a.e. on the set $u^{-1}(A) = \{x \in U; u(x) \in A\}$,*
2. *if $F\colon \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous, then $F(u) \in W^{1,p}(U)$ and we have the chain rule*

$$\nabla F(u) = F'(u)\nabla u \tag{7.99}$$

a.e in U.

One may wonder what is the exact meaning on $F'(u)\nabla u$ in (7.99). Indeed $F'(s)$ is only defined a.e., say for $s \in \mathbb{R} \setminus A$, where A is a Borel negligible set, by Rademacher's theorem. If $u(x) \in A$, then we cannot give a meaning to $F'(u(x))$. However, the first point *1.* ensures that $\nabla u = 0$ on the set $\{u \in A\}$. Consequently, we can use the convention $F'(u)\nabla u = 0$ on $u^{-1}(A)$. We will see then that (7.99) is satisfied for good. Theorem 7.19 can be applied with $F(u) = u^+$ in particular. We obtain $u^+ \in W^{1,p}(U)$, with

$$\nabla u^+ = \mathbf{1}_{u \geq 0}\nabla u = \mathbf{1}_{u>0}\nabla u.$$

The theorem is not true for Sobolev functions $u \in W^{k,p}(U)$ with $k \geq 2$. This can be expected from the computation

$$\Delta F(u) = \operatorname{div}(F'(u)\nabla u) = F''(u)|\nabla u|^2 + F'(u)\Delta u$$

One may restrict to functions in $W^{1,kp}(U) \cap W^{k,p}(U)$ for instance, [Bou91, Hof13].

Proof of Theorem 7.19. We first want to prove that

$$|A| = 0 \Rightarrow \mu(A) = 0, \quad \mu(A) := \nu(u^{-1}(A)), \quad \nu(B) := \int_{U \cap B} |\nabla u|^p dx. \tag{7.100}$$

In (7.100), μ is a bounded Borel measure, so μ is inner regular:

$$\mu(A) = \sup\{\mu(K); K \subset A\}.$$

Therefore it is sufficient to prove (7.100) for K compact. If K is compact with $|K| = 0$, we set

$$\chi_\varepsilon(x) = (1 - \varepsilon^{-1}d(K, x))^+, \quad G_\varepsilon(x) = \int_{-\infty}^{\varepsilon} \chi_\varepsilon(y)dy.$$

Note that χ_ε is continuous, bounded, with $\operatorname{supp}(\chi_\varepsilon) \subset V_\varepsilon(K)$. We have $G_\varepsilon \in C^1_b(\mathbb{R})$, hence $G_\varepsilon(u) \in W^{1,p}(U)$ and $\nabla G_\varepsilon(u) = G'_\varepsilon(u)\nabla u = \chi_\varepsilon(u)\nabla u$, as seen in exercises class. Let $v \in \mathcal{D}(U)$. For $1 \leq i \leq d$, we can write

$$\left|\int_U \chi_\varepsilon(u)\partial_{x_i} u v dx\right| = \left|\int_U G_\varepsilon(u)\partial_{x_i} v dx\right| \leq C(v)|U|\|G_\varepsilon\|_{L^\infty(\mathbb{R})}. \tag{7.101}$$

When we pass to the limit $\varepsilon \to 0$ in (7.101), the left-hand side converges by dominated convergence and

$$\|G_\varepsilon\|_{L^\infty(\mathbb{R})} \leq \|\chi_\varepsilon\|_{L^1(\mathbb{R})} \to |K| = 0,$$

so

$$\int_U \mathbf{1}_K(u)\partial_{x_i} uv dx = 0, \tag{7.102}$$

and, (7.102) being satisfied for all $v \in \mathcal{D}(U)$, $\mathbf{1}_K(u)\nabla u = 0$ a.e. in U. Let us establish (7.99) now. Let (ρ_ε) be an approximation of the unit as in (2.139), when $d = 1$. Let $F_\varepsilon = F * \rho_\varepsilon$. By the chain-rule with functions in $C_b^1(\mathbb{R})$, we have

$$\int_U F_\varepsilon(u)\partial_{x_i} v dx = \int_U F'_\varepsilon(u)\partial_{x_i} uv dx, \quad v \in \mathcal{D}(U), \quad 1 \le i \le d. \tag{7.103}$$

We have $F_\varepsilon(s) \to F(s)$ for all $s \in \mathbb{R}$, so the left-hand side of (7.103) converges to the integral of $F(u)\partial_{x_i} v$ by dominated convergence. Let A denote the complementary in $\mathbb{R}$ of the set of Lebesgue points of F'. If $s \notin A$, then $F'_\varepsilon(s) = (F' * \rho_\varepsilon)(s) \to F'(s)$. By dominated convergence, we obtain

$$\int_U F'_\varepsilon(u)\partial_{x_i} uv dx = \int_{U\setminus A} F'_\varepsilon(u)\partial_{x_i} uv dx \to \int_{U\setminus A} F'(u)\partial_{x_i} uv dx,$$

and thus

$$\int_U F(u)\partial_{x_i} v dx = \int_U F'(u)\partial_{x_i} uv dx,$$

with the convention $F'(u)\partial_{x_i} u = 0$ if $u \in A$. □

A Surface measure

In this section, we give the proof of Proposition 7.6. We will first need the following result.

Lemma A.1 (Convergence of distributions of finite order). *Let (α_n) be a sequence in $\mathcal{D}'(U)$ converging to a distribution $\alpha \in \mathcal{D}'(U)$. Assume that all distributions α_n and α are uniformly of finite order k in the sense that, for each compact $L \subset U$, there exists $A \ge 0$ such that*

$$|\langle \alpha, u\rangle| \le A p_{L,k}(u), \quad |\langle \alpha_n, u\rangle| \le A p_{L,k}(u), \tag{A.1}$$

for all $u \in C_L^\infty(U)$. Then α_n and α can be extended as linear functional on $C_c^k(U)$, and we have

$$\langle \alpha_n, u\rangle \to \langle \alpha, u\rangle \tag{A.2}$$

when $n \to +\infty$, for all $u \in C_c^k(U)$.

Remark A.1 (Uniform condition (A.1)). The convergence $\alpha_n \to \alpha$ in $\mathcal{D}'(\mathbb{R}^d)$ is not sufficient to ensure the uniform condition (A.1). Consider for instance an approximation of the unit (ρ_n), a sequence (ε_n) of real numbers converging to 0 not too rapidly ($\varepsilon_n^{-1} = o(n)$ will do) and $\alpha_n = \varepsilon_n(\rho_n)'$. Then (A.1) will not be satisfied for $k = 0$ (but will be for $k = 1$). Nevertheless, it is clear that assuming simply

$$|\langle \alpha_n, u\rangle| \le A p_{L,k}(u), \tag{A.3}$$

for all $u \in C_L^\infty(U)$, will give $|\langle \alpha, u\rangle| \le A p_{L,k}(u)$, and thus (A.1) for all $u \in C_L^\infty(U)$.

Proof of Lemma A.1. We use the density of $\mathcal{D}(U)$ in $C_c^k(U)$, simply obtained by convolution with a compactly supported approximation of the unit (ρ_m) (as in (6.34) for instance). For each $u \in C_c^k(U)$, we have $\rho_m * u \in \mathcal{D}(U)$ for m large enough (such that $\mathrm{supp}(u) + \bar{B}(0, 1/m) \subset U$) and $p_{K,k}(\rho_m * u - u) \to 0$ for each semi-norm $p_{K,k}$, K compact subset of U. For such a fixed K, there is a compact $L \supset K$ such that $\mathrm{supp}(\rho_m * u) \subset L$ for all m large enough. Since

$p_{L,k}(\rho_m * u - \rho_{m'} * u) \to 0$ when $m, m' \to +\infty$, (A.1) shows that the sequence $(\langle \alpha, \rho_m * u \rangle)$ is Cauchy. We denote by $\langle \tilde{\alpha}, u \rangle$ its limit. It is easy to check that we define indeed a linear functional $\tilde{\alpha}$ on $C_c^k(U)$. Since $|\langle \tilde{\alpha}, u \rangle| \le A p_{L,k}(u)$, this functional is continuous. We proceed similarly with α_n. We have then

$$|\langle \tilde{\alpha}, u \rangle - \langle \tilde{\alpha}_n, u \rangle| \le |\langle \tilde{\alpha}, \rho_m * u - u \rangle| + |\langle \tilde{\alpha}_n, \rho_m * u - u \rangle| + |\langle \alpha, \rho_m * u \rangle - \langle \alpha_n, \rho_m * u \rangle|,$$

which we can bound by

$$2A p_{L,k}(\rho_m * u - u) + |\langle \alpha, \rho_m * u \rangle - \langle \alpha_n, \rho_m * u \rangle|.$$

Given $\varepsilon > 0$, we choose first a fixed m such that $2A p_{L,k}(\rho_m * u - u) < \varepsilon$, and then N such that $|\langle \alpha, \rho_m * u \rangle - \langle \alpha_n, \rho_m * u \rangle| < \varepsilon$ for all $n \ge N$ to obtain

$$|\langle \tilde{\alpha}, u \rangle - \langle \tilde{\alpha}_n, u \rangle| < 2\varepsilon,$$

for all $n \ge N$. This gives the convergence (A.2). □

Proof of Proposition 7.6. There are several steps in the proof of this result.

Step 1: Existence of σ. We will not try to establish the intrinsic character of the measure σ (but σ is indeed the normalized volume form of the orientable manifold Γ). Instead, we first consider the distribution

$$\alpha = -\operatorname{div}(\mathbf{1}_U \nabla \rho) + \mathbf{1}_U \Delta \rho. \tag{A.4}$$

Note that, in virtue of the second formula in (6.47), α is formally given by

$$\alpha = -\nabla \mathbf{1}_U \cdot \nabla \rho. \tag{A.5}$$

We will show that α is a measure supported in Γ and then set $\sigma = \frac{1}{|\nabla \rho|}\alpha$. Then (A.5) gives the formal expression

$$\sigma = -\nabla \mathbf{1}_U \cdot \nu \tag{A.6}$$

for σ. Note first that α is well defined as a distribution since $\mathbf{1}_U \nabla \rho$ and $\mathbf{1}_U \Delta \rho$ are distributions, by injection of $L^1_{\text{loc}}(\mathbb{R}^d)$ in $\mathcal{D}'(\mathbb{R}^d)$. Let us prove that the distribution α is supported in Γ. Let $u \in \mathcal{D}(\mathbb{R}^d \setminus \Gamma)$, let $K = \operatorname{supp}(u) \cap \bar{U}$. Since $d(\operatorname{supp}(u), \Gamma) > 0$, we have $K \subset U$. Let χ be a bump function such that $K \prec \chi \prec U$. We have $\chi u = u$ and $\nabla(\chi u) = \chi \nabla(\chi u)$, so

$$\langle \alpha, u \rangle = \int_U \nabla \rho \nabla(\chi u) dx + \int_U \chi u \Delta \rho dx = \int_{\mathbb{R}^d} \nabla \rho \nabla(\chi u) dx + \int_{\mathbb{R}^d} \chi u \Delta \rho dx,$$

and integration by parts gives $\langle \alpha, u \rangle = 0$. Let us now show that α is a non-negative distribution. We approximate $\mathbf{1}_U$ by the sequence $H_\varepsilon = \theta_\varepsilon \circ \rho$, where θ_ε is a bump function such that $]-\infty, -\varepsilon] \prec \theta_\varepsilon \prec \mathbb{R}_-$ and θ_ε is monotone non-increasing. Since $H_\varepsilon \to \mathbf{1}_U$ in $L^1_{\text{loc}}(\mathbb{R}^d)$, we have

$$H_\varepsilon \nabla \rho \to \mathbf{1}_U \nabla \rho, \quad H_\varepsilon \Delta \rho \to \mathbf{1}_U \Delta \rho$$

in $L^1_{\text{loc}}(\mathbb{R}^d)$, hence in $\mathcal{D}'(\mathbb{R}^d)$, so $\alpha_\varepsilon \to \alpha$ in $\mathcal{D}'(\mathbb{R}^d)$, where

$$\alpha_\varepsilon = -\operatorname{div}(H_\varepsilon \nabla \rho) + H_\varepsilon \Delta \rho.$$

By (6.47), we have

$$\alpha_\varepsilon = -\nabla H_\varepsilon \cdot \nabla \rho = -\theta'_\varepsilon(\rho)|\nabla \rho|^2 \ge 0, \tag{A.7}$$

so $\alpha \geq 0$ in the sense of distributions (*i.e.* $\langle \alpha_\varepsilon, u\rangle \geq 0$ if $u \in \mathcal{D}(\mathbb{R}^d)$ and $u \geq 0$). We have seen (in exercise class) that a non-negative distribution is represented by a non-negative Borel measure, finite on compact sets. It now makes sense to define σ as the product of the continuous function $|\nabla\rho|^{-1}$ with α. To establish the Green formula (7.55), we use again the approximation of $\mathbf{1}_U$ by H_ε: by integration by parts,

$$\int_{\mathbb{R}^d} \mathrm{div}(\Psi) H_\varepsilon dx = -\int_{\mathbb{R}^d} \Psi \cdot \nabla H_\varepsilon dx = -\int_{\mathbb{R}^d} \Psi \cdot \nabla\rho\theta'_\varepsilon(\rho)dx.$$

Since $\nabla\rho = \nu|\nabla\rho|$, we have

$$\int_{\mathbb{R}^d} \mathrm{div}(\Psi) H_\varepsilon dx = -\int_{\mathbb{R}^d} \Psi \cdot \nu\frac{1}{|\nabla\rho|}\theta'_\varepsilon(\rho)|\nabla\rho|^2 dx, \tag{A.8}$$

which is $\langle \alpha_\varepsilon, \varphi\rangle$ by (A.7), where $\varphi = \Psi \cdot \nu|\nabla\rho|^{-1}$. We know that $\alpha_\varepsilon \to \alpha$ in the sense of distributions. We also have, for L compact in $\mathbb{R}^d$ and $u \in C^\infty_L(\mathbb{R}^d)$,

$$|\langle \alpha_\varepsilon, u\rangle| \leq \|\alpha_\varepsilon\|_{L^1(\mathbb{R}^d)} p_{L,0}(u) \leq A p_{L,0}(u),$$

since

$$\|\alpha_\varepsilon\|_{L^1(\mathbb{R}^d)} = -\int_{\mathbb{R}^d} \nabla H_\varepsilon \cdot \nabla\rho dx = \int_{\mathbb{R}^d} H_\varepsilon \Delta\rho dx \leq A := \|\Delta\rho\|_{L^1(V_1(\Gamma))},$$

where we use the fact that H_ε is supported in $V_1(\Gamma)$ for ε small enough. By Lemma A.1 and Remark A.1, we can then justify the convergence $\langle \alpha_\varepsilon, \varphi\rangle \to \langle \alpha, \varphi\rangle$ since φ is continuous (note that we should add here a discussion on the support of φ and α_ε since φ is well-defined in a neighbourhood of Γ only, but there is no difficulty in this point). We can now pass to the limit in (A.8) to obtain (7.55).

Step 2: Expression of σ in local coordinates. In a second step, we give the demonstration of (7.57). So let W be an open set such that U and $\Gamma = \partial U$ admit a parametrization by local graph in W as in (7.56). Note well that expression of ν on Γ at the point $g(x') = (x', \psi(x'))$ is

$$\nu(g(x')) = \frac{1}{\sqrt{1 + |\nabla\psi(x')|^2}} \begin{bmatrix} \nabla\psi(x') \\ -1 \end{bmatrix}. \tag{A.9}$$

Let $v \in C^1(\mathbb{R}^d)$ be supported in W. Our aim is to compute the expression $\langle \sigma, v\rangle$. Without loss of generality, we can assume that $\nabla\rho \neq 0$ in W so that $\nu = \frac{\nabla\rho}{|\nabla\rho|}$ is well defined on W. By the Green Formula (7.55), we have

$$\langle \sigma, v\rangle = \int_\Gamma v(z)d\sigma(z) = \int_{W\cap U} \mathrm{div}(\theta(x)v \cdot \nu(x))dx, \tag{A.10}$$

where θ is any function in $C^1(W)$ such that $\theta = 1$ on $\Gamma \cap W$. On the other hand, for any function $F \in C^1(\mathbb{R}^d; \mathbb{R}^d)$ supported in W, (7.56) yields the expression

$$\int_U \mathrm{div}(F(x))dx = \int_V \int_0^\infty \mathrm{div}(F)(x', x_d + \psi(x'))dx_d dx'.$$

By the chain-rule,

$$\mathrm{div}(F)(x', x_d + \psi(x')) = \mathrm{div}[F(x', x_d + \psi(x'))] - \nabla\psi(x') \cdot (\partial_{x_d} F')(x', x_d + \psi(x')),$$

where $F' = (F)_{1\leq i\leq d-1}$. By explicit integration, the expression of the integral of the divergence in coordinates is

$$\int_U \operatorname{div}(F(x))dx = -\int_V F_d(g(x'))dx' - \int_V \int_0^\infty \nabla\psi(x')\cdot(\partial_{x_d}F')(x', x_d + \psi(x'))dx_d dx'. \quad (A.11)$$

Let (ζ_η) be an approximation of the unit on $\mathbb{R}$ such that ζ_η is supported in $(0,\eta)$ and define, for $t \in \mathbb{R}$, $x = (x', x_d + \psi(x'))) \in U \cap W$,

$$\xi_\eta(t) = \int_t^\infty \zeta_\eta(s)ds, \quad \omega_\eta(x) = \xi_\eta(x_d - \psi(x')).$$

We will take $\theta = \omega_\eta$ in (A.10), which means that we will apply (A.11) with $F^\eta = v\nu\omega_\eta$. Note that ω_η tends to 0 in $L^1(U\cap W)$ by dominated convergence, since $0 \leq \omega_\eta \leq 1$ and $\omega_\eta(x) = 0$ as soon as $x_d - \psi(x') > \eta$. Consequently,

$$\nabla\psi(x')\cdot(\partial_{x_d}F')(x', x_d + \psi(x')) = -\nabla\psi(x')\cdot\nu'(g)\zeta_\eta(x_d) + o(1)$$

in $L^1(U\cap W)$, which gives

$$\langle\sigma, v\rangle = \lim_{\eta\to 0}\int_U \operatorname{div}(F^\eta(x))dx = -\int_V F_d^0(g(x'))dx' - \int_V \nabla\psi(x')\cdot\partial_{x_d}(F^0)'(g((x'))dx',$$

where $F^0(x) = v(x)\nu(x)$. We can compute then, with (A.9), that

$$\langle\sigma, v\rangle = \int_V \frac{1}{\sqrt{1+|\nabla\psi(x')|^2}}(1+|\nabla\psi(x')|^2)v(g(x'))dx',$$

and obtain (7.57).

Step 3: proof of (7.53)**.** By use of an appropriate partition of unity, it is sufficient to establish (7.53) when u is a continuous function supported in an open set W such that U and $\Gamma = \partial U$ admit a parametrization by local graph in W as in (7.56). We will then derive the formula

$$\int_{V_\varepsilon(\Gamma)} u(x)dx = \int_{\mathbb{R}^{d-1}}\int_{-\varepsilon}^{\varepsilon} u\circ\Phi(g(y'),t)\pi(g(y'),t)\sqrt{1+|\nabla\psi(y')|^2}dy'dt, \quad (A.12)$$

where $g(y') := (y', \psi(y'))$ and $\pi(g(y'),t)$ is given by (7.54) with $z = g(y')$. The local expression (7.57) of σ shows that (A.12) is precisely (7.53) under the local parametrization given by g. The identity (A.12) results from the change of variable $x = f(y',t) := \Phi(g(y'),t)$, once we show that the Jacobian determinant Jf of f is given by

$$Jf(y',t) = \pi(g(y'),t)\sqrt{1+|\nabla\psi(y')|^2}. \quad (A.13)$$

Let $z = g(y')$. By the chain rule, the linear map $T := d_{(y',t)}f$ is given by $T = S\circ h$, where $S\colon \mathbb{R}^{d+1}\to\mathbb{R}$ and $h\colon \mathbb{R}^d\to\mathbb{R}^{d+1}$ are given by

$$S(x,s) = d_{(z,t)}\Phi(x,s) = x - s\nu(z) - td_z\nu(x), \quad h(x',s) = (d_{y'}g(x'),s) = (x', \nabla\psi(y')\cdot x', s),$$

for $z \in \mathbb{R}^d$, $x' \in \mathbb{R}^{d-1}$. We have used the formula (7.51) to express the differential of Φ. We will also use the fact that there is an orthonormal basis $(\varepsilon_i)_{1,d}$ of $\mathbb{R}^d$ such that

$$\varepsilon_d = \nu(z), \quad d_z\nu(\varepsilon_i) = \lambda_i\varepsilon_i,\ 1\leq i\leq d,$$

where we have set $\lambda_d = 0$. This was mentioned after (7.48) and the product in (7.54) involves the principal curvatures λ_i, $i = 1, \ldots, d-1$. We denote by (e_i) the canonical basis of $\mathbb{R}^d$. We expand ε_i as $\varepsilon_i = (\varepsilon_i', \varepsilon_i^d) \in \mathbb{R}^{d-1} \times \mathbb{R}$ and compute, for $1 \leq i \leq d-1$,

$$0 = \varepsilon_i \cdot \nu(z) = \frac{1}{\sqrt{1 + |\nabla\psi(y')|^2}} \left[\nabla\psi(y') \cdot \varepsilon_i' - \varepsilon_i^d\right],$$

hence $\varepsilon_i^d = \nabla\psi(y') \cdot \varepsilon_i'$ and $h(\varepsilon_i) = (\varepsilon_i', \nabla\psi(y') \cdot \varepsilon_i', \varepsilon_i^d) = (\varepsilon_i, \varepsilon_i^d)$, from which follows the first identity

$$T(\varepsilon_i) = S(\varepsilon_i, \varepsilon_i^d) = (1 - t\lambda_i)\varepsilon_i - \varepsilon_i^d \nu(z). \tag{A.14}$$

We also compute $h(e_d) = (0, e_d) \in \mathbb{R}^d \times \mathbb{R}$ and get

$$T(e_d) = S(0, 1) = -\nu(z) = -\varepsilon_d. \tag{A.15}$$

We will extract the value of $|\det(T)|$ from the identity

$$|\det(T(\varepsilon_1), \cdots, T(\varepsilon_{d-1}), T(e_d))| = |\det(T)||\det(\varepsilon_1, \cdots, \varepsilon_{d-1}, e_d)|.$$

Since

$$|\det(T(\varepsilon_1), \cdots, T(\varepsilon_{d-1}), T(e_d))| = \prod_{i=1}^{d-1}(1 - t\lambda_i)|\det(\varepsilon_1, \ldots, \varepsilon_d)| = \prod_{i=1}^{d-1}(1 - t\lambda_i) = \pi(g(y'), t),$$

we will obtain the desired conclusion if we show that

$$|\det(\varepsilon_1, \cdots, \varepsilon_{d-1}, e_d)| = \frac{1}{\sqrt{1 + |\nabla\psi(y')|^2}}. \tag{A.16}$$

To compute the determinant in (A.16), we use the formula

$$|\det(\varepsilon_1, \cdots, \varepsilon_{d-1}, e_d)| = \sqrt{\mathrm{Gram}(\varepsilon_1, \cdots, \varepsilon_{d-1}, e_d)},$$

where the Gram determinant $\mathrm{Gram}(u_1, \cdots, u_d)$ of d vectors $u_i \in \mathbb{R}^d$ is the determinant of the Gram matrix $G = (u_i \cdot u_j)_{1 \leq i,j \leq d}$. We compute

$$\varepsilon_i \cdot \varepsilon_j = \delta_{ij}, \quad \varepsilon_i \cdot e_d = \varepsilon_i^d,$$

so

$$\mathrm{Gram}(\varepsilon_1, \cdots, \varepsilon_{d-1}, e_d) = \det \begin{bmatrix} 1 & & & \varepsilon_1^d \\ & \ddots & & \vdots \\ & & 1 & \varepsilon_{d-1}^d \\ \varepsilon_1^d & \ldots & \varepsilon_{d-1}^d & 1 \end{bmatrix}.$$

We develop the determinant along the last column to obtain

$$\mathrm{Gram}(\varepsilon_1, \cdots, \varepsilon_{d-1}, e_d) = 1 - \left[|\varepsilon_1^d|^2 + \cdots + |\varepsilon_{d-1}^d|^2\right].$$

By decomposition of the vector e_d in the orthonormal basis $(\varepsilon_i)_{1,d}$, we have

$$1 = \sum_{i=1}^{d-1} |\varepsilon_i^d|^2 + \frac{1}{1 + |\nabla\psi(y')|^2},$$

so

$$\mathrm{Gram}(\varepsilon_1, \cdots, \varepsilon_{d-1}, e_d) = \frac{1}{1 + |\nabla\psi(y')|^2},$$

and (A.16) follows.

□

References

[Ada75] Robert A. Adams. *Sobolev spaces*. Pure and Applied Mathematics, Vol. 65. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975.

[Bil95] P. Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, third edition, 1995. A Wiley-Interscience Publication.

[Bil99] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[Bou91] Gérard Bourdaud. Le calcul fonctionnel dans les espaces de Sobolev. (Functional calculus in Sobolev spaces). *Invent. Math.*, 104(2):435–446, 1991.

[Bre11] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.

[CDL08] Gianluca Crippa and Camillo De Lellis. Estimates and regularity results for the DiPerna-Lions flow. *J. Reine Angew. Math.*, 616:15–46, 2008.

[Cla36] James A. Clarkson. Uniformly convex spaces. *Trans. Amer. Math. Soc.*, 40(3):396–414, 1936.

[DPZ14] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 2014.

[Dro01] Jérôme Droniou. Quelques Résultats sur les Espaces de Sobolev. working paper or preprint, April 2001, https://hal.archives-ouvertes.fr/hal-01382370/file/sobolev.pdf.

[DS58] Nelson Dunford and Jacob T. Schwartz. *Linear Operators. I. General Theory*. Pure and Applied Mathematics, Vol. 7. Interscience Publishers, Inc., New York; Interscience Publishers, Ltd., London, 1958. With the assistance of W. G. Bade and R. G. Bartle.

[EG92] L.C. Evans and R.F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[EG15] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions*. Textbooks in Mathematics. CRC Press, Boca Raton, FL, revised edition, 2015.

[Eva10] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[Fol99] Gerald B. Folland. *Real analysis. Modern techniques and their applications*. New York, NY: Wiley, 1999.

[Giu84] E. Giusti. Minimal surfaces and functions of bounded variation. Birkhäuser Verlag, 1984.

[Haj99] Piotr Hajlasz. Pointwise Hardy inequalities. *Proc. Am. Math. Soc.*, 127(2):417–423, 1999.

[Hof13] M. Hofmanová. Strong solutions of semilinear stochastic partial differential equations. *NoDEA Nonlinear Differential Equations Appl.*, 20(3):757–778, 2013.

[Lax02] Peter D. Lax. *Functional analysis*. Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York, 2002.

[Lim88] Elon L. Lima. The Jordan-Brouwer separation theorem for smooth hypersurfaces. *Amer. Math. Monthly*, 95(1):39–42, 1988.

[LL01] Elliott H. Lieb and Michael Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2001.

[McG16] Peter McGrath. On the smooth Jordan Brouwer separation theorem. *Amer. Math. Monthly*, 123(3):292–295, 2016.

[Rud73] Walter Rudin. *Functional analysis*. McGraw-Hill Series in Higher Mathematics. McGraw-Hill Book Co., New York-Düsseldorf-Johannesburg, 1973.

[Rud87] W. Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[Sim83] Leon Simon. *Lectures on geometric measure theory*, volume 3 of *Proceedings of the Centre for Mathematical Analysis, Australian National University*. Australian National University, Centre for Mathematical Analysis, Canberra, 1983.

[SS05] Elias M. Stein and Rami Shakarchi. *Real analysis*, volume 3 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ, 2005. Measure theory, integration, and Hilbert spaces.

[Vil03] Cédric Villani. Analyse II, 2003, https://www.cedricvillani.org/sites/dev/files/old_images/2013/03/ana2.pdf.

[Yos80] K. Yosida. *Functional analysis*, volume 123 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin-New York, sixth edition, 1980.

[Zui02] Claude Zuily. *Eléments de distributions et d'équations aux dérivées partielles*. Sciences Sup, 2002.