Chapter

2

A review of analysis

2.1 A few basic function spaces

Let us rapidly review the most basic function spaces that we will need. In the sequel, Ω denotes an open subset of $\mathbb{R}^d$. The canonical scalar product of two vectors in $\mathbb{R}^d$ will be denoted $x \cdot y$ and the associated Euclidean norm $\|x\|$. We recall the multiindex notation for partial derivatives. Let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d) \in \mathbb{N}^d$ be a multiindex. The integer $|\alpha| = \sum_{i=1}^d \alpha_i$ is the length of α and we set

$$\partial^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_d^{\alpha_d}}.$$

whenever the function u is $|\alpha|$-times differentiable and the partial derivatives commute. The space $C^0(\Omega)$ is the space of real-valued, continuous functions on Ω, and for all $k \in \mathbb{N}$, we define

$$C^k(\Omega) = \{u; \text{for all } \alpha \in \mathbb{N}^d, |\alpha| \leq k, \partial^\alpha u \in C^0(\Omega)\}$$

to be the space of k-times continuously differentiable functions on Ω. The space of indefinitely differentiable functions on Ω is defined by

$$C^\infty(\Omega) = \bigcap_{k \in \mathbb{N}} C^k(\Omega).$$

We do not specify the natural topology of these vector spaces as we will not need it. Beware however that they are not normed vector spaces.

The *support* of a function u, $\operatorname{supp} u$, is the complement of the largest open subset of Ω on which u vanishes. It is thus a closed subset of Ω. We recall that a compact subset K of Ω is a closed subset that "does not touch the boundary" in the sense that $d(K, \complement_{\mathbb{R}^d}\Omega) > 0$. There is a "security strip" between K and $\partial\Omega$. Functions with compact support play an important role and deserve a notation of their own:

$$C_c^k(\Omega) = \mathscr{D}^k(\Omega) = \{u \in C^k(\Omega); \operatorname{supp} u \text{ is compact}\}$$

and

$$C_c^\infty(\Omega) = \mathscr{D}(\Omega) = \bigcap_{k\in\mathbb{N}} \mathscr{D}^k(\Omega).$$

Again, these vector spaces are endowed with natural topologies that we will not describe. We will return to these spaces later when talking about distributions.

Let $\bar{\Omega}$ be the closure of Ω in $\mathbb{R}^d$. The space $C^0(\bar{\Omega})$ is the space of continuous functions on $\bar{\Omega}$. If $\bar{\Omega}$ is *compact*, that is to say, when Ω is bounded, this space is normed by

$$\|u\|_{C^0(\bar{\Omega})} = \sup_{x\in\bar{\Omega}} |u(x)| = \max_{x\in\bar{\Omega}} |u(x)|.$$

The convergence associated to this normed topology is just uniform convergence. From now on, we will assume that Ω is bounded.

Likewise, we define $C^k(\bar{\Omega})$ to be the space of functions in $C^k(\Omega)$, all the partial derivatives of which up to order k have a continuous extension to $\bar{\Omega}$. Keeping the same symbol for this extension, the natural norm of this space is

$$\|u\|_{C^k(\bar{\Omega})} = \max_{|\alpha|\leq k} \|\partial^\alpha u\|_{C^0(\bar{\Omega})}.$$

All these spaces are Banach spaces, *i.e.*, they are complete for the metric defined by their norm. We also define

$$C^\infty(\bar{\Omega}) = \bigcap_{k\in\mathbb{N}} C^k(\bar{\Omega}),$$

which is again not a normed space.

For $0 < \beta \leq 1$, we define the spaces of Hölder functions (Lipschitz for $\beta = 1$) by

$$C^{0,\beta}(\bar{\Omega}) = \Big\{u \in C^0(\bar{\Omega}); \sup_{x\neq y} \frac{|u(x)-u(y)|}{\|x-y\|^\beta} < +\infty\Big\}$$

and

$$C^{k,\beta}(\bar{\Omega}) = \{u \in C^k(\bar{\Omega}); \partial^\alpha u \in C^{0,\beta}(\bar{\Omega}) \text{ for all } |\alpha| = k\}.$$

When equipped with the norms

$$\|u\|_{C^{k,\beta}(\bar\Omega)} = \|u\|_{C^k(\bar\Omega)} + \max_{|\alpha|=k}\Bigl(\sup_{x\neq y}\frac{|\partial^\alpha u(x)-\partial^\alpha u(y)|}{\|x-y\|^\beta}\Bigr),$$

these spaces also are Banach spaces. There are continuous injections $C^{k,\beta}(\bar\Omega) \hookrightarrow C^{k,\beta'}(\bar\Omega) \hookrightarrow C^k(\bar\Omega) \hookrightarrow C^{k-1,\gamma}(\bar\Omega)$ which are compact for $\beta' < \beta$ and $\gamma < 1$ by Ascoli's theorem (the compactness of the first embedding requires some regularity on Ω, see Section 2.2). We recall that a linear mapping f from a normed space E to a normed space F is continuous if and only if there exists a constant C such that for all $x \in E$, $\|f(x)\|_F \le C\|x\|_E$. In the continuous injections above, f is just the identity, $f(u) = u$. A mapping is compact if it transforms bounded sets into relatively compact sets.

The other major family of function spaces that will be useful to us is that of the Lebesgue spaces. We recall that

$$L^p(\Omega) = \Bigl\{u \text{ measurable}; \int_\Omega |u(x)|^p\,dx < +\infty\Bigr\}$$

for $1 \le p < +\infty$ and

$$L^\infty(\Omega) = \Bigl\{u \text{ measurable}; \operatorname*{ess\,sup}_\Omega |u| < +\infty\Bigr\}.$$

Now in these definitions, u is not strictly speaking a function, but an equivalence class of functions that are equal almost everywhere with respect to the Lebesgue measure. Likewise, the essential supremum of the second definition has a slightly convoluted definition to accommodate the equivalence classes. However, in practice and outside of very specific circumstances, it is harmless to think of u as just a function, not an equivalence class. We just need to keep this fact at the back of our mind, just in case.

When equipped with the norms

$$\|u\|_{L^p(\Omega)} = \Bigl(\int_\Omega |u(x)|^p\,dx\Bigr)^{\frac{1}{p}}$$

for $1 \le p < +\infty$ and

$$\|u\|_{L^\infty(\Omega)} = \operatorname*{ess\,sup}_\Omega |u|,$$

the Lebesgue spaces are Banach spaces. For $p = 2$, the space $L^2(\Omega)$ is a Hilbert space for the scalar product

$$(u|v)_{L^2(\Omega)} = \int_\Omega u(x)v(x)\,dx.$$

We recall Hölder's inequality

$$\int_\Omega |u(x)v(x)|\,dx \le \Big(\int_\Omega |u(x)|^p\,dx\Big)^{\frac{1}{p}} \Big(\int_\Omega |v(x)|^{p'}\,dx\Big)^{\frac{1}{p'}}$$

when p, p' are conjugate exponents, $\frac{1}{p}+\frac{1}{p'}=1$ (the integrals do not need to be finite). In particular, if $u \in L^p(\Omega)$ and $v \in L^{p'}(\Omega)$, then $uv \in L^1(\Omega)$ and

$$\Big|\int_\Omega u(x)v(x)\,dx\Big| \le \|uv\|_{L^1(\Omega)} \le \|u\|_{L^p(\Omega)}\|v\|_{L^{p'}(\Omega)}.$$

For $p=2$, we get the Cauchy-Schwarz inequality, which is actually a Hilbert space property

$$|(u|v)_{L^2(\Omega)}| \le \|u\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)}.$$

Since Ω is assumed to be bounded, there are continuous injections $C^k(\bar{\Omega}) \hookrightarrow L^p(\Omega) \hookrightarrow L^q(\Omega)$ whenever $q \le p$.[1]

The Lebesgue spaces admit local versions

$$L^p_{\text{loc}}(\Omega) = \{u; u_{|K} \in L^p(K) \text{ for all compact } K \subset \Omega\}.$$

These vector spaces have a natural topology which is not a normed topology.

Clearly, in view of Hölder's inequality, we have $L^p_{\text{loc}}(\Omega) \subset L^q_{\text{loc}}(\Omega)$ whenever $q \le p$. In particular, the space $L^1_{\text{loc}}(\Omega)$ is the largest of all these spaces, and actually the largest of all function spaces introduced up to now, which are all continuously embedded in it.

The following result is of importance.

Proposition 2.1.1 *Let $u \in L^1_{\text{loc}}(\Omega)$ be such that $\int_\Omega u\varphi\,dx = 0$ for all $\varphi \in \mathscr{D}(\Omega)$. Then $u = 0$ almost everywhere.*

Proof. Note first that since φ has support in a compact subset K of Ω, so does the product $u\varphi$. Since φ is bounded, it follows that $u\varphi \in L^1(K)$ and the integral is well-defined.

Let $x_0 \in \Omega$ and n be large enough so that $B\big(x_0, \frac{1}{n}\big) \subset \Omega$. It is possible (exercise) to construct a sequence $\varphi_k \in \mathscr{D}(\Omega)$ such that $\operatorname{supp}\varphi_k \subset B\big(x_0, \frac{1}{n}\big)$ and for all $x \in B\big(x_0, \frac{1}{n}\big)$, $\varphi_k(x) \to 1$. Consequently, by the Lebesgue dominated convergence theorem, we have

$$0 = \int_{B(x_0,\frac{1}{n})} u\varphi_k\,dx \underset{k\to+\infty}{\longrightarrow} \int_{B(x_0,\frac{1}{n})} u\,dx.$$

Hence, since

$$0 = \frac{1}{\operatorname{meas} B\big(x_0,\frac{1}{n}\big)} \int_{B(x_0,\frac{1}{n})} u\,dx \underset{n\to+\infty}{\longrightarrow} u(x_0)$$

for almost all x_0 by the Lebesgue points theorem, we obtain the result. □

[1] This is doubly false if Ω is not bounded.

Remark 2.1.1 Here we see again at work the idea a testing a function u with a test-function φ in order to obtain information on u. The general concept behind it is that of *duality* and it will be used in much larger generality in the context of distributions and variational formulations that we will see later on. □

2.2 Regularity of open subsets of $\mathbb{R}^d$

The structure of the open subsets of $\mathbb{R}^d$ for the usual topology is very simple for $d = 1$, since every open set is a union of an at most countable family of disjoint open intervals. The situation is more complicated in higher dimensions.

People tend to think of a connected open set of $\mathbb{R}^d$ as a potato-shaped object drawn in $\mathbb{R}^2$. This geometrical intuition is basically correct as far as the open set itself is concerned. It is misleading when the boundary of the open set is concerned. In fact, the boundary of an open set in $\mathbb{R}^d$, $d > 1$, can be more or less regular, more or less smooth.

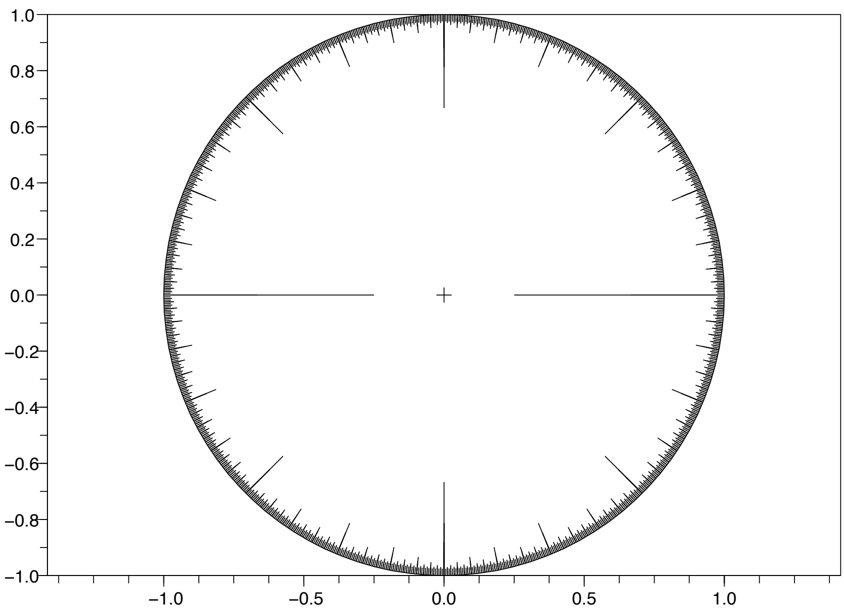

Figure 1. An open set in $\mathbb{R}^2$ with a relatively wild boundary (imagine an infinity of little spikes pointing inward the disk).

There is worse: the Mandelbrot set is compact, its complement is open with a very convoluted boundary.

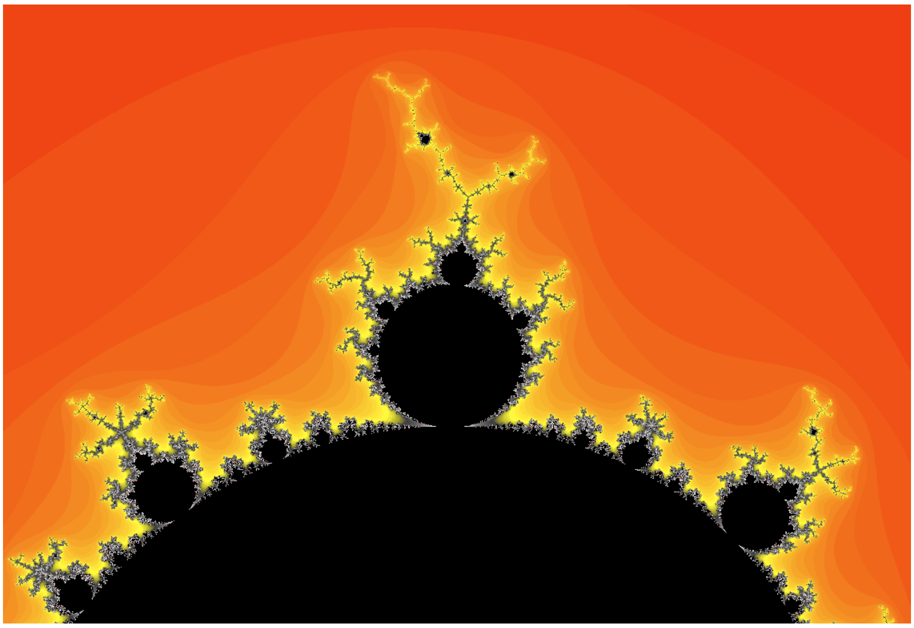

Figure 2. A zoom on the complement (in color) of the Mandelbrot set (in black).

It is even possible to construct open sets in $\mathbb{R}^2$ (or in $\mathbb{R}^d$ for any d for that matter), the boundary of which has strictly positive Lebesgue measure, *i.e.*, a strictly positive area! PDE problems are posed in open subsets of $\mathbb{R}^d$ and we often need a certain amount of regularity for the boundary of such open sets in order to deal with boundary conditions.

There are several ways of quantifying the regularity of an open set boundary, or in short the regularity of that open set. Let us give the definition that is the most adequate for our purposes here. You may encounter other definitions—equivalent or not—in the literature.

Definition 2.2.1 *We say that a bounded open subset of $\mathbb{R}^d$ is Lipschitz* (resp. *of class $C^{k,\beta}$*) *if its boundary $\partial\Omega$ can be covered by a finite number of open hypercubes C_j, $j = 1,\dots,m$, with an attached system of orthonormal Cartesian coordinates, $y^j = (y_1^j, y_2^j, \dots, y_d^j)$, in such a way that*

$$C_j = \{y \in \mathbb{R}^d; |y_i^j| < a_j \text{ for } i = 1,\dots,d\},$$

and there exists Lipschitz functions (resp. *of class $C^{k,\beta}$*) $\varphi_j \colon \mathbb{R}^{d-1} \to \mathbb{R}$ *such that*

$$\Omega \cap C_j = \{y \in C_j; y_d^j < \varphi_j((y^j)')\},$$

using the notation $\mathbb{R}^{d-1} \ni (y^j)' = (y_1^j, y_2^j, \dots, y_{d-1}^j)$.

The meaning of Definition 2.2.1 is that locally in C_j, Ω consists of those points located strictly under the graph of φ_j, in the other words, the *hypograph* of φ_j, see Figure 3. In particular, such an open set is situated on just one side of its boundary, which consists of pieces of graphs, since

$$\partial\Omega \cap C_j = \{y \in C_j; y_d^j = \varphi_j((y^j)')\}.$$

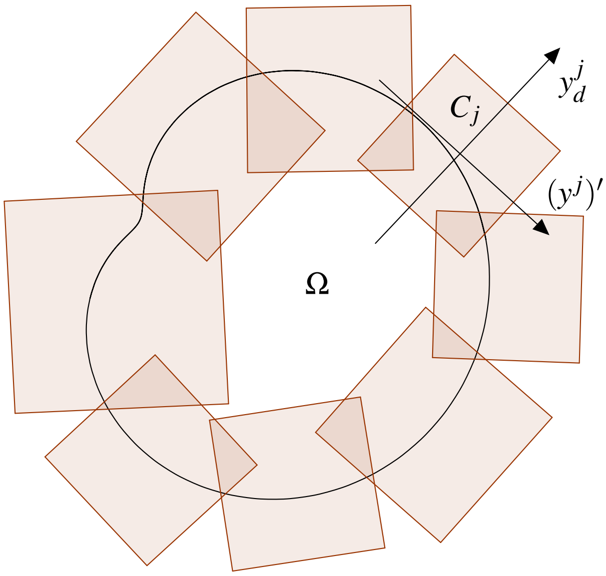

Figure 3. Covering the boundary with hypercubes.

Remark 2.2.1 It is fairly clear that a bounded polygon is a Lipschitz open set in dimension 2. None of the wild examples considered before is of class $C^{0,\beta}$.

On the other hand, there also are perfectly nice open sets that are not Lipschitz in the previous sense. Here is an example.

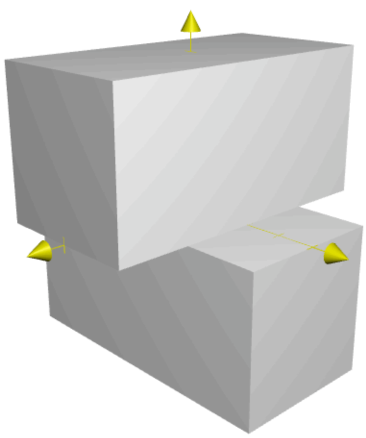

Figure 4. Simple, however not Lipschitz.

This example is obtained by gluing together two parallelepipeds one on top of the other, adding the open square of contact. It is impossible to describe the resulting set as a hypograph at each vertex of that square. The open set is nonetheless perfectly tame, it is a polyhedron. □

The boundary of a Lipschitz open set, and a fortiori that of an open set of class $C^{k,\alpha}$, $k \geq 1$, possesses a certain number of useful properties.

Proposition 2.2.1 *Let Ω be a Lipschitz open set. There exists a normal unit exterior vector n, defined almost everywhere on $\partial\Omega$.*

Normal means orthogonal to the boundary, exterior means that it points toward the complement of Ω. We will go back to the meaning of almost everywhere later.
Proof. Let us work in C_j and drop all j indices and exponents to simplify notation. We will admit *Rademacher's theorem*, a nontrivial results that says that a Lipschitz function on $\mathbb{R}^{d-1}$ is differentiable in the classical sense, almost everywhere with respect to the Lebesgue measure in $\mathbb{R}^{d-1}$.

Let y' be a point of differentiability of φ. At this point, the differentiability implies that the graph of φ has a tangent hyperplane generated by the $d-1$ vectors

$$a_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \partial_1\varphi(y') \end{pmatrix}, a_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ \partial_2\varphi(y') \end{pmatrix}, \cdots, a_{d-1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \partial_{d-1}\varphi(y') \end{pmatrix},$$

(slightly different notation for partial derivatives here, $\partial_i\varphi = \frac{\partial\varphi}{\partial y_i}$). The orthogonal straight line is generated by the vector

$$N = \begin{pmatrix} -\partial_1\varphi(y') \\ -\partial_2\varphi(y') \\ \vdots \\ -\partial_{d-1}\varphi(y') \\ 1 \end{pmatrix}.$$

which is clearly orthogonal to all a_i. To conclude, we just need to normalize it and notice that it points outwards due to the strictly positive last component and Ω lying under the graph

$$n = \frac{1}{\sqrt{1+\|\nabla\varphi(y')\|^2}} \begin{pmatrix} -\partial_1\varphi(y') \\ -\partial_2\varphi(y') \\ \vdots \\ -\partial_{d-1}\varphi(y') \\ 1 \end{pmatrix},$$

with $\|\nabla\varphi(y')\|^2 = \sum_{i=1}^{d-1}(\partial_i\varphi(y'))^2$. □

Remark 2.2.2 It should be noted that the normal vector n is an object of purely geometric nature that does not depend on the particular system of coordinates used to compute it. In particular, if we take another admissible covering of the boundary, the same formulas apply and compute the same vector in different coordinate systems. This geometrically obvious remark can also be checked by direct computation in two different coordinate systems.

The "almost everywhere" is meant in the sense of the space $\mathbb{R}^{d-1}$ associated with a local coordinate system. It will shortly be given an intrinsic meaning. □

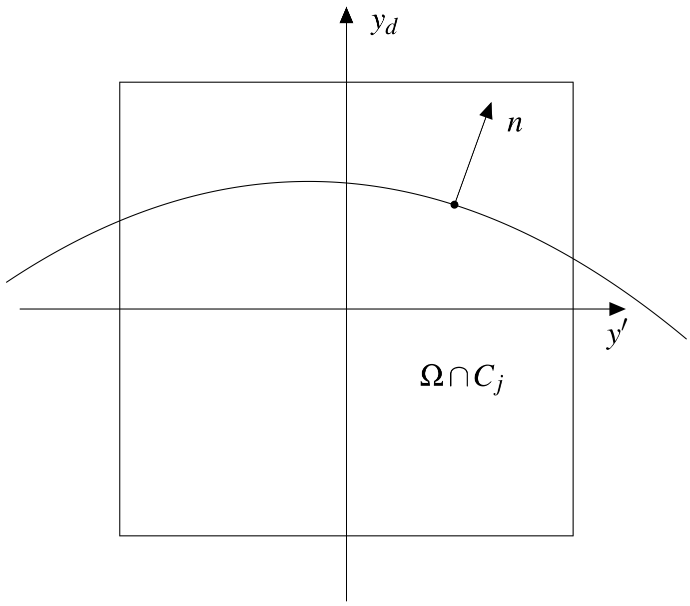

Figure 5. Local aspect of the boundary of Lipschitz open set and the normal unit exterior vector.

If Ω is a Lipschitz subset of $\mathbb{R}^d$, there is a natural measure on $\partial\Omega$ that is inherited in a sense from the Lebesgue measure in $\mathbb{R}^d$, that we will call the *boundary measure*. We will not go into all the detail but give a few ideas on how this measure can be computed.

Let $A \subset \partial\Omega$ be a Borel subset of $\partial\Omega$. Since the open sets C_j cover the boundary, we can partition A with Borel sets $A_j \subset C_j$. Let Π_j be the orthogonal projection from C_j onto $\mathbb{R}^{d-1}$ according to the coordinate system associated with C_j. The restriction of the projection to the graph of φ_j is a homeomorphism, therefore $\Pi_j(A_j)$ is a Borel subset of $\mathbb{R}^{d-1}$.

We set

$$\mathscr{H}_{d-1}(A_j) = \int_{\Pi_j(A_j)} \sqrt{1+\|\nabla\varphi_j((y^j)')\|^2}\, d(y^j)' \text{ and } \mathscr{H}_{d-1}(A) = \sum_{j=1}^{m} \mathscr{H}_{d-1}(A_j).$$

It can be checked, although it is quite tedious, that this formula does not depend on the covering and coordinates chosen to compute it, and that it defines a Borel measure on $\partial\Omega$.

In the case when $\nabla\varphi_j$ is constant, that is to say if the graph is portion of a hyperplane, it also easy to check that the formula above gives the $(d-1)$-dimensional Lebesgue measure on the hyperplane, using the same unit of length as in $\mathbb{R}^d$. In this sense, the boundary measure is inherited from $\mathbb{R}^d$.

For $d = 2$, the boundary of Ω consists of curves and if these curves are regular, we recognize the length of the parametric curve $y_1 \mapsto (y_1, \varphi(y_1))$. Same thing for $d = 3$ with the area of a parametric surface. The notation $\mathscr{H}_{d-1}$ alludes to the $(d-1)$-Hausdorff measure, a much more general and complicated object that coincides here with our hand-crafted measure.

It is now clear that the normal vector is defined almost everywhere with respect to the boundary measure. In addition, we can now define $L^p(\partial\Omega)$ spaces and compute all sorts of integrals on the boundary, using this measure. In order to have a more economical notation, we will write it $d\Gamma$ in the integrals. Thus, if g is a function on the boundary with support in C_j, we have

$$\int_{\partial\Omega} g\, d\Gamma = \int_{\Pi_j(C_j)} g(\Pi_j^{-1}(y')) \sqrt{1+\|\nabla\varphi_j((y^j)')\|^2}\, d(y^j)'.$$

The formula is extended to all functions without condition of support by a partition of unity, see below.

2.3 Partitions of unity

Partitions of unity are a basic tool that is used in many contexts whenever the need arises to localize a function. In what follows, Ω will be a bounded open subset of $\mathbb{R}^d$ with a finite covering C_j, $j = 0,\dots,m$, of its boundary $\partial\Omega$.[2]

Proposition 2.3.1 *Let C_0 be an open set such that $\bar{C}_0 \subset \Omega$ and $\Omega \subset \cup_{j=0}^{m} C_j$. There exist $m+1$ functions $\psi_j \colon \mathbb{R}^d \to [0,1]$ of class C^∞ such that* $\operatorname{supp}\psi_j \subset \bar{C}_j$ *and $\sum_{j=0}^{m} \psi_j = 1$ in Ω.*

[2] This particular assumption is only because this is the context in which we will use partitions of unity here. It should be clear from the proof, that the result extends to more general covers.

Proof. Recall first that for any closed set A, the function

$$x \mapsto d(x,A) = \inf_{y\in A} \|x-y\|$$

is a continuous function from $\mathbb{R}^d$ into $\mathbb{R}_+$ that vanishes exactly on A.

We can choose $\eta > 0$ small enough so that:

1. The sets $C_j^\eta = \{x \in C_j; d(x, \partial C_j) > \eta\}$ still form an open cover of Ω in the sense that $\Omega \subset \cup_{j=0}^m C_j^\eta$.

2. We can take an open set C_{m+1}^η such that $\bar{C}_{m+1}^\eta \subset \mathbb{R}^d \setminus \Omega$ in order to cover the whole of $\mathbb{R}^d = \cup_{j=0}^{m+1} C_j^\eta$, and such that $d(\bar{C}_{m+1}^\eta, \bar{\Omega}) > \eta)$.

This is possible by compactness of $\bar{\Omega}$ but we omit the (tedious) details.

The functions

$$\psi_j^\eta(x) = \frac{d(x, \mathbb{R}^d \setminus C_j^\eta)}{\sum_{k=0}^{m+1} d(x, \mathbb{R}^d \setminus C_k^\eta)} \tag{2.1}$$

are continuous on $\mathbb{R}^d$, indeed the denominator never vanishes because of the covering property. They are $[0,1]$-valued and ψ_j^η has support C_j^η. Finally, it is clear that $\sum_{j=0}^{m+1} \psi_j(x) = 1$ on $\mathbb{R}^d$, with ψ_{m+1}^η identically zero on the set $\{x; d(x,\bar{\Omega}) \le \eta\}$ which contains Ω.

This family of functions has all the desired properties except that the functions are not smooth. We thus use the convolution with a mollifier ρ_η with support in the ball $B(0,\eta)$.[3] We have

$$1 = 1 \star \rho_\eta = \Big(\sum_{j=0}^{m+1} \psi_j^\eta\Big) \star \rho_\eta = \sum_{j=0}^{m+1} (\psi_j^\eta \star \rho_\eta).$$

Each function $\psi_j^\eta \star \rho_\eta$ has support in C_j for $j = 0, \ldots, m+1$, with $\psi_{m+1}^\eta \star \rho_\eta = 0$ on Ω (this is the reason why we shrank all open sets by η in the beginning since the convolution spreads supports by an amount η), and is of class C^∞. □

Let us give an example in dimension 1, without the final smoothing step. We take $\Omega =]0,1[$, $C_0 =]\frac{1}{8}, \frac{7}{8}[$, $C_1 =]-\frac{1}{4}, \frac{1}{4}[$, $C_2 =]\frac{3}{4}, \frac{5}{4}[$, $C_3 =]-\frac{1}{8}, \frac{3}{8}[$ and $C_4 =]-\infty, 0[\cup]1, +\infty[$. All functions can be computed explicitly. Thus, denoting $\xi_j(x) = d(x, \mathbb{R} \setminus C_j)$, we have

$$\xi_0(x) = \min\left\{ \left(x - \frac{1}{8}\right)_+, \left(\frac{7}{8} - x\right)_+ \right\},$$

and so on.

[3] For more details, see for instance http://www.ann.jussieu.fr/~ledret/OBAAultime.pdf in French, or any standard analysis textbook.

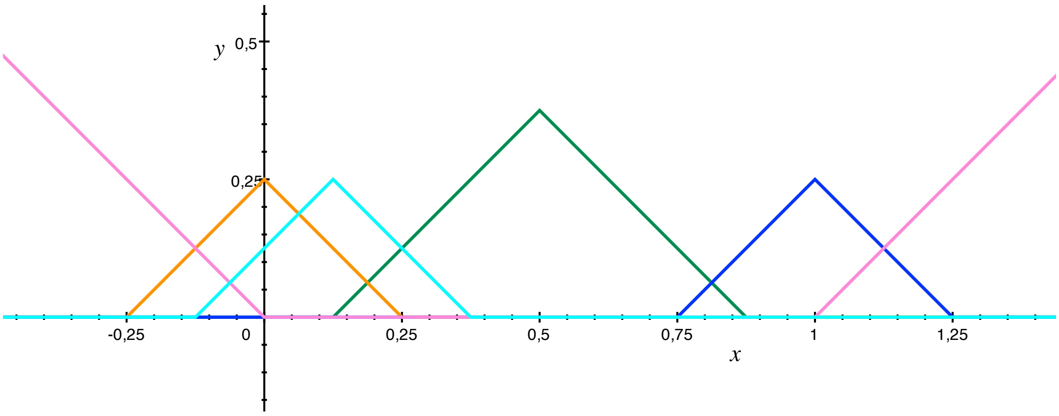

Figure 6. The five functions ξ_j, ξ_0 in green, ξ_1 in orange, ξ_2 in blue ξ_3 in turquoise and ξ_4 in pink.

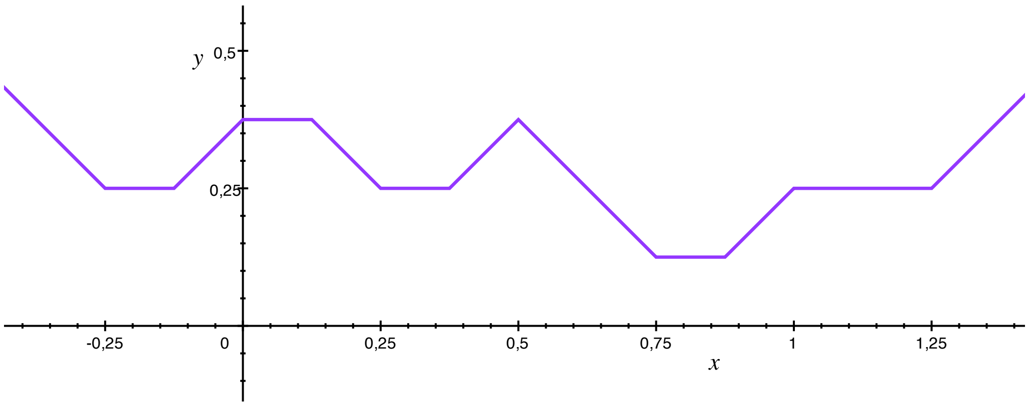

Figure 7. Their sum, *i.e.*, the denominator of (2.1), which never vanishes, in violet.

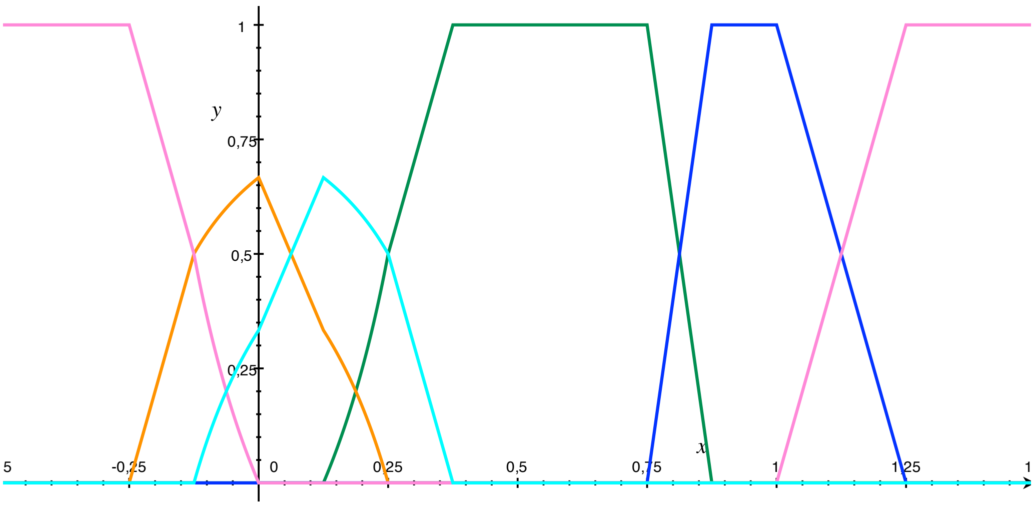

Figure 8. The partition of unity $\psi_j, j = 0, \ldots, 4$, with the same color convention.

We note that the set C_3 is unnecessary to have a covering of Ω. We just added it to have a nicer picture. If we had not added it, the partition of unity would have been piecewise affine and it is a mistake to think the partitions of unity derived from formula (2.1) are always piecewise affine!

Let us also illustrate an example in dimension 2, Ω is the unit disk covered, by three squares of side 2.5, centered at 1, j and j^2 (identifying $\mathbb{R}^2$ and $\mathbb{C}$) and rotated so as to form a covering of the boundary as required. There is no C_0, since the three squares already cover Ω.

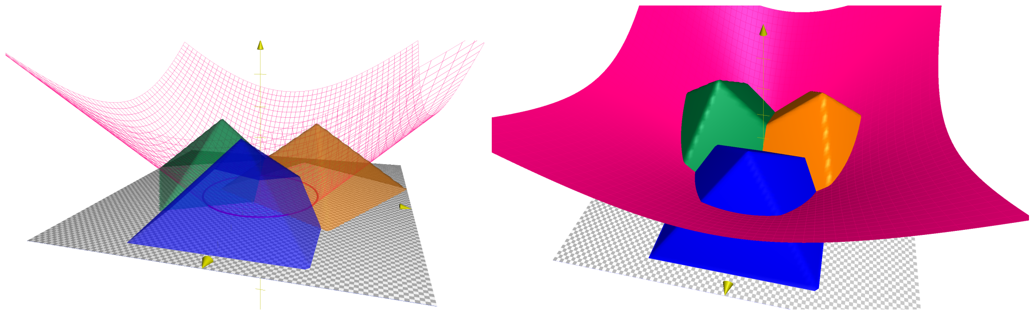

Figure 9. The four functions ξ_j.

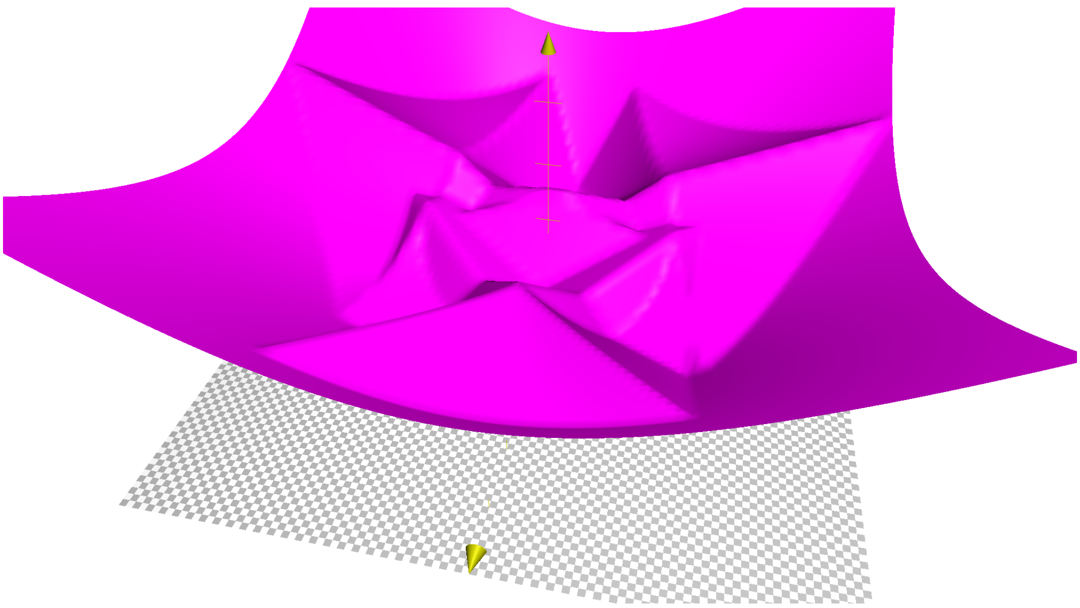

Figure 10. Their sum.

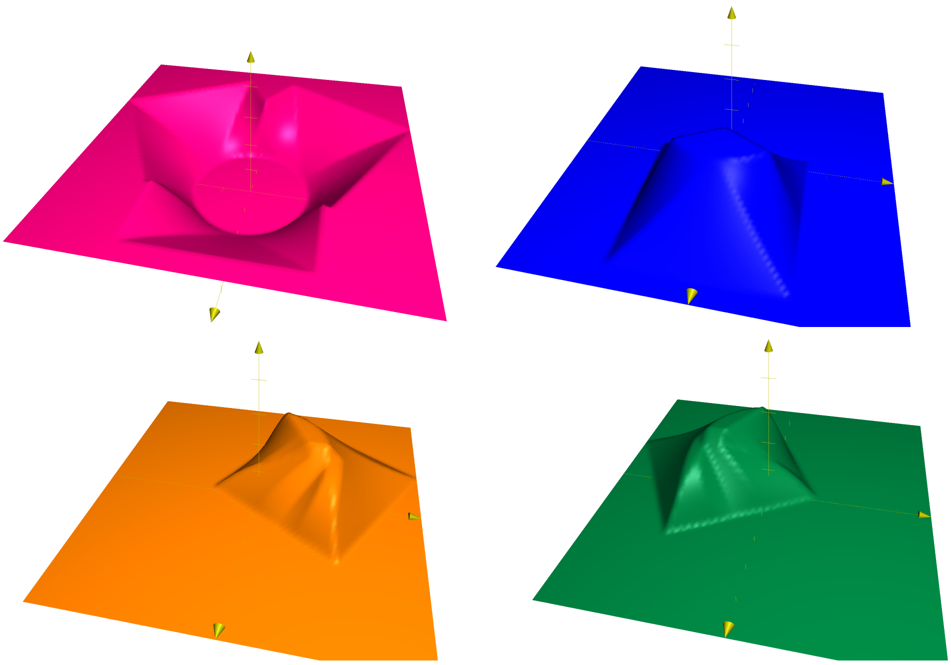

Figure 11. The four functions ψ_j, drawn separately.

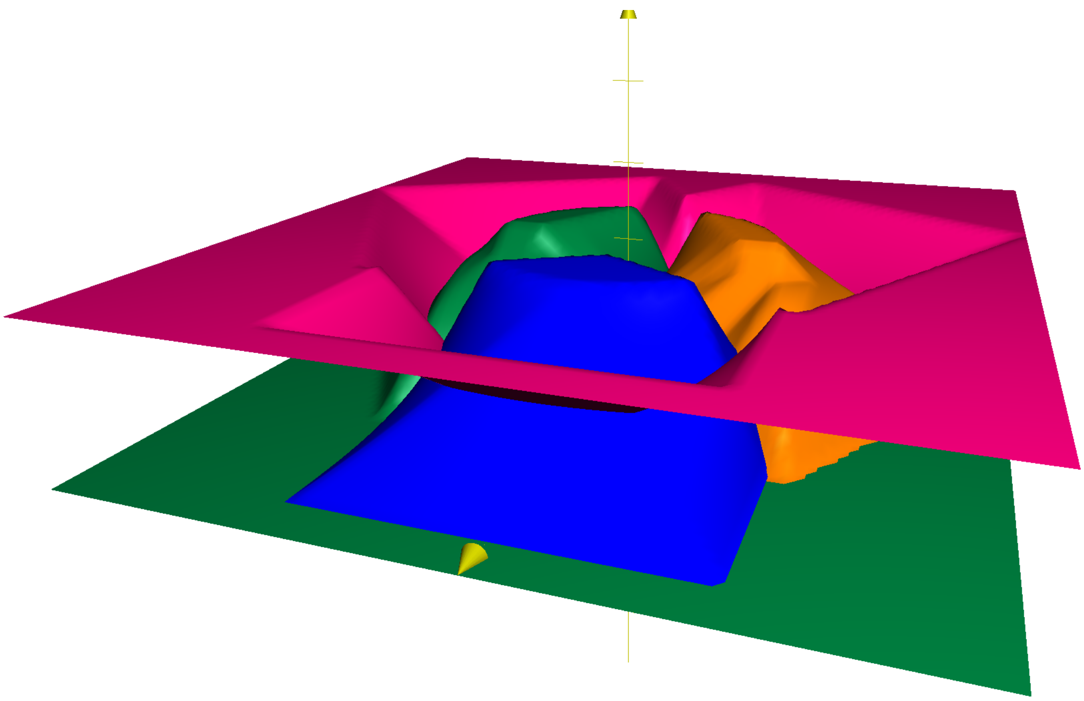

Figure 12. The whole partition of unity.

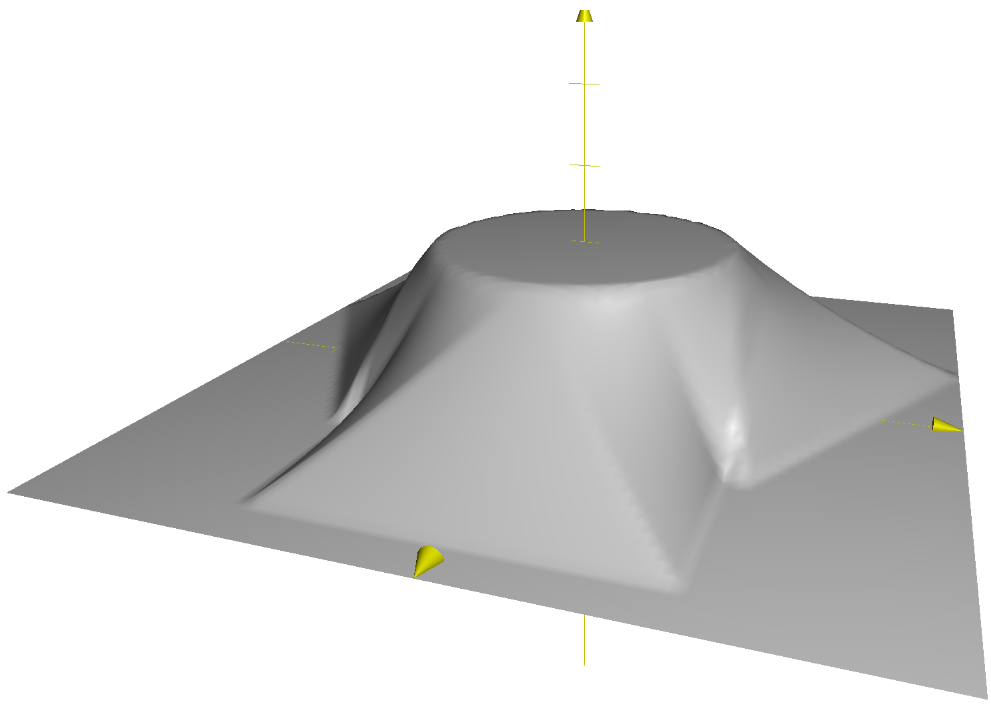

Figure 13. $\psi_1 + \psi_2 + \psi_3 = 1$ on Ω.

Corollary 2.3.1 *Let Ω be a bounded open set in $\mathbb{R}^d$ and u be a function on Ω belonging to one of the function spaces we have seen (and that we will see later). Let C_j be an open cover as in Proposition 2.3.1. Then we can write $u = \sum_{j=0}^{m} u_j$ with $\operatorname{supp} u_j \subset C_j$ and u_j has the same smoothness or integrability as u.*

Proof. We use the partition of unity ψ_j. Since $1 = \sum_{j=0}^{m} \psi_j$ on Ω, we can write

$$u = u \times 1 = u \sum_{j=0}^{m} \psi_j = \sum_{j=0}^{m} u \psi_j$$

and set $u_j = u\psi_j$. As $\operatorname{supp} \psi_j \subset C_j$, it follows that u_j vanishes outside of C_j, and since ψ_j is C^∞, u_j is as differentiable or as integrable as u is already. □

It is in this sense that partitions of unity are used to localize a function u. Such a function is decomposed into a sum of functions, each with support in a given open set of a covering. It is often easier to work with the localized parts u_j than with the function u itself. A prime example is integration by parts in the next section.

2.4 Integration by parts in dimension d and applications

Integration by parts in $\mathbb{R}^d$ is a basic formula, that is hardly ever entirely proved. It has to be said that the proof is not exactly a lot of fun. In what follows, Ω will be an at least Lipschitz open subset of $\mathbb{R}^d$. The most crucial integration by parts formula, from which all the others follow, is given in the next Theorem.

Theorem 2.4.1 *Let Ω be a Lipschitz open set in $\mathbb{R}^d$ and $u \in C^1(\bar{\Omega})$. Then we have*

$$\int_\Omega \frac{\partial u}{\partial x_i}\,dx = \int_{\partial\Omega} u n_i\,d\Gamma, \tag{2.2}$$

where n_i is the ith *component of the exterior unit normal vector.*

Proof. We will only write the proof in dimension $d = 2$, which is not a real restriction as the general case $d \geq 2$ follows from exactly the same arguments, and only in the case when Ω is of class C^1. This is a real restriction: there are additional technical difficulties in the Lipschitz case due to only almost everywhere differentiability.

We start with the partition of unity associated with the given covering C_j of the boundary completed by an open set C_0 to cover the interior. We have $u = \sum_{j=0}^m u_j$ with $u_j = u\psi_j$, and each u_j belongs to $C^1(\bar{\Omega})$ and has support in $\bar{C}_j$. Consequently, since formula (2.2) is linear with respect to u, it is sufficient to prove it for each u_j.

Let us start with the case $j = 0$. In this case, u_0 is compactly supported in Ω since $\bar{C}_0 \subset \Omega$. In particular, it vanishes on $\partial\Omega$, so that $\int_{\partial\Omega} u_0 n_i\,d\Gamma = 0$.

We extend u_0 by 0 to $\mathbb{R}^2$, thus yielding a $C^1(\mathbb{R}^2)$ function $\tilde{u}_0$. Since Ω is bounded, we choose a square that contains it, $\Omega \subset Q =]-M,M[^2$, for some M. Letting $i' = 1$ if $i = 2$, $i' = 2$ if $i = 1$, we obtain

$$\int_\Omega \frac{\partial u_0}{\partial x_i}\,dx = \int_Q \frac{\partial \tilde{u}_0}{\partial x_i}\,dx = \int_{-M}^{M}\Big(\int_{-M}^{M} \frac{\partial \tilde{u}_0}{\partial x_i}\,dx_i\Big)dx_{i'} = \int_{-M}^{M} [\tilde{u}_0]_{x_i=-M}^{x_i=M} dx_{i'} = 0,$$

by Fubini's theorem and the fact that $\tilde{u}_0 = 0$ sur ∂Q. Formula (2.2) is thus established for u_0.

The case $j > 0$ is a little more complicated. To simplify the notation, we omit all j indices and exponents. We thus have a function u with support in $\bar{C}$. In particular, $u = 0$ on $\partial C \cap \bar{\Omega}$. We recall that

$$\Omega \cap C = \{y \in C; y_2 < \varphi(y_1)\},$$

see also Figure 5. We first establish formula (2.2) in the (y_1, y_2) coordinate system in which $C =]-a,a[^2$ for some a. We let $n_{y,i}$, $i = 1,2$, denote the components of the normal vector in this coordinate system. There are two different computations depending on the coordinate under consideration.

Case $i = 1$. We first use Fubini's theorem

$$\int_\Omega \frac{\partial u}{\partial y_1}\,dy = \int_{-a}^{a}\Big(\int_{-a}^{\varphi(y_1)} \frac{\partial u}{\partial y_1}(y_1,y_2)\,dy_2\Big)\,dy_1,$$

see again Figure 5. Now it is well-known from elementary calculus that

$$\frac{d}{dy_1}\Big(\int_{-a}^{\varphi(y_1)} u(y_1,y_2)\,dy_2\Big) = \int_{-a}^{\varphi(y_1)} \frac{\partial u}{\partial y_1}(y_1,y_2)\,dy_2 + u(y_1,\varphi(y_1))\varphi'(y_1),$$

(this is where the fact that φ is C^1 intervenes and where it would be a little harder to have φ only Lipschitz). Consequently,

$$\int_\Omega \frac{\partial u}{\partial y_1}\,dy = \int_{-a}^{a}\frac{d}{dy_1}\Big(\int_{-a}^{\varphi(y_1)} u(y_1,y_2)\,dy_2\Big)\,dy_1 - \int_{-a}^{a} u(y_1,\varphi(y_1))\varphi'(y_1)\,dy_1.$$

In the first integral, we integrate a derivative, so that

$$\int_{-a}^{a}\frac{d}{dy_1}\Big(\int_{-a}^{\varphi(y_1)} u(y_1,y_2)\,dy_2\Big)\,dy_1$$
$$= \int_{-a}^{\varphi(a)} u(a,y_2)\,dy_2 - \int_{-a}^{\varphi(-a)} u(-a,y_2)\,dy_2 = 0,$$

since $u = 0$ on ∂C. We thus see that

$$\begin{aligned}\int_\Omega \frac{\partial u}{\partial y_1}\,dy &= \int_{-a}^{a} u(y_1,\varphi(y_1))\frac{-\varphi'(y_1)}{\sqrt{1+\varphi'(y_1)^2}}\sqrt{1+\varphi'(y_1)^2}\,dy_1\\ &= \int_{-a}^{a} u(y_1,\varphi(y_1))n_{y,1}(y_1)\sqrt{1+\varphi'(y_1)^2}\,dy_1\\ &= \int_{C\cap\partial\Omega} un_{y,1}\,d\Gamma,\end{aligned}$$

by the formulas established in Section 2.2 for the normal vector components and the boundary measure. Hence formula (2.2) in this case.

Case $i = 2$. We start again with Fubini's theorem

$$\begin{aligned}\int_\Omega \frac{\partial u}{\partial y_2}\,dy &= \int_{-a}^{a}\Big(\int_{-a}^{\varphi(y_1)} \frac{\partial u}{\partial y_2}(y_1,y_2)\,dy_2\Big)\,dy_1\\ &= \int_{-a}^{a} u(y_1,\varphi(y_1))\,dy_1\\ &= \int_{-a}^{a} u(y_1,\varphi(y_1))\frac{1}{\sqrt{1+\varphi'(y_1)^2}}\sqrt{1+\varphi'(y_1)^2}\,dy_1\\ &= \int_{C\cap\partial\Omega} un_{y,2}\,d\Gamma,\end{aligned}$$

since $u(y_1,-a)=0$. This proves the integration by parts formula in the (y_1,y_2) system attached to the cube C covering a part of the boundary.

We need to go back to the original coordinate system (x_1,x_2). Let us write the coordinate change formulas

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = R\begin{pmatrix} x_1-c_1 \\ x_2-c_2 \end{pmatrix} \text{ or again } \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = R^T\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},$$

where R is an orthogonal matrix and (c_1,c_2) are the (x_1,x_2) coordinates of the center of C. Similarly, if v_x and v_y denote the column-vectors of the components of the same vector $v\in\mathbb{R}^2$ in each of the coordinate systems, we have

$$v_y = Rv_x \Longleftrightarrow v_x = R^T v_y.$$

This is true in particular for the normal vecteur n, $n_x = R^T n_y$. Let us note $\nabla_x u$ and $\nabla_y u$ the components of the gradient of u in the two coordinate systems, we see by the chain rule that

$$(\nabla_x u)_i = \frac{\partial u}{\partial x_i} = \sum_{j=1}^{2} \frac{\partial u}{\partial y_j}\frac{\partial y_j}{\partial x_i} = \sum_{j=1}^{2} R_{ji}\frac{\partial u}{\partial y_j} = (R^T\nabla_y u)_i,$$

hence the final result by linearity of the integrals. □

Once the basic formula is available, a whole bunch of other formulas are derived pretty cheaply, that bear various names in the literature. We do not specify the regularity of the functions below, it is understood that they are sufficiently differentiable for all derivatives to make sense.

Corollary 2.4.2 *We have*

i) Integration by parts strictly speaking

$$\int_\Omega \frac{\partial u}{\partial x_i} v\,dx = -\int_\Omega u\frac{\partial v}{\partial x_i}\,dx + \int_{\partial\Omega} uvn_i\,d\Gamma, \tag{2.3}$$

ii) Green's formula

$$\int_\Omega (\Delta u)v\,dx = -\int_\Omega \nabla u\cdot\nabla v\,dx + \int_{\partial\Omega}\frac{\partial u}{\partial n} v\,d\Gamma, \tag{2.4}$$

where $\frac{\partial u}{\partial n} = \nabla u\cdot n = \sum_{i=1}^{d}\frac{\partial u}{\partial x_i} n_i$ *denotes the* normal derivative *of u on $\partial\Omega$.*

iii) A slightly more symmetrical version of Green's formula

$$\int_\Omega (\Delta u)v\,dx = \int_\Omega u(\Delta v)\,dx + \int_{\partial\Omega}\Big(\frac{\partial u}{\partial n}v - u\frac{\partial v}{\partial n}\Big)\,d\Gamma, \tag{2.5}$$

iv) Stokes formula

$$\int_\Omega \operatorname{div} U\, dx = \int_{\partial\Omega} U \cdot n\, d\Gamma, \tag{2.6}$$

where $U\colon \Omega \to \mathbb{R}^d$ *is a vector field, its divergence is* $\operatorname{div} U = \sum_{i=1}^d \frac{\partial U_i}{\partial x_i}$ *and* $U \cdot n = \sum_{i=1}^d U_i n_i$ *is the flux of the vector field through the boundary of* Ω.

Proof. For i), we apply the basic formula (2.2) to the product uv, and so on. □

2.5 Distributions

In this section, Ω is an arbitrary open subset of $\mathbb{R}^d$.

It turns out that functions that are differentiable in the classical sense are not sufficient to work with PDEs. A more general concept is needed, which is called *distributions*. As we will see, distributions are a lot more general than functions. They can always be differentiated indefinitely, even when they correspond to functions that are not differentiable in the classical sense, and their derivatives are distributions. This is why distributional solutions to linear PDEs of arbitrary order make sense (with technical conditions on their coefficients). We will also use distributions to define an important class of function spaces for PDEs, the Sobolev spaces.

Let us first go back to the space of indefinitely differentiable functions with compact support $\mathscr{D}(\Omega)$ encountered in Section 2.1. It is trivial, but crucial for the sequel, that the space $\mathscr{D}(\Omega)$ is stable by differentiation of arbitrary order, *i.e.*, if $\varphi \in \mathscr{D}(\Omega)$ then $\partial^\alpha \varphi \in \mathscr{D}(\Omega)$ for any multiindex α.

As mentioned before, this vector space has a natural topology that is a little difficult to understand (technically, it is an *LF*-space, a strict inductive limit of a sequence of Fréchet spaces and it is not metrizable) and it is not very useful to master the details of this topology for the applications we have in mind. So we will just skip it.

The convergence of a sequence in $\mathscr{D}(\Omega)$ is on the other hand quite easy to characterize.

Proposition 2.5.1 *A sequence* $\varphi_n \in \mathscr{D}(\Omega)$ *converges to* $\varphi \in \mathscr{D}(\Omega)$ *in the sense of* $\mathscr{D}(\Omega)$ *if and only if*

i) There exists a compact subset K *of* Ω *such that* $\operatorname{supp} \varphi_n \subset K$ *for all* n.

ii) For all $\alpha \in \mathbb{N}^d$, $\partial^\alpha \varphi_n \to \partial^\alpha \varphi$ *uniformly.*

When a (real or complex) vector space is equipped with a topology that makes its vector space operations continuous, that is when it is a *topological vector space*, it makes sense to look at the vector space of continuous linear forms, that is the space of real or complex valued linear mappings that are continuous for the aforementioned topology. This space is called the *topological dual*, or in short dual space.

Definition 2.5.1 *The* space of distributions on Ω, $\mathscr{D}'(\Omega)$, *is the dual of the space* $\mathscr{D}(\Omega)$.

We will indifferently use the notations $T(\varphi) = \langle T, \varphi \rangle$ to denote the value of a distribution T on a test-function φ and the duality pairing between the two. Of course, since $\mathscr{D}'(\Omega)$ is a vector space, we can add distributions and multiply them by a scalar in the obvious way.

Now, not knowing the topology of $\mathscr{D}(\Omega)$ makes it a little difficult to decide which linear forms on $\mathscr{D}(\Omega)$ are continuous and which are not. Fortunately, even though the topology in question is not metrizable, the usual sequential criterion happens to still work in this particular case.

Proposition 2.5.2 *A linear form T on $\mathscr{D}(\Omega)$ is a distribution if and only if we have $T(\varphi_n) \to 0$ for all sequences $\varphi_n \in \mathscr{D}(\Omega)$ such that $\varphi_n \to 0$ in the sense of $\mathscr{D}(\Omega)$.*

Proof. We admit Proposition 2.5.2. □

Remark 2.5.1 Let us note that the property $T(\varphi_n) \to 0$ for all sequences $\varphi_n \in \mathscr{D}(\Omega)$ such that $\varphi_n \to 0$ immediately implies that $T(\varphi_n) \to T(\varphi)$ for all sequences φ_n such that $\varphi_n \to \varphi$ in the sense of $\mathscr{D}(\Omega)$, by linearity, hence the sequential continuity of the linear form T. The difficulty is that in a non metrizable topological space, sequential continuity does not imply continuity in general, even though, in this particular case, it does. □

Let us now see in which sense distributions generalize the usual notion of function.

Proposition 2.5.3 *For all $f \in L^1_{\mathrm{loc}}(\Omega)$ there exists a distribution $\imath(f)$ on Ω defined by the formula*

$$\langle \imath(f), \varphi \rangle = \int_\Omega f \varphi \, dx$$

for all $\varphi \in \mathscr{D}(\Omega)$. The mapping $\imath \colon L^1_{\mathrm{loc}}(\Omega) \to \mathscr{D}'(\Omega)$ is one-to-one.

Proof. That the integral is well-defined has already been seen. It clearly defines a linear form on $\mathscr{D}(\Omega)$ by the linearity of the integral. What remains to be established for $\imath(f)$ to be a distribution, is its continuity. Let us thus be given a sequence $\varphi_n \to 0$ in the sense of $\mathscr{D}(\Omega)$, and K the associated compact set. We have

$$|\langle \imath(f), \varphi_n \rangle| = \Big|\int_\Omega f\varphi_n\,dx\Big| = \Big|\int_K f\varphi_n\,dx\Big|$$
$$\leq \int_K |f||\varphi_n|\,dx \leq \max_K |\varphi_n| \int_K |f|\,dx \to 0$$

since $f \in L^1(K)$ and φ_n tends to 0 uniformly on K.

Let us now show that the mapping $\imath$ is one-to-one. Since it is clearly linear, it suffices to show that its kernel is reduced to the zero vector. Let $f \in \ker \imath$, which means that $\imath(f)$ is the zero linear form, or in other words that $\int_\Omega f\varphi\,dx = 0$ for all φ in $\mathscr{D}(\Omega)$. By Proposition 2.1.1, it follows that $f = 0$, and the proof is complete. □

Remark 2.5.2 The mapping $\imath$ is not only one-to-one, it is also continuous (for the topology of $\mathscr{D}'(\Omega)$ as topological dual of $\mathscr{D}(\Omega)$, which we also keep shrouded in mystery). The mapping $\imath$ thus provides a faithful representation of one type of object, L^1_{loc} functions, as objects of a completely different nature, distributions. It is so faithful that in day-to-day practice, we say that an L^1_{loc} function f *is* a distribution and dispense with the notation $\imath$ altogether, that is we just write $\langle f, \varphi \rangle$ for the duality bracket.

Conversely, when a distribution T belongs to the image of $\imath$, that is to say when there exists f in L^1_{loc} such that $\langle T, \varphi \rangle = \int_\Omega f\varphi\,dx$ for all φ in $\mathscr{D}(\Omega)$, we just say that T *is* a function and we write $T = f$. Beware however that most distributions are not functions and that the notation $\int_\Omega T\varphi\,dx$ is *unacceptable* for such distributions.

Proposition 2.5.3 is all the more important as it shows that the elements of all the function spaces introduced up to now actually are distributions, since $L^1_{\text{loc}}(\Omega)$ is the largest of all such spaces. □

Remark 2.5.3 The characterization of convergence in $\mathscr{D}(\Omega)$ of Proposition 2.5.1 is implied by the topology of $\mathscr{D}(\Omega)$. In the proof of Proposition 2.5.3, we can see the importance of having a fixed compact K containing all the supports. If it was not the case, the final estimate would break down. □

Proposition 2.5.3 gives us our first examples of distributions. There are however many others which are not functions. Let us describe a couple of them.

We choose a point $a \in \Omega$ and define

$$\langle \delta_a, \varphi \rangle = \varphi(a)$$

for all $\varphi \in \mathscr{D}(\Omega)$. This is clearly a linear form on $\mathscr{D}(\Omega)$ and we just need to check its continuity. Let us thus be given again a sequence $\varphi_n \to 0$ in the sense of $\mathscr{D}(\Omega)$. In particular, it converges to 0 uniformly on Ω, hence pointwise. Therefore $\varphi_n(a) \to 0$ and we are done: $\delta_a \in \mathscr{D}'(\Omega)$. This distribution is called the *Dirac mass* or *Dirac distribution* at point a. It is an interesting exercise to show that it does not belong to the image of ι, *i.e.*, loosely speaking, that it is not a function. When $a = 0$, it is often simply denoted δ.

Let us give a second example with $\Omega = \mathbb{R}$. The function $x \mapsto 1/x$ almost everywhere is not in $L^1_{\text{loc}}(\mathbb{R})$ because it is not integrable in a neighborhood of 0. Therefore, it is not a distribution, which is rather unfortunate for such a simple function and a concept claiming to widely generalize functions. The distribution defined by

$$\left\langle \text{vp}\,\frac{1}{x}, \varphi \right\rangle = \lim_{\varepsilon \to 0^+} \left(\int_{-\infty}^{-\varepsilon} \frac{\varphi(x)}{x}\, dx + \int_{\varepsilon}^{+\infty} \frac{\varphi(x)}{x}\, dx \right)$$

is called the *principal value of* $1/x$ and replaces the function $x \mapsto 1/x$ for all intents and purposes (exercise: show that it is a distribution). It is however *not* a function.

We have hinted at a topology on the space of distributions. Here again, it is not too important to know the details of this topology. The convergence of sequences is more than enough and is surprisingly simple.

Proposition 2.5.4 *A sequence $T_n \in \mathscr{D}'(\Omega)$ converges to $T \in \mathscr{D}'(\Omega)$ in the sense of $\mathscr{D}'(\Omega)$ if and only if $\langle T_n, \varphi \rangle \to \langle T, \varphi \rangle$ for all $\varphi \in \mathscr{D}(\Omega)$.*

Proof. We admit Proposition 2.5.4. □

Since distributions are real-valued functions defined on the space $\mathscr{D}(\Omega)$, we see that convergence in the sense of distributions is actually nothing but simple or pointwise convergence on $\mathscr{D}(\Omega)$. This makes it very easy to handle (and unfortunately, very easy to abuse. Remember, it is not magic!).

This notion of convergence agrees with all previous notions defined on smaller function spaces. In particular, we have

Proposition 2.5.5 *Let $u_n \to u$ in $L^p(\Omega)$ for some $p \in [1, +\infty]$. Then $u_n \to u$ in the sense of $\mathscr{D}'(\Omega)$.*

Proof. For all $\varphi \in \mathscr{D}(\Omega)$, we have

$$|\langle u_n, \varphi \rangle - \langle u, \varphi \rangle| \le \int_\Omega |u_n - u||\varphi|\, dx \le \|u_n - u\|_{L^p(\Omega)} \|\varphi\|_{L^{p'}(\Omega)} \to 0$$

by Hölder's inequality. □

We have said earlier that distributions can be differentiated indefinitely, however in a specific sense.

Definition 2.5.2 *Let T be a distribution on Ω. The formula*

$$\langle S, \varphi \rangle = -\Big\langle T, \frac{\partial \varphi}{\partial x_i} \Big\rangle \tag{2.7}$$

for all $\varphi \in \mathscr{D}(\Omega)$, defines a distribution which is called the (distributional) partial derivative of T with respect to x_i and is denoted $\frac{\partial T}{\partial x_i}$.

Proof. This definition needs a proof. Formula (2.7) clearly defines a linear form on $\mathscr{D}(\Omega)$. Let us see that it is continuous. Let us be given a sequence $\varphi_n \to 0$ in $\mathscr{D}(\Omega)$. It is apparent that $\frac{\partial \varphi_n}{\partial x_i} \to 0$ in $\mathscr{D}(\Omega)$. Indeed, the support condition is the same, since the support of the partial derivative of a function is included in the support of this function, and the uniform convergence of all derivatives trivially holds true as all derivatives of $\frac{\partial \varphi_n}{\partial x_i}$ are derivatives of φ_n. Therefore,

$$\langle S, \varphi_n \rangle = -\Big\langle T, \frac{\partial \varphi_n}{\partial x_i} \Big\rangle \to 0$$

for all sequences $\varphi_n \to 0$ in $\mathscr{D}(\Omega)$. □

For example, the derivative of the Dirac mass δ in dimension one is the distribution

$$\langle \delta', \varphi \rangle = -\langle \delta, \varphi' \rangle = -\varphi'(0),$$

for all $\varphi \in \mathscr{D}(\mathbb{R})$.

The reason why it is reasonable to call this new distribution a partial derivative is in the next proposition.

Proposition 2.5.6 *Let u be a function in $C^1(\Omega)$. Then its distributional partial derivatives coincide with its classical partial derivatives.*

Proof. If $u \in C^1(\Omega)$ and $\varphi \in \mathscr{D}(\Omega)$, then it is clear that $u\varphi \in C^1(\bar{\Omega})$. An inspection of the proof[4] shows that this is enough to apply the integration by parts formula (2.3). Since $u\varphi$ vanishes on $\partial\Omega$, the result follows from Proposition 2.5.3. □

Remark 2.5.4 Be careful that the same result is false for functions that are only almost everywhere differentiable. Let us show an example, which is also a showcase example of how to compute a distributional derivative. Let H be the *Heaviside function* defined on $\mathbb{R}$ by $H(x) = 0$ for $x \leq 0$, $H(x) = 1$ for $x > 0$. This function is

[4]Hint: extend u in a $C^1(\bar{\Omega})$ fashion out of the support of φ.

classically differentiable with zero derivative for $x \neq 0$ and has a discontinuity of the first kind at $x = 0$. It is also in $L^\infty(\mathbb{R})$, hence in $L^1_{\text{loc}}(\mathbb{R})$, hence a distribution. Let us compute its distributional derivative. For all $\varphi \in \mathscr{D}(\mathbb{R})$, we have

$$\langle H', \varphi \rangle = -\langle H, \varphi' \rangle = -\int_0^{+\infty} \varphi'(s)\, ds = \varphi(0) - \lim_{x \to +\infty} \varphi(x) = \varphi(0),$$

since φ is compactly supported in $\mathbb{R}$, hence vanishes for x large enough. Therefore we see that

$$H' = \delta$$

even though the almost everywhere classical derivative of H is 0. This is an example of a function that is not differentiable in the classical sense, but that is also a distribution, hence has a distributional derivative and this derivative is not a function. The example also shows that H is a distributional primitive of the Dirac mass. □

Once a distribution is known to have partial derivatives of order one which are again distributions, it is obvious that the same operation can be repeated indefinitely and we have, for any distribution T and any multiindex α,

$$\langle \partial^\alpha T, \varphi \rangle = (-1)^{|\alpha|} \langle T, \partial^\alpha \varphi \rangle$$

for all $\varphi \in \mathscr{D}(\Omega)$, by induction on the length of α.

Differentiation in the sense of distributions is continuous, which is violently false in most function spaces. We just show here the sequential continuity, which is amply sufficient for the applications.

Proposition 2.5.7 *Let $T_n \to T$ in the sense of $\mathscr{D}'(\Omega)$. Then, for all multiindices α, we have $\partial^\alpha T_n \to \partial^\alpha T$ in the sense of $\mathscr{D}'(\Omega)$.*

Proof. For all $\varphi \in \mathscr{D}(\Omega)$, we have

$$\langle \partial^\alpha T_n, \varphi \rangle = (-1)^{|\alpha|} \langle T_n, \partial^\alpha \varphi \rangle \to (-1)^{|\alpha|} \langle T, \partial^\alpha \varphi \rangle = \langle \partial^\alpha T, \varphi \rangle,$$

hence the result. □

This continuity provides another reason why the partial derivative terminology is adequate for distributions. Indeed, it can be shown that $C^\infty(\Omega)$ functions are dense in $\mathscr{D}'(\Omega)$. For any distribution T, there exists a sequence of indefinitely differentiable functions ψ_n that tend to T in the sense of distributions. Therefore, their distributional partial derivatives of arbitrary order, which coincide with their classical partial derivatives, also converge in the sense of distributions. So the distributional partial derivatives of a distribution appear as distributional limits of approximating classical partial derivatives.

Many other operations usually performed on functions can be extended to distributions using the same transposition trick as for partial derivatives. Let us just mention here the multiplication by a smooth functions.

Definition 2.5.3 *Let T be a distribution on Ω and $f \in C^\infty(\Omega)$. The formula*

$$\langle fT, \varphi \rangle = \langle T, f\varphi \rangle \tag{2.8}$$

for all $\varphi \in \mathscr{D}(\Omega)$, defines a distribution.

We leave the easy proof as an exercise. Of course, when $T \in L^1_{\text{loc}}(\Omega)$, fT coincides with the classical pointwise product and the mapping $T \mapsto fT$ is sequentially continuous on $\mathscr{D}'(\Omega)$. Note that it is not possible to define such a product in all generality by a function that is less smooth than C^∞. In particular, there is no product of two distributions with the reasonable properties to be expected from a product—a famous theorem by L. Schwartz that limits the usefulness of general distributions in dealing with nonlinear PDEs.

The partial derivatives of a distribution multiplied by a smooth function follow the classical Leibniz rule.

Proposition 2.5.8 *Let T be a distribution on Ω and $f \in C^\infty(\Omega)$. For all multi-indices α such that $|\alpha| = 1$, we have*

$$\partial^\alpha(fT) = f\partial^\alpha T + \partial^\alpha f\, T. \tag{2.9}$$

Proof. We just use the definitions. For all $\varphi \in \mathscr{D}(\Omega)$, we have

$$\begin{aligned}\langle \partial^\alpha(fT), \varphi \rangle &= -\langle fT, \partial^\alpha \varphi \rangle = -\langle T, f\partial^\alpha \varphi \rangle = -\langle T, \partial^\alpha(f\varphi) \rangle + \langle T, \partial^\alpha f\, \varphi \rangle \\ &= \langle \partial^\alpha T, f\varphi \rangle + \langle \partial^\alpha f\, T, \varphi \rangle = \langle f\partial^\alpha T, \varphi \rangle + \langle \partial^\alpha f\, T, \varphi \rangle \\ &= \langle f\partial^\alpha T + \partial^\alpha f\, T, \varphi \rangle\end{aligned}$$

by the Leibniz formula for smooth functions. □

We conclude this very brief review of distribution theory with the following result.

Proposition 2.5.9 *Let Ω be a connected open set of $\mathbb{R}^d$ and T a distribution on Ω such that $\frac{\partial T}{\partial x_i} = 0$ for $i = 1, \dots, d$. Then, there exists a constant $c \in \mathbb{R}$ such that $T = c$.*

We omit the proof of this proposition which shows that distributions behave the same as functions when their gradient vanishes. There is nothing exotic added in this respect when generalizing from functions to distributions. Just note that $T = c$ means that for all $\varphi \in \mathscr{D}(\Omega)$, $\langle T, \varphi \rangle = c \int_\Omega \varphi\, dx$.

2.6 Sobolev spaces

We now introduce and briefly study an important class of function spaces for PDEs, the Sobolev spaces. As we have seen, every function in $L^p(\Omega)$ is actually a distribution, therefore it has distributional partial derivatives. In general, these derivatives are not functions, of course. There are however some functions whose distributional derivative also are functions, even though they may not differentiable in the classical sense. These are the functions we are going to be interested in.

Definition 2.6.1 *Let $m \in \mathbb{N}$ and $p \in [1,+\infty]$. We define the Sobolev space*

$$W^{m,p}(\Omega) = \{u \in L^p(\Omega); \partial^\alpha u \in L^p(\Omega) \text{ for all } \alpha \text{ such that } |\alpha| \leq m\}.$$

When $p = 2$, we use the notation $W^{m,2}(\Omega) = H^m(\Omega)$.

Note the special case $m = 0$, where $W^{0,p}(\Omega) = L^p(\Omega)$ and $H^0(\Omega) = L^2(\Omega)$. So the notation is hardly ever used for $m = 0$. In these notes, we will mainly use the $H^m(\Omega)$ spaces, with special emphasis on $H^1(\Omega)$. The natural Sobolev norms are as follows

$$\|u\|_{W^{m,p}(\Omega)} = \Big(\sum_{|\alpha|\leq m} \|\partial^\alpha u\|^p_{L^p(\Omega)}\Big)^{\frac{1}{p}}$$

for $p < +\infty$ and

$$\|u\|_{W^{m,\infty}(\Omega)} = \max_{|\alpha|\leq m} \|\partial^\alpha u\|_{L^\infty(\Omega)}.$$

In particular, for $p = 2$, we have

$$\|u\|_{H^m(\Omega)} = \Big(\sum_{|\alpha|\leq m} \|\partial^\alpha u\|^2_{L^2(\Omega)}\Big)^{\frac{1}{2}}.$$

This latter norm is clearly a prehilbertian norm associated with the scalar product

$$(u|v)_{H^m(\Omega)} = \sum_{|\alpha|\leq m} (\partial^\alpha u|\partial^\alpha u)_{L^2(\Omega)}.$$

The notations $\|u\|_{m,p}$ and $\|u\|_m$ for the $W^{m,p}$ and H^m norms are also encountered in the literature if the context is clear.

Remark 2.6.1 It follows from the definition that $W^{m+1,p}(\Omega) \subset W^{m,p}(\Omega)$ for all m, p. Moreover, if Ω is bounded $W^{m,p}(\Omega) \subset W^{m,q}(\Omega)$ whenever $q \leq p$. Also if Ω is bounded, we have $C^m(\bar{\Omega}) \subset W^{m,p}(\Omega)$. If Ω is Lipschitz, we have in addition that $C^m(\bar{\Omega})$ is dense in $W^{m,p}(\Omega)$ (we will prove it later on for $m = 1$, $p = 2$). Of course, there are functions in $W^{m,p}(\Omega)$ that are not of class C^m. For example, the function $x \mapsto x_+ = \max(x,0)$ is in $H^1(]-1,1[)$ since $(x_+)' = H(x)$ in the sense of $\mathscr{D}'$ (exercise), but it is not differentiable in the classical sense at $x = 0$.

Similarly, there are functions in L^p that are not in $W^{1,p}$, such as the Heaviside function H whose derivative is δ which is not a function. □

Theorem 2.6.1 *The spaces $W^{m,p}(\Omega)$ are Banach spaces. In particular, the spaces $H^m(\Omega)$ are Hilbert spaces.*

Proof. We need to show that $W^{m,p}(\Omega)$ is complete for its norm. Let us thus be given a Cauchy sequence $(u_n)_{n\in\mathbb{N}}$ in $W^{m,p}(\Omega)$. In view of the definition of the norm, it follows that for each multiindex α, $|\alpha| \le m$, the sequence of partial derivatives $\partial^\alpha u_n$ is a Cauchy sequence in $L^p(\Omega)$. We know that $L^p(\Omega)$ is complete, therefore there exists $g_\alpha \in L^p(\Omega)$ such that $\partial^\alpha u_n \to g_\alpha$ in $L^p(\Omega)$. By Proposition 2.5.5, it follows that $\partial^\alpha u_n \to g_\alpha$ in the sense of $\mathscr{D}'(\Omega)$. Now by Proposition 2.5.7, we also know that $\partial^\alpha u_n \to \partial^\alpha u$ in the sense of $\mathscr{D}'(\Omega)$, where $u = g_{(0,0,\dots,0)}$ is the limit of the sequence in $L^p(\Omega)$. Therefore $\partial^\alpha u = g_\alpha \in L^p(\Omega)$ since the space of distributions is separated and thus a converging sequence can only have one limit. This shows that u belongs to $W^{m,p}(\Omega)$ on the one hand, and that $u_n \to u$ in $W^{m,p}(\Omega)$ since

$$\|u_n - u\|^p_{W^{m,p}(\Omega)} = \sum_{|\alpha|\le m} \|\partial^\alpha u_n - \partial^\alpha u\|^p_{L^p(\Omega)} = \sum_{|\alpha|\le m} \|\partial^\alpha u_n - g_\alpha\|^p_{L^p(\Omega)} \to 0,$$

for $p < +\infty$ and the same for $p = +\infty$. Therefore $W^{m,p}(\Omega)$ is complete, and so is the proof. □

From now on, we will mostly consider the case $p = 2$. Let us introduce an important subset of $H^m(\Omega)$.

Definition 2.6.2 *The closure of $\mathscr{D}(\Omega)$ in $H^m(\Omega)$ is denoted $H_0^m(\Omega)$.*

In other words, $H_0^m(\Omega)$ consists exactly of those functions u of $H^m(\Omega)$ which can be approximated in the sense of $H^m(\Omega)$ by indefinitely differentiable functions with compact support, *i.e.*, such that there exists a sequence $\varphi_n \in \mathscr{D}(\Omega)$ with $\|\varphi_n - u\|_{H^m(\Omega)} \to 0$. By definition, it is a closed vector subspace of $H^m(\Omega)$ and thus a Hilbert space for the scalar product of $H^m(\Omega)$.

The following is a very important result. We introduce the semi-norm

$$|u|_{H^m(\Omega)} = \Big(\sum_{|\alpha|=m} \|\partial^\alpha u\|^2_{L^2(\Omega)}\Big)^{\frac{1}{2}}.$$

This semi-norm just keep the partial derivatives of the higher order compared with the norm.

Theorem 2.6.2 (Poincaré's inequality) *Let Ω be a bounded open subset of $\mathbb{R}^d$. There exists a constant C which only depends on Ω such that for all $u \in H_0^1(\Omega)$,*

$$\|u\|_{L^2(\Omega)} \le C|u|_{H^1(\Omega)}.$$

Proof. Since Ω is assumed to be bounded, it is included in a strip[5] that we may assume to be of the form

$$\Omega \subset S_{a,b} = \{(x', x_d); x' \in \mathbb{R}^{d-1}, a < x_d < b\}$$

for some a and b, without loss of generality.

We argue by density. First let $\varphi \in \mathscr{D}(\Omega)$. We extend it by 0 to the whole of $\mathbb{R}^d$ and still call the extension φ. Let $\alpha_d = (0, 0, \dots, 0, 1)$ so that $\partial^{\alpha_d}\varphi = \frac{\partial \varphi}{\partial x_d}$. Since $\varphi(x', a) = 0$ for all $x' \in \mathbb{R}^{d-1}$ and φ is C^1 with respect to x_d, we can write

$$\varphi(x', x_d) = \int_a^{x_d} \partial^{\alpha_d}\varphi(x', s)\, ds$$

for all (x', x_d). In particular, for $a \le x_d \le b$, we obtain

$$\varphi(x', x_d)^2 \le (x_d - a)\int_a^{x_d} \big(\partial^{\alpha_d}\varphi(x', s)\big)^2 ds \le (b - a)\int_a^{b} \big(\partial^{\alpha_d}\varphi(x', s)\big)^2 ds$$

by the Cauchy-Schwarz inequality. We integrate the above inequality with respect to x'

$$\int_{\mathbb{R}^{d-1}} \varphi(x', x_d)^2\, dx' \le (b - a)\int_{S_{a,b}} \big(\partial^{\alpha_d}\varphi(x)\big)^2 dx$$

by Fubini's theorem. Now, because the support of φ is included in $\Omega \subset S_{a,b}$, it follows that

$$\int_{S_{a,b}} \big(\partial^{\alpha_d}\varphi(x)\big)^2 dx = \|\partial^{\alpha_d}\varphi\|^2_{L^2(\Omega)}.$$

We integrate again with respect to x_d between a and b and obtain

$$\|\varphi\|^2_{L^2(\Omega)} \le (b - a)^2 \|\partial^{\alpha_d}\varphi\|^2_{L^2(\Omega)},$$

for the same reasons (Fubini and support of φ). Now by definition of the semi-norm, it follows that

$$\|\partial^{\alpha_d}\varphi\|^2_{L^2(\Omega)} \le \sum_{|\alpha|=1} \|\partial^{\alpha}\varphi\|^2_{L^2(\Omega)} = |\varphi|^2_{H^1(\Omega)},$$

hence Poincaré's inequality for a function $\varphi \in \mathscr{D}(\Omega)$ with constant $C = (b - a)$.

We complete the proof by a density argument. Let $u \in H^1_0(\Omega)$. By definition, there exists a sequence $\varphi_n \in \mathscr{D}(\Omega)$ such that $\varphi_n \to u$ in $H^1(\Omega)$. Inspection of the definition of the H^1 norm reveals that this is equivalent to $\varphi_n \to u$ in $L^2(\Omega)$ and $\partial^\alpha \varphi_n \to \partial^\alpha u$ for all $|\alpha| = 1$ also in $L^2(\Omega)$. Since all the L^2 norms then converge, we obtain in the limit

$$\|u\|_{L^2(\Omega)} \le (b - a)|u|_{H^1(\Omega)},$$

which is Poincaré's inequality on $H^1_0(\Omega)$. □

[5] It is enough for Poincaré's inequality to be valid that Ω be included in such a strip but not necessarily bounded.

Remark 2.6.2 Poincaré's inequality shows that $H_0^1(\Omega)$ is a strict subspace of $H^1(\Omega)$ when Ω is bounded. Indeed, the constant function $u = 1$ is in $H^1(\Omega)$ but does not satisfy the inequality, since all its partial derivatives vanish. It follows that it is impossible to approximate a non zero constant by a sequence in $\mathscr{D}(\Omega)$ in the norm of $H^1(\Omega)$. □

From now on, we will use the gradient notation ∇u to denote the vector of all first order distributional partial derivatives of u. Individual first order derivatives will be denoted $\partial_i = \frac{\partial}{\partial x_i}$ instead of with the multiindex notation, and second order derivatives $\partial_{ij} = \frac{\partial^2}{\partial x_i \partial x_j}$. When $u \in H^1(\Omega)$, then we have $\nabla u \in L^2(\Omega;\mathbb{R}^d)$. Poincaré's inequality has an important corollary.

Corollary 2.6.3 *Let Ω be a bounded subset of $\mathbb{R}^d$. The H^1 semi-norm $|\cdot|_{H^1(\Omega)}$ is a norm on $H_0^1(\Omega)$ that is equivalent to the H^1 norm. It is also a hilbertian norm associated with the scalar product*

$$(u|v)_{H_0^1(\Omega)} = \int_\Omega \nabla u \cdot \nabla v \, dx.$$

Proof. First of all, it is clear that $|u|_{H^1(\Omega)} \leq \|u\|_{H^1(\Omega)}$ for all $u \in H^1(\Omega)$, hence all $u \in H_0^1(\Omega)$, since the norm squared is the semi-norm squared plus the L^2 norm squared.

The converse inequality follows from Poincaré's inequality. Indeed, for $u \in H_0^1(\Omega)$, we have $\|u\|_{L^2(\Omega)} \leq C|u|_{H^1(\Omega)}$. Therefore

$$\|u\|_{H^1(\Omega)} = (\|u\|_{L^2(\Omega)}^2 + |u|_{H^1(\Omega)}^2)^{\frac{1}{2}} \leq (C^2+1)^{\frac{1}{2}} |u|_{H^1(\Omega)}.$$

This shows both that the semi-norm is a norm on $H_0^1(\Omega)$ and that it is equivalent to the full H^1 norm on $H_0^1(\Omega)$. This also shows that the bilinear form above is positive definite, hence a scalar product. □

Remark 2.6.3 The fact that the two norms are equivalent implies that $H_0^1(\Omega)$ is also complete for the semi-norm. Hence, it is also a Hilbert space for the scalar product corresponding to the semi-norm. Beware however that this is a different Hilbert structure from the one obtained by restricting the H^1 scalar product to H_0^1. Indeed, we now have two different notions of orthogonality, and (at least) two different ways of identifying the dual of H_0^1. □

Remark 2.6.4 The above results generalize to $H_0^m(\Omega)$ on which the semi-norm $|\cdot|_{H^m(\Omega)}$ is equivalent to the full H^m norm. They also generalize to the spaces $W_0^{m,p}(\Omega)$ defined in a obvious way. □

Let us give yet another way of identifying the dual of $H_0^1(\Omega)$.

Definition 2.6.3 *Let*

$$H^{-1}(\Omega) = \{T \in \mathscr{D}'(\Omega); \exists C, \forall \varphi \in \mathscr{D}(\Omega), |\langle T, \varphi\rangle| \leq C|\varphi|_{H^1(\Omega)}\}, \qquad (2.10)$$

equipped with the norm

$$\|T\|_{H^{-1}(\Omega)} = \inf\{C \text{ appearing in formula (2.10)}\} = \sup \frac{|\langle T, \varphi\rangle|}{|\varphi|_{H^1(\Omega)}}.$$

Then $H^{-1}(\Omega)$ is isometrically isomorphic to $(H_0^1(\Omega))'$.

Proof. Since $\mathscr{D}(\Omega) \subset H_0^1(\Omega)$ by definition, any linear form ℓ on $H_0^1(\Omega)$ defines a linear form on $\mathscr{D}(\Omega)$ by restriction. Moreover, if $\varphi_n \to 0$ in $\mathscr{D}(\Omega)$, we obviously have $\varphi_n \to 0$ in $H_0^1(\Omega)$ as well. Hence, if ℓ is continuous, that is $\ell \in (H_0^1(\Omega))'$, its restriction to $\mathscr{D}(\Omega)$ is a distribution $T \in \mathscr{D}'(\Omega)$. This distribution trivially belongs to $H^{-1}(\Omega)$.

Conversely, let us be given an element T of $H^{-1}(\Omega)$. By definition, it is a linear form defined on a dense subspace of $H_0^1(\Omega)$ and continuous with respect to the H_0^1-norm. Therefore, it extends to an element ℓ of the dual space $(H_0^1(\Omega))'$, with the same norm. □

In order to explain de -1 exponent in the notation, we note the following.

Proposition 2.6.1 *Let $f \in L^2(\Omega)$, then $\partial_i f \in H^{-1}(\Omega)$ and $\langle \partial_i f, \varphi\rangle = -\int_\Omega f \partial_i \varphi \, dx$ for all $\varphi \in \mathscr{D}(\Omega)$.*

Proof. By definition of distributional derivatives, $\langle \partial_i f, \varphi\rangle = -\langle f, \partial_i \varphi\rangle = -\int_\Omega f \partial_i \varphi \, dx$ since f is locally integrable. Thus

$$|\langle \partial_i f, \varphi\rangle| \leq \|f\|_{L^2(\Omega)} \|\partial_i \varphi\|_{L^2(\Omega)} \leq \|f\|_{L^2(\Omega)} |\varphi|_{H_0^1(\Omega)},$$

by Cauchy-Schwarz, hence $\partial_i f \in H^{-1}(\Omega)$ with $\|\partial_i f\|_{H^{-1}(\Omega)} \leq \|f\|_{L^2(\Omega)}$. □

Remark 2.6.5 This shows that the operator ∂_i is linear continuous from $L^2(\Omega)$ $(= H^0(\Omega))$ into $H^{-1}(\Omega)$, just as it is linear continuous from $H^1(\Omega)$ into $L^2(\Omega)$. Conversely, when Ω is regular, a distribution in $H^{-1}(\Omega)$ whose first order partial derivatives are all in $H^{-1}(\Omega)$ is in fact a function in $L^2(\Omega)$. The latter result is known as Lions's lemma. □

In the same vein:

Corollary 2.6.4 *The operator $-\Delta$ is linear continuous from $H_0^1(\Omega)$ into $H^{-1}(\Omega)$ and for all $u, v \in H_0^1(\Omega)$, we have*

$$\langle -\Delta u, v\rangle = \int_\Omega \nabla u \cdot \nabla v \, dx.$$

The dual of $H_0^m(\Omega)$ is likewise identified with a subspace $H^{-m}(\Omega)$ of $\mathscr{D}'(\Omega)$.

2.7 Sobolev spaces in one dimension

For simplicity, we mostly consider $H^1(\Omega)$ where $\Omega =]a,b[$ is a bounded open interval of $\mathbb{R}$. Let us admit a density result that we will prove later in arbitrary dimension.

Proposition 2.7.1 *The space $C^1([a,b])$ is dense in $H^1(]a,b[)$.*

We have already seen examples of functions in dimension one that are H^1 but not C^1. All one-dimensional H^1 functions however are continuous, in the sense that each equivalence class contains one continuous representative. The density above is also meant in the sense that the equivalence classes of elements of $C^1([a,b])$ are dense in $H^1(]a,b[)$. There is even a more precise embedding.

Theorem 2.7.1 *We have that $H^1(]a,b[) \hookrightarrow C^{0,1/2}([a,b])$.*

Recall that the hooked arrow means that there is an injection between the two spaces and that this injection is continuous.
Proof. We argue by density again. Let $v \in C^1([a,b])$. For all $x, y \in [a,b]$, we can write

$$v(y) - v(x) = \int_x^y v'(s)\,ds. \tag{2.11}$$

Passing $v(x)$ on the right-hand side in equation (2.11) and taking the square, we obtain

$$\begin{aligned} v(y)^2 &= \Big(v(x) + \int_x^y v'(s)\,ds\Big)^2 \\ &\le 2\Big[v(x)^2 + \Big(\int_x^y v'(s)\,ds\Big)^2\Big] \\ &\le 2\Big[v(x)^2 + |y-x|\int_x^y (v'(s))^2\,ds\Big] \\ &\le 2\Big[v(x)^2 + (b-a)\int_a^b (v'(s))^2\,ds\Big], \end{aligned}$$

by the Cauchy-Schwarz inequality for the third line. Integrating with respect to x, we obtain

$$v(y)^2 \le \frac{2}{b-a}\|v\|^2_{L^2(a,b)} + 2(b-a)\|v'\|^2_{L^2(a,b)} \le C^2\|v\|^2_{H^1(]a,b[)}$$

with $C = \sqrt{\max\big(\frac{2}{b-a}, 2(b-a)\big)}$. Since this is true for all $y \in [a,b]$, it follows that

$$\max_{y\in[a,b]} |v(y)| = \|v\|_{C^0([a,b])} \le C\|v\|_{H^1(]a,b[)}, \tag{2.12}$$

for all $v \in C^1([a,b])$.

We now go back to equation (2.11) and square it.

$$(v(y)-v(x))^2 = \left(\int_x^y v'(s)\,ds\right)^2 \le |y-x| \int_a^b (v'(s))^2\,ds \le |y-x| \|v\|^2_{H^1(]a,b[)},$$

by Cauchy-Schwarz again. Therefore, for all $x \neq y$, we obtain

$$\frac{|v(y)-v(x)|}{|y-x|^{1/2}} \le \|v\|_{H^1(]a,b[)}.$$

Now the right-hand side above does not depend on x or y, thus

$$\sup_{x\neq y} \frac{|v(y)-v(x)|}{|y-x|^{1/2}} \le \|v\|_{H^1(]a,b[)}, \tag{2.13}$$

for all $v \in C^1([a,b])$.

Putting estimates (2.12) and (2.13) together, we see that for all $v \in C^1([a,b])$, we have

$$\|v\|_{C^{0,1/2}([a,b])} \le C'\|v\|_{H^1(]a,b[)}, \tag{2.14}$$

with $C' = C+1$.

Let us now conclude the density argument. For all $u \in H^1(]a,b[)$ there exists a sequence $v_n \in C^1([a,b])$ such that $v_n \to u$ in $H^1(]a,b[)$. It is therefore a Cauchy sequence in $H^1(]a,b[)$, and applying estimate (2.14) to $(v_n - v_m) \in C^1([a,b])$, we see that it is also a Cauchy sequence in $C^{0,1/2}([a,b])$. But $C^{0,1/2}([a,b])$ is a Banach space, thus there exists $v \in C^{0,1/2}([a,b])$ such that $v_n \to v$ in $C^{0,1/2}([a,b])$.

Of course, $v_n \to u$ in $H^1(]a,b[)$ implies $v_n \to u$ in $\mathscr{D}'(]a,b[)$, and $v_n \to v$ in $C^{0,1/2}([a,b])$ implies $v_n \to v$ in $\mathscr{D}'(]a,b[)$, so that $u = v$.[6] This shows that $u \in C^{0,1/2}([a,b])$ and, passing to the limit in estimate (2.14), that the injection is continuous. □

That all H^1 functions are continuous is specific to dimension one, as we will see later. Note also that not all $C^{0,1/2}$ functions are H^1 (consider $x \mapsto \sqrt{x}$ on $]0,1[$). The injection above is nonetheless optimal since for each $\beta > 1/2$, there is an H^1 function that is not $C^{0,\beta}$ (consider $x \mapsto x^{\frac{2\beta+1}{4}}$ on $]0,1[$).

An important feature of the one-dimensional case is that pointwise values of a H^1 function are unambiguously defined as the pointwise value of its continuous representative. Moreover, such pointwise values depend continuously on the function in the H^1 norm by estimate (2.14). This is in particular true of the endpoint values at a and b, which can be surprising because the Sobolev space definition is based on the open set $]a,b[$.

[6]This is in the sense that the equivalence class u contains one (and only one) continuous representative, which is v.

Corollary 2.7.2 *The linear mapping $H^1(]a,b[) \to \mathbb{R}^2$, $u \mapsto (u(a),u(b))$ is continuous.*

Proof. Obviously $\max(|u(a)|,|u(b)|) \leq \|u\|_{C^{0,1/2}([a,b])} \leq C\|u\|_{H^1(]a,b[)}$. □

Remark 2.7.1 This is the one-dimensional version of the *trace theorem* that we will prove in all dimensions later on. The linear mapping in question is called the *trace mapping*. The result also shows that Dirichlet boundary conditions make sense for functions of $H^1(]a,b[)$, a fact that was not evident from the start.

Because of the continuity of the trace, it is clear that $H^1_0(]a,b[)$ is included in the kernel of the trace $\{u \in H^1(]a,b[); u(a) = u(b) = 0\}$. It suffices to take a sequence φ_n of $\mathscr{D}(]a,b[)$ that tends to u in $H^1(]a,b[)$. Actually, the reverse inclusion holds true so that

$$H^1_0(]a,b[) = \{u \in H^1(]a,b[); u(a) = u(b) = 0\}.$$

The space $H^1_0(]a,b[)$ is thus adequate for homogeneous Dirichlet conditions for second order boundary value problems. □

To conclude the one-dimensional case, let us mention the Rellich compact embedding theorem.

Theorem 2.7.3 *The injection $H^1(]a,b[) \to L^2(a,b)$, $u \mapsto u$ is compact.*

Proof. We recall that a mapping is compact if it transforms bounded sets into relatively compact sets. Here, it is enough to take the unit ball of $H^1(]a,b[)$ by linearity. By estimate (2.14), it is a bounded subset of $C^{0,1/2}([a,b])$. Bounded sets of $C^{0,1/2}([a,b])$ are equicontinuous, therefore relatively compact in $C^0([a,b])$ by Ascoli's theorem. Finally the embedding $C^0([a,b]) \hookrightarrow L^2(a,b)$ is continuous, thus transforms relatively compact sets into relatively compact sets. □

Remark 2.7.2 We also have $H^m(]a,b[) \hookrightarrow C^{m-1,1/2}([a,b])$, the trace on $H^m(]a,b[)$ is $u \mapsto (u(a),u'(a),\dots,u^{(m-1)}(a),u(b),u'(b),\dots,u^{(m-1)}(b))$ and $H^m_0(]a,b[)$ is the set of u such that $u(a) = u'(a) = \dots = u^{(m-1)}(a) = u(b) = u'(b) = \dots = u^{(m-1)}(b) = 0$. Similar results can be written for the $W^{m,p}(]a,b[)$ spaces, not with the same Hölder exponent, though (exercise). □

2.8 Density of smooth functions and trace in dimension d

We have seen that Sobolev functions in dimension one are continuous. This is no longer true in dimensions 2 and higher. We will concentrate on the space $H^1(\Omega)$.

Let D be the unit disk in $\mathbb{R}^2$. It can be checked (exercise) that the function $u\colon x \mapsto \ln(|\ln(\|x\|/e)|)$ is in $H_0^1(D)$. This function tends to $+\infty$ at the origin, thus there is no continuous function in its equivalence class.

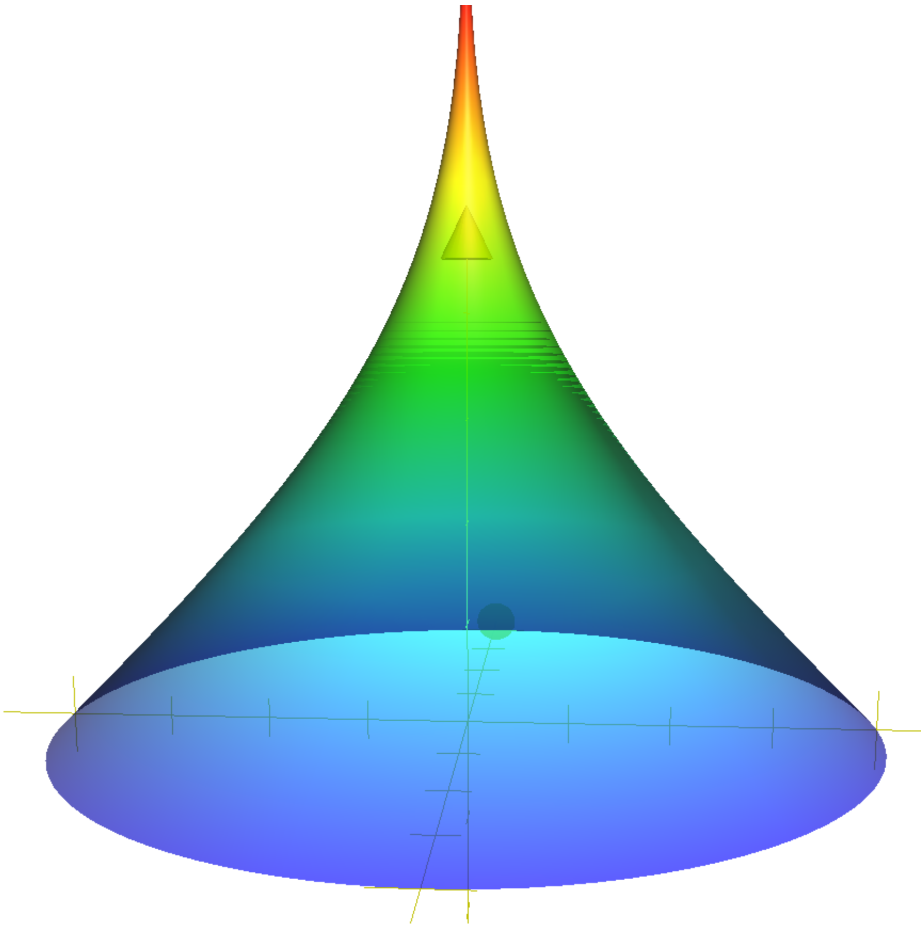

Figure 13. A discontinuous H^1-function.

Now we can do much worse! We extend u by 0 to $\mathbb{R}^2$, which still is a function in $H^1(\mathbb{R}^2)$. Next, let $(x_i)_{i\in\mathbb{N}}$ be countable, dense set of points in $\mathbb{R}^2$. Then the function $v(x) = \sum_{i=0}^{+\infty} 2^{-i} u(x - x_i)$ is in $H^1(\mathbb{R}^2)$, since $\|u(\cdot - x_i)\|_{H^1(\mathbb{R}^2)} = \|u\|_{H^1(\mathbb{R}^2)}$ and we have a normally convergent series, but this function tends to $+\infty$ at all points x_i, which are dense. Therefore, v is not locally bounded: there is no open set on which it is bounded. This sounds pretty bad, even though it is a perfectly legitimate, although hard to mentally picture, function of $H^1(\mathbb{R}^2)$.

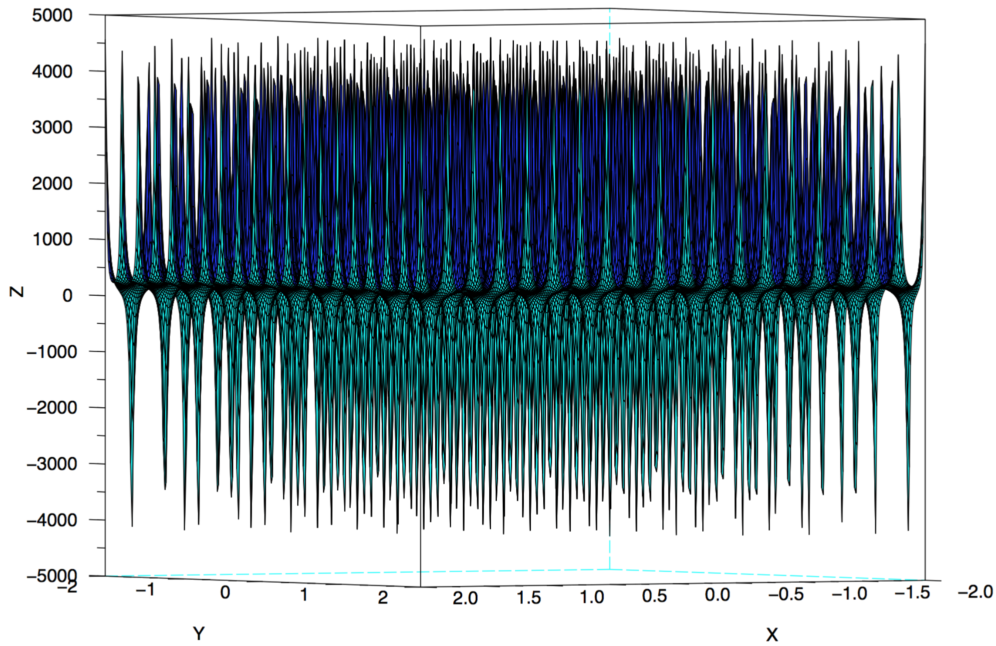

Figure 14. An attempt to draw a very bad H^1-function (graphics cheat: the spikes should be thinner, (infinitely) higher and (infinitely) denser).

In higher dimensions, we can picture such singularities occurring on a dense set of curves or hypersurfaces of dimension $d-2$. In view of this state of things, ascribing some kind of boundary value to a H^1 function that would be a reasonably continuous extension from Ω seems difficult. In PDE problems, we nonetheless need boundary values, to write Dirichlet conditions for example.

The definition of a good boundary value for H^1 functions is by means of a mapping called the *trace mapping*. This mapping is defined by density of smooth functions, so let us deal with that first. Besides, as should already be quite clear, density arguments are very useful in Sobolev spaces.

Theorem 2.8.1 *Let Ω be a Lipschitz open subset of $\mathbb{R}^d$. Then the space $C^1(\bar{\Omega})$ is dense in $H^1(\Omega)$.*

Proof. We use the same partition of unity as before. It suffices to construct a $C^1(\bar{\Omega})$ approximation for each part $u_j = \psi_j u$ of u. Indeed, we have $u_j \in H^1(\Omega)$ for all j by Proposition 2.5.8.

We start with the case $j = 0$. Since u_0 is compactly supported in Ω, its extension by 0 to the whole of $\mathbb{R}^d$ belongs to $H^1(\mathbb{R}^d)$ as is easily checked. Let us take a mollifier ρ, that is to say a C^∞ function with compact support in the unit ball B and such that $\int_B \rho(y)\,dy = 1$.

For all integers $n \geq 1$, we set $\rho_n(y) = n^d \rho(ny)$ and $u_{0,n} = \rho_n \star u_0$, where the star denotes the convolution. By the general properties of convolution, we have $u_{0,n} \in C^\infty(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ and $u_{0,n} \to u_0$ in $L^2(\mathbb{R}^d)$ when $n \to +\infty$. Moreover, since $\partial_i u_{0,n} = \rho_n \star \partial_i u_0$, the same argument shows that $\partial_i u_{0,n} \to \partial_i u_0$ in $L^2(\mathbb{R}^d)$, whence $u_{0,n} \to u_0$ in $H^1(\mathbb{R}^d)$ when $n \to +\infty$. This settles the case $j = 0$ because the restriction of $u_{0,n}$ to $\bar{\Omega}$ is of class C^1 (in fact, it is even compactly supported as soon as n is large enough) and the H^1 norm on Ω is smaller than the H^1 norm on $\mathbb{R}^d$.

The regularity of Ω comes into play for $j > 0$, in the hypercubes that cover the boundary. We drop again all subscripts or superscripts j for brevity. The difficulty compared with $j = 0$ is that we cannot extend u by 0 to $\mathbb{R}^d$ and remain in $H^1(\mathbb{R}^d)$. For example, it is easy to see that the function equal to 1 in Ω and 0 outside is not in $H^1(\mathbb{R}^d)$. We will use a two step process, first a translation, then a convolution.

Let $n \in \mathbb{N}^*$. We set $u_n(y) = u(y', y_d - 1/n)$, which is a function defined on the translated set $\Omega_n = \{y \in \mathbb{R}^d; (y', y_d - 1/n) \in \Omega \cap C\}$, see Figure 14 below. We extend u_n by 0 to $\mathbb{R}^d$ and let $\widetilde{u}_n$ denote this extension. Since u is supported in C, and the translation shifts it upwards, the restriction of $\widetilde{u}_n$ to $\Omega \cap C$ is still in $H^1(\Omega \cap C)$.

It can be shown[7] that the translation is continuous on $L^2(\mathbb{R}^d)$ in the sense that $\widetilde{u}_n \to \widetilde{u}$ in $L^2(\mathbb{R}^d)$ when $n \to +\infty$, thus by restriction we have $\widetilde{u}_{n|\Omega\cap C} \to u_{|\Omega\cap C}$ in $L^2(\Omega \cap C)$. Computing the partial derivatives in the sense of distributions shows that $\partial_i(\widetilde{u}_{n|\Omega\cap C}) = (\partial_i u)_{n|\Omega\cap C}$ using the same notation for the translation. Therefore, we have the same convergence for the partial derivatives, which shows that $\widetilde{u}_{n|\Omega\cap C} \to u_{|\Omega\cap C}$ in $H^1(\Omega \cap C)$ when $n \to +\infty$. To conclude, we just need to approximate $\widetilde{u}_{n|\Omega\cap C}$ for any given n by a $C^1(\bar{\Omega})$ function, and use a double limit argument.

We now use the convolution by a mollifier again and set $\widetilde{u}_{n,p} = \widetilde{u}_n \star \rho_p$. By construction, $\widetilde{u}_{n,p}$ is of class C^∞ on $\mathbb{R}^d$ and $\widetilde{u}_{n,p} \to \widetilde{u}_n$ in $L^2(\mathbb{R}^d)$ when $p \to +\infty$. Now for the subtle point. We do not have L^2 convergence of the gradients, because in general $\partial_i \widetilde{u}_n$ is not a function, let alone in $L^2(\mathbb{R}^d)$. Take for example $\varphi = 0$ and $u = 1$, then $\partial_d \widetilde{u}_n$ is a Dirac mass on the hyperplane $y_d = \frac{1}{n}$, *cf.* the one-dimensional case. However, since we have shifted the discontinuity outside of Ω by the translation, there is hope that the restrictions to Ω still converge.

To see that this is the case, we let $\widetilde{\partial_i u_n}$ denote the extension of $\partial_i u_n$ to $\mathbb{R}^d$ by 0. We have $\widetilde{\partial_i u_n} \in L^2(\mathbb{R}^d)$ and $\widetilde{\partial_i u_n} \star \rho_p \to \widetilde{\partial_i u_n}$ in $L^2(\mathbb{R}^d)$ when $p \to +\infty$ by the properties of convolution again. Of course, as already noted, $\widetilde{\partial_i u_n} \neq \partial_i \widetilde{u}_n$ so that

[7] The fairly easy proof uses the density of continuous, compactly supported functions in $L^2(\mathbb{R}^d)$.

$\widetilde{\partial_i u_n} \star \rho_p \neq \partial_i \widetilde{u}_{n,p}$. We will show that, for p large enough, we nonetheless have $(\widetilde{\partial_i u_n} \star \rho_p)_{|\Omega\cap C} = (\partial_i \widetilde{u}_{n,p})_{|\Omega\cap C}$. As we have just seen that $\widetilde{\partial_i u_n} \star \rho_p$ converges in L^2, this will lead to the conclusion that $(\partial_i \widetilde{u}_{n,p})_{|\Omega\cap C} \to \partial_i u_n$ in $L^2(\Omega\cap C)$, hence $\widetilde{u}_{n,p|\Omega\cap C} \to u_{n|\Omega\cap C}$ in $H^1(\Omega\cap C)$ when $p \to +\infty$. Since $\widetilde{u}_{n,p|\overline{\Omega\cap C}} \in C^1(\overline{\Omega\cap C})$, we will have our approximation.

To show that the restrictions are equal, we go back to the convolution formula

$$\widetilde{\partial_i u_n} \star \rho_p(x) = \int_{\mathbb{R}^d} \rho_p(x-y)\widetilde{\partial_i u_n}(y)\,dy = \int_{B(x,1/p)} \rho_p(x-y)\widetilde{\partial_i u_n}(y)\,dy$$

since ρ has support in the unit ball. Now, if for all $x \in \Omega\cap C$, we had $B\big(x, \frac{1}{p}\big) \subset \Omega_n$, then the only values of $\widetilde{\partial_i u_n}$ in the integral would coincide with those of $\partial_i \widetilde{u}_n$, see Figure 14. Hence the equality of the restrictions since $\partial_i \widetilde{u}_{n,p} = \rho_p \star \partial_i \widetilde{u}_n$.

We are thus down to a geometry question, where the regularity of Ω intervenes (at last). We need to estimate the distance between the graph of φ, denoted G, and the same graph translated upwards by $\frac{1}{n}$, denoted G_n. Let L be the Lipschitz constant of φ and take two points $x \in G$ and $y \in G_n$. We have

$$\|y-x\|^2 = \|y'-x'\|^2 + \Big(\varphi(y') - \varphi(x') + \frac{1}{n}\Big)^2,$$

using the prime notation to denote the projection on $\mathbb{R}^{d-1}$ as usual. Now

$$\varphi(y') - \varphi(x') + \frac{1}{n} \geq \frac{1}{n} - L\|y'-x'\|,$$

therefore if $\|y'-x'\| \leq \frac{1}{2nL}$, then $\|y-x\| \geq \frac{1}{2n}$. On the other hand, if $\|y'-x'\| \geq \frac{1}{2nL}$, it follows trivially that $\|y-x\| \geq \frac{1}{2nL}$. We thus see that

$$\|y-x\| \geq \min\Big(\frac{1}{2n}, \frac{1}{2nL}\Big).$$

If we choose

$$p > \frac{1}{\min\big(\frac{1}{2n}, \frac{1}{2nL}\big)},$$

then $B\big(x, \frac{1}{p}\big) \subset \Omega_n$, hence the final result. □

Remark 2.8.1 There is a slight cheat in the geometric part of the above proof, in that we have ignored what happens on the lateral sides of C. Indeed, there is actually no problem, since u vanishes there. □

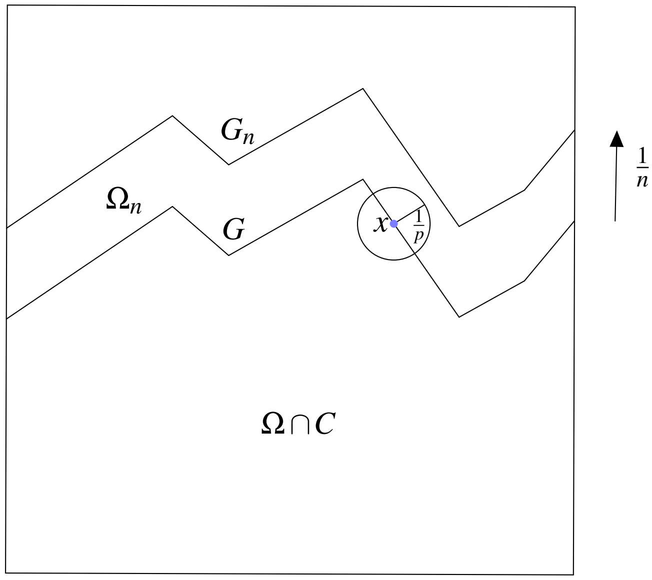

Figure 15. The translated open set Ω_n and the ball of radius $\frac{1}{p}$ used to compute the convolution at point x. For x in $\bar{\Omega} \cap C$, the ball remains included in Ω_n uniformly for $p \to +\infty$, n fixed.

Remark 2.8.2 Warning: there are open sets less regular than Lipschitz on which not only does the above proof not work, but the density result is false (exercise, find a simple example). It is however always true that $C^1(\Omega) \cap H^1(\Omega)$ is dense in $H^1(\Omega)$, a weaker result (Meyers-Serrin theorem) which is sometimes sufficient. But not here for the trace. □

Once the density of $C^1(\bar{\Omega})$ is established, we can prove the trace theorem.

Theorem 2.8.2 *Let Ω be a Lipschitz open subset of $\mathbb{R}^d$. There exists a unique continuous linear mapping $\gamma_0 \colon H^1(\Omega) \to L^2(\partial\Omega)$ such that for all $u \in C^1(\bar{\Omega})$, we have*

$$\gamma_0(u) = u_{|\partial\Omega}.$$

In other words, the trace is the unique reasonable way of defining a boundary value for H^1 functions, as the continuous extension of the restriction to the boundary for functions for which this restriction makes sense.

Proof. We write the proof only in dimension $d = 2$, but the general case is strictly identical, up to heavier notation.

Let $u \in C^1(\bar{\Omega})$. By partition of unity, we consider $u\psi$ which is supported in one of the $C_j = C$ for $j = 1, \dots, m$. Let $G = \partial\Omega \cap C$ be the part of the boundary included in C. By definition of the boundary measure, we have

$$\|u\psi\|^2_{L^2(G)} = \int_{-a}^{a} (u\psi)(y_1,\varphi(y_1))^2\sqrt{1+\varphi'(y_1)^2}\,dy_1$$
$$= \int_{-a}^{a}\Big(\int_{-a}^{\varphi(y_1)} \frac{\partial(u\psi)}{\partial y_2}(y_1,y_2)\,dy_2\Big)^2\sqrt{1+\varphi'(y_1)^2}\,dy_1.$$

By the Cauchy-Schwarz inequality, we have

$$\Big(\int_{-a}^{\varphi(y_1)} \frac{\partial(u\psi)}{\partial y_2}(y_1,y_2)\,dy_2\Big)^2 \le |\varphi(y_1)+a|\int_{-a}^{\varphi(y_1)}\Big(\frac{\partial(u\psi)}{\partial y_2}(y_1,y_2)\Big)^2 dy_2,$$

with $|\varphi(y_1)+a| \le 2a$. Let us set $M = \max_{[-a,a]}\sqrt{1+\varphi'(y_1)^2}$. We obtain

$$\|u\psi\|^2_{L^2(G)} \le 2aM\int_{-a}^{a}\int_{-a}^{\varphi(y_1)}\Big(\frac{\partial(u\psi)}{\partial y_2}\Big)^2 dy_2dy_1 = 2aM\int_{\Omega\cap C}\Big(\frac{\partial(u\psi)}{\partial y_2}\Big)^2 dx.$$

Now $\frac{\partial(u\psi)}{\partial y_2} = \psi\frac{\partial u}{\partial y_2} + u\frac{\partial\psi}{\partial y_2}$, so that

$$\|u\psi\|^2_{L^2(G)} \le 4aM\Big(\int_{\Omega\cap C}\psi^2\Big(\frac{\partial u}{\partial y_2}\Big)^2 dx + \int_{\Omega\cap C}u^2\Big(\frac{\partial\psi}{\partial y_2}\Big)^2 dx\Big)$$
$$\le 4aM\Big[\Big\|\frac{\partial u}{\partial y_2}\Big\|^2_{L^2(\Omega\cap C)} + \max_{\Omega\cap C}\Big(\frac{\partial\psi}{\partial y_2}\Big)^2\|u\|^2_{L^2(\Omega\cap C)}\Big]$$
$$\le C^2\|u\|^2_{H^1(\Omega)},$$

(recall that ψ is $[0,1]$-valued) for some constant C that we could make more explicit if we wanted to.

We put all the estimates together by partition of unity and the triangle inequality:

$$\|u\|_{L^2(\partial\Omega)} \le \sum_{j=1}^{m}\|u\psi_j\|_{L^2(G_j)} \le mC\|u\|_{H^1(\Omega)},$$

for all $u \in C^1(\bar\Omega)$. The linear mapping $u \mapsto u_{|\partial\Omega}$ defined on $C^1(\bar\Omega)$ is thus continuous in the $H^1(\Omega)$ and $L^2(\partial\Omega)$ norms. Since $C^1(\bar\Omega)$ is dense in $H^1(\Omega)$, the mapping has a unique continuous extension to $H^1(\Omega)$ with values in $L^2(\partial\Omega)$, which is called the trace mapping γ_0. □

Remark 2.8.3 Again, there are open sets less regular than Lipschitz on which no trace mapping can be defined. □

We now are in a position to extend the integration by parts formula(s) to elements of Sobolev spaces.

Theorem 2.8.3 *Let Ω be a Lipschitz open set and $u, v \in H^1(\Omega)$. Then we have*

$$\int_\Omega \frac{\partial u}{\partial x_i} v\,dx = -\int_\Omega u \frac{\partial v}{\partial x_i}\,dx + \int_{\partial\Omega} \gamma_0(u)\gamma_0(v) n_i\,d\Gamma. \tag{2.15}$$

Proof. We argue by density. We already know that formula (2.16) holds true on $C^1(\bar{\Omega})$. Let u_n, v_n be sequences in $C^1(\bar{\Omega})$ such that $u_n \to u$ and $v_n \to v$ in $H^1(\Omega)$ when $n \to +\infty$. This means that $u_n \to u$, $v_n \to v$, $\partial_i u_n \to \partial_i u$ and $\partial_i v_n \to \partial_i u$ in $L^2(\Omega)$. Therefore, $\partial_i u_n v_n \to \partial_i u v$ and $u_n \partial_i v_n \to u \partial_i v$ in $L^1(\Omega)$ and the left-hand side integral and the first integral in the right-hand side pass to the limit. Secondly, we have $\gamma_0(u_n) \to \gamma_0(u)$ and $\gamma_0(v_n) \to \gamma_0(v)$ in $L^2(\partial\Omega)$ since the trace mapping is continuous, hence $\gamma_0(u_n)\gamma_0(v_n) \to \gamma_0(u)\gamma_0(v)$ in $L^1(\partial\Omega)$ and the second integral in the right-hand side also passes to the limit. □

The various corollaries of the integration by parts formula also hold true, provided all the integrals make sense. For instance, for all $u \in H^1$,

$$\int_\Omega \frac{\partial u}{\partial x_i}\,dx = \int_{\partial\Omega} \gamma_0(u) n_i\,d\Gamma, \tag{2.16}$$

as is seen from taking $v = 1$.

The formulas that entail second derivatives should be applied to H^2 functions. Such functions u are in H^1, thus have a trace $\gamma_0(u)$ and they also have a second trace $\gamma_1(u)$, called the normal trace, that plays the role of the normal derivative for a regular function. Indeed, $\partial_i u \in H^1(\Omega)$ therefore $\gamma_1(u) = \sum_{i=1}^d \gamma_0(\partial_i u) n_i$ is well defined and continuous from $H^2(\Omega)$ into $L^2(\partial\Omega)$. Furthermore, if $u \in C^2(\bar{\Omega})$, then $\gamma_1(u) = \frac{\partial u}{\partial n}$. We thus establish the Green formula

$$\int_\Omega (\Delta u) v\,dx = -\int_\Omega \nabla u \cdot \nabla v\,dx + \int_{\partial\Omega} \gamma_1(u)\gamma_0(v)\,d\Gamma, \tag{2.17}$$

for all $u \in H^2(\Omega)$ and all $v \in H^1(\Omega)$ by density of $C^2(\bar{\Omega})$ in $H^2(\Omega)$ and of $C^1(\bar{\Omega})$ in $H^1(\Omega)$, starting from formula (2.5).

Proposition 2.8.1 *Let Ω be a Lipschitz open set. Then we have*

$$H_0^1(\Omega) = \ker \gamma_0.$$

Proof. One inclusion is easy. The space $H_0^1(\Omega)$ is by definition the closure of the space $\mathscr{D}(\Omega)$ in $H^1(\Omega)$. Thus, if $u \in H_0^1(\Omega)$, then there exist $\varphi_n \in \mathscr{D}(\Omega)$ such that $\varphi_n \to u$ in $H^1(\Omega)$. It is clear that $\gamma_0(\varphi_n) = 0$, thus $u \in \ker \gamma_0$ by continuity of the trace, or in other words $H_0^1(\Omega) \subset \ker \gamma_0$.

We just give the idea for the reverse inclusion. Take a zero trace function u, use a partition of unity adapted to the boundary, extend all the parts to the whole of $\mathbb{R}^d$ by 0 (the integration by parts formula (2.15) shows that the extension remains in H^1 this time), translate each function downwards by a small amount in its cube then perform the convolution step which provides a compactly supported, C^∞ approximation. We leave the details to the reader, all the necessary technical elements have already been introduced previously. □

Remark 2.8.4 Similar arguments show that $H_0^2(\Omega) = \ker \gamma_0 \cap \ker \gamma_1$ and so on. It should be noted that the trace γ_0 is not onto. Its image is called $H^{1/2}(\Omega)$, it is a strict dense subspace of $L^2(\partial\Omega)$. Everything we have said in terms of traces can also naturally be done in the spaces $W^{m,p}(\Omega)$. □