Lecture notes for 18.155: distributions, elliptic regularity, and applications to PDEs

December 10, 2022

Semyon Dyatlov

Contents

Preface

These are the lecture notes for the course 18.155 (Differential Analysis I) taught at MIT in Fall 2022. The topics include:

- basics of the theory of distributions,
- fundamental solutions to some constant coefficient PDEs,
- Fourier transform on distributions and Sobolev spaces,
- elliptic regularity and Fredholm mapping properties of elliptic operators on Sobolev spaces,
- and applications to PDEs such as discreteness of the spectrum of the Laplacian and Hodge's theorem.

There are several sources used (acknowledged in more detail at the end of each section), including:

- Hörmander's book, volume I [**Hör03**]: a classical and comprehensive treatment of distribution theory. The text is quite dense, so it is not an easy source to learn about distrbutions for the first time, but a lot of the arguments in the distribution theory part of these notes are taken from there.
- Friedlander–Joshi's book [**FJ98**]: a much shorter book which does the essentials of distribution theory. This one can be used by a beginner (familiar with topics such as Lebesgue integration). Some of other arguments in the distribution theory part of these notes are from this book.
- Melrose's lecture notes for 18.155 [**Mel**]: these inspired some of the arguments in the later part of the course. I also tried to model the structure of my version of 18.155 roughly after the version taught by Prof. Melrose.

To comfortably read the entire notes, a reader would find it helpful to be familiar with some fundamentals of analysis and differential geometry. Some of these are reviewed briefly in the notes, as a reminder and to fix notation, but a lot of proofs and explanations are replaced by references to the literature. The topics we will need are:

- Real analysis (18.100B at MIT): basics of metric space topology, the theory of differentiation and integration, and the Arzelà–Ascoli Theorem. This one is a definite prerequisite to taking 18.155.

- Lebesgue integration (covered in 18.102, 18.103, 18.125, or 18.675 at MIT): Lebesgue integral, its convergence properties such as the Dominated Convergence Theorem, metric spaces, the spaces L^p and their completeness, and the change of variables formula. This one is used from the beginning but largely as a black box.
- Functional analysis (partially covered in 18.102 at MIT, these notes provide the pieces which are not covered): Hilbert spaces and their basic properties (orthogonal projections, Riesz representation theorem), Banach spaces and their basic properties (including Banach–Steinhaus theorem), compact and Fredholm operators, spectra of self-adjoint compact operators on Hilbert spaces (Hilbert–Schmidt theorem), and the Fredholm alternative.
- Manifolds (18.101 at MIT): for the latter part of the course the reader should be familiar with the concept of an abstract C^∞ manifold, tangent and cotangent bundles, differential forms and Stokes' theorem, and basic Riemannian geometry.
- Complex analysis (18.112 at MIT): we will occasionally use a bit of the basics e.g. unique continuation of analytic functions.

To help the reader get the most of these notes, I use the following superscripts for section names/theorems/etc.:

- $^{\mathbf{R}}$: review, a topic which I would say should be in the prerequisites for this course rather than the course itself (regardless of whether prerequisite courses at MIT actually cover this – if they don't then I develop this in more detail);
- $^{\mathbf{S}}$: straightforward once you have enough understanding of the concepts involved. If you feel comfortable with the material you might be able to skip some of the details there;
- $^{\mathbf{X}}$: extra, will help deepen your understanding of the material but you might be able to skip it at first reading.

I have been supported by the NSF CAREER grant DMS-1749858.

CHAPTER 1

Prologue: motivation and background

1.1. A bit of motivation

In this course we (among other things) develop the theory of *distributions* and show various forms of *elliptic regularity*. These both take a while to set up, so let us first look at a couple of applications to PDEs (partial differential equations).

1.1.1. Solving Poisson's equation. To keep things simple, let us restrict to the case of dimension 3. Consider the Laplace operator[1] on $\mathbb{R}^3$

$$\Delta := \partial_{x_1}^2 + \partial_{x_2}^2 + \partial_{x_3}^2$$

and let us study *Poisson's equation*

$$\Delta u = f, \tag{1.1}$$

where f is a given function on $\mathbb{R}^3$ and u is the unknown function. If you took a physics course, you might have encountered (1.1) in electrostatics (u = electric potential, f = density of charge) or in Newtonian gravity (u = gravitational potential, f = density of mass).

Perhaps you also learned that one solution to (1.1) is given by the integral formula

$$u(x) = \int_{\mathbb{R}^3} E(x-y) f(y)\, dy, \quad x \in \mathbb{R}^3, \tag{1.2}$$

where the *Coulomb potential* E is defined by

$$E(x) = -\frac{1}{4\pi|x|}, \quad x \in \mathbb{R}^3 \setminus \{0\}. \tag{1.3}$$

This is something that can be checked directly for sufficiently nice f, see Exercise 1.1 below. However, this leaves open some questions, which can be conceptually addressed by developing the theory of distributions.

One question is: the formula (1.2) makes sense, for example, for any bounded compactly supported (Lebesgue measurable) function f. But in this case u might not

[1]There are two conventions in the literature: $\Delta = \partial_{x_1}^2 + \partial_{x_2}^2 + \partial_{x_3}^2$ and $\Delta = -\partial_{x_1}^2 - \partial_{x_2}^2 - \partial_{x_3}^2$. Physicists often use the first one and geometers often use the second one. We will use the first convention here, for no good reason other than the author's personal preference. However, the Hodge Laplacian in §17.3.3 will use the opposite sign convention.

be twice differentiable. Can we still say that u solves Poisson's equation in a certain sense?

An answer to this question is given by *weak solutions*. Assume first that $u \in C^2(\mathbb{R}^3)$ solves the equation (1.1) (we call such u a *classical solution* because it has enough derivatives to make sense of the equation at each point). Take any smooth compactly supported function $\varphi \in C_c^\infty(\mathbb{R}^3)$ (see §1.2.4 below), which we call a *test function*. Integrating by parts twice using the Divergence Theorem (where the boundary terms do not appear since φ is compactly supported – see Theorem 1.17 below), we see that

$$\int_{\mathbb{R}^3} f(x)\varphi(x)\,dx = \int_{\mathbb{R}^3} (\Delta u(x))\varphi(x)\,dx$$
$$= -\int_{\mathbb{R}^3} \sum_{j=1}^{3} \big(\partial_{x_j} u(x)\big)\big(\partial_{x_j}\varphi(x)\big)\,dx = \int_{\mathbb{R}^3} u(x)\Delta\varphi(x)\,dx,$$

that is we have

$$\int_{\mathbb{R}^3} f\varphi\,dx = \int_{\mathbb{R}^3} u(\Delta\varphi)\,dx \quad \text{for all } \varphi \in C_c^\infty(\mathbb{R}^3). \tag{1.4}$$

The latter makes sense for any u, f which are locally integrable. For such u, f we say that u is a *weak solution* to Poisson's equation (1.1) if (1.4) holds. Two comments are in order:

- as the calculation above shows, if u is a classical solution, then it is a weak solution as well;
- if u is a weak solution and we also know that $u \in C^2(\mathbb{R}^3)$ then u is a classical solution.

Weak solutions are thus a superset of classical solutions. The second comment above leads to the following general strategy for studying solutions of linear partial differential equations:

- understand all weak solutions to the equation;
- assuming regularity of the right-hand side, establish regularity of the weak solution. If the weak solution is regular enough, then it is also a classical solution.

Another question raised by the formula (1.2) is related to the following computation:

$$\Delta u(x) = \int_{\mathbb{R}^3} (\Delta E)(x-y) f(y)\,dy = 0. \tag{1.5}$$

Here in the first equality we differentiate (in x) under the integral sign. In the second equality we use the fact that $\Delta E = 0$ on $\mathbb{R}^3 \setminus \{0\}$ (see Exercise 1.1(a)) and thus the integral is 0. This seems to contradict our expectation that $\Delta u = f$.

From the point of view of classical (Lebesgue) integration, the computation (1.5) is invalid because the gradient $\nabla E(x)$ blows up too fast at $x = 0$ to be able to differentiate under the integral sign. But the theory of distributions gives another way of thinking about this computation, which also gives a proof that $\Delta u = f$. Namely, the first equality in (1.5) is valid if we treat ΔE as a distribution and think of the integral as a distributional pairing. In distributions, we do not have $\Delta E = 0$, instead

$$\Delta E = \delta_0 \tag{1.6}$$

where δ_0 is the *Dirac delta* at the origin, which is not a function, but a distribution with the following property:

$$\int_{\mathbb{R}^3} \delta_0(x)\varphi(x)\,dx = \varphi(0) \quad \text{for all} \quad \varphi \in C_c^\infty(\mathbb{R}^3).$$

Then (1.5) can be made correct as follows:

$$\begin{aligned}\Delta u(x) &= \int_{\mathbb{R}^3} (\Delta E)(x-y)f(y)\,dy = \int_{\mathbb{R}^3} \delta_0(x-y)f(y)\,dy \\ &= \int_{\mathbb{R}^3} \delta_0(y)f(x-y)\,dy = f(x).\end{aligned} \tag{1.7}$$

Here the integral signs no longer denote literal integrals; instead they are distributional pairings. It will take us some time to develop the theory of distributions enough to rigorously justify statements such as (1.6) and (1.7), but these formulas have an immediate physical interpretation which predates the development of distribution theory in mathematics: (1.6) tells us that the Coulomb potential E corresponds to a *point charge* (with δ_0 being the 'density' of a point charge) and (1.7) is a version of *superposition principle* (i.e. for linear equations, taking linear combinations or more generally parametric integrals of solutions gives other solutions).

1.1.2. Examples of elliptic regularity. For the next example, let us work in $\mathbb{R}^2$. Consider the following three PDEs:

$$(\partial_{x_1}^2 + \partial_{x_2}^2)u = 0, \tag{1.8}$$

$$(\partial_{x_1} + i\partial_{x_2})u = 0, \tag{1.9}$$

$$(\partial_{x_1}^2 - \partial_{x_2}^2)u = 0. \tag{1.10}$$

Here is a question:

Is it true that every solution u to these equations is a smooth (C^∞) function?

Here we can restrict ourselves to classical solutions (e.g. for (1.8), if $u \in C^2$ solves the equation, then $u \in C^\infty$), or we can define weak solutions similarly to (1.4); the answer will be the same either way.

The answer to the question above has been known a long time ago:

- functions solving the Laplace equation (1.8) are called *harmonic* and they are always smooth;
- functions solving the Cauchy–Riemann equation (1.9) are *analytic* functions of the complex variable $z = x_1 + ix_2$ and they are also always smooth;
- but the wave equation (1.10) has some nonsmooth solutions, for example $u(x_1, x_2) = f(x_1 + x_2)$ for any C^2 function f.

However, the XXth century analysis that we study in this course will give a more systematic point of view on understanding what is different between the equations (1.8), (1.9), and (1.10). To give a preview of it, let P be the differential operator such that the equation studied is $Pu = 0$. Define the homogeneous polynomial $p(\xi_1, \xi_2)$ by replacing ∂_{x_1} by ξ_1 and ∂_{x_2} by ξ_2, so that (1.8)–(1.10) correspond to the polynomials

$$\xi_1^2 + \xi_2^2, \tag{1.11}$$

$$\xi_1 + i\xi_2, \tag{1.12}$$

$$\xi_1^2 - \xi_2^2. \tag{1.13}$$

We say that the polynomial p is *elliptic*, if the equation $p(\xi_1, \xi_2) = 0$ has only one solution on $\mathbb{R}^2$, namely $\xi_1 = \xi_2 = 0$.

One of the main results of this course is *elliptic regularity* which in particular says that if the polynomial p is elliptic, then all solutions to the equation $Pu = 0$ are smooth. This applies to the equations (1.8)–(1.9), since the corresponding polynomials are elliptic, but not to (1.10).

We will study three versions of elliptic regularity. The third version has many applications, three of which we present in the setting of compact manifolds without boundary:

- Fredholm mapping property of elliptic differential operators on Sobolev spaces;
- discreteness of spectrum of self-adjoint elliptic operators;
- and Hodge's Theorem, giving a bijection between de Rham cohomology classes (an algebraic topological invariant) and harmonic forms (a Riemann geometric/spectral theoretic object).

1.2. Functional spaces$^{\mathbf{R}}$

We start by giving a very brief review of the spaces L^p and C^k. We then introduce the space of smooth compactly supported functions C_c^∞, which is important to us since the space of distributions will be its dual.

For now we will work with subsets of $\mathbb{R}^n$. The definition below collects some useful notation.

DEFINITION 1.1. *Let $\mathcal{M}$ be a metric space and $U \subset V \subset \mathcal{M}$ be two sets.*

(1) *We write*
$$U \Subset V$$
if U *is a relatively open subset of* V*.*

(2) *We say that* U *is* compactly contained *in* V*, and write*
$$U \Subset V$$
if there exists a compact set K *such that* $U \subset K \subset V$*.*

An alternative definition of compact containment is that the closure of U be compact and contained inside V.

Recall that any open set $U \Subset \mathbb{R}^n$ can be exhausted by compact subsets:
$$U = \bigcup_{j=1}^{\infty} K_j \quad \text{where } K_j \Subset U, \quad K_j \Subset K_{j+1}. \tag{1.14}$$
Indeed, one can for example let K_j consist of all points x such that $|x| \leq j$ and the open ball $B^\circ(x, 1/j)$ is contained in U. Moreover, any $K \Subset U$ is contained in one of the sets K_j.

When U is a set, by default a *function* on U is a map $f : U \to \mathbb{C}$. That is, functions are assumed complex valued unless stated otherwise.

1.2.1. Lebesgue integral and the spaces L^p. A theory of Lebesgue measure and integral on $\mathbb{R}^n$ produces:

- the notion of which subsets of $\mathbb{R}^n$ are *measurable* (in practice, any set you can construct without, say, using the Axiom of Choice will be measurable so we will not worry about checking measurability in these notes);
- the *Lebesgue measure*, which maps each measurable subset $A \subset \mathbb{R}^n$ to its 'volume' $\mathrm{vol}(A) \in [0, \infty]$;
- the *Lebesgue integral*, which defines for certain functions $f : \mathbb{R}^n \to \mathbb{C}$ their integral $\int_{\mathbb{R}^n} f(x)\, dx$. More precisely, we always need f to be measurable, that is for each $a \in \mathbb{R}$ the set $\{x \in \mathbb{R}^n \mid f(x) \leq a\}$ should be measurable. If f is nonnegative, then nothing else is needed and the integral $\int_{\mathbb{R}^n} f(x)\, dx$ is always defined as a possibly infinite number in $[0, \infty]$. For general f we impose the additional condition that $\int_{\mathbb{R}^n} |f(x)|\, dx < \infty$ (we call such functions *Lebesgue integrable*) and the integral $\int_{\mathbb{R}^n} f(x)\, dx$ is defined as a complex number.

We refer to the books [**Bea02**, **Jon01**, **Rud64**, **Rud87**, **Str11**] for constructions of the above objects and their standard properties, which we will freely use in these notes. We in particular note that:

- If f is Riemann integrable (in the proper sense), then it is also Lebesgue integrable and the Riemann integral is the same as Lebesgue integral. In other words, the Lebesgue theory does not give a different value of the integral, instead it lets us integrate more functions.

- We have *Fubini's/Tonelli's Theorem*: if we write elements of $\mathbb{R}^{n+m}$ as (x, y) where $x \in \mathbb{R}^n, y \in \mathbb{R}^m$, the function $f : \mathbb{R}^{n+m} \to \mathbb{C}$ is measurable and f is either nonnegative or Lebesgue integrable, then
$$\int_{\mathbb{R}^{n+m}} f(x, y)\, dxdy = \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} f(x, y)\, dy \right) dx.$$
In particular, this lets us reduce (at least in principle) integrals over $\mathbb{R}^n$ to integrals over $\mathbb{R}$, which one can hope to compute using the Fundamental Theorem of Calculus.
- We also have the *Dominated Convergence Theorem*: if a sequence of measurable functions $f_k : \mathbb{R}^n \to \mathbb{C}$ converges to some function f for almost every x (see below) and there exists an integrable function g such that $|f_k| \le g$ for all k, then $\int_{\mathbb{R}^n} f_k(x)\, dx \to \int_{\mathbb{R}^n} f(x)\, dx$.

For a logical statement $S(x)$ with one free variable $x \in \mathbb{R}^n$, we say that it holds (Lebesgue) *almost everywhere* (often abbreviated to 'a.e.'), if the set $\{x \in \mathbb{R}^n \mid S(x) \text{ is false}\}$ has Lebesgue measure 0. A measurable function f is equal to 0 almost everywhere if and only if $\int_{\mathbb{R}^n} |f(x)|\, dx = 0$.

We can now define the spaces L^p. We start with the case $p < \infty$:

DEFINITION 1.2. *Let $1 \le p < \infty$. For a measurable function $f : \mathbb{R}^n \to \mathbb{C}$, its L^p-norm is*
$$\|f\|_{L^p} := \left(\int_{\mathbb{R}^n} |f(x)|^p\, dx \right)^{1/p} \in [0, \infty].$$
We define $L^p(\mathbb{R}^n)$ as the quotient space
$$L^p(\mathbb{R}^n) := \frac{\{f : \|f\|_{L^p} < \infty\}}{\{f : f = 0 \ a.e.\}}. \tag{1.15}$$

In the above definition we identify two L^p functions f and g if $f = g$ almost everywhere. This is important because otherwise $\| \bullet \|_{L^p}$ is not a norm on the space L^p (as there are nonzero elements of the space which have norm zero). It also corresponds well to the theory of distributions: two L^p functions are equal as distributions if and only if they are equal almost everywhere as functions (see Theorem 1.16 below). So if, say, a solution to some differential equation is given by the indicator function of a ball, then we will not be worrying about what the values of this function on the boundary of the ball are since that boundary has measure 0. Note however that none of this matters for continuous functions: two continuous functions are equal a.e. if and only if they are equal everywhere.

One of the main advantages of the Lebesgue integral over the Riemann integral is that the space L^p with the norm $\| \bullet \|_{L^p}$ is a *Banach space*, namely it is a normed vector space which is complete. This is very useful in the study of PDE since quite often solutions of differential equations are constructed as limits of sequences of approximate solutions (even though in the modern theory this aspect is somewhat hidden).

We can also define the space L^∞ which carries a version of the sup-norm adjusted for measure zero sets. Namely, we put

$$\|f\|_{L^\infty} := \inf\{a \geq 0 \colon |f(x)| \leq a \text{ for a.e. } x\}$$

and define the Banach space $L^\infty(\mathbb{R}^n)$ as a quotient similarly to (1.15).

For later use we recall *Hölder's inequality*: if $f \in L^p(\mathbb{R}^n)$, $g \in L^q(\mathbb{R}^n)$, and $\frac{1}{p}+\frac{1}{q} = 1$, then

$$\|fg\|_{L^1} \leq \|f\|_{L^p}\|g\|_{L^q}. \tag{1.16}$$

More generally, we can define the spaces $L^p(U)$ where $U \Subset \mathbb{R}^n$. More precisely, $L^p(U)$ consists of measurable functions $f : U \to \mathbb{C}$ such that $\mathbf{1}_U f \in L^p(\mathbb{R}^n)$, where for a set $A \subset \mathbb{R}^n$ we denote by $\mathbf{1}_A : \mathbb{R}^n \to \mathbb{R}$ its *indicator function*:

$$\mathbf{1}_A(x) = \begin{cases} 0, & x \in A, \\ 1, & x \notin A. \end{cases} \tag{1.17}$$

We also define the spaces of *locally* L^p functions

$$L^p_{\text{loc}}(U) := \frac{\{f : U \to \mathbb{C} \colon \mathbf{1}_K f \in L^p(U) \quad \text{for all compact } K \subset U\}}{\{f \colon f = 0 \text{ a.e.}\}}$$

and *compactly supported* L^p functions

$$L^p_{\text{c}}(U) := \{f \in L^p(U) \mid \text{there exists compact } K \subset U \text{ such that } f = \mathbf{1}_K f \text{ a.e.}\}.$$

(Strictly speaking, $L^p_{\text{c}}(U)$ is the space of functions of compact *essential support*, whose definition is different from Definition 1.5 below by adding 'almost everywhere'.)

From Hölder's inequality we can see that

$$L^p_{\text{loc}}(U) \subset L^r_{\text{loc}}(U), \quad L^p_{\text{c}}(U) \subset L^r_{\text{c}}(U) \quad \text{for all } p \geq r. \tag{1.18}$$

1.2.2. More on the space L^2. Let $U \Subset \mathbb{R}^n$. For us the most convenient L^p space will often be the one with $p = 2$. This is because the space $L^2(U)$ is a *Hilbert space*, whose norm is induced by the L^2 Hermitian inner product (with $\overline{\bullet}$ denoting complex conjugation)

$$\langle f, g\rangle_{L^2(U)} := \int_U f(x)\overline{g(x)}\, dx. \tag{1.19}$$

Note that $\|f\|^2_{L^2} = \langle f, f\rangle_{L^2}$. We have the *Cauchy–Schwarz* inequality (following e.g. from Hölder's inequality)

$$|\langle f, g\rangle_{L^2}| \leq \|f\|_{L^2}\|g\|_{L^2}. \tag{1.20}$$

We now list several important properties of $L^2(U)$, which are actually true for general Hilbert spaces (relying crucially on completeness). See for example [DS88, §IV.4], [**Lax02**, Chapter 6], [**RS81**, Chapter II], or [**Rud87**, Chapter 4] for the proofs. We start with

THEOREM 1.3 (Orthogonal Complement Theorem). *Assume that $W \subset L^2(U)$ is a closed subspace. Define its* orthogonal complement *as*

$$W^\perp := \{f \in L^2(U) \mid \text{for all } g \in L^2(U) \text{ we have } \langle f, g\rangle_{L^2} = 0\}.$$

Then $L^2(U) = W \oplus W^\perp$.

Recall from functional analysis that a (linear) *functional* on a Banach space X is a linear operator $X \to \mathbb{C}$. The *dual space* X' to X is the space of bounded linear functionals on X, and it is a Banach space when taken with the operator norm

$$\|T\|_{X'} = \sup_{f \in X\setminus\{0\}} \frac{|T(f)|}{\|f\|_X}.$$

The next theorem establishes a canonical isomorphism between $L^2(U)$ and its dual space $L^2(U)'$.

THEOREM 1.4 (Riesz Representation Theorem for L^2). *1. Let $g \in L^2(U)$. Then*

$$T_g : f \in L^2(U) \;\mapsto\; \langle f, g\rangle_{L^2} \in \mathbb{C}$$

is a bounded linear functional on $L^2(U)$, and $\|T_g\|_{L^2(U)'} = \|g\|_{L^2(U)}$.

2. Assume that T is a bounded linear functional on $L^2(U)$. Then there exists unique $g \in L^2(U)$ such that $T = T_g$.

More generally, the spaces $L^p(U)$ and $L^q(U)$ are dual to each other when $p, q \in (1, \infty)$ and $\frac{1}{p} + \frac{1}{q} = 1$, but we will not use this fact in this course.

1.2.3. The spaces C^k. For $U \Subset \mathbb{R}^n$, define the space of continuous functions

$$C^0(U) := \{f : U \to \mathbb{C} \mid f \text{ is continuous}\}.$$

The natural norm to use on the space of continuous functions would be the sup-norm

$$\|f\|_{C^0} := \sup_{x \in U} |f(x)|.$$

However, since U is open rather than compact, the sup-norm $\|f\|_{C^0}$ is infinite for some $f \in C^0(U)$. This is a common theme for many of the spaces of functions and distributions that we will be using in the study of PDEs: we do not make any a priori assumptions on the growth of $f(x)$ as x approaches the boundary of U.

To fix this we can consider the space of *compactly supported* continuous functions:

DEFINITION 1.5. *Let $U \Subset \mathbb{R}^n$ and $f : U \to \mathbb{C}$. Define the* support *of f, denoted* $\operatorname{supp} f$, *as the closure of the set $\{x \in U \mid f(x) \neq 0\}$ in U. We say that f is* compactly supported *if* $\operatorname{supp} f$ *is a compact subset of U.*

For example, if $U = B^\circ(0,1)$ is the open unit ball, then the function $f(x) \equiv 1$ is not compactly supported since its support is the whole U. But the indicator function $f = \mathbf{1}_{B(0,1/2)}$ is compactly supported. We typically use Definition 1.5 for continuous functions only.

Denote by $C^0_{\mathrm{c}}(U)$ the space of compactly supported functions in $C^0(U)$. Then $\|\bullet\|_{C^0}$ defines a norm on $C^0_{\mathrm{c}}(U)$, though $C^0_{\mathrm{c}}(U)$ is not complete with respect to this norm (see Exercise 1.2 below). Moreover, we have the inclusion

$$C^0_{\mathrm{c}}(U) \ \subset \ L^p(U) \quad \text{for all } p.$$

Finally, any function $f \in C^0_{\mathrm{c}}(U)$ is uniformly continuous, thus it has a modulus of continuity:

$$\omega_f(\varepsilon) := \sup\big\{|f(x) - f(y)| : x, y \in U,\ |x-y| \le \varepsilon\big\} \to 0 \quad \text{as} \quad \varepsilon \to 0+\,. \tag{1.21}$$

We next define the spaces C^k of k times continuously differentiable functions:

Definition 1.6. *Let $U \Subset \mathbb{R}^n$. Define the space $C^1(U)$ consisting of continuously differentiable functions:*

$$C^1(U) := \{f \in C^0(U) \mid \partial_{x_1} f, \dots, \partial_{x_n} f \text{ exist and lie in } C^0(U)\}.$$

For an integer $k \ge 2$, define the space $C^k(U)$ inductively by

$$C^k(U) := \{f \in C^{k-1}(U) \mid \partial_{x_1} f, \dots, \partial_{x_n} f \text{ exist and lie in } C^{k-1}(U)\}.$$

Denote by $C^k_{\mathrm{c}}(U)$ the space of compactly supported functions in $C^k(U)$.

Remark 1.7.$^{\mathbf{X}}$ *Even though we used partial derivatives with respect to the given coordinates in the definition above, the resulting spaces are independent of the choice of linear coordinates on $\mathbb{R}^n$. A conceptual way to see this would be to give an alternative definition of C^1 in terms of the (Fréchet) differential $df : U \to (\mathbb{R}^n)'$, define C^2 using the differential $d(df) : U \to (\mathbb{R}^n)' \otimes (\mathbb{R}^n)'$ etc. but we do not develop this here since for our purposes working with coordinates is perfectly fine.*

Remark 1.8. *The closure in the definition of support is important in particular because this makes support well-behaved under differentiation:*

$$\operatorname{supp}(\partial_{x_j} f) \ \subset \ \operatorname{supp} f \quad \text{for all } f \in C^1(U). \tag{1.22}$$

Indeed, if $x \in U \setminus \operatorname{supp} f$, then f vanishes on some ball $B(x,\varepsilon)$ centered at x, so $\partial_{x_j} f(x) = 0$.

The set $\{f \ne 0\}$ is not closed under differentiation: consider for example the function $f(x) = x$ on $\mathbb{R}$.

To work with higher order derivatives, we introduce the *multiindex notation*. A multiindex in $\mathbb{R}^n$ is a vector $\alpha = (\alpha_1, \dots, \alpha_n)$ whose entries are nonnegative integers.

We denote (recalling that for functions in C^k the order of differentiation k times does not matter)

$$|\alpha| := \alpha_1 + \cdots + \alpha_n, \quad \partial_x^\alpha := \partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n}.$$

A norm on $C_c^k(U)$ (which is coordinate dependent but the resulting topology is canonical) is given by

$$\|f\|_{C^k} := \max_{|\alpha| \le k} \sup_{x \in U} |\partial_x^\alpha f(x)|. \tag{1.23}$$

For later use, we also introduce the notation

$$x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n. \tag{1.24}$$

1.2.4. The spaces C^∞ and C_c^∞. We now define the spaces C^∞ and C_c^∞ that are ubiquitous in the rest of these notes.

DEFINITION 1.9. *Let $U \Subset \mathbb{R}^n$. We say that a function $f : U \to \mathbb{C}$ lies in $C^\infty(U)$ if it lies in $C^k(U)$ for all k, that is*

$$C^\infty(U) := \bigcap_{k \ge 0} C^k(U).$$

The elements of $C^\infty(U)$ are called smooth functions*.*

We say that a function f lies in $C_c^\infty(U)$ if it lies in $C^\infty(U)$ and is compactly supported. We often call functions in $C_c^\infty(U)$ test functions *because of the way they are used to define the space of distributions later.*

Note that the partial differential operators ∂_{x_j} act

$$\partial_{x_j} : C^\infty(U) \to C^\infty(U), \quad C_c^\infty(U) \to C_c^\infty(U). \tag{1.25}$$

REMARK 1.10.**X** *In principle, most of the results involving the spaces C^∞ can be proved for the spaces C^k where k is large enough. But this means we will have to keep track of the value of k, which changes from place to place (as seen already from* (1.25)*: the operator ∂_{x_j} does not map C^k to itself). The space C^∞ provides a much cleaner way to develop the basic theory and, as we see very soon, it still contains a lot of functions.*

It is easy to give plenty of examples of functions in $C^\infty(\mathbb{R}^n)$; one can for example take any polynomial. Nontrivial functions in $C_c^\infty(\mathbb{R}^n)$ are a bit harder to construct because a lot of basic formulas produce functions which are real analytic and thus cannot be compactly supported (or even vanish on any ball) unless they are identically zero. A standard example of a function in $C_c^\infty(\mathbb{R}^n)$, and one which is used in the next section to construct many more functions in this space, is given by the 'bump function'

$$f(x) = \begin{cases} \exp\left(-\frac{1}{1-|x|^2}\right), & |x| < 1; \\ 0, & |x| \ge 1. \end{cases} \tag{1.26}$$

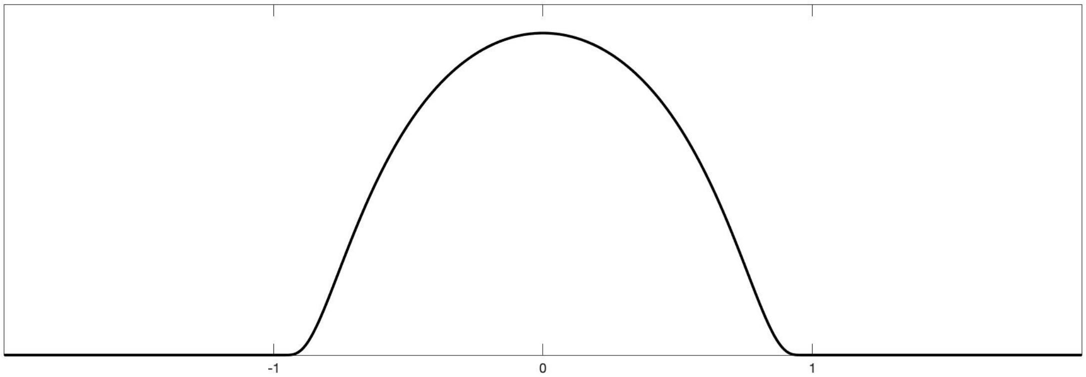

FIGURE 1.1. The 'bump function' (1.26).

The function (1.26) is plotted on Figure 1.1. See for example [**Hör03**, Lemma 1.2.3] for a proof that this function does indeed lie in $C_c^\infty(\mathbb{R}^n)$.

1.3. Convolution and approximation by smooth functions

We now discuss how to approximate 'rough' functions (e.g. those in C^0 or L^p) by 'smooth' functions (those in C_c^∞). One of the goals is to prove the following

THEOREM 1.11. *Let $U \Subset \mathbb{R}^n$. Then the space $C_c^\infty(U)$ is* dense *in the space $C_c^0(U)$, more precisely for each $f \in C_c^0(U)$ there exists a sequence $f_k \in C_c^\infty(U)$ such that $f_k \to f$ uniformly on U and all the supports* $\operatorname{supp} f_k$ *are contained in some k-independent compact subset of U.*

We will also show the L^p version of this statement, see Theorem 1.14 below.

1.3.1. Convolution. To show Theorem 1.11 we take a function $f \in C_c^0(U)$ and mollify it to get a function $f_n \in C_c^\infty(U)$. This is typically done using *convolution*, which is an operation on functions on $\mathbb{R}^n$ important in its own right. Thus we start with introducing convolution and studying its basic properties. We assume for now that the convolved functions are in $C_c^0(\mathbb{R}^n)$ but the integral below makes sense under much weaker assumptions. In fact, convolution appeared in these notes already in (1.2). See [**Hör03**, §1.3] and [**Str11**, §6.3.2] for more information about convolution.

DEFINITION 1.12. *Assume that $f, g \in C_c^0(\mathbb{R}^n)$. Define their* convolution $f * g \in L^\infty(\mathbb{R}^n)$ *by*

$$f * g(x) = \int_{\mathbb{R}^n} f(y) g(x-y)\, dy, \quad x \in \mathbb{R}^n. \tag{1.27}$$

Some standard properties of convolution are collected in

THEOREM 1.13. *1. For $f, g \in C^0_{\mathrm{c}}(\mathbb{R}^n)$ we have $f * g \in C^0_{\mathrm{c}}(\mathbb{R}^n)$ and*

$$\operatorname{supp}(f * g) \;\subset\; \operatorname{supp} f + \operatorname{supp} g \;:=\; \{x + y \mid x \in \operatorname{supp} f,\; y \in \operatorname{supp} g\}. \tag{1.28}$$

2. $f * (g * h) = (f * g) * h$ and $f * g = g * f$, that is convolution is associative and commutative.

3. If $f \in C^0_{\mathrm{c}}(\mathbb{R}^n)$ and $g \in C^1_{\mathrm{c}}(\mathbb{R}^n)$, then $f * g \in C^1_{\mathrm{c}}(\mathbb{R}^n)$ and

$$\partial_{x_j}(f * g) = f * (\partial_{x_j} g).$$

4. If $f \in C^0_{\mathrm{c}}(\mathbb{R}^n)$ and $g \in C^k_{\mathrm{c}}(\mathbb{R}^n)$, then $f * g \in C^k_{\mathrm{c}}(\mathbb{R}^n)$ and

$$\partial^\alpha_x (f * g) = f * (\partial^\alpha_x g) \quad \text{for all } \alpha,\; |\alpha| \le k.$$

PROOF. 1. We first check that $f * g$ is continuous. Let $x, \tilde{x} \in \mathbb{R}^n$. We compute

$$\begin{aligned} |(f * g)(x) - (f * g)(\tilde{x})| &= \left| \int_{\mathbb{R}^n} f(y)\big(g(x-y) - g(\tilde{x}-y)\big)\, dy \right| \\ &\le \|f\|_{L^1(\mathbb{R}^n)} \sup_y \big|g(x-y) - g(\tilde{x}-y)\big|. \end{aligned}$$

Since g lies in $C^0_{\mathrm{c}}(\mathbb{R}^n)$, it has a modulus of continuity ω_g, see (1.21). We then estimate

$$|(f * g)(x) - (f * g)(\tilde{x})| \le \|f\|_{L^1(\mathbb{R}^n)}\, \omega_g(|x - \tilde{x}|)$$

which shows that the function $f * g$ is continuous.

For the support property, we note first that the set $\{x \in \mathbb{R}^n \mid f * g(x) \neq 0\}$ is contained in $\operatorname{supp} f + \operatorname{supp} g$ since in order for $f * g(x)$ to be nonzero there must exist some $y \in \mathbb{R}^n$ such that $f(y) \neq 0$ and $g(x - y) \neq 0$. Next, the set $\operatorname{supp} f + \operatorname{supp} g$ is compact (as the image of the compact set $\operatorname{supp} f \times \operatorname{supp} g$ under the map $(x, y) \mapsto x + y$) and thus closed, so $\operatorname{supp}(f * g) \subset \operatorname{supp} f + \operatorname{supp} g$.

2. For associativity, we compute

$$\begin{aligned} f * (g * h)(x) &= \int_{\mathbb{R}^{2n}} f(y) g(z) h(x - y - z)\, dydz, \\ (f * g) * h(x) &= \int_{\mathbb{R}^{2n}} f(p) g(q - p) h(x - q)\, dpdq \end{aligned}$$

and make the change of variables $y = p, z = q - p$. Commutativity follows similarly by using the change of variables $y \mapsto x - y$.

3. The fact that $f * g$ is compactly supported already follows from (1.28), and we also know that $f * (\partial_{x_j} g)$ is continuous. Thus it remains to show that $\partial_{x_j}(f * g) = f * (\partial_{x_j} g)$. Denoting by $e_1, \ldots, e_n$ the canonical basis of $\mathbb{R}^n$, we compute for $x \in \mathbb{R}^n$ and $t \in \mathbb{R} \setminus \{0\}$

$$\frac{(f * g)(x + te_j) - f * g(x)}{t} = \int_{\mathbb{R}^n} f(y)\, \frac{g(x - y + te_j) - g(x - y)}{t}\, dy.$$

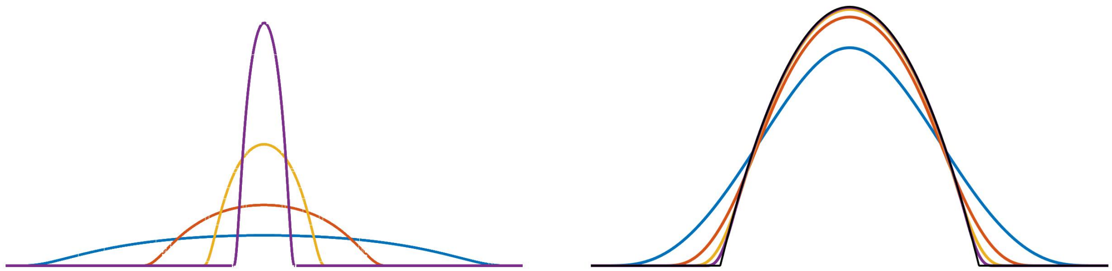

FIGURE 1.2. Left: the mollifying kernels χ_ε from (1.31) for $n = 1$ and several values of ε. Right: a function f on $\mathbb{R}$ (in black) and its successive mollifications f_ε. We have $f_\varepsilon \to f$ uniformly in x.

From the Mean Value Theorem and the fact that $\partial_{x_j} g$ is uniformly continuous we get

$$\frac{g(z + te_j) - g(z)}{t} \to \partial_{x_j} g(z) \quad \text{as } t \to 0 \quad \text{uniformly in } z \in \mathbb{R}^n. \tag{1.29}$$

Since f is bounded and compactly supported, we can pass to the limit under the integral sign and get

$$\frac{(f * g)(x + te_j) - f * g(x)}{t} \to \int_{\mathbb{R}^n} f(y) \partial_{x_j} g(x - y)\, dy = f * (\partial_{x_j} g)(x) \quad \text{as } t \to 0$$

which means that $\partial_{x_j}(f * g)(x) = f * (\partial_{x_j} g)(x)$.

4. This follows from the previous property by induction on k. □

From part 4 of Theorem 1.13 we see that

$$f \in C^0_c(\mathbb{R}^n),\ g \in C^\infty_c(\mathbb{R}^n) \quad \Longrightarrow \quad f * g \in C^\infty_c(\mathbb{R}^n). \tag{1.30}$$

That is, convolving a rough function with a smooth one produces a smooth result.

1.3.2. Mollification and the density theorems. We are now ready to give

PROOF OF THEOREM 1.11. 1. Fix a ‘bump function’

$$\chi \in C^\infty_c(\mathbb{R}^n), \quad \operatorname{supp} \chi \subset B(0, 1), \quad \int_{\mathbb{R}^n} \chi(x)\, dx = 1.$$

One way to construct it is to multiply the function (1.26) by a constant.

For $\varepsilon > 0$, define the rescaling (see Figure 1.2)

$$\chi_\varepsilon(x) := \varepsilon^{-n} \chi\Big(\frac{x}{\varepsilon}\Big), \quad \chi_\varepsilon \in C^\infty_c(\mathbb{R}^n), \quad \operatorname{supp} \chi_\varepsilon \subset B(0, \varepsilon). \tag{1.31}$$

Let $U \Subset \mathbb{R}^n$. Take arbitrary $f \in C^0_c(U)$ and extend it by 0 to a function in $C^0_c(\mathbb{R}^n)$ (which we still denote by f). Define the *mollifications* of f as

$$f_\varepsilon := f * \chi_\varepsilon. \tag{1.32}$$

(See Figure 1.2.) From (1.30) and (1.28), we see that for each $\varepsilon > 0$

$$f_\varepsilon \in C_c^\infty(\mathbb{R}^n), \quad \operatorname{supp} f_\varepsilon \subset \operatorname{supp} f + B(0,\varepsilon).$$

Since $\operatorname{supp} f \subset U$ is compact, for ε small enough (that is, smaller than the distance between $\operatorname{supp} f$ and $\mathbb{R}^n \setminus U$) we have $\operatorname{supp} f_\varepsilon \subset U$ and thus we can think of f_ε as a function in $C_c^\infty(U)$. We claim that

$$f_\varepsilon \to f \quad \text{as } \varepsilon \to 0+ \quad \text{uniformly on } \mathbb{R}^n. \tag{1.33}$$

Once we show (1.33), the proof of Theorem 1.11 is finished (we can for example take $f_k := f_\varepsilon$ with $\varepsilon = 1/k$ for large k).

2. To show (1.33), let ω_f be the modulus of continuity of f defined in (1.21). Take $x \in \mathbb{R}^n$ and estimate

$$\begin{aligned} |f(x) - f_\varepsilon(x)| &= \left| \int_{\mathbb{R}^n} \big(f(x) - f(x-y)\big)\chi_\varepsilon(y)\,dy \right| \\ &\leq \|\chi_\varepsilon\|_{L^1(\mathbb{R}^n)} \sup_{y\in B(0,\varepsilon)} \big|f(x) - f(x-y)\big| \\ &\leq \|\chi\|_{L^1(\mathbb{R}^n)}\,\omega_f(\varepsilon) \end{aligned} \tag{1.34}$$

Here in the first line we use the definition of convolution and the fact that $\int \chi_\varepsilon = 1$. In the second line we use that $\operatorname{supp}\chi_\varepsilon \subset B(0,\varepsilon)$. In the last line we use that $\|\chi_\varepsilon\|_{L^1} = \|\chi\|_{L^1}$. Now the expression on the last line of (1.34) is independent of x and converges to 0 as $\varepsilon \to 0+$, which gives the uniform convergence statement (1.33) and finishes the proof. □

We now give an L^p version of Theorem 1.11:

THEOREM 1.14. *Let $U \Subset \mathbb{R}^n$ and $1 \leq p < \infty$. Then the space $C_c^\infty(U)$ is dense in the space $L^p(U)$, more precisely for each $f \in L^p(U)$ there exists a sequence $f_k \in C_c^\infty(U)$ such that $f_k \to f$ in $L^p(U)$.*

PROOF. We do not give a detailed proof to avoid going too deep into the details of Lebesgue theory of integration. But here is a scheme of a proof:

- Using Theorem 1.11 and the fact that $C_c^0(U)$-convergence in that theorem implies convergence in $L^p(U)$, we see that it suffices to show that $C_c^0(U)$ is dense in $L^p(U)$.
- A standard fact in the theory of Lebesgue integral is that the space of simple L^p functions is dense in $L^p(U)$, where 'simple' means that the function only takes finitely many different values. So it remains to show that any simple function can be approximated in $L^p(U)$ by functions in $C_c^0(U)$, and this immediately reduces to approximating indicator functions $\mathbf{1}_A$ where $A \subset U$ is measurable of finite measure.

- For A as above, the regularity property of Lebesgue measure implies that for each $\varepsilon > 0$ there exists a compact set K and an open set V such that $K \subset A \subset V \subset U$. There exists a function $g \in C^0_c(U)$ such that $g = 1$ on K, $0 \leq g \leq 1$ everywhere, and $\operatorname{supp} g \subset V$ (it can be constructed for example as a function of distance to K). Then $\|\mathbf{1}_A - g\|_{L^p(U)} \leq \varepsilon^{1/p}$ and, since ε can be chosen arbitrarily small, we can approximate $\mathbf{1}_A$ in L^p by functions in $C^0_c(U)$.

□

1.3.3. More on smooth compactly supported functions. We finally give two more statements about the spaces C^∞_c. The first one is the existence of smooth partitions of unity:

THEOREM 1.15. *Let $U_1, \dots, U_m \Subset \mathbb{R}^n$ and $K \subset U_1 \cup \dots \cup U_m$ be a compact set. Then there exist functions*

$$\chi_j \in C^\infty_c(U_j),\ j = 1, \dots, m, \quad \chi_j \geq 0, \quad \chi_1 + \dots + \chi_m \leq 1,$$
$$\chi_1 + \dots + \chi_m = 1 \quad \text{in a neighborhood of } K.$$

The last statement above can be alternatively written as $\operatorname{supp}(1 - \chi_1 - \dots - \chi_m) \cap K = \emptyset$.

For the proof of Theorem 1.15, see for example [**Hör03**, Theorem 1.4.5]. One of the key points of the proof is that one can construct a function in $C^\infty_c(\mathbb{R}^n)$ approximating the indicator function of a set A by taking the convolution $\mathbf{1}_A * \chi_\varepsilon$ for small $\varepsilon > 0$ similarly to the proof of Theorem 1.11.

The second statement, which is crucial for the development of the theory of distributions, tells us that a function $f \in L^1_{\text{loc}}(U)$ is determined uniquely by the integrals $\int_U f\varphi$ for all the functions $\varphi \in C^\infty_c(U)$:

THEOREM 1.16. *Let $U \Subset \mathbb{R}^n$, $f \in L^1_{\text{loc}}(U)$, and assume that*

$$\int_U f(x)\varphi(x)\,dx = 0 \quad \text{for all } \varphi \in C^\infty_c(U). \tag{1.35}$$

Then $f(x) = 0$ for almost every $x \in U$.

PROOF. As with Theorem 1.14 we do not give a detailed proof to avoid going too much into Lebesgue integration theory. But here are the sketches of two different proofs:

- The equation (1.35) actually holds for all $\varphi \in C^0_c(U)$. Indeed, by Theorem 1.11 we can take a sequence $\varphi_k \in C^\infty_c(U)$ which converges to φ. We have $\int_U f\varphi_k = 0$ for all k and we can pass to the limit under the integral to get $\int_U f\varphi = 0$ as well. Now, the uniqueness part of the Riesz representation theorem for measures gives that $f = 0$ almost everywhere.

- It suffices to show that $\psi f = 0$ almost everywhere for any $\psi \in C^\infty_c(U)$. Consider a mollifying kernel χ_ε as in (1.31) and define the convolution $(\psi f) * \chi_\varepsilon$, where we extend ψf by zero to a function on $\mathbb{R}^n$. Using (1.35) for the function $\varphi(y) = \psi(y)\chi_\varepsilon(x-y)$, we see that $(\psi f) * \chi_\varepsilon(x) = 0$ for all $x \in \mathbb{R}^n$. On the other hand, we have $(\psi f) * \chi_\varepsilon(x) \to \psi f(x)$ as $\varepsilon \to 0+$ for almost every x:

$$
\begin{aligned}
|\psi f(x) - (\psi f) * \chi_\varepsilon(x)| &= \left| \int_{\mathbb{R}^n} \big(\psi f(x) - \psi f(y)\big) \chi_\varepsilon(x-y)\, dy \right| \\
&\leq \sup|\chi| \cdot \varepsilon^{-n} \int_{B(x,\varepsilon)} \big|\psi f(x) - \psi f(y)\big|\, dy \to 0 \quad \text{for a.e. } x
\end{aligned}
$$

where the last step follows from the Lebesgue Differentiation Theorem. Thus $\psi f = 0$ almost everywhere. □

We finish this section by reviewing a simple yet very powerful tool, integration by parts.

THEOREM 1.17. *Let $U \Subset \mathbb{R}^n$. Assume that $f \in C^1(U)$ and $g \in C^1_c(U)$. Then we have for all j,*

$$
\int_U (\partial_{x_j} f(x)) g(x)\, dx = -\int_U f(x)(\partial_{x_j} g(x))\, dx. \tag{1.36}
$$

PROOF. We will show that

$$
\int_U \partial_{x_j} h(x)\, dx = 0 \quad \text{for all } h \in C^1_c(U). \tag{1.37}
$$

The identity (1.36) follows by applying (1.37) to the function $h := fg$.

To show (1.37), extend h by zero to a function in $C^\infty_c(\mathbb{R}^n)$, which we still denote by h. For notational convenience assume that $j = 1$ and write $x = (x_1, x')$ where $x' \in \mathbb{R}^n$. Now by Fubini's Theorem

$$
\int_{\mathbb{R}^n} \partial_{x_1} h(x)\, dx = \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \partial_{x_1} h(x_1, x')\, dx_1 dx' = 0
$$

since $\int_{\mathbb{R}} \partial_{x_1}\varphi(x_1)\, dx_1 = 0$ for any $\varphi \in C^1_c(\mathbb{R})$ by the Fundamental Theorem of Calculus. □

REMARK 1.18. *If U is bounded with a smooth boundary, then the Divergence Theorem for the vector field $h(x)e_j$, where e_j is the j-th coordinate vector on $\mathbb{R}^n$, gives the following version of (1.37) for $h \in C^1(\overline{U})$ (that is, C^1 up to the boundary of U) which is not necessarily compactly supported:*

$$
\int_U \partial_{x_j} h(x)\, dx = \int_{\partial U} h(x) n_j(x)\, dS(x). \tag{1.38}
$$

Here $n_j(x)$ is the j-th coordinate of the outward unit normal vector to ∂U at x and dS is the area measure on ∂U. This in turn gives the integration by parts identity for $f, g \in C^1(\overline{U})$

$$\int_U (\partial_{x_j} f(x)) g(x)\, dx = \int_{\partial U} f(x) g(x) n_j(x)\, dS(x) - \int_U f(x)(\partial_{x_j} g(x))\, dx. \tag{1.39}$$

1.4. Notes and exercises

For a quick review of classical differential calculus, see [**Hör03**, Chapter 1]. This is in particular where our proof of Theorem 1.16 comes from, see [**Hör03**, Theorem 1.2.5].

EXERCISE 1.1. (4 = 1 + 1 + 2 pts) *Assume that $f \in C_c^\infty(\mathbb{R}^3)$ (see §1.2.4 below). Let u be defined by the formula (1.2). Show that u solves the equation (1.1), following the steps below:*

(a) *Show that $\Delta E(x) = 0$ for all $x \in \mathbb{R}^3 \setminus \{0\}$.*

(b) *Show that $u \in C^\infty(\mathbb{R}^3)$ and*

$$\Delta u(x) = \int_{\mathbb{R}^3} E(x-y)\Delta f(y)\, dy.$$

(Hint: make the change of variables $y \mapsto x - y$ in the integral.)

(c) *Fix $x \in \mathbb{R}^3$ and let $\Omega_\varepsilon := \{y \in \mathbb{R}^3 : \varepsilon \le |x-y| \le \varepsilon^{-1}\}$ for small $\varepsilon > 0$. Write*

$$\Delta u(x) = \lim_{\varepsilon \to 0+} \int_{\Omega_\varepsilon} E(x-y)\Delta f(y)\, dy$$

Use the Divergence Theorem/integration by parts twice to write $\Delta u(x)$ as an integral over the sphere $\partial B(x, \varepsilon)$. Letting $\varepsilon \to 0+$, show that $\Delta u = f$.

EXERCISE 1.2. (1 pt) *Let $U := (-1, 1) \subset \mathbb{R}$. Show that the space $C_c^0(U)$ is not complete with respect to the sup-norm.*

EXERCISE 1.3. (1 pt) *Let $U := (-1, 1) \subset \mathbb{R}$. Show that $C_c^\infty(U)$ is not dense in $L^\infty(U)$.*

CHAPTER 2

Basics of distribution theory

2.1. Definition of distributions

We are now ready to introduce distributions, which are one of the central objects of this course. The definition below is somewhat technical and some philosophical explanations are provided later. But the general idea is: the space of distributions on an open set U is the dual to the space of smooth compactly supported functions $C^\infty_c(U)$ (the latter also known as *test functions*), i.e. the space of *continuous linear functionals* on $C^\infty_c(U)$. The notion of convergence on $C^\infty_c(U)$ is complicated (we only study sequential convergence, see §2.2.1 below), so we first define a distribution as a *bounded* linear functional, with the boundedness made precise in

DEFINITION 2.1. *Let $U \Subset \mathbb{R}^n$ and assume that*

$$u : C^\infty_c(U) \to \mathbb{C}$$

is a linear functional. We say that u is a distribution *on U if for each compact set $K \subset U$ there exist constants C, N such that*

$$\big|u(\varphi)\big| \leq C\|\varphi\|_{C^N} \quad \textit{for all } \varphi \in C^\infty_c(U) \quad \textit{such that } \operatorname{supp}\varphi \subset K. \tag{2.1}$$

Here the C^N norm $\|\bullet\|_{C^N}$ is defined in (1.23) above.

We denote the set of all distributions on U by

$$\mathcal{D}'(U).$$

This notation goes back to Laurent Schwartz, the inventor of the theory of distributions: he denoted $\mathcal{D}(U) := C^\infty_c(U)$, and $\mathcal{D}'(U)$ was its dual space.

PROPOSITION 2.2.$^{\mathbf{S}}$ *$\mathcal{D}'(U)$ is a vector space.*

The relation of distributions to functions comes from the following embedding of locally integrable functions into distributions:

PROPOSITION 2.3. *Let $U \Subset \mathbb{R}^n$ and $f \in L^1_{\mathrm{loc}}(U)$. Define the linear functional*

$$\widetilde{f} : C^\infty_c(U) \to \mathbb{C}, \quad \widetilde{f}(\varphi) = \int_U f(x)\varphi(x)\,dx \quad \textit{for all } \varphi \in C^\infty_c(U). \tag{2.2}$$

Then $\widetilde{f}$ is a distribution in $\mathcal{D}'(U)$ and the map $f \mapsto \widetilde{f}$ is linear and injective.

PROOF. The functional $\widetilde{f}$ is a distribution because it satisfies the following bound of type (2.1) for any compact $K \subset U$ and $\varphi \in C_c^\infty(U)$ such that $\operatorname{supp}\varphi \subset K$:

$$|\widetilde{f}(\varphi)| \leq \|\mathbf{1}_K f\|_{L^1} \cdot \|\varphi\|_{C^0}.$$

Linearity of the map $f \mapsto \widetilde{f}$ is immediate, and injectivity follows from Theorem 1.16. □

We now introduce important notation to be used throughout the rest of these notes:

- For $f \in L^1_{\mathrm{loc}}(U)$, we identify the function f with the distribution $\widetilde{f}$ from Proposition 2.3;
- For $f \in L^1_{\mathrm{loc}}(U)$ and $\varphi \in C_c^\infty(U)$, we define the pairing
$$(f, \varphi) := \int_U f(x)\varphi(x)\,dx; \tag{2.3}$$
- For $u \in \mathcal{D}'(U)$ and $\varphi \in C_c^\infty(U)$, we define the pairing
$$(u, \varphi) := u(\varphi). \tag{2.4}$$
Sometimes we may even write (which is something you will see in textbooks and papers using distribution theory, so you might as well get used to it)
$$\int_U u(x)\varphi(x)\,dx := u(\varphi)$$
which still means the distributional pairing and not an actual Lebesgue integral since u might not even be a function.

This notation might be confusing at first, but it makes the presentation much cleaner. It is also represents the following philosophical point underlying the theory of distributions. To specify a function $f : U \to \mathbb{C}$, we need to answer the following question:

$$\text{For any } x \in U, \text{ what is the value of } f \text{ at the point } x? \tag{2.5}$$

To specify a distribution $u \in \mathcal{D}'(U)$, we need to answer a different question:

$$\text{For any test function } \varphi \in C_c^\infty(U), \text{ what is the integral } \int_U u(x)\varphi(x)\,dx\ ? \tag{2.6}$$

The question (2.6) provides weaker information than (2.5), which corresponds to the fact that there are plenty of distributions which are not functions (as we will see shortly). In fact, we cannot even answer the question (2.5) for a function $f \in L^1_{\mathrm{loc}}(U)$ because the space L^1 is defined modulo equality almost everywhere.

Moreover, the question (2.6) more physically relevant because if u is a physical quantity (for example, the temperature in some reservoir) then, since a physical sensor has positive size, any measurement of u will produce an integral featuring u instead of the value of u at a single point. (Not to mention that if we go to a subatomic scale then the notion of the temperature at a given point does not make sense – all we can really define is the average temperature in a macroscopic region, which is a rougher version of (2.6).)

A standard example of a distribution which is not a function is given by the *Dirac delta function*:

Definition 2.4. *Let $U \Subset \mathbb{R}^n$ and $y \in U$ be a point. Define the distribution $\delta_y \in \mathcal{D}'(U)$ by*

$$(\delta_y, \varphi) := \varphi(y) \quad \textit{for all } \varphi \in C_c^\infty(U).$$

This is something that has been used in physics much earlier than the mathematically rigorous development of the theory of distributions: if we think of a distribution in $\mathcal{D}'(U)$ as, say, the density of electric charge, then δ_y is the density of the point charge centered at y.

To see that δ_y is not a function, assume the contrary: $\delta_y = f$ for some $f \in L^1_{\mathrm{loc}}(U)$. Take arbitrary $\chi \in C_c^\infty(\mathbb{R}^n)$ with $\chi(0) = 1$ and consider the test function $\varphi_\varepsilon(x) := \chi((x-y)/\varepsilon)$ which lies in $C_c^\infty(U)$ for sufficiently small $\varepsilon > 0$. Then $(f, \varphi_\varepsilon) \to 0$ as $\varepsilon \to 0+$ by the Dominated Convergence Theorem but $(\delta_y, \varphi_\varepsilon) = 1$.

2.2. Distributions and convergence

2.2.1. Sequential convergence of test functions. If X is a Banach space, then a functional $u : X \to \mathbb{C}$ is bounded if and only if it is continuous. We now give an analog of this statement for distributions. A proper way to do this would be to define a topology on $C_c^\infty(U)$ but it turns out to be an *inductive limit topology* which is a bit complicated to describe (related to property (1) in Definition 2.5 below). So we instead settle for defining convergence of sequences of elements in $C_c^\infty(U)$, which is enough for our applications. See for example [**RS81**, §V.4] or [**Rud91**, §6.2] for the definition of the inductive limit topology on $C_c^\infty(U)$.

Definition 2.5. *Let $U \Subset \mathbb{R}^n$ and assume that $\varphi_k \in C_c^\infty(U)$ is a sequence and $\varphi \in C_c^\infty(U)$. We say that*

$$\varphi_k \to \varphi \quad \textit{as } k \to \infty \quad \textit{in } C_c^\infty(U)$$

if the following two conditions hold:

1. *there exists compact $K \subset U$ such that $\operatorname{supp} \varphi_k \subset K$ for all k, and*
2. *we have $\|\varphi_k - \varphi\|_{C^N} \to 0$ as $k \to \infty$ for all N.*

We use the sequential notion of continuity to establish the equivalence of boundedness and continuity for functionals on $C_c^\infty(U)$:

Proposition 2.6. *Let $u : C_c^\infty(U) \to \mathbb{C}$ be a linear functional. Then the following are equivalent:*

(1) *u is a distribution, that is it satisfies the norm bounds (2.1);*
(2) *for each sequence $\varphi_k \in C_c^\infty(U)$, if $\varphi_k \to 0$ in $C_c^\infty(U)$, then $(u, \varphi_k) \to 0$.*

PROOF. (1) ⇒ (2): Assume that $\varphi_k \to 0$ in $C_c^\infty(U)$. Then in particular there exists $K \Subset U$ such that $\operatorname{supp} \varphi_k \subset K$ for all K. The norm bound (2.1) implies that there exist C, N such that for all k

$$|(u, \varphi_k)| \leq C\|\varphi_k\|_{C^N}.$$

The right-hand side goes to 0 as $k \to \infty$, so $(u, \varphi_k) \to 0$ as needed.

(2) ⇒ (1): We argue by contradiction. Assume that u does not satisfy the norm bounds (2.1), that is there exists $K \Subset U$ such that for any choice of C, N there exists $\varphi \in C_c^\infty(U)$ such that $\operatorname{supp} \varphi \subset K$ and $|(u, \varphi)| \geq C\|\varphi\|_{C^N}$. Choosing $C = N = k$ and dividing φ by $u(\varphi)$, we construct a sequence

$$\varphi_k \in C_c^\infty(U), \quad \operatorname{supp} \varphi_k \subset K, \quad (u, \varphi_k) = 1, \quad \|\varphi_k\|_{C^k} \leq \tfrac{1}{k}.$$

The sequence φ_k converges to 0 in $C_c^\infty(U)$ since for all $k \geq m$ we have $\|\varphi_k\|_{C^m} \leq \|\varphi_k\|_{C^k} \leq \frac{1}{k}$. Thus u does not satisfy the sequential continuity property (2). □

2.2.2. Weak convergence of distributions. We next discuss convergence of sequences of distributions. This is a very weak notion of convergence, in contrast with convergence in $C_c^\infty(U)$ which is very strong.

DEFINITION 2.7. *Let $U \Subset \mathbb{R}^n$, $u_k \in \mathcal{D}'(U)$ be a sequence, and $u \in \mathcal{D}'(U)$. We say that*

$$u_k \to u \quad \text{as } k \to \infty \quad \text{in } \mathcal{D}'(U)$$

if we have

$$(u_k, \varphi) \to (u, \varphi) \quad \text{as } k \to \infty \quad \text{for all } \varphi \in C_c^\infty(U).$$

We give a few examples of weak convergence:

PROPOSITION 2.8. *If $u_k, u \in L^1_{\text{loc}}(U)$ satisfy $u_k(x) \to u(x)$ for almost every $x \in U$, and there exists $g \in L^1_{\text{loc}}(U)$ such that $|u_k(x)| \leq g(x)$ for all k, then $u_k \to u$ in $\mathcal{D}'(U)$.*

PROOF. This follows immediately from the Dominated Convergence Theorem. □

PROPOSITION 2.9. *Define the functions $u_k \in L^1_{\text{loc}}(\mathbb{R})$ by*

$$u_k(x) := k\mathbf{1}_{[-1/k,1/k]}(x).$$

(This is a classical example of a sequence which does not satisfy the assumptions of the Dominated Convergence Theorem.) Then $u_k \to 2\delta_0$ in $\mathcal{D}'(\mathbb{R})$.

PROOF. Take arbitrary $\varphi \in C_c^\infty(\mathbb{R})$. Then

$$(u_k, \varphi) = k \int_{-1/k}^{1/k} \varphi(x)\, dx \;\to\; 2\varphi(0) = (2\delta_0, \varphi)$$

by the continuity of φ. □

PROPOSITION 2.10. *Define the functions $u_k \in L^1_{\mathrm{loc}}(\mathbb{R})$ by*

$$u_k(x) = e^{ikx}.$$

Then $u_k \to 0$ in $\mathcal{D}'(\mathbb{R})$.

PROOF. Take arbitrary $\varphi \in C_c^\infty(\mathbb{R})$. Then

$$(u_k, \varphi) = \int_{\mathbb{R}} e^{ikx} \varphi(x)\, dx = \widehat{\varphi}(-k)$$

where $\widehat{\varphi}$ is the Fourier transform of φ. We will study the Fourier transform in detail later but for now we note that integration by parts using $e^{ikx} = -(i/k)\partial_x e^{ikx}$ gives $|\widehat{\varphi}(-k)| = \mathcal{O}(1/k)$ and thus $(u_k, \varphi) \to 0$. □

While the convergence in Definition 2.7 is indeed very weak, it does imply a (weak) uniform bound on the sequence u_k – see Theorem 4.16 below.

2.3. Localization

We now discuss how the space $\mathcal{D}'(U)$ depends on the open set U and the related question of localization of distributions. This is easy for functions (when we have access to values at points) and takes more effort for distributions.

As with many concepts later in these notes, we start by reviewing what happens for functions and then generalize to distributions. Let $V \Subset U \Subset \mathbb{R}^n$. Then we have the restriction operator

$$L^1_{\mathrm{loc}}(U) \to L^1_{\mathrm{loc}}(V), \quad f \mapsto f|_V.$$

Its generalization to distributions is given by

DEFINITION 2.11. *Let $V \Subset U \Subset \mathbb{R}^n$. For $u \in \mathcal{D}'(U)$, define its* restriction *$u|_V \in \mathcal{D}'(V)$ as follows:*

$$(u|_V, \varphi) := (u, \varphi) \quad \textit{for all } \varphi \in C_c^\infty(V).$$

Here $C_c^\infty(V)$ is considered a subset of $C_c^\infty(U)$ as follows: for $\varphi \in C_c^\infty(V)$ we extend it by 0 to produce an element of $C_c^\infty(U)$ (owing to compactness of support).

PROPOSITION 2.12.[S] *The restriction map* $r_{V,U} : \mathcal{D}'(U) \to \mathcal{D}'(V)$ *from Definition 2.11 is linear and satisfies*

$$r_{U,U} = I,$$
$$r_{W,V} r_{V,U} = r_{W,U} \quad \text{for all } W \Subset V \Subset U \Subset \mathbb{R}^n.$$

(In algebraic terms, we have obtained a presheaf *– but if you don't know what this means, do not worry since we won't be using this terminology later.)*

The next theorem states that if we have an open cover of U, then a distribution on U is uniquely determined from its restrictions to the elements of the cover. That is, if we construct a distribution locally (i.e. on each set of the cover) then we can recover it globally. The proof would be straightforward for functions but takes more effort for distributions, using in a key way partitions of unity.

THEOREM 2.13 (Sheaf property of distributions). *Assume that* $\mathcal{J}$ *is an arbitrary set and*

$$U_j \Subset \mathbb{R}^n \quad \text{for } j \in \mathcal{J}, \qquad U = \bigcup_{j \in \mathcal{J}} U_j.$$

Assume next that we are given $u_j \in \mathcal{D}'(U_j)$, $j \in \mathcal{J}$, *satisfying the compatibility conditions*

$$u_j|_{U_j \cap U_\ell} = u_\ell|_{U_j \cap U_\ell} \quad \text{for all } j, \ell \in \mathcal{J}. \tag{2.7}$$

Then there exists unique $u \in \mathcal{D}'(U)$ *such that*

$$u|_{U_j} = u_j \quad \text{for all } j \in \mathcal{J}. \tag{2.8}$$

PROOF. 1. We first show uniqueness of u, which can be reformulated as follows:

$$u \in \mathcal{D}'(U), \quad u|_{U_j} = 0 \quad \text{for all } j \in \mathcal{J} \quad \Longrightarrow \quad u = 0. \tag{2.9}$$

Recalling Definition 2.11, we can reformulate (2.9) as follows:

$$(u, \psi) = 0 \quad \text{for all } j \in \mathcal{J}, \psi \in C^\infty_{\mathrm{c}}(U_j) \quad \Rightarrow \quad (u, \varphi) = 0 \quad \text{for all } \varphi \in C^\infty_{\mathrm{c}}(U). \tag{2.10}$$

Take abritrary $\varphi \in C^\infty_{\mathrm{c}}(U)$. We can decompose it as

$$\varphi = \sum_{j \in \mathcal{J}} \varphi_j, \quad \varphi_j \in C^\infty_{\mathrm{c}}(U_j), \tag{2.11}$$

where only finitely many of φ_j are nonzero. Indeed, since $\operatorname{supp}\varphi$ is compact and covered by the open sets U_j, there exists a finite set $\mathcal{J}' \subset \mathcal{J}$ such that $\operatorname{supp}\varphi \subset \bigcup_{j \in \mathcal{J}'} U_j$. Using Theorem 1.15, we take a partition of unity

$$\chi_j \in C^\infty_{\mathrm{c}}(U_j),\ j \in \mathcal{J}', \quad \sum_{j \in \mathcal{J}'} \chi_j = 1 \quad \text{on } \operatorname{supp}\varphi.$$

Multiplying the last identity by φ, we get (2.11) if we put $\varphi_j := \chi_j \varphi$ for $j \in \mathcal{J}'$ and $\varphi_j := 0$ otherwise.

Pairing (2.11) with u, we get

$$(u, \varphi) = \sum_{j \in \mathcal{J}} (u, \varphi_j).$$

If the assumption in (2.10) holds, then $(u, \varphi_j) = 0$ for all j, which gives $(u, \varphi) = 0$ as needed.

2. It remains to show that given $u_j \in \mathcal{D}'(U_j)$ satisfying the compatibility conditions (2.7), there exists $u \in \mathcal{D}'(U)$ satisfying (2.8). To define u, we need to specify (u, φ) for each $\varphi \in C_c^\infty(U)$. Take such φ and decompose it as in (2.11):

$$\varphi = \sum_{j \in \mathcal{J}} \chi_j \varphi, \quad \chi_j \in C_c^\infty(U_j), \tag{2.12}$$

where only finitely many χ_j are nonzero. We then put

$$(u, \varphi) := \sum_{j \in \mathcal{J}} (u_j, \chi_j \varphi). \tag{2.13}$$

The rest of the proof proceeds in several steps:

- The value of (u, φ) from (2.13) does not depend on the choice of the decomposition (2.12). Indeed, assume that we have a different decomposition

$$\varphi = \sum_{j' \in \mathcal{J}} \widetilde{\chi}_{j'} \varphi, \quad \widetilde{\chi}_{j'} \in C_c^\infty(U_{j'}).$$

  We write

$$\begin{aligned} \sum_{j \in \mathcal{J}} (u_j, \chi_j \varphi) &= \sum_{j, j' \in \mathcal{J}} (u_j, \chi_j \widetilde{\chi}_{j'} \varphi) \\ &= \sum_{j, j' \in \mathcal{J}} (u_{j'}, \chi_j \widetilde{\chi}_{j'} \varphi) \\ &= \sum_{j' \in \mathcal{J}} (u_{j'}, \widetilde{\chi}_{j'} \varphi) \end{aligned}$$

  giving the required independence. Here in the first equality above we use that $\chi_j \varphi = \sum_{j' \in \mathcal{J}} \chi_j \widetilde{\chi}_{j'} \varphi$. In the second equality we use the compatibility conditions (2.7): we have $\chi_j \widetilde{\chi}_{j'} \varphi \in C_c^\infty(U_j \cap U_{j'})$ and the restrictions of u_j and $u_{j'}$ to $U_j \cap U_{j'}$ are equal. Finally, in the last equality we use that $\widetilde{\chi}_{j'} \varphi = \sum_{j \in \mathcal{J}} \chi_j \widetilde{\chi}_{j'} \varphi$.
- ($^{\mathbf{S}}$) The map $\varphi \mapsto (u, \varphi)$ is linear. Indeed, take any $\varphi^{(1)}, \varphi^{(2)} \in C_c^\infty(U)$ and $a_1, a_2 \in \mathbb{C}$. Take a partition of unity (with only finitely many nonzero elements as before)

$$\chi_j \in C_c^\infty(U_j), \quad \sum_{j \in \mathcal{J}} \chi_j = 1 \quad \text{on } \operatorname{supp} \varphi^{(1)} \cup \operatorname{supp} \varphi^{(2)}.$$

Then (2.12) holds for $\varphi^{(1)}$, $\varphi^{(2)}$, and their linear combination $a_1\varphi^{(1)} + a_2\varphi^{(2)}$. By (2.13) and since the maps $\varphi \to (u_j, \varphi)$ are linear, we have

$$\begin{aligned}(u, a_1\varphi^{(1)} + a_2\varphi^{(2)}) &= \sum_{j\in\mathcal{J}} \big(u_j, \chi_j(a_1\varphi^{(1)} + a_2\varphi^{(2)})\big)\\ &= a_1\sum_{j\in\mathcal{J}}(u_j, \chi_j\varphi^{(1)}) + a_2\sum_{j\in\mathcal{J}}(u_j, \chi_j\varphi^{(2)})\\ &= a_1(u,\varphi^{(1)}) + a_2(u,\varphi^{(2)})\end{aligned}$$

which shows linearity.

- ($^{\mathbf{S}}$) The linear map u satisfies the bounds (2.1) and thus defines a distribution in $\mathcal{D}'(U)$. Indeed, take any $K \Subset U$. Fix a partition of unity

$$\chi_j \in C_c^\infty(U_j), \quad \sum_{j\in\mathcal{J}'} \chi_j = 1 \quad \text{on } K,$$

where $\mathcal{J}' \subset \mathcal{J}$ is a finite set. Since each of the finitely many distributions u_j, $j \in \mathcal{J}'$, satisfies the bounds (2.1), we can find C, N such that

$$|(u_j,\psi)| \le C\|\psi\|_{C^N} \quad \text{for all } j \in \mathcal{J}',\ \psi \in C_c^\infty(U_j),\ \operatorname{supp}\psi \subset \operatorname{supp}\chi_j.$$

Then for each $\varphi \in C_c^\infty(U)$ with $\operatorname{supp}\varphi \subset K$ we have by (2.13)

$$\begin{aligned}|(u,\varphi)| &\le \sum_{j\in\mathcal{J}'} |(u_j,\chi_j\varphi)|\\ &\le C\sum_{j\in\mathcal{J}'} \|\chi_j\varphi\|_{C^N}\\ &\le C'\|\varphi\|_{C^N}\end{aligned}$$

for some constant C' depending only on C, N, and the functions χ_j. This gives the bounds (2.1).

- We have $u|_{U_j} = u_j$ for all $j \in \mathcal{J}$, that is $(u,\varphi) = (u_j,\varphi)$ for each $\varphi \in C_c^\infty(U_j)$. For such φ we have the decomposition (2.12) if we choose $\chi_j \in C_c^\infty(U_j)$ with $\chi_j = 1$ on $\operatorname{supp}\varphi$ and put $\chi_{j'} = 0$ for all $j' \neq j$. Then (2.13) gives

$$(u,\varphi) = (u_j,\chi_j\varphi) = (u_j,\varphi)$$

finishing the proof. □

2.4. Notes and exercises

The modern theory of distributions was developed by Schwartz in the 1950s (and was included in the citation for his Fields medal), see [**Sch50, Sch57**]. There have been various precursors to this theory, most notably the definition of weak derivatives of functions by Sobolev in [**Sob36**] in the context of existence and uniqueness theorems for hyperbolic equations.

Our presentation largely follows [**Hör03**, §§2.1–2.2] and [**FJ98**, §§1.3–1.4].

EXERCISE 2.1. (1 pt) *Let* $U \Subset \mathbb{R}^n$ *and assume that* $u \in \mathcal{D}'(U)$ *satisfies the bound*

$$|(u, \varphi)| \leq C\|\varphi\|_{L^2(U)}$$

for some constant C *and all* $\varphi \in C_c^\infty(U)$. *Show that* $u \in L^2(U)$. *(Hint: use the Continuous Linear Extension theorem from functional analysis.)*

EXERCISE 2.2. (1 pt) *Let* $\chi \in C_c^\infty(\mathbb{R}^n)$ *satisfy* $\int_{\mathbb{R}^n} \chi = 1$. *Define*

$$\chi_\varepsilon(x) := \varepsilon^{-n}\chi(x/\varepsilon), \quad \varepsilon > 0.$$

Show that $\chi_\varepsilon \to \delta_0$ *in* $\mathcal{D}'(\mathbb{R}^n)$ *as* $\varepsilon \to 0+$.

EXERCISE 2.3. (2 pts) *Assume that the sequence* $\{a_k\}_{k\in\mathbb{Z}}$ *satisfies*

$$|a_k| \leq C(1 + |k|)^N \quad \text{for some constants} \quad C, N.$$

Show that the Fourier series

$$\sum_{k\in\mathbb{Z}} a_k e^{ikx}$$

converges in $\mathcal{D}'(\mathbb{R})$.

EXERCISE 2.4. (2 pts) *Let* $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^{n'}$ *be open and consider a linear operator*

$$A : C_c^\infty(U) \to C_c^\infty(V).$$

Show that the following two definitions of continuity of A *are equivalent:*

(1) *the following two conditions both hold:*
   (a) *for every* $K \Subset U$ *there exists* $K' \Subset V$ *such that for all* $\varphi \in C_c^\infty(U)$ *with* $\operatorname{supp}\varphi \subset K$, *we have* $\operatorname{supp}(A\varphi) \subset K'$ *(we can call this 'uniform control on compact support'); and*
   (b) *for every* $K \Subset U$ *and* $N \in \mathbb{N}$ *there exist* $C > 0$, $N' \in \mathbb{N}$ *such that we have the seminorm bound*

$$\|A\varphi\|_{C^N(V)} \leq C\|\varphi\|_{C^{N'}(U)} \quad \text{for all } \varphi \in C_c^\infty(U) \text{ with } \operatorname{supp}\varphi \subset K;$$

(2) *for each sequence* $\varphi_k \in C_c^\infty(U)$ *such that* $\varphi_k \to 0$ *in* $C_c^\infty(U)$, *we have* $A\varphi_k \to 0$ *in* $C_c^\infty(V)$ *(this is called 'sequential continuity').*

(Hint: for the direction (2) $\Rightarrow$ *(1) you can argue by contradiction: if either 1(a) or 1(b) fails then construct a sequence* φ_k *which violates sequential continuity. In case of 1(a) it helps to take a sequence of compact subsets* K_ℓ *exhausting* V *(see (1.14)): if 1(a) fails then there exists* $K \subset U$ *such that neither of the sets* K_ℓ *will work as* K'.*)*

Exercise 2.5. (1 pt) *Show that*

$$u(\varphi) = \sum_{k=1}^{\infty} \partial_x^k \varphi(1/k), \quad \varphi \in C_{\mathrm{c}}^{\infty}((0,\infty))$$

defines a distribution on $(0,\infty)$ *but this distribution does not extend to* $\mathbb{R}$*, that is there exists no* $v \in \mathcal{D}'(\mathbb{R})$ *such that* $u = v|_{(0,\infty)}$*. (Hint: pair* u *with a dilated cutoff function whose support contains* $1/k$ *but no other points of the form* $1/j$*,* $j \in \mathbb{N}$*.)*

CHAPTER 3

Operations with distributions

To quote from [**Hör03**], *"In differential calculus one encounters immediately the unpleasant fact that every function is not differentiable. The purpose of distribution theory is to remedy this flaw; indeed, the space of distributions is essentially the smallest extension of the space of continuous functions where differentiation is always well defined."*

In this chapter we learn how to differentiate distributions and also how to multiply them by smooth functions. This will follow two general principles:

- Uniqueness of extension from a dense set: for any operator on the space of smooth functions there is at most one continuous extension of this operator to distributions, because any distribution can be approximated in $\mathcal{D}'(U)$ by functions in $C^\infty_{\mathrm{c}}(U)$.
- Duality: one can extend many operators $A : C^\infty(U) \to C^\infty(U)$ to distributions by defining $(Au, \varphi) = (u, A^t\varphi)$ for all $u \in \mathcal{D}'(U)$, $\varphi \in C^\infty_{\mathrm{c}}(U)$, and a correct choice of the transpose operator A^t. (We make this strategy into a theorem in §7.3 below.) That is, one defines operations on distributions by defining the dual operation on test functions.

Once we define the two fundamental operations above, we can apply to a distribution any differential operator with smooth coefficients, and thus we can pose PDEs in distributions. We are not yet ready to study any 'serious' PDE, but in this chapter we will solve two 'baby' ODEs: $u' = 0$ and $xu = 0$.

3.1. Differentiation

3.1.1. Definition. Before giving the definition of a derivative of a distribution, let us first discuss which properties this operation should satisfy:

(1) We are looking for a linear operator $\widetilde{\partial}_{x_j} : \mathcal{D}'(U) \to \mathcal{D}'(U)$ where $U \Subset \mathbb{R}^n$.
(2) This operator should agree with the usual (classical) partial derivative on nice functions: if $f \in C^1(U)$ then $\widetilde{\partial}_{x_j} f = \partial_{x_j} f$. Here $C^1(U) \subset L^1_{\mathrm{loc}}(U)$ is embedded into $\mathcal{D}'(U)$ in the standard way (2.2).
(3) This operator should also be (sequentially) continuous: if $u_k \in \mathcal{D}'(U)$ is a sequence converging to 0 in $\mathcal{D}'(U)$, then $\widetilde{\partial}_{x_j} u_k \to 0$ in $\mathcal{D}'(U)$ as well.

Remark 3.1. *If an operator satisfying (1)–(3) above exists, then it is unique. Indeed, we will show later (see Theorem 6.10) that $C^\infty_c(U)$ is dense in $\mathcal{D}'(U)$, that is for each $u \in \mathcal{D}'(U)$ there exists a sequence $f_k \in C^\infty_c(U)$ which converges to u in $\mathcal{D}'(U)$. Then $\widetilde{\partial}_{x_j} u$ has to be the limit in $\mathcal{D}'(U)$ of the classical derivatives $\partial_{x_j} f_k$ and thus is uniquely determined.*

This proof applies to other operators on distributions that we define below, and makes it possible to shorten proofs of various identities featuring these operations: by density it is enough to verify these identities for 'nice' functions only.

As mentioned above, we will define the operator $\widetilde{\partial}_{x_j}$ by duality:

(1) First, let us take a 'nice' $f \in C^1(U)$, and a test function $\varphi \in C^\infty_c(U)$. Using integration by parts (Theorem 1.17), we see that
$$(\partial_{x_j} f, \varphi) = \int_U (\partial_{x_j} f)\varphi\, dx = -\int_U f(\partial_{x_j}\varphi)\, dx = -(f, \partial_{x_j}\varphi). \tag{3.1}$$
(2) Now we take (3.1) as the *definition* of $\widetilde{\partial}_{x_j}$. More precisely, if $u \in \mathcal{D}'(U)$ and $\varphi \in C^\infty_c(U)$, then we define
$$(\widetilde{\partial}_{x_j} u, \varphi) := -(u, \partial_{x_j}\varphi). \tag{3.2}$$
Here we use that $\partial_{x_j}\varphi \in C^\infty_c(U)$.

(3) It is direct to see (as ∂_{x_j} is a linear operator on $C^\infty_c(U)$) that the formula (3.2) defines a linear functional $\widetilde{\partial}_{x_j} u : C^\infty_c(U) \to \mathbb{C}$. We now show that this functional satisfies the bound (2.1) and thus gives a distribution $\widetilde{\partial}_{x_j} u \in \mathcal{D}'(U)$. Fix arbitrary $K \Subset U$. Since u is a distribution, it satisfies the bound (2.1): there exist C, N such that
$$|(u,\psi)| \le C\|\psi\|_{C^N} \quad \text{for all } \psi \in C^\infty_c(U) \quad \text{such that } \operatorname{supp}\psi \subset K.$$
If $\varphi \in C^\infty_c(U)$ and $\operatorname{supp}\varphi \subset K$, then we apply the above bound with $\psi := \partial_{x_j}\varphi$ to get
$$|(\widetilde{\partial}_{x_j} u, \varphi)| = |(u, \partial_{x_j}\varphi)| \le C\|\partial_{x_j}\varphi\|_{C^N} \le C\|\varphi\|_{C^{N+1}},$$
that is the bound (2.1) does indeed hold for $\widetilde{\partial}_{x_j} u$ with N replaced by $N+1$. (Alternatively, we could use Proposition 2.6 and the fact that if $\varphi_k \to 0$ in $C^\infty_c(U)$, then $\partial_{x_j}\varphi_k \to 0$ in $C^\infty_c(U)$ as well.)

(4) From (3.1) we see that if $f \in C^1(U)$, then $\widetilde{\partial}_{x_j} f = \partial_{x_j} f$. Moreover, the operator $\widetilde{\partial}_{x_j}$ is sequentially continuous on $\mathcal{D}'(U)$. Indeed, if $u_k \in \mathcal{D}'(U)$ converges to 0 in $\mathcal{D}'(U)$, then for each $\varphi \in C^\infty_c(U)$ we have
$$(\widetilde{\partial}_{x_j} u_k, \varphi) = -(u_k, \partial_{x_j}\varphi) \to 0$$
and thus $\widetilde{\partial}_{x_j} u_k \to 0$ in $\mathcal{D}'(U)$ as well.

We now constructed the operator $\widetilde{\partial}_{x_j}$ that satisfies the properties (1)–(3) above. By a slight abuse of notation, we will henceforth forget about the tilde and just write

$$\partial_{x_j} u := \widetilde{\partial}_{x_j} u \quad \text{for all } u \in \mathcal{D}'(U).$$

We remark that we still have $\partial_{x_j}\partial_{x_\ell} = \partial_{x_\ell}\partial_{x_j}$ in distributions, so we can define $\partial_x^\alpha : \mathcal{D}'(U) \to \mathcal{D}'(U)$ for any multiindex α.

3.1.2. Examples. Having defined derivatives of distributions, we look at a few examples with $U = \mathbb{R}$:

- $u(x) = |x|$. To compute $u' = \partial_x u \in \mathcal{D}'(\mathbb{R})$, take arbitrary $\varphi \in C_c^\infty(\mathbb{R})$ and write

$$\begin{aligned} (u', \varphi) = -(u, \varphi') &= -\int_{\mathbb{R}} |x|\varphi'(x)\,dx \\ &= \int_{-\infty}^0 x\varphi'(x)\,dx - \int_0^\infty x\varphi'(x)\,dx \\ &= -\int_{-\infty}^0 \varphi(x)\,dx + \int_0^\infty \varphi(x)\,dx \end{aligned}$$

  where in the last equality we use integration by parts, with the boundary terms being zero. This shows that $\partial_x|x|$ is given by the locally integrable function $\operatorname{sgn} x$:

$$\partial_x|x| = \operatorname{sgn} x := \begin{cases} 1, & x > 0, \\ -1, & x < 0. \end{cases}$$

- The Heaviside function:

$$H(x) := \begin{cases} 1, & x > 0, \\ 0, & x < 0. \end{cases} \tag{3.3}$$

  Take arbitrary $\varphi \in C_c^\infty(\mathbb{R})$ and compute

$$(H', \varphi) = -(H, \varphi') = -\int_0^\infty \varphi'(x)\,dx = \varphi(0)$$

  where the last equality follows from the Fundamental Theorem of Calculus. Thus the distributional derivative of the Heaviside function is the Dirac delta 'function':

$$H'(x) = \delta_0(x). \tag{3.4}$$

- The delta function δ_0. Take arbitrary $\varphi \in C_c^\infty(\mathbb{R})$ and compute

$$(\delta_0', \varphi) = -(\delta_0, \varphi') = -\varphi'(0).$$

  There isn't a better way to write this down, so we just denote the derivative of the delta function by δ_0'.

3.1.3. A differential equation. We are now ready to solve our first differential equation, $u' = 0$. In this case being in distributions does not produce any new solutions:

PROPOSITION 3.2. *Assume that $U \subset \mathbb{R}$ is an open interval, $u \in \mathcal{D}'(U)$, and $u' = 0$. Then u is a constant function.*

PROOF. The statement that $u' = 0$ in distributions is equivalent to

$$(u, \psi') = 0 \quad \text{for all } \psi \in C_c^\infty(U).$$

Let us rewrite this as follows:

$$(u, \varphi) = 0 \quad \text{for all } \varphi \quad \text{in the space } \mathcal{V} := \{\psi' \mid \psi \in C_c^\infty(U)\}. \tag{3.5}$$

The space $\mathcal{V}$ has codimension 1 inside $C_c^\infty(U)$, indeed it can be characterized as

$$\mathcal{V} = \left\{\varphi \in C_c^\infty(U) \,\middle|\, \int_U \varphi(x)\,dx = 0\right\}.$$

To check this, we have to show that each $\varphi \in C_c^\infty(U)$ which integrates to 0 can be written as ψ' for some $\psi \in C_c^\infty(U)$, which can be done by putting $\psi(x) = \int_a^x \varphi(t)\,dt$ where $a \in U$ lies to the left of $\operatorname{supp} \varphi$.

Now, fix $\chi_0 \in C_c^\infty(U)$ such that $\int_U \chi_0(x)\,dx = 1$. Then for each $\varphi \in C_c^\infty(U)$ we have

$$\varphi - (1, \varphi)\chi_0 \in \mathcal{V} \quad \text{where } (1, \varphi) = \int_U \varphi(x)\,dx.$$

Then by (3.5) we have for all $\varphi \in C_c^\infty(U)$

$$(u, \varphi) = (1, \varphi)(u, \chi_0) = \big((u, \chi_0)1, \varphi\big),$$

that is $u = (u, \chi_0)1$ is a constant function. □

3.2. Multiplication by smooth functions

3.2.1. Definition and basic properties. The next operation we extend to distributions is the multiplication operator

$$f \in L^1_{\text{loc}}(U) \ \mapsto \ af$$

where $a \in C^\infty(U)$ is given. For each $f \in L^1_{\text{loc}}(U)$ and a test function $\varphi \in C_c^\infty(U)$ we have

$$(af, \varphi) = \int_U a(x)f(x)\varphi(x)\,dx = (f, a\varphi).$$

Thus we define for $u \in \mathcal{D}'(U)$ and $a \in C^\infty(U)$ the product $au \in \mathcal{D}'(U)$ as follows:

$$(au, \varphi) := (u, a\varphi) \quad \text{for all } \varphi \in C_c^\infty(U). \tag{3.6}$$

This gives the usual pointwise multiplication when $u \in L^1_{\text{loc}}(U)$. Arguing similarly to §3.1 we see that au is indeed a distribution and the map $u \mapsto au$ is sequentially continuous.

REMARK 3.3. *The definition (3.6) uses crucially that $a \in C^\infty(U)$ and thus $a\varphi \in C^\infty_{\rm c}(U)$. In general it is not possible to define the product au when u is an arbitrary distribution and a is a non-smooth function. Similarly, we generally cannot define the product of two distributions. Indeed, let u_k be the step function from Proposition 2.9, with $u_k \to 2\delta_0$ in $\mathcal{D}'(\mathbb{R})$. If we could define products of distributions, we would expect that $u_k^2 \to 4\delta_0^2$, but*

$$u_k^2(x) = k^2 \mathbf{1}_{[-1/k,1/k]}(x)$$

does not have a limit in $\mathcal{D}'(\mathbb{R})$ since $(u_k^2, \chi) \to \infty$ for any $\chi \in C^\infty_{\rm c}(\mathbb{R})$ such that $\chi(0) > 0$.

As one would expect, the Leibniz rule still applies in distributions:

PROPOSITION 3.4. *Assume that $u \in \mathcal{D}'(U)$ and $a \in C^\infty(U)$. Then*

$$\partial_{x_j}(au) = (\partial_{x_j} a)u + a(\partial_{x_j} u). \tag{3.7}$$

REMARK 3.5. *Note that (3.7) features the distributional derivatives defined in (3.2) and the distributional multiplication by a smooth function a defined in (3.6). If we denote the first of these operators by $\widetilde{\partial}_{x_j}$ and the second one by M_a, then a more pedantic way to write (3.7) would be*

$$\widetilde{\partial}_{x_j}(M_a(u)) = M_{\partial_{x_j} a}(u) + M_a(\widetilde{\partial}_{x_j}(u)).$$

PROOF OF PROPOSITION 3.4. First proof: We can do this by direct computation. Let $\varphi \in C^\infty_{\rm c}(U)$, then

$$\begin{aligned}
(\partial_{x_j}(au), \varphi) &= -(au, \partial_{x_j}\varphi) = -(u, a(\partial_{x_j}\varphi)),\\
((\partial_{x_j} a)u, \varphi) &= (u, (\partial_{x_j} a)\varphi),\\
(a(\partial_{x_j} u), \varphi) &= (\partial_{x_j} u, a\varphi) = -(u, \partial_{x_j}(a\varphi))
\end{aligned}$$

which gives (3.7) since $\partial_{x_j}(a\varphi) = (\partial_{x_j} a)\varphi + a(\partial_{x_j}\varphi)$.

Second proof: This one relies on the density of $C^\infty_{\rm c}$ in $\mathcal{D}'$ that we have not proved yet, but it is more robust than the first one. By Theorem 6.10 below, there exists a sequence $f_k \in C^\infty_{\rm c}(U)$ converging to u in $\mathcal{D}'(U)$. By the usual Leibniz rule we have for all k

$$\partial_{x_j}(af_k) = (\partial_{x_j} a)f_k + a(\partial_{x_j} f_k).$$

We now pass to the limit in $\mathcal{D}'(U)$, using that the operations $u \mapsto \partial_{x_j} u$ and $u \mapsto au$ are sequentially continuous, and obtain (3.7). □

As a basic example of multiplication of distributions and smooth functions we have the following formula featuring the delta function: if $y \in U$ and $a \in C^\infty(U)$, then

$$a(x)\delta_y(x) = a(y)\delta_y(x). \tag{3.8}$$

3.2.2. Another differential equation. The next proposition solves the differential equation $xu = 0$ in distributions. This time there are interesting solutions which are not functions, namely constant multiples of the Dirac delta function δ_0, where $x\delta_0(x) = 0$ by (3.8).

PROPOSITION 3.6. *Let $U \Subset \mathbb{R}$ be an interval containing 0. Assume that $u \in \mathcal{D}'(U)$ and $xu = 0$. Then $u = c\delta_0$ for some $c \in \mathbb{C}$.*

The proof of Proposition 3.6 uses the following lemma from classical analysis which is important in its own right:

LEMMA 3.7. *Assume that $U \Subset \mathbb{R}$ is an interval containing 0 and $\varphi \in C_c^\infty(U)$ satisfies $\varphi(0) = 0$. Then there exists $\psi \in C_c^\infty(U)$ such that $\varphi(x) = x\psi(x)$.*

PROOF. It is tempting to just define $\psi(x) := \varphi(x)/x$ and compute derivatives of all orders to see that they extend continuously to $x = 0$. But a faster strategy is to apply the Fundamental Theorem of Calculus to the function $t \mapsto \varphi(tx)$ on the interval $[0, 1]$ and get for $x \in U$

$$\varphi(x) = \int_0^1 \partial_t(\varphi(tx))\,dt = x\psi(x) \quad \text{where } \psi(x) := \int_0^1 \varphi'(tx)\,dt.$$

Differentiating under the integral sign, we get that $\psi \in C^\infty(U)$, and it is compactly supported since $\varphi(x) = x\psi(x)$ and φ is compactly supported. □

We can now give

PROOF OF PROPOSITION 3.6. We follow a similar scheme to Proposition 3.2. The statement that $xu = 0$ in distributions is equivalent to

$$(u, \varphi) = 0 \quad \text{for all } \varphi \quad \text{in the space } \mathcal{V} := \{x\psi \mid \psi \in C_c^\infty(U)\}. \tag{3.9}$$

From Lemma 3.7 we see that $\mathcal{V}$ has codimension 1, in fact

$$\mathcal{V} = \{\varphi \in C_c^\infty(U) \mid \varphi(0) = 0\}.$$

Fix $\chi_0 \in C_c^\infty(U)$ such that $\chi_0(0) = 1$. Then for each $\varphi \in C_c^\infty(U)$ we have

$$\varphi(x) - \varphi(0)\chi_0(x) \in \mathcal{V}.$$

By (3.9) this implies that for all $\varphi \in C_c^\infty(U)$

$$(u, \varphi) = \varphi(0)(u, \chi_0) = \big((u, \chi_0)\delta_0, \varphi\big),$$

that is $u = (u, \chi_0)\delta_0$ is a multiple of the delta function. □

3.3. Notes and exercises

A natural generalization of Proposition 3.2 is that for a linear ODE

$$u^{(m)}(x) + a_{m-1}(x)u^{(m-1)}(x) + \cdots + a_0(x)u(x) = f \ \in \ C^0(U)$$

on an interval $U \Subset \mathbb{R}$ with coefficients $a_j \in C^\infty(U)$, all distributional solutions are classical, i.e. they lie in $C^m(U)$. See [**Hör03**, Corollary 3.1.6]. If $f \in C^\infty(U)$ then we can also see this as a corollary of elliptic regularity, proved in Theorem 14.2 below.

Our presentation follows [**Hör03**, §3.1] and [**FJ98**, Chapter 2].

Exercise 3.1. (1 pt) *Consider a function $f : \mathbb{R} \to \mathbb{C}$ such that f lies in C^1 on $(-\infty, a)$ and (a, ∞) for some $a \in \mathbb{R}$ and the derivative $f' \in C^0(\mathbb{R} \setminus \{a\})$ is locally integrable on $\mathbb{R}$. The latter implies the existence of one-sided limits $f(a + 0)$ and $f(a - 0)$. Show that*

$$\partial_x f = f' + (f(a + 0) - f(a - 0))\delta_a$$

where $\partial_x f$ denotes the distributional derivative of $f \in \mathcal{D}'(\mathbb{R})$.

Exercise 3.2. (1 pt) *Assume that $u, v \in C^0(\mathbb{R})$ and $\partial_x u = v$ in the sense of distributions in $\mathcal{D}'(\mathbb{R})$. Show that $u \in C^1(\mathbb{R})$ and $u' = v$ in the sense of the ordinary derivative. That is, if the distributional derivative is continuous, then it is the ordinary derivative.*

Exercise 3.3. ($3 = 1 + 2$ pts) ***(a)*** *For $m \in \mathbb{N}$, write $x\partial_x^m \delta_0 \in \mathcal{D}'(\mathbb{R})$ as a linear combination of $\delta_0, \partial_x\delta_0, \dots, \partial_x^{m-1}\delta_0$.*

(b) *Show that the space of solutions to the equation $x^m u = 0$, $u \in \mathcal{D}'(\mathbb{R})$, is the span of $\delta_0, \partial_x\delta_0, \dots, \partial_x^{m-1}\delta_0$. (Hint: for $m = 1$ this was done in class. The $m = 1$ result can be iterated to get the general case.)*

Exercise 3.4. (2 pts) *Find all $u \in \mathcal{D}'(\mathbb{R})$ such that $u \sin x = 0$.*

Exercise 3.5. (2 pts) *This exercise gives a higher dimensional version of the Division Lemma 3.7. Let $U \Subset \mathbb{R}^n$ contain 0. Define*

$$\mathcal{V} := \{x_1\psi_1 + \cdots + x_n\psi_n \mid \psi_1, \dots, \psi_n \in C_c^\infty(U)\} \ \subset \ C_c^\infty(U).$$

(In algebraic terms, at least if we forget about the compact support condition, $\mathcal{V}$ is the ideal generated by $x_1, \dots, x_n$.) Show that

$$\mathcal{V} = \{\varphi \in C_c^\infty(U) \mid \varphi(0) = 0\}.$$

(Hint: first show that $\mathcal{V}$ contains $C_c^\infty(U \setminus \{0\})$, by taking a partition of unity subordinate to covering by the sets $U \cap \{x_j \neq 0\}$. Next, take arbitrary $\varphi \in C_c^\infty(U)$ such that $\varphi(0) = 0$ and use the Fundamental Theorem of Calculus for the function $t \mapsto \varphi(tx)$ to write φ as the sum of an element of $\mathcal{V}$ and an element of $C_c^\infty(U \setminus \{0\})$.)

EXERCISE 3.6. (1 pt) *Let* $U \Subset \mathbb{R}^n$ *contain 0. Using Exercise 3.5, find all solutions* $u \in \mathcal{D}'(U)$ *to the system of equations*

$$x_1 u = x_2 u = \cdots = x_n u = 0.$$

CHAPTER 4

Distributions and support

In this chapter we define the support of a distribution. We next consider the space $\mathcal{E}'$ of distributions with compact support and give an alternative characterization of it as the dual to C^∞. The latter is a Fréchet space, which lets us prove a Banach–Steinhaus Theorem for distributions. Finally, we give a complete description of the space of distributions supported at a single point.

4.1. Support of a distribution

Recall from Definition 1.5 that the support of a continuous function f is the closure of the set $\{x \mid f(x) \neq 0\}$. If f is instead a distribution, then we cannot define the support this way since we cannot evaluate f at a point. Luckily, all we actually need is to know when f vanishes identically on an open subset, which makes sense in distributions thanks to the restriction operator from Definition 2.11.

Definition 4.1. *Let $U \Subset \mathbb{R}^n$ and $u \in \mathcal{D}'(U)$. We say a point $x \in U$ does **not** lie in $\operatorname{supp} u$ if there exists $V \Subset U$ containing x and such that $u|_V = 0$. This defines a subset*

$$\operatorname{supp} u \ \subset\ U.$$

From the definition above we see that $\operatorname{supp} u$ is a relatively closed subset of U since its complement is open. We also see that for $f \in C^0(U)$, Definitions 1.5 and 4.1 give the same set $\operatorname{supp} f$. As another example, the support of a delta function consists of a single point:

$$\operatorname{supp} \delta_y = \{y\}.$$

The next statement is trivial for continuous functions, but it needs a proof for distributions since support of the latter is not defined in a pointwise way.

Proposition 4.2. *Let $u \in \mathcal{D}'(U)$. Then*

$$u|_{U \setminus \operatorname{supp} u} = 0.$$

That is, if $\varphi \in C_c^\infty(U)$ and $\operatorname{supp} u \cap \operatorname{supp} \varphi = \emptyset$, then $(u, \varphi) = 0$.

Proof. Here is a short proof: for each $x \in U \setminus \operatorname{supp} u$, there exists $V_x \Subset U \setminus \operatorname{supp} u$ containing x and such that $u|_{V_x} = 0$. The sets V_x cover $U \setminus \operatorname{supp} u$, so by the uniqueness part of Theorem 2.13 applied to $u|_{U\setminus \operatorname{supp} u}$, we see that $u|_{U\setminus \operatorname{supp} u} = 0$.

Alternatively we can repeat part of the proof of Theorem 2.13. Assume that $\varphi \in C_c^\infty(U)$ and $\operatorname{supp} u \cap \operatorname{supp} \varphi = \emptyset$. Then for each $x \in \operatorname{supp} \varphi$ there exists $V_x \Subset U$ containing x and such that $u|_{V_x} = 0$. Using a partition of unity we can write $\varphi = \varphi_1 + \cdots + \varphi_m$ where each $\varphi_j \in C_c^\infty(U)$ is supported in one of the sets V_x. Then $(u, \varphi_j) = 0$ and thus $(u, \varphi) = 0$. □

Other properties of the support of a distribution are given in

PROPOSITION 4.3.$^{\mathbf{S}}$ *Let $U \Subset \mathbb{R}^n$. For all $u, v \in \mathcal{D}'(U)$ and $a \in C^\infty(U)$ we have:*

(1) $u = 0$ *if and only if* $\operatorname{supp} u = \emptyset$;
(2) $\operatorname{supp}(u + v) \subset \operatorname{supp} u \cup \operatorname{supp} v$;
(3) $\operatorname{supp}(au) \subset \operatorname{supp} a \cap \operatorname{supp} u$;
(4) $\operatorname{supp}(\partial_{x_j} u) \subset \operatorname{supp} u$;
(5) *if* $au = 0$ *then* $\operatorname{supp} u \subset \{x \in U \mid a(x) = 0\}$;
(6) *if* $V \Subset U$ *then* $\operatorname{supp}(u|_V) = \operatorname{supp} u \cap V$;
(7) *if* $u_k \to u$ *in* $\mathcal{D}'(U)$, *then* $\operatorname{supp} u$ *is contained in the closure (in U) of the union* $\bigcup_k \operatorname{supp} u_k$.

We omit the proofs since they are straightforward; some of the above properties are assigned as exercises below.

Later in §8.3 we will study the related notion of *singular support* which will be essential for Elliptic Regularity.

4.2. Distributions with compact support

Let $U \Subset \mathbb{R}^n$. We previously defined $\mathcal{D}'(U)$ as the dual to the space $C_c^\infty(U)$. One can alternatively consider the space $\mathcal{E}'(U)$ which is dual to the space $C^\infty(U)$ of all smooth functions, not necessarily compactly supported. (The notation $\mathcal{E}'$ goes back to Schwartz who denoted $\mathcal{E} := C^\infty$.) In this section we define the space $\mathcal{E}'(U)$ and identify it with the space of compactly supported distributions in $\mathcal{D}'(U)$.

We start by defining convergence of sequences in $C^\infty(U)$, by requiring uniform convergence of all derivatives on every compact set. For $K \Subset U$ and $\varphi \in C^\infty(U)$, define the seminorm

$$\|\varphi\|_{C^N(U,K)} := \max_{|\alpha| \le N} \sup_{x \in K} |\partial_x^\alpha \varphi(x)|. \tag{4.1}$$

DEFINITION 4.4. *Let $\varphi_k \in C^\infty(U)$ be a sequence and $\varphi \in C^\infty(U)$. We say that $\varphi_k \to \varphi$ in $C^\infty(U)$ if*

$$\|\varphi_k - \varphi\|_{C^N(U,K)} \to 0 \quad \textit{as } k \to \infty \quad \textit{for all } K \Subset U \textit{ and } N.$$

We define the space $\mathcal{E}'(U)$ in a way analogous to the alternative definition of $\mathcal{D}'(U)$ from Proposition 2.6:

DEFINITION 4.5. *Let $u : C^\infty(U) \to \mathbb{C}$ be a linear functional. We say that u lies in $\mathcal{E}'(U)$ if it is sequentially continuous, namely for each sequence φ_k converging to 0 in $C^\infty(U)$ we have $u(\varphi_k) \to 0$.*

As in the case of $\mathcal{D}'(U)$, we use the notation $(u, \varphi) := u(\varphi)$ when $u \in \mathcal{E}'(U)$ and $\varphi \in C^\infty(U)$.

We next discuss the relationship between $\mathcal{D}'(U)$ and $\mathcal{E}'(U)$. Let $u \in \mathcal{E}'(U)$. Then the functional $u : C^\infty(U) \to \mathbb{C}$ can be restricted to $C^\infty_\mathrm{c}(U)$, which yields a distribution in $\mathcal{D}'(U)$ by Proposition 2.6 and since $\varphi_k \to 0$ in $C^\infty_\mathrm{c}(U)$ implies that $\varphi_k \to 0$ in $C^\infty(U)$ as well. This yields the operator

$$\iota : \mathcal{E}'(U) \to \mathcal{D}'(U), \quad (\iota(u), \varphi) = (u, \varphi) \quad \text{for all } u \in \mathcal{E}'(U),\ \varphi \in C^\infty_\mathrm{c}(U). \tag{4.2}$$

The next theorem shows that ι is injective and its range is exactly the space of distributions in $\mathcal{D}'(U)$ with compact support. Thus ι gives an identification

$$\mathcal{E}'(U) \simeq \{u \in \mathcal{D}'(U) \mid \operatorname{supp} u \Subset U\}.$$

Once the theorem is proved, we will drop ι in the notation and treat $\mathcal{E}'(U)$ as a subspace of $\mathcal{D}'(U)$.

THEOREM 4.6. *1. Assume that $u \in \mathcal{E}'(U)$ and $\iota(u) = 0$. Then $u = 0$.*

2. Assume that $u \in \mathcal{E}'(U)$. Then $\operatorname{supp} \iota(u) \Subset U$.

3. Assume that $v \in \mathcal{D}'(U)$ and $\operatorname{supp} v \Subset U$. Then there exists $u \in \mathcal{E}'(U)$ such that $\iota(u) = v$.

PROOF. 1. Take arbitrary $\varphi \in C^\infty(U)$. Then there exists a sequence $\varphi_k \in C^\infty_\mathrm{c}(U)$ which converges to φ in $C^\infty(U)$. Indeed, using (1.14), take a sequence of compact subsets exhausting U:

$$U = \bigcup_{k=1}^\infty K_k, \quad K_k \Subset U. \tag{4.3}$$

Take cutoff functions

$$\chi_k \in C^\infty_\mathrm{c}(U), \quad \operatorname{supp}(1 - \chi_k) \cap K_k = \emptyset \tag{4.4}$$

and put $\varphi_k := \chi_k \varphi \in C^\infty_\mathrm{c}(U)$. Then $\varphi_k \to \varphi$ in $C^\infty(U)$, since for each $K \Subset U$ there exists k_0 such that for all $k \geq k_0$ we have $K \subset K_k$ and thus $\|\varphi_k - \varphi\|_{C^N(U,K)} = 0$.

Now, since $\iota(u) = 0$ we have $(u, \varphi_k) = 0$ for all k. Passing to the limit, we see that $(u, \varphi) = 0$ as well, which shows that $u = 0$.

2. We argue by contradiction. Assume that $\operatorname{supp} \iota(u)$ is not compactly contained in U. Take a sequence K_k as in (4.3), then we have $\operatorname{supp} \iota(u) \not\subset K_k$ for each k. Then there exists a test function $\varphi_k \in C^\infty_\mathrm{c}(U \setminus K_k)$ such that $(u, \varphi_k) = 1$. Similarly to part 1 of this proof, we have $\varphi_k \to 0$ in $C^\infty(U)$, which gives a contradiction with the sequential continuity of u as a functional on $C^\infty(U)$.

3. Fix a cutoff $\chi \in C_c^\infty(U)$ such that $\operatorname{supp}(1-\chi) \cap \operatorname{supp} v = \emptyset$. For $\varphi \in C^\infty(U)$, define

$$(u, \varphi) := (v, \chi\varphi).$$

It is straightforward to check that this defines $u \in \mathcal{E}'(U)$; indeed, if $\varphi_k \to 0$ in $C^\infty(U)$ then $\chi\varphi_k \to 0$ in $C_c^\infty(U)$. Moreover, if $\varphi \in C_c^\infty(U)$ then by Proposition 4.2 applied to v and $(1-\chi)\varphi$ we have

$$(v, \varphi) - (u, \varphi) = (v, (1-\chi)\varphi) = 0$$

which shows that $\iota(u) = v$. □

We used in §2.3 that any test function $\varphi \in C_c^\infty(V)$ can be extended by zero to a test function in any open set containing V. The next statement is a version of this for distributions with compact support. Its proof is left as an exercise below.

Proposition 4.7. *Let $V \Subset U \Subset \mathbb{R}^n$ and $v \in \mathcal{E}'(V)$. Then there exists unique $u \in \mathcal{E}'(U)$ such that $u|_V = v$ and $\operatorname{supp} u \subset V$. In fact, we have $\operatorname{supp} u = \operatorname{supp} v$.*

Similarly to the space $\mathcal{D}'(U)$, we define weak convergence in $\mathcal{E}'(U)$:

Definition 4.8. *Let $U \Subset \mathbb{R}^n$, $u_k \in \mathcal{E}'(U)$ be a sequence, and $u \in \mathcal{E}'(U)$. We say that*

$$u_k \to u \quad \text{as } k \to \infty \quad \text{in } \mathcal{E}'(U)$$

if we have

$$(u_k, \varphi) \to (u, \varphi) \quad \text{as } k \to \infty \quad \text{for all } \varphi \in C^\infty(U).$$

As we see in Proposition 4.15 below, this convergence can be characterized in terms of convergence in $\mathcal{D}'(U)$.

Finally, let us give here the analog of Proposition 4.2 for the space $\mathcal{E}'$. Note that when writing $\operatorname{supp} u$ for $u \in \mathcal{E}'(U)$, we technically mean $\operatorname{supp} \iota(u)$.

Proposition 4.9. *Let $U \Subset \mathbb{R}^n$, $u \in \mathcal{E}'(U)$, $\varphi \in C^\infty(U)$, and $\operatorname{supp} u \cap \operatorname{supp} \varphi = \emptyset$. Then $(u, \varphi) = 0$.*

Proof. Let $\varphi_k = \chi_k\varphi \in C_c^\infty(U)$ be the sequence constructed in Step 1 of the proof of Theorem 4.6, converging to φ in $C^\infty(U)$. Then $\operatorname{supp} u \cap \operatorname{supp} \varphi_k = \emptyset$, so by Proposition 4.2 we have $(u, \varphi_k) = 0$. Since $u \in \mathcal{E}'(U)$, we have $(u, \varphi_k) \to (u, \varphi)$, giving that $(u, \varphi) = 0$. □

4.3. Fréchet metric and Banach–Steinhaus for distributions

4.3.1. A metric on $C^\infty(U)$. Unlike $C^\infty_{\mathrm{c}}(U)$, convergence of sequences in the space $C^\infty(U)$ from Definition 4.4 corresponds to a metric topology, which we introduce now. The proofs in this section are left as exercises below.

Let $U \Subset \mathbb{R}^n$. Using (1.14), take a sequence of compact subsets exhausting U:

$$U = \bigcup_{j=1}^{\infty} K_j, \quad K_j \Subset U, \quad K_j \Subset K_{j+1}.$$

For $N \in \mathbb{N}$, define the N-th seminorm $\|\bullet\|_N$ on $C^\infty(U)$ using (4.1):

$$\|\varphi\|_N := \|\varphi\|_{C^N(U,K_N)}.$$

These seminorms depend on the choice of the exhausting sets K_N, and they are not coordinate invariant, but the convergence they define is independent of these choices, in fact it is the convergence of Definition 4.4:

PROPOSITION 4.10. *Let $\varphi_k \in C^\infty(U)$ be a sequence and $\varphi \in C^\infty(U)$. Then $\varphi_k \to \varphi$ in $C^\infty(U)$ if and only if $\|\varphi_k - \varphi\|_N \to 0$ as $k \to \infty$ for each N.*

The set of seminorms $\|\bullet\|_N$ makes $C^\infty(U)$ into a complete space:

PROPOSITION 4.11. *Assume that $\varphi_k \in C^\infty(U)$ is a Cauchy sequence in the following sense: for each N we have*

$$\sup_{k,\ell \geq r} \|\varphi_k - \varphi_\ell\|_N \to 0 \quad \textit{as } r \to \infty.$$

Then there exists $\varphi \in C^\infty(U)$ such that $\varphi_k \to \varphi$ in $C^\infty(U)$.

As an application of the seminorms $\|\bullet\|_N$, arguing similarly to Proposition 2.6 we can reformulate the statement that $u \in \mathcal{E}'(U)$ in terms of a seminorm bound:

PROPOSITION 4.12. *Let $u : C^\infty(U) \to \mathbb{C}$ be a linear map. Then $u \in \mathcal{E}'(U)$ if and only if there exist $K \Subset U$ and constants C, N such that*

$$|(u,\varphi)| \leq C\|\varphi\|_{C^N(U,K)} \quad \textit{for all } \varphi \in C^\infty(U). \tag{4.5}$$

PROOF.$^{\mathbf{S}}$ If (4.5) holds, then it is immediate that $\varphi_k \to 0$ in $C^\infty(U)$ implies that $(u, \varphi_k) \to 0$, and thus $u \in \mathcal{E}'(U)$.

Assume now that $u \in \mathcal{E}'(U)$. We show the bound (4.5) by contradiction. If (4.5) does not hold, then for each N there exists

$$\varphi_N \in C^\infty(U), \quad (u, \varphi_N) = 1, \quad \|\varphi_N\|_N \leq \tfrac{1}{N}.$$

Then $\varphi_N \to 0$ as $N \to \infty$ in $C^\infty(U)$, which contradicts the sequential continuity of $u : C^\infty(U) \to \mathbb{C}$. □

We now follow the standard construction (coming from the theory of Fréchet spaces) to define a metric on $C^\infty(U)$. Namely, for $\varphi, \psi \in C^\infty(U)$, put

$$d_{C^\infty}(\varphi, \psi) := \sum_{N=1}^{\infty} 2^{-N} \frac{\|\varphi - \psi\|_N}{1 + \|\varphi - \psi\|_N}. \tag{4.6}$$

The fundamental properties of d_{C^∞} are collected in

PROPOSITION 4.13. *1. d_{C^∞} defines a metric on $C^\infty(U)$.*

2. For a sequence $\varphi_k \in C^\infty(U)$, we have $\varphi_k \to \varphi$ in $C^\infty(U)$ (in the sense of Definition 4.4) if and only if $d_{C^\infty}(\varphi_k, \varphi) \to 0$.

3. The metric space $(C^\infty(U), d_{C^\infty})$ is complete.

4.3.2. Banach–Steinhaus for distributions. The next theorem shows in particular that if a sequence of distributions in $\mathcal{E}'(U)$ converges weakly, then it satisfies a uniform bound. The proof is analogous to the Banach–Steinhaus theorem for operators on Banach spaces and can be skipped at first reading.

THEOREM 4.14 (Banach–Steinhaus for $\mathcal{E}'(U)$). *Let $U \subseteq \mathbb{R}^n$ and assume that a sequence of compactly supported distributions $u_k \in \mathcal{E}'(U)$ is weakly bounded in the following sense:*

$$\text{for each } \varphi \in C^\infty(U) \quad \text{there exists } C_\varphi \quad \text{such that for all } k \quad |(u_k, \varphi)| \le C_\varphi. \tag{4.7}$$

Then there exist $K \Subset U$ and constants C, N such that for all k we have:

$$\operatorname{supp} u_k \subset K, \tag{4.8}$$

$$|(u_k, \varphi)| \le C\|\varphi\|_{C^N(U,K)} \quad \text{for all } \varphi \in C^\infty(U). \tag{4.9}$$

PROOF.$^{\mathbf{X}}$ 1. We use the notation of §4.3.1. For $L \in \mathbb{N}$, define the subset of $C^\infty(U)$

$$A_L := \{\varphi \in C^\infty(U) \colon \text{for all } k,\ |(u_k, \varphi)| \le L\}.$$

Each set A_L is closed in $(C^\infty(U), d_{C^\infty})$. Indeed, assume that $\varphi_m \to \varphi$ in $C^\infty(U)$ and $\varphi_m \in A_L$ for all m. For each k we have $u_k \in \mathcal{E}'(U)$, so $(u_k, \varphi_m) \to (u_k, \varphi)$ as $m \to \infty$. Thus $|(u_k, \varphi_m)| \le L$ implies that $|(u_k, \varphi)| \le L$ which shows that $\varphi \in A_L$.

By the weak bound (4.7) we have

$$C^\infty(U) = \bigcup_{L \ge 1} A_L.$$

Then by the Baire Category Theorem for the complete metric space $(C^\infty(U), d_{C^\infty})$ we can fix L such that the interior of A_L is nonempty, that is A_L contains a metric ball:

$$B_{d_{C^\infty}}(\psi, \varepsilon) \subset A_L \quad \text{for some} \quad \psi \in C^\infty(U),\ \varepsilon > 0. \tag{4.10}$$

From (4.10) we get

$$B_{d_{C^\infty}}(0, \varepsilon) \subset A_{2L}. \tag{4.11}$$

Indeed, take arbitrary $\varphi \in B_{d_{C^\infty}}(0,\varepsilon)$. Then both $\varphi+\psi$ and ψ lie in $B_{d_{C^\infty}}(\psi,\varepsilon)$, which is contained in A_L; thus $\varphi \in A_{2L}$.

2. Recalling the definition (4.6) of d_{C^∞}, and using that the seminorms $\|\varphi\|_N$ are a monotone increasing sequence, we see that for any N

$$d(\varphi,0) \le \|\varphi\|_N + 2^{-N}.$$

Thus there exist $N,\delta>0$ such that for all $\varphi \in C^\infty(U)$

$$\|\varphi\|_N \le \delta \quad\Longrightarrow\quad d(\varphi,0)\le\varepsilon.$$

Putting $C := 2L/\delta$, we get from (4.11) that

$$\text{for all } \varphi \in C^\infty(U) \quad \text{and } k, \quad |(u_k,\varphi)| \le C\|\varphi\|_N.$$

This implies the bound (4.9) with $K := K_N$. It also shows that $\operatorname{supp} u_k \subset K_N$ for all k: indeed, if $\varphi \in C_c^\infty(U \setminus K_N)$ then $\|\varphi\|_N = 0$ and thus $(u_k,\varphi)=0$. □

As a consequence of Theorem 4.14, we obtain a characterization of the convergence in $\mathcal{E}'(U)$ (see Definition 4.8) in terms of convergence in $\mathcal{D}'(U)$ (see Definition 2.7). To make the proof easier to read, we make explicit the use of the embedding ι from (4.2).

PROPOSITION 4.15. *Assume that $u_k \in \mathcal{E}'(U)$. Then we have $u_k \to u$ in $\mathcal{E}'(U)$ if and only if both of the conditions below hold:*

(1) *there exists $K \Subset U$ such that for all k we have $\operatorname{supp} u_k \subset K$, and*
(2) *$\iota(u_k) \to \iota(u)$ in $\mathcal{D}'(U)$.*

PROOF.[S] Assume first that the conditions (1) and (2) hold. Without loss of generality we have $\operatorname{supp} u \subset K$. Fix $\chi \in C_c^\infty(U)$ such that $\operatorname{supp}(1-\chi)\cap K = \emptyset$. Then by Proposition 4.9 we have for each $\varphi \in C^\infty(U)$

$$(u_k,\varphi) = (u_k,\chi\varphi), \quad (u,\varphi) = (u,\chi\varphi).$$

Now, since $\chi\varphi \in C_c^\infty(U)$ and $\iota(u_k)\to\iota(u)$ in $\mathcal{D}'(U)$, we have $(u_k,\chi\varphi)\to(u,\chi\varphi)$. Thus $u_k \to u$ in $\mathcal{E}'(U)$.

Now, assume that $u_k \to u$ in $\mathcal{E}'(U)$. Then condition (1) above follows from Theorem 4.14 and condition (2) above it immediate since $C_c^\infty(U) \subset C^\infty(U)$. □

Another consequence of Theorem 4.14 is the Banach–Steinhaus theorem in the space $\mathcal{D}'$:

THEOREM 4.16 (Banach–Steinhaus for $\mathcal{D}'(U)$). *Let $U \Subset \mathbb{R}^n$ and assume that $u_k \in \mathcal{D}'(U)$ is a sequence of distributions which is weakly bounded in the following sense:*

$$\textit{for each } \varphi \in C_c^\infty(U) \quad \textit{there exists } C_\varphi \quad \textit{such that for all } k, \quad |(u_k,\varphi)| \le C_\varphi. \tag{4.12}$$

Then u_k satisfies a uniform version of the norm bound (2.1), namely for each compact $K \subset U$ there exist C, N such that

$$|(u_k, \varphi)| \leq C\|\varphi\|_{C^N} \quad \textit{for all } k \textit{ and } \varphi \in C_c^\infty(U) \quad \textit{such that } \operatorname{supp}\varphi \subset K. \tag{4.13}$$

Proof. Take arbitrary $K \Subset U$ and fix a cutoff function $\chi \in C_c^\infty(U)$ such that $\operatorname{supp}(1-\chi) \cap K = \emptyset$. Then χu_k lies in $\mathcal{E}'(U)$ and is weakly bounded in the sense of (4.7). By Theorem 4.14 we see that there exist C, N, and $K' \Subset U$ such that for all k

$$|(\chi u_k, \varphi)| \leq C\|\psi\|_{C^N(U,K')} \quad \text{for all } \varphi \in C^\infty(U). \tag{4.14}$$

Now, take any $\varphi \in C_c^\infty(U)$ such that $\operatorname{supp}\varphi \subset K$. Then $\varphi = \chi\varphi$ and thus by (4.14) we have for all k

$$|(u_k, \varphi)| = |(\chi u_k, \varphi)| \leq C\|\varphi\|_{C^N(U,K')} \leq C\|\varphi\|_{C^N}.$$

This gives the required estimate (4.13). □

One of the corollaries of Theorems 4.14 and 4.16 is that if a sequence of distributions converges weakly, then the limit is always a distribution. We state it in the space $\mathcal{D}'(U)$:

Proposition 4.17. *Assume that $u_k \in \mathcal{D}'(U)$ is a sequence of distributions and $u : C_c^\infty(U) \to \mathbb{C}$ is a map such that*

$$(u_k, \varphi) \to u(\varphi) \quad \textit{as } k \to \infty \quad \textit{for all } \varphi \in C_c^\infty(U).$$

Then $u \in \mathcal{D}'(U)$ is a distribution as well.

Proof. Since each u_k is linear, u is also a linear map. Passing to the limit in the estimate provided by Theorem 4.16, we see that u satisfies the bounds (2.1) and thus $u \in \mathcal{D}'(U)$. □

Finally, we show here that the map $u \in \mathcal{D}'(U), \varphi \in C_c^\infty(U) \mapsto (u, \varphi) \in \mathbb{C}$ is sequentially continuous. An analogous statement holds for $u \in \mathcal{E}'(U)$ and $\varphi \in C^\infty(U)$.

Proposition 4.18. *Assume that $u_k \in \mathcal{D}'(U)$ and $\varphi_k \in C_c^\infty(U)$ are sequences such that*

$$u_k \to u \quad \textit{in } \mathcal{D}'(U), \qquad \varphi_k \to \varphi \quad \textit{in } C_c^\infty(U).$$

Then $(u_k, \varphi_k) \to (u, \varphi)$.

Proof. We estimate

$$\big|(u_k, \varphi_k) - (u, \varphi)\big| \leq \big|(u_k, \varphi_k - \varphi)\big| + \big|(u_k - u, \varphi)\big|.$$

The second term on the right-hand side converges to 0 since $u_k \to u$ in $\mathcal{D}'(U)$. As for the first term, we take $K \Subset U$ containing $\operatorname{supp}\varphi_k$ for all k; Theorem 4.16 shows that there exist constants C, N such that for all k

$$\big|(u_k, \varphi_k - \varphi)\big| \leq C\|\varphi_k - \varphi\|_{C^N}.$$

Since $\|\varphi_k - \varphi\|_{C^N} \to 0$, we have $|(u_k, \varphi_k - \varphi)| \to 0$ as well, finishing the proof. □

4.4. Distributions supported at one point

Here we discuss distributions whose support consists of a single point. The next theorem provides their complete description as linear combinations of the delta function and its derivatives:

THEOREM 4.19. *Assume that $u \in \mathcal{E}'(\mathbb{R}^n)$ and $\operatorname{supp} u \subset \{y\}$ for some $y \in \mathbb{R}^n$. Then we have*

$$u = \sum_{|\alpha| \leq N} c_\alpha \, \partial_x^\alpha \delta_y \tag{4.15}$$

for some $N \in \mathbb{N}_0$ and some coefficients $c_\alpha \in \mathbb{C}$.

To simplify the notation in the proof, we assume that $y = 0$. We use the following

DEFINITION 4.20. *Let $\varphi \in C^\infty(\mathbb{R}^n)$ and $m \in \mathbb{N}_0$. We say that φ* vanishes at 0 with m derivatives *if*

$$\partial_x^\alpha \varphi(0) = 0 \quad \text{for all } \alpha,\ |\alpha| \leq m.$$

The basic properties of vanishing are collected in

PROPOSITION 4.21.[S] *1. Assume that φ vanishes at 0 with m derivatives. Then $\varphi(x) = \mathcal{O}(|x|^{m+1})$ as $x \to 0$.*

2. Assume that φ vanishes at 0 with m derivatives, and $|\alpha| \leq m$. Then $\partial_x^\alpha \varphi$ vanishes at 0 with $m - |\alpha|$ derivatives.

The key ingredient in the proof of Theorem 4.19 is the following

LEMMA 4.22. *Assume that $u \in \mathcal{E}'(\mathbb{R}^n)$ and $\operatorname{supp} u \subset \{0\}$. Then there exists N such that for each $\varphi \in C^\infty(\mathbb{R}^n)$ which vanishes at 0 with N derivatives, we have $(u, \varphi) = 0$.*

PROOF. 1. Since $\operatorname{supp} u \subset \{0\}$, we have by Proposition 4.9

$$(u, \psi) = 0 \quad \text{for all } \psi \in C^\infty(\mathbb{R}^n) \quad \text{such that } 0 \notin \operatorname{supp} \psi. \tag{4.16}$$

Fix a cutoff function

$$\chi \in C_c^\infty(B(0,1)), \quad 0 \notin \operatorname{supp}(1 - \chi).$$

For $\varphi \in C^\infty(\mathbb{R}^n)$ and $0 < \varepsilon < 1$, define the function

$$\varphi_\varepsilon(x) := \chi\Big(\frac{x}{\varepsilon}\Big)\varphi(x), \quad \varphi_\varepsilon \in C_c^\infty(\mathbb{R}^n).$$

Applying (4.16) to $\psi := \varphi - \varphi_\varepsilon$, we see that

$$(u, \varphi) = (u, \varphi_\varepsilon) \quad \text{for all } \varepsilon \in (0, 1). \tag{4.17}$$

2. Since u is a distribution on $\mathbb{R}^n$, by (2.1) there exists constants C, N such that

$$|(u,\psi)| \le C\|\psi\|_{C^N} \quad \text{for all } \psi \in C_c^\infty(B(0,1)). \tag{4.18}$$

Assume that $\varphi \in C^\infty(\mathbb{R}^n)$ vanishes at 0 with N derivatives. We will show that

$$\|\varphi_\varepsilon\|_{C^N} \to 0 \quad \text{as } \varepsilon \to 0+\,. \tag{4.19}$$

This finishes the proof since by (4.17) and (4.18) we have for all $\varepsilon \in (0,1)$

$$|(u,\varphi)| = |(u,\varphi_\varepsilon)| \le C\|\varphi_\varepsilon\|_{C^N}.$$

3. It remains to prove (4.19). To make the proof more readable, we first consider the simpler case $N = 0$. Since φ_ε is supported in $B(0,\varepsilon)$, we estimate

$$\|\varphi_\varepsilon\|_{C^0} \;=\; \sup|\varphi_\varepsilon| \;\le\; \|\chi\|_{C^0} \sup_{B(0,\varepsilon)} |\varphi| \;=\; \mathcal{O}(\varepsilon)$$

since $\varphi(0) = 0$.

The case of general N is handled similarly. Fix a multiindex α with $|\alpha| \le N$; we need to show that $\|\partial_x^\alpha \varphi_\varepsilon(x)\|_{C^0} \to 0$ as $\varepsilon \to 0+$. By the Leibniz rule, the derivative $\partial_x^\alpha \varphi_\varepsilon$ is a linear combination with constant coefficients of terms of the form

$$\varepsilon^{-|\beta|}\, \partial_x^\beta \chi\Big(\frac{x}{\varepsilon}\Big)\, \partial_x^\gamma \varphi(x)$$

where the multiindices β, γ satisfy $\alpha = \beta + \gamma$. By Proposition 4.21 and since φ vanishes at 0 with N derivatives, we have $\partial_x^\gamma \varphi(x) = \mathcal{O}(|x|^{N+1-|\gamma|})$ as $x \to 0$. Thus

$$\sup_x \Big|\varepsilon^{-|\beta|}\, \partial_x^\beta \chi\Big(\frac{x}{\varepsilon}\Big)\, \partial_x^\gamma \varphi(x)\Big| \le \varepsilon^{-|\beta|} \sup|\partial_x^\beta \chi| \sup_{B(0,\varepsilon)} |\partial_x^\gamma \varphi| = \mathcal{O}(\varepsilon^{N+1-|\beta|-|\gamma|})$$

which finishes the proof since $|\beta| + |\gamma| = |\alpha| \le N$. □

We are now ready to give

Proof of Theorem 4.19. Assume that $u \in \mathcal{E}'(\mathbb{R}^n)$, $\operatorname{supp} u \subset \{0\}$, and let N be the number in Lemma 4.22. For any $\varphi \in C^\infty(\mathbb{R}^n)$ we have the Taylor expansion

$$\varphi(x) = \sum_{|\alpha| \le N} \frac{x^\alpha}{\alpha!} \partial_x^\alpha \varphi(0) + \psi(x)$$

where $\psi \in C^\infty(\mathbb{R}^n)$ vanishes at 0 with N derivatives and $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ are monomials. By Lemma 4.22 we have $(u,\psi) = 0$ and thus

$$(u,\varphi) = \sum_{|\alpha| \le N} \frac{(u, x^\alpha)}{\alpha!} \partial_x^\alpha \varphi(0).$$

This shows that u has the form (4.15) with $c_\alpha := (-1)^{|\alpha|}(u, x^\alpha)/\alpha!$. □

4.5. Notes and exercises

Our presentation largely follows [**Hör03**, §§2.2–2.3] and [**FJ98**, §§1.4,3.1–3.2]. These books do not prove the Banach–Steinhaus theorems in §4.3.2, sending the reader instead to functional analysis textbooks such as [**Rud91**, Theorem 2.6 and §6.16].

EXERCISE 4.1. $(3 = 0.75 + 0.75 + 0.75 + 0.75$ pts) *Prove parts (3)–(6) of Proposition 4.3.*

EXERCISE 4.2. (1 pt) *Prove Proposition 4.7.*

EXERCISE 4.3. (0.5 pts) *Prove Proposition 4.10. (You will have to use that each compact subset of U is contained in one of the sets K_j.)*

EXERCISE 4.4. (1 pt) *Prove Proposition 4.11. (This is similar to how we prove completeness of the spaces C^k in an analysis course.)*

EXERCISE 4.5. $(3 = 1 + 1 + 1$ pts) *Prove Proposition 4.13.*

CHAPTER 5

Homogeneous distributions

One of the goals of the next few sections is to prove the following fact: if P is a constant coefficient differential operator on $\mathbb{R}^n$ (for example, the Laplacian Δ) and $f \in \mathcal{E}'(\mathbb{R}^n)$ is a compactly supported distribution, then a solution to the differential equation $Pu = f$ is given by the distribution

$$u = E * f \tag{5.1}$$

where '$*$' denotes convolution of distributions (defined in Chapter 8 below) and $E \in \mathcal{D}'(\mathbb{R}^n)$ is a *fundamental solution* of P, namely $PE = \delta_0$. (We will deliver (5.1) in Chapter 9 below.)

To make (5.1) work it is important to find a fundamental solution E. Quite often these fundamental solutions are homogeneous distributions on $\mathbb{R}^n$. In this section we study general homogeneous distributions and give important examples of homogeneous distributions on $\mathbb{R}$.

5.1. Basic properties

5.1.1. Homogeneous functions. We first review the definition of a homogeneous function:

Definition 5.1. *A function $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{C}$ is called* homogeneous *of degree $a \in \mathbb{C}$ if*

$$f(tx) = t^a f(x) \quad \text{for all } t > 0,\ x \in \mathbb{R}^n \setminus \{0\}. \tag{5.2}$$

Here to make sense of t^a when $t > 0$ and a is complex, we define $t^a := \exp(a \log t)$ where $\log t \in \mathbb{R}$.

We collect some basic properties of homogeneous functions in

Proposition 5.2.**R** *1. A function f is homogeneous of degree a if and only if it can be written in polar coordinates as*

$$f(x) = r^a g(\theta), \quad x = r\theta,\ r > 0,\ \theta \in \mathbb{S}^{n-1} \tag{5.3}$$

for a function g on $\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n \colon |x| = 1\}$.

2. If $f \in C^1(\mathbb{R}^n \setminus \{0\})$ then f is homogeneous of degree a if and only if it satisfies Euler's equation

$$x \cdot \partial_x f = af \quad \textit{where } x \cdot \partial_x := \sum_{j=1}^n x_j \partial_{x_j}. \tag{5.4}$$

3. If $f \in C^1(\mathbb{R}^n \setminus \{0\})$ is homogeneous of degree a, then $\partial_{x_j} f$ is homogeneous of degree $a - 1$.

5.1.2. Homogeneous distributions. To define homogeneity in distributions, we rewrite the definition (5.2) in terms of the distributional pairing $(\bullet, \bullet)$. Assume that $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ is a homogeneous function of degree a (where strictly speaking, we should require (5.2) to hold for almost every x). Take $\varphi \in C^\infty_{\mathrm{c}}(\mathbb{R}^n)$ and $t > 0$ and pair the equation (5.2) with φ:

$$t^a(f, \varphi) = \int_{\mathbb{R}^n} f(tx)\varphi(x)\, dx = t^{-n} \int_{\mathbb{R}^n} f(x)\varphi(t^{-1}x)\, dx.$$

Here in the last equality we use the change of variables formula. Thus, if we define the dilated function

$$\Lambda_t \varphi \in C^\infty_{\mathrm{c}}(\mathbb{R}^n), \quad \Lambda_t\varphi(x) = \varphi(tx), \tag{5.5}$$

then we have the identity

$$(f, \Lambda_t \varphi) = t^{-a-n}(f, \varphi) \quad \text{for all } t > 0. \tag{5.6}$$

Conversely, if (5.6) holds for all $\varphi \in C^\infty_{\mathrm{c}}(\mathbb{R}^n)$, then f is homogeneous of degree a. Thus we can give the following definition of homogeneity for distributions:

Definition 5.3. *Assume that $u \in \mathcal{D}'(\mathbb{R}^n)$ and $a \in \mathbb{C}$. We say that u is* homogeneous of degree a *if*

$$(u, \Lambda_t \varphi) = t^{-a-n}(u, \varphi) \quad \textit{for all } t > 0,\ \varphi \in C^\infty_{\mathrm{c}}(\mathbb{R}^n). \tag{5.7}$$

This extends the usual definition of homogeneity, so for example the constant function 1 is homogeneous of degree 0, and more generally any homogeneous polynomial of degree $k \in \mathbb{N}_0$ is a homogeneous distribution of degree k in the sense of Definition 5.3. A genuinely distributional example is given by the delta function (with the proof by a direct computation):

Proposition 5.4. *The delta distribution $\delta_0 \in \mathcal{D}'(\mathbb{R}^n)$ is homogeneous of degree $-n$.*

Some properties of homogeneous distributions are collected in

Proposition 5.5. *1. If $u \in \mathcal{D}'(\mathbb{R}^n)$ is homogeneous of degree a, then $x_j u$ is homogeneous of degree $a + 1$ and $\partial_{x_j} u$ is homogeneous of degree $a - 1$.*

2. If $u \in \mathcal{D}'(\mathbb{R}^n)$, then u is homogeneous of degree a if and only if it solves Euler's equation (5.4) in the sense of distributions on $\mathbb{R}^n$.

We leave part 1 as an exercise below. For a proof of part 2, see for example [**Hör03**, (3.2.19)′].

5.1.3. Extending homogeneous distributions through the origin. In §5.1.2, we considered homogeneous distributions on $\mathbb{R}^n$. One could alternatively define homogeneous distributions on $\mathbb{R}^n\setminus\{0\}$, following Definition 5.3 but with $\varphi\in C_c^\infty(\mathbb{R}^n\setminus\{0\})$. The latter class turns out to be isomorphic to distributions on the sphere $\mathbb{S}^{n-1}$ by a distributional version of the formula (5.3) (something we cannot do here as we have not yet introduced distributions on manifolds). However, for applications to PDE we will often need homogeneous distributions on $\mathbb{R}^n$ with the origin included. Thus it is reasonable to ask the following question:

$$\begin{gathered}\text{Given } v\in\mathcal{D}'(\mathbb{R}^n\setminus\{0\}) \text{ which is homogeneous of degree } a\in\mathbb{C},\\ \text{does there exist unique } u\in\mathcal{D}'(\mathbb{R}^n) \text{ homogeneous of degree } a\\ \text{such that } u|_{\mathbb{R}^n\setminus\{0\}}=v?\end{gathered}$$

Proposition 5.4 shows that in general the answer is 'No': the delta function δ_0 is homogeneous of degree $-n$ and it restricts to 0 on $\mathbb{R}^n\setminus\{0\}$ (existence also fails for $a=-n$ though it is a bit harder to see). It turns out that the answer to the question above is 'Yes' unless a is a negative integer $\leq -n$, and for such integers one can give a precise description of non-existence and non-uniqueness – see [**Hör03**, Theorems 3.2.3 and 3.2.4]. In these notes we only present a simpler special case when $\operatorname{Re} a>-n$:

Theorem 5.6. *Assume that $a\in\mathbb{C}$ satisfies $\operatorname{Re} a>-n$, and $v\in\mathcal{D}'(\mathbb{R}^n\setminus\{0\})$ is homogeneous of degree a. Then there exists unique $u\in\mathcal{D}'(\mathbb{R}^n)$ homogeneous of degree a and such that $u|_{\mathbb{R}^n\setminus\{0\}}=v$.*

Proof. 1. We first show uniqueness. Assume that $u\in\mathcal{D}'(\mathbb{R}^n)$ is homogeneous of degree a and $u|_{\mathbb{R}^n\setminus\{0\}}=0$. Then $\operatorname{supp} u\subset\{0\}$, so Theorem 4.19 gives

$$u=\sum_{0\leq k\leq N}u_k,\quad u_k=\sum_{|\alpha|=k}c_\alpha\partial_x^\alpha\delta_0.$$

for some N and $c_\alpha\in\mathbb{C}$. By Propositions 5.4 and 5.5, each u_k is homogeneous of degree $-n-k$. Since u is homogeneous of degree a, we have for each $\varphi\in C_c^\infty(\mathbb{R}^n)$ and $t>0$

$$t^{-a-n}(u,\varphi)=(u,\Lambda_t\varphi)=\sum_{0\leq k\leq N}t^k(u_k,\varphi). \tag{5.8}$$

Since $\operatorname{Re}(-a-n)<0$, we see that the left-hand side of (5.8) converges to 0 as $t\to\infty$. This implies that $(u_k,\varphi)=0$ for all k and φ, which shows that $u=0$.

2. To motivate the proof of existence, assume first that v is a function in $L^1_{\mathrm{loc}}(\mathbb{R}^n\setminus\{0\})$. Then by (5.3) we can write $v(r\theta)=r^ag(\theta)$ for some $w\in L^1(\mathbb{S}^{n-1})$. Since $\operatorname{Re} a>-n$, the function $|x|^a$ is integrable near the origin, and we see that v actually lies in

$L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and thus defines a distribution u on $\mathbb{R}^n$. To present this argument in a more distributional way, we take any $\varphi \in C^\infty_{\mathrm{c}}(\mathbb{R}^n)$ and write (u, φ) using integration in polar coordinates:

$$(u, \varphi) = \int_0^\infty \int_{\mathbb{S}^{n-1}} r^a g(\theta)\varphi(r\theta) r^{n-1}\, dS(\theta) dr = \int_{\mathbb{S}^{n-1}} g(\theta) R_a\varphi(\theta)\, dS(\theta) \tag{5.9}$$

where we define

$$R_a\varphi(x) := \int_0^\infty t^{a+n-1}\varphi(tx)\, dt \quad \text{for any } x \in \mathbb{R}^n \setminus \{0\}.$$

Since $\operatorname{Re}(a+n-1) > -1$ and φ is compactly supported, the integral $R_a\varphi(x)$ converges for each $x \in \mathbb{R}^n \setminus \{0\}$. Moreover, we can differentiate under the integral sign to see that $R_a\varphi \in C^\infty(\mathbb{R}^n \setminus \{0\})$, in fact the derivatives are given by the integrals

$$\partial_x^\alpha R_a\varphi(x) = \int_0^\infty t^{a+n-1+|\alpha|}(\partial_x^\alpha \varphi)(tx)\, dt$$

which still converge as $\operatorname{Re}(a + n - 1 + |\alpha|) > -1$. Since $g \in L^1(\mathbb{S}^{n-1})$, and $R_a\varphi$ is bounded on $\mathbb{S}^{n-1}$, the right-hand side of (5.9) converges and gives a way to define u as a distribution.

We now show existence in the case when v is a general distribution on $\mathbb{R}^n \setminus \{0\}$. It is tempting to still define the extension u by (5.9) since $R_a\varphi$ restricts to a smooth function on $\mathbb{S}^{n-1}$ and g is a distribution on $\mathbb{S}^{n-1}$. This is perfectly valid but we cannot do this here since we do not know distributions on manifolds yet (and accordingly have not shown the distributional analog of (5.3)).

3.$^{\mathbf{X}}$ So instead we fix a radial function $\psi \in C^\infty_{\mathrm{c}}(\mathbb{R}^n \setminus \{0\})$ as follows:

$$\psi(x) = \chi(|x|), \quad \chi \in C^\infty_{\mathrm{c}}((0,\infty)), \quad \int_0^\infty \frac{\chi(t)}{t}\, dt = \int_0^\infty \frac{\chi(1/t)}{t}\, dt = 1.$$

Note that

$$\int_0^\infty \frac{\psi(x/t)}{t}\, dt = 1 \quad \text{for all } x \in \mathbb{R}^n \setminus \{0\}. \tag{5.10}$$

Let $v \in \mathcal{D}'(\mathbb{R}^n \setminus 0)$ be homogeneous of degree a. Define the linear map $u : C^\infty_{\mathrm{c}}(\mathbb{R}^n) \to \mathbb{C}$ as follows:

$$(u, \varphi) := (v, \psi R_a\varphi) \quad \text{for all } \varphi \in C^\infty_{\mathrm{c}}(\mathbb{R}^n).$$

Here, as discussed above, $R_a\varphi \in C^\infty(\mathbb{R}^n \setminus \{0\})$ and thus the product $\psi R_a\varphi$ lies in $C^\infty_{\mathrm{c}}(\mathbb{R}^n \setminus \{0\})$ and can be paired with v. The operator $\varphi \mapsto \psi R_a\varphi$ is sequentially continuous $C^\infty_{\mathrm{c}}(\mathbb{R}^n) \to C^\infty_{\mathrm{c}}(\mathbb{R}^n \setminus \{0\})$, so $u \in \mathcal{D}'(\mathbb{R}^n)$ is a distribution. We claim that u is the extension of v we are looking for.

We first show that $u|_{\mathbb{R}^n \setminus 0} = v$, which is where the homogeneity of v is exploited. Take arbitrary $\varphi \in C^\infty_{\mathrm{c}}(\mathbb{R}^n \setminus 0)$; we need to show that

$$(v, \psi R_a\varphi) = (v, \varphi). \tag{5.11}$$

If $v \in L^1_{\mathrm{loc}}(\mathbb{R}^n \setminus 0)$ is a function, then this follows from (5.3) by a direct computation. In general, we write (with integral converging in $C^\infty_c(\mathbb{R}^n \setminus \{0\})$)

$$\psi R_a \varphi = \int_0^\infty t^{a+n-1} \psi \Lambda_t \varphi \, dt.$$

Since v is continuous $C^\infty_c(\mathbb{R}^n \setminus 0) \to \mathbb{C}$, we can pair it with both sides to get

$$(v, \psi R_a \varphi) = \int_0^\infty t^{a+n-1} (v, \psi \Lambda_t \varphi) \, dt \tag{5.12}$$

It does require some work to justify putting pairing with u inside the integral – this can be done using Riemann sums similarly to Lemma 6.8 below.

Now, since v is homogeneous of degree a, we have

$$(v, \psi \Lambda_t \varphi) = (v, \Lambda_t(\Lambda_{t^{-1}} \psi \cdot \varphi)) = t^{-a-n} (v, \Lambda_{t^{-1}} \psi \cdot \varphi)$$

and thus

$$(v, \psi R_a \varphi) = \int_0^\infty t^{-1} (v, \Lambda_{t^{-1}} \psi \cdot \varphi) \, dt.$$

However, (5.10) shows that, with the integral converging in $C^\infty_c(\mathbb{R}^n \setminus \{0\})$ (with the support property coming from our assumption on $\operatorname{supp} \varphi$),

$$\int_0^\infty t^{-1} \Lambda_{t^{-1}} \psi \cdot \varphi \, dt = \varphi.$$

Pairing this with v and again putting the pairing inside the integral, we get (5.11). Thus u is indeed an extension of v.

It remains to show that u is homogeneous of degree a. Take arbitrary $\varphi \in C^\infty_c(\mathbb{R}^n)$. Then for any $t > 0$ and $x \in \mathbb{R}^n \setminus \{0\}$ we have

$$R_a \Lambda_t \varphi(x) = \int_0^\infty s^{a+n-1} \varphi(tsx) \, ds = t^{-a-n} \int_0^\infty s^{a+n-1} \varphi(sx) \, ds = t^{-a-n} R_a \varphi(x).$$

Thus

$$(u, \Lambda_t \varphi) = (v, \psi R_a \Lambda_t \varphi) = t^{-a-n} (v, \psi R_a \varphi) = t^{-a-n} (u, \varphi),$$

showing that u is homogeneous of degree a. □

5.2. Homogeneous distributions on ℝ

We now introduce an important family of homogeneous distributions $x^a_+ \in \mathcal{D}'(\mathbb{R})$, where $a \in \mathbb{C}$ is a complex parameter. Let us first assume that $\operatorname{Re} a > -1$ and define x^a_+ as the locally integrable function

$$x^a_+ := \begin{cases} x^a, & x > 0, \\ 0, & x < 0. \end{cases} \tag{5.13}$$

Then x^a_+ is a homogeneous distribution of degree a. Note that the Heaviside function is the special case with $a = 0$.

5.2.1. Non-exceptional values of a. We now want to consider the case of general $a \in \mathbb{C}$. When $\operatorname{Re} a \le -1$, the function (5.13) is no longer locally integrable, so we need a different definition. We use the following

LEMMA 5.7. *If* $\operatorname{Re} a > 0$ *then, with derivatives understood in* $\mathcal{D}'(\mathbb{R})$,

$$\partial_x x_+^a = a x_+^{a-1}. \tag{5.14}$$

PROOF.[S] Take arbitrary $\varphi \in C_c^\infty(\mathbb{R})$ and compute

$$(\partial_x x_+^a, \varphi) = -(x_+^a, \varphi') = -\int_0^\infty x^a \varphi'(x)\, dx = -\lim_{\varepsilon \to 0+} \int_\varepsilon^\infty x^a \varphi'(x)\, dx.$$

Integrating by parts, we see that this is equal to

$$\lim_{\varepsilon \to 0+} \left(\varepsilon^a \varphi(\varepsilon) + \int_\varepsilon^\infty a x^{a-1} \varphi(x)\, dx \right) = \int_0^\infty a x^{a-1} \varphi(x)\, dx,$$

giving (5.14). □

We can now extend x_+^a to $\operatorname{Re} a > -2$ except $a = -1$ as follows:

$$x_+^a := \frac{\partial_x x_+^{a+1}}{a+1} \quad \text{for } \operatorname{Re} a > -2,\ a \neq -1. \tag{5.15}$$

Here ∂_x is the distributional derivative, so (5.15) means that

$$(x_+^a, \varphi) := -\int_0^\infty \frac{x^{a+1}}{a+1} \varphi'(x)\, dx \quad \text{for all } \varphi \in C_c^\infty(\mathbb{R}).$$

By Lemma 5.7 the definitions (5.15) and (5.13) agree for $\operatorname{Re} a > -1$.

The process can be repeated: for any $j \ge 0$ we can put

$$x_+^a := \frac{\partial_x^j x_+^{a+j}}{(a+j)\cdots(a+2)(a+1)} \quad \text{for } \operatorname{Re} a > -j-1,\ a \notin -\mathbb{N} \tag{5.16}$$

and by Lemma 5.7 these definitions agree for different j (as long as $\operatorname{Re} a > -j-1$).

We have obtained *an* extension of x_+^a to all $a \in \mathbb{C}$ except at the points $a \in -\mathbb{N} = \{-1, -2, \dots\}$. In a natural way it is *the* extension because of holomorphy in a. Indeed, for each $\varphi \in C_c^\infty(\mathbb{R})$ the map $a \mapsto (x_+^a, \varphi)$ is holomorphic in $\{\operatorname{Re} a > -1\}$ as can be seen by differentiating under the integral sign; in fact

$$\partial_a x_+^a = x_+^a \log x$$

is still a locally integrable function for $\operatorname{Re} a > -1$. Using (5.16) we see that in fact the map $a \mapsto (x_+^a, \varphi)$ is holomorphic in $a \in \mathbb{C} \setminus -\mathbb{N}$. By the unique continuation property of holomorphic functions, we see that the formula (5.16) provides the *unique* holomorphic continuation of x_+^a, defined in $\{\operatorname{Re} a > -1\}$ by the formula (5.13), to $a \in \mathbb{C} \setminus -\mathbb{N}$.

We record some standard properties of x_+^a (which can be checked directly for $\operatorname{Re} a > -1$ and follow in general by analytic continuation, or using (5.16)) in

PROPOSITION 5.8. *For $a \in \mathbb{C} \setminus -\mathbb{N}$ the identity (5.14) still holds and*

$$x \cdot x_+^a = x_+^{a+1}, \tag{5.17}$$

$$\operatorname{supp} x_+^a = [0, \infty). \tag{5.18}$$

Moreover, $x_+^a \in \mathcal{D}'(\mathbb{R})$ is homogeneous of degree a.

5.2.2. Exceptional values of a. We now briefly discuss what happens at the exceptional values $a \in -\mathbb{N}$. Looking at (5.16), we see that x_+^a is meromorphic in $a \in \mathbb{C}$ with poles at $-\mathbb{N}$ (in the sense that for each $\varphi \in C_c^\infty(\mathbb{R})$, the map $a \mapsto (x_+^a, \varphi)$ is meromorphic). A typical way to get rid of the singularities of x_+^a is to divide by the Gamma function, looking at the distribution

$$\chi_+^a := \frac{x_+^a}{\Gamma(a+1)} \tag{5.19}$$

which is holomorphic in $a \in \mathbb{C}$. Note that it satisfies the identities for all $a \in \mathbb{C}$

$$\partial_x \chi_+^a = \chi_+^{a-1}, \tag{5.20}$$

$$x \cdot \chi_+^a = (a+1)\chi_+^{a+1}. \tag{5.21}$$

We also have $\operatorname{supp} \chi_+^a \subset [0, \infty)$ and χ_+^a is homogeneous of degree a. In a way it is more natural to consider χ_+^a than x_+^a, defining it similarly to (5.13) for $\operatorname{Re} a > -1$ and extending to general $a \in \mathbb{C}$ using the identity (5.20) similarly to (5.16).

Given that χ_+^a makes sense for all $a \in \mathbb{C}$, it is irresistible to compute what it is when a lies in the exceptional set $-\mathbb{N}$. For $a = -1$, we use (5.20): $\chi_+^0 = x_+^0 = H(x)$ is the Heaviside function, so by (3.4)

$$\chi_+^{-1} = \delta_0. \tag{5.22}$$

Using (5.20) repeatedly, we then get

PROPOSITION 5.9. *For $k \in \mathbb{N}$ we have*

$$\chi_+^{-k} = \partial_x^{k-1} \delta_0. \tag{5.23}$$

REMARK 5.10. *So far we studied homogeneous distributions of degree a which are supported on $[0, \infty)$. We can alternatively define their analogs supported on $(-\infty, 0]$, starting with the locally integrable function when* $\operatorname{Re} a > -1$

$$(-x)_+^a := \begin{cases} 0, & x > 0, \\ (-x)^a, & x < 0 \end{cases}$$

and repeating the above constructions to obtain distributions $(-x)_+^a \in \mathcal{D}'(\mathbb{R})$ when $a \in \mathbb{C} \setminus -\mathbb{N}$ and $(-\chi)_+^a \in \mathcal{D}'(\mathbb{R})$ for all $a \in \mathbb{C}$. Another useful family of homogeneous distributions, $(x \pm i0)^a$, are defined in Exercise 5.4 below.

5.2.3. A division problem. In §3.2.2 we found all solutions $u \in \mathcal{D}'(\mathbb{R})$ to the equation $xu = 0$. We now look at the equation

$$xu = 1 \quad \text{where } u \in \mathcal{D}'(\mathbb{R}). \tag{5.24}$$

If we restrict this equation to $\mathbb{R} \setminus \{0\}$ and multiply both sides by $\frac{1}{x} \in C^\infty(\mathbb{R} \setminus \{0\})$, then we get

$$u|_{\mathbb{R}\setminus\{0\}} = \frac{1}{x}. \tag{5.25}$$

So we can think of (5.24) as the problem of extending the function $\frac{1}{x}$ to a distribution on $\mathbb{R}$. This is nontrivial since $\frac{1}{x}$ is not locally integrable on $\mathbb{R}$ and Theorem 5.6 does not apply since $\frac{1}{x}$ is homogeneous of degree -1; since $x\delta_0 = 0$ the extension is also not going to be unique.

As in §5.2.1 a solution is to define a solution u as a distributional derivative. The standard antiderivative of $\frac{1}{x}$ on $\mathbb{R} \setminus \{0\}$ is given by the function $\log|x|$, which is locally integrable on $\mathbb{R}$. We now define the *principal value distribution*

$$\text{p.v.}\frac{1}{x} := \partial_x \log|x| \ \in \ \mathcal{D}'(\mathbb{R}). \tag{5.26}$$

To justify the term 'principal value', we compute for $\varphi \in C_c^\infty(\mathbb{R})$ using integration by parts

$$\begin{aligned} \Big(\text{p.v.}\frac{1}{x}, \varphi\Big) &= -(\log|x|, \varphi') = -\int_{\mathbb{R}} \varphi'(x) \log|x|\, dx \\ &= -\lim_{\varepsilon\to 0+} \int_{\mathbb{R}\setminus(-\varepsilon,\varepsilon)} \varphi'(x)\log|x|\, dx \\ &= \lim_{\varepsilon\to 0+} \Big(\int_{\mathbb{R}\setminus(-\varepsilon,\varepsilon)} \frac{\varphi(x)}{x}\, dx + (\varphi(\varepsilon) - \varphi(-\varepsilon))\log\varepsilon \Big) \\ &= \lim_{\varepsilon\to 0+} \int_{\mathbb{R}\setminus(-\varepsilon,\varepsilon)} \frac{\varphi(x)}{x}\, dx. \end{aligned} \tag{5.27}$$

One can see directly that the limit on the last line of (5.27) exists by writing $\varphi(x) = \varphi(0) + \mathcal{O}(x)$. This limit is known as the principal value integral of $\varphi(x)/x$.

Some properties of the distribution $\text{p.v.}\frac{1}{x}$ (including (5.24) and (5.25)) are collected in

PROPOSITION 5.11. *We have*

$$\text{p.v.}\frac{1}{x}\Big|_{\mathbb{R}\setminus\{0\}} = \frac{1}{x}, \tag{5.28}$$

$$x \cdot \text{p.v.}\frac{1}{x} = 1. \tag{5.29}$$

Moreover, $\text{p.v.}\frac{1}{x}$ *is a homogeneous distribution of order* -1*.*

The proof is left as an exercise below.

5.3. Notes and exercises

Our presentation follows [**Hör03**, §3.2] and [**FJ98**, §§2.2–2.3], and to a lesser extent [**FJ98**, §4.2]. Our proof of Theorem 5.6 differs slightly from [**Hör03**], in particular in how we justify the identity (5.12) – Hörmander is careful to not use anything that has not been proved yet but our approach is perhaps more direct and thus easier to comprehend at first reading.

EXERCISE 5.1. (1 pt) *Prove part 1 of Proposition 5.5.*

EXERCISE 5.2. (2 pts) *Show that* $x_+^{-1} - (-x)_+^{-1} = \mathrm{p.v.}\frac{1}{x}$ *in the following sense (where you can restrict to* $a \in \mathbb{R}$ *in the limit):*

$$x_+^a - (-x)_+^a \to \mathrm{p.v.}\frac{1}{x} \quad \text{in } \mathcal{D}'(\mathbb{R}) \quad \text{as } a \to -1+0.$$

EXERCISE 5.3. (1 pt) *Prove Proposition 5.11, using the definition of* $\mathrm{p.v.}\frac{1}{x}$ *provided by the limit on the last line of* (5.27) *(and without using explicitly the distributional definition* (5.26)*).*

EXERCISE 5.4. (3 = 1+1+1 pts) *This exercise explores homogeneous distributions on* $\mathbb{R}$ *which are alternatives to* x_+^a *and* $(-x)_+^a$*.*

(a) *For* $\varepsilon > 0$ *and* $a \in \mathbb{C}$*, define* $(x + i\varepsilon)^a \in C^\infty(\mathbb{R})$ *by the formula* $(x + i\varepsilon)^a := \exp(a \log(x + i\varepsilon))$ *where we use the branch of log on* $\mathbb{C} \setminus (-\infty, 0]$ *which sends* $(0, \infty)$ *to reals. Similarly we can define* $(x - i\varepsilon)^a$*. Show that there exist limits in* $\mathcal{D}'(\mathbb{R})$

$$(x \pm i0)^a = \lim_{\varepsilon \to 0+} (x \pm i\varepsilon)^a \;\in\; \mathcal{D}'(\mathbb{R}).$$

(Hint: for $\operatorname{Re} a > -1$ *this is direct and* $(x \pm i0)^a$ *are locally integrable functions. For* $a = -1$*, write* $(x \pm i\varepsilon)^{-1} = \partial_x \log(x \pm i\varepsilon)$ *and note that* $\log(x \pm i\varepsilon)$ *has a distributional limit which is in* $L^1_{\mathrm{loc}}(\mathbb{R})$*. For general* $a \neq -1$*, reduce to the case of* $a + 1$ *by antidifferentiation, similarly to what was done for* x_+^a *in lecture.)*

(b) *For* $a \in \mathbb{C} \setminus -\mathbb{N}$*, express* $(x \pm i0)^a$ *as a linear combination of* x_+^a *and* $(-x)_+^a$*. (Hint: it is enough to consider the case* $\operatorname{Re} a > -1$ *by analytic continuation.)*

(c) *Show the identities*

$$(x - i0)^{-1} - (x + i0)^{-1} = 2\pi i \delta_0,$$

$$(x - i0)^{-1} + (x + i0)^{-1} = 2\,\mathrm{p.v.}\frac{1}{x}$$

(Hint: write $(x \pm i0)^{-1} = \partial_x \log(x \pm i0)$*. Note that* $\log(x \pm i0) = \log x$ *for* $x > 0$ *and* $\log(x \pm i0) = \log(-x) \pm i\pi$ *for* $x < 0$*.)*

CHAPTER 6

Convolution I

In §1.3.1 we introduced the notion of *convolution* of two functions,

$$f * g(x) = \int_{\mathbb{R}^n} f(y)g(x-y)\,dy. \tag{6.1}$$

We would like to extend this notion to distributions. In this chapter, we define the convolution when one of the factors is a distribution and another one is a smooth function, with the result which is a smooth function. We next use this notion to show that smooth functions are dense in the space of distributions.

6.1. Convolution of a distribution and a smooth function

For a function $\varphi \in C^\infty(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$, define the function

$$\varphi(x-\bullet) \in C^\infty(\mathbb{R}^n), \quad \varphi(x-\bullet)(y) = \varphi(x-y).$$

Then (6.1) can be rewritten in terms of the pairing $(\bullet,\bullet)$ as

$$f * \varphi(x) = (f, \varphi(x-\bullet)).$$

Thus we can extend the operation of convolution to the case when f is a distribution as follows:

Definition 6.1. *Assume that $u \in \mathcal{D}'(\mathbb{R}^n)$, $\varphi \in C^\infty(\mathbb{R}^n)$, and either u or φ is compactly supported. Define the function $u * \varphi : \mathbb{R}^n \to \mathbb{C}$ by*

$$u * \varphi(x) := (u, \varphi(x-\bullet)), \quad x \in \mathbb{R}^n. \tag{6.2}$$

As an example, we compute the convolution with a delta function:

Proposition 6.2. *For any $\varphi \in C^\infty(\mathbb{R}^n)$ we have*

$$\delta_0 * \varphi = \varphi. \tag{6.3}$$

Proof. Let $x \in \mathbb{R}^n$. Then

$$\delta_0 * \varphi(x) = (\delta_0, \varphi(x-\bullet)) = \varphi(x-\bullet)(0) = \varphi(x).$$

□

6.1.1. Smoothness of convolution. We now want to show that the function $u * \varphi$ defined in (6.2) is smooth. This is an application of the following general fact on the pairing of a distribution with a test function depending on a parameter, which we will use several more times later:

PROPOSITION 6.3. *Assume that* $U \Subset \mathbb{R}^n$, $V \Subset \mathbb{R}^m$, *and*

$$v \in \mathcal{E}'(V), \quad \psi \in C^\infty(U \times V). \tag{6.4}$$

For $x \in U$, *define* $\psi(x, \bullet) \in C^\infty(V)$ *by* $\psi(x, \bullet)(y) = \psi(x, y)$. *Then the function*

$$f(x) := (v, \psi(x, \bullet)), \quad x \in U \tag{6.5}$$

lies in $C^\infty(U)$, *with derivatives given by* $\partial_x^\alpha f(x) = (v, \partial_x^\alpha \psi(x, \bullet))$.

The same conclusion holds if instead of (6.4) *we assume that* $v \in \mathcal{D}'(V)$, $\psi \in C^\infty(U \times V)$, *and the restriction of the projection* $\pi_x : U \times V \to U$ *to* $\operatorname{supp} \psi$ *is proper, namely for each compact subset* $K_U \subset U$, *the preimage* $\pi_x^{-1}(K_U) \cap \operatorname{supp} \psi$ *is compact.*

PROOF. 1. Assume that (6.4) holds. Let us first show that $f \in C^0(U)$. Since $v \in \mathcal{E}'(V)$, we have the bound (4.5): there exists $K_V \Subset V$ and constants C, N such that

$$|(v, \varphi)| \le C \|\varphi\|_{C^N(V, K_V)} \quad \text{for all } \varphi \in C^\infty(V). \tag{6.6}$$

Fix $x \in U$ and estimate for $\tilde{x} \in U$ close to x

$$|f(x) - f(\tilde{x})| = \big|(v, \psi(x, \bullet) - \psi(\tilde{x}, \bullet))\big| \le C \|\psi(x, \bullet) - \psi(\tilde{x}, \bullet)\|_{C^N(V, K_V)}. \tag{6.7}$$

Since $\psi \in C^\infty(U \times V)$, we have as $\tilde{x} \to x$

$$\|\psi(x, \bullet) - \psi(\tilde{x}, \bullet)\|_{C^N(V, K_V)} = \max_{|\beta| \le N} \sup_{y \in K_V} |\partial_y^\beta \psi(x, y) - \partial_y^\beta \psi(\tilde{x}, y)| \to 0 \tag{6.8}$$

thus $f(\tilde{x}) \to f(x)$ as $\tilde{x} \to x$, which shows that f is indeed continuous.

Assume now that the alternative condition to (6.4) (in the last paragraph of the statement of the proposition) holds. Fix $x \in U$ and put $K_U := B(x, \varepsilon)$ where $\varepsilon > 0$ is small enough so that $K_U \Subset U$. Since $\pi_x|_{\operatorname{supp} \psi}$ is proper, there exists $K_V \Subset V$ such that $\operatorname{supp} \psi(\tilde{x}, \bullet) \subset K_V$ for all $\tilde{x} \in K_U$. Using the bound (2.1) for v with this set K_V, we get similarly to (6.7) and (6.8) that there exist C, N such that for all $\tilde{x} \in K_U$

$$|f(x) - f(\tilde{x})| \le C \|\psi(x, \bullet) - \psi(\tilde{x}, \bullet)\|_{C^N} \to 0 \quad \text{as } \tilde{x} \to x$$

giving again the continuity of f.

2. We now show that f is differentiable (in the classical sense) and

$$\partial_{x_j} f(x) = (v, \partial_{x_j} \psi(x, \bullet)). \tag{6.9}$$

Iterating this statement, we get that $f \in C^\infty(U)$ and the formula for the derivatives of f.

We assume that (6.4) holds; the argument can be modified for the alternative assumption in the same way as for the continuity of f above. Fix $x \in U$ and estimate for small $t \in \mathbb{R}$ (where e_j denotes the j-th basis vector in $\mathbb{R}^n$ and we use the bound (6.6))

$$\begin{aligned}\big|f(x+te_j) - f(x) - t(v, \partial_{x_j}\psi(x,\bullet))\big| &= \big|\big(v, \psi(x+te_j,\bullet) - \psi(x,\bullet) - t\partial_{x_j}\psi(x,\bullet)\big)\big| \\ &\leq C\|\psi(x+te_j,\bullet) - \psi(x,\bullet) - t\partial_{x_j}\psi(x,\bullet)\|_{C^N(V,K_V)}.\end{aligned}$$

Since $\psi \in C^\infty$, we have as $t \to 0$

$$\|\psi(x+te_j,\bullet) - \psi(x,\bullet) - t\partial_{x_j}\psi(x,\bullet)\|_{C^N(V,K_V)} = o(t),$$

which shows that

$$\lim_{t\to 0}\frac{f(x+te_j) - f(x)}{t} = (v, \partial_{x_j}\psi(x,\bullet))$$

and gives (6.9). □

Armed with Proposition 6.3, we can now prove

THEOREM 6.4. *Assume that $u \in \mathcal{D}'(\mathbb{R}^n)$, $\varphi \in C^\infty(\mathbb{R}^n)$, and either u or φ is compactly supported. Then $u * \varphi \in C^\infty(\mathbb{R}^n)$ and*

$$\partial_x^\alpha(u * \varphi) = u * (\partial_x^\alpha \varphi) = (\partial_x^\alpha u) * \varphi. \tag{6.10}$$

PROOF. Define $\psi \in C^\infty(\mathbb{R}^{2n})$ by the formula

$$\psi(x,y) = \varphi(x-y).$$

Then we can rewrite (6.2) as

$$u * \varphi(x) = (u, \psi(x,\bullet)).$$

Applying Proposition 6.3, we see that $u * \varphi \in C^\infty(\mathbb{R}^n)$ and

$$\partial_x^\alpha(u * \varphi) = (u, \partial_x^\alpha \psi(x,\bullet)) = u * \partial_x^\alpha \varphi.$$

Here if $u \in \mathcal{E}'(\mathbb{R}^n)$, then the assumption (6.4) holds. If instead $\varphi \in C_c^\infty(\mathbb{R}^n)$, then the projection $\pi_x|_{\operatorname{supp}\psi}$ is proper. Indeed, if $K \subset \mathbb{R}^n$ is compact, then we have

$$\pi_x^{-1}(K) \cap \operatorname{supp}\psi = \{(x,y) \in \mathbb{R}^{2n} \mid x \in K,\ x - y \in \operatorname{supp}\varphi\} \tag{6.11}$$

which is a compact set as it is the image of $K \times \operatorname{supp}\varphi$ by a continuous map (see also Figure 6.1).

Finally, the last equality in (6.10) follows from the definition (3.2) of distributional derivatives:

$$\begin{aligned}u * (\partial_x^\alpha \varphi)(x) &= \big(u, (\partial_x^\alpha\varphi)(x-\bullet)\big) = (-1)^{|\alpha|}\big(u, \partial_y^\alpha(\varphi(x-\bullet))\big) \\ &= \big(\partial_y^\alpha u, \varphi(x-\bullet)\big) = (\partial_x^\alpha u) * \varphi(x).\end{aligned}$$

□

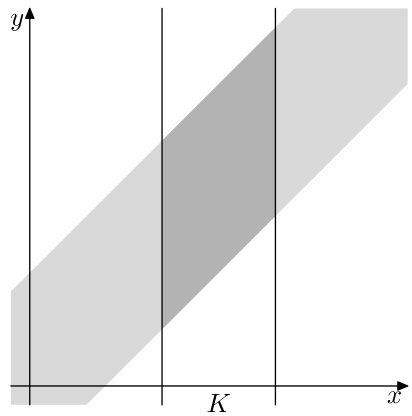

FIGURE 6.1. An illustration of the proper projection condition in the proof of Theorem 6.4. The lighter shaded area is $\operatorname{supp}\psi$ which is not compact. The darker shaded area is the compact set in (6.11).

6.1.2. Further properties of convolution. The bilinear map $(u, \varphi) \mapsto u * \varphi$ is sequentially continuous:

PROPOSITION 6.5. *Assume that* $u_k \in \mathcal{D}'(\mathbb{R}^n)$, $\varphi_k \in C_c^\infty(\mathbb{R}^n)$ *satisfy*

$$u_k \to u \quad \text{in } \mathcal{D}'(\mathbb{R}^n), \qquad \varphi_k \to \varphi \quad \text{in } C_c^\infty(\mathbb{R}^n). \tag{6.12}$$

Then $u_k * \varphi_k \to u * \varphi$ *in* $C^\infty(\mathbb{R}^n)$. *Same is true if instead* $u_k \in \mathcal{E}'(\mathbb{R}^n)$, $\varphi_k \in C^\infty(\mathbb{R}^n)$ *and we revise* (6.12) *accordingly.*

PROOF. By the formula (6.10) for derivatives of convolution, we see that it suffices to show that for each $K \subset \mathbb{R}^n$ we have

$$\sup_K |u_k * \varphi_k - u * \varphi| \to 0. \tag{6.13}$$

We assume that (6.12) holds; the argument for the case $u_k \in \mathcal{E}'$, $\varphi_k \in C^\infty$ is similar.

To show (6.13) it suffices to check that for each sequence $x_k \in K$ converging to some $x_\infty \in K$ we have

$$u_k * \varphi_k(x_k) \to u * \varphi(x_\infty). \tag{6.14}$$

By the definition (6.2) we have

$$u_k * \varphi_k(x_k) = (u_k, \varphi_k(x_k - \bullet)).$$

Since $\varphi_k \to \varphi \in C^\infty(\mathbb{R}^n)$ and $x_k \to x_\infty$, we have

$$\varphi_k(x_k - \bullet) \to \varphi(x_\infty - \bullet) \quad \text{in } C_c^\infty(\mathbb{R}^n).$$

Now the convergence statement (6.14) follows from Proposition 4.18, which itself is a corollary of the Banach–Steinhaus theorem for distributions. □

We collect some further properties of convolution in

PROPOSITION 6.6. *1. If $u \in \mathcal{D}'(\mathbb{R}^n)$, $\varphi \in C^\infty(\mathbb{R}^n)$, and either u or φ is compactly supported, then*

$$\operatorname{supp}(u * \varphi) \subset \operatorname{supp} u + \operatorname{supp} \varphi. \tag{6.15}$$

In particular, if both u and φ are compactly supported, then so is their convolution.

2. If $u \in \mathcal{D}'(\mathbb{R}^n)$ and $\varphi, \psi \in C_c^\infty(\mathbb{R}^n)$, then

$$(u * \varphi) * \psi = u * (\varphi * \psi). \tag{6.16}$$

The proofs are given as exercises below.

6.2. Approximation of distributions by smooth functions

In this section, we show that the space $C_c^\infty(U)$ is dense in $\mathcal{D}'(U)$, see Theorem 6.10 below. (It is also dense in $\mathcal{E}'(U)$, which is shown by a similar argument.)

6.2.1. The case of $\mathbb{R}^n$. Before giving the density statement for a general open set $U \Subset \mathbb{R}^n$, we consider the case $U = \mathbb{R}^n$. We follow the same mollification procedure as in §1.3.2. Fix a 'bump function'

$$\chi \in C_c^\infty(\mathbb{R}^n), \quad \operatorname{supp} \chi \subset B(0,1), \quad \int_{\mathbb{R}^n} \chi(x)\, dx = 1,$$

and define the rescaling for $\varepsilon > 0$

$$\chi_\varepsilon(x) := \varepsilon^{-n} \chi\Big(\frac{x}{\varepsilon}\Big) \ \in \ C_c^\infty(\mathbb{R}^n). \tag{6.17}$$

The next theorem implies in particular that $C^\infty(\mathbb{R}^n)$ is dense in $\mathcal{D}'(\mathbb{R}^n)$.

THEOREM 6.7. *Let $u \in \mathcal{D}'(\mathbb{R}^n)$ and define for $\varepsilon > 0$*

$$u_\varepsilon := u * \chi_\varepsilon \tag{6.18}$$

which lies in $C^\infty(\mathbb{R}^n)$ by Theorem 6.4. Then

$$u_\varepsilon \to u \quad \text{as } \varepsilon \to 0+ \quad \text{in } \mathcal{D}'(\mathbb{R}^n).$$

To prove Theorem 6.7, we need to represent (u_ε, φ) for any $\varphi \in C_c^\infty(\mathbb{R}^n)$ as the result of applying the original distribution u to some function. This is done by the following

LEMMA 6.8. *Let $u \in \mathcal{D}'(\mathbb{R}^n)$, u_ε be defined in (6.18), $\varphi \in C_c^\infty(\mathbb{R}^n)$, and define $\varphi_\varepsilon \in C_c^\infty(\mathbb{R}^n)$ by*

$$\varphi_\varepsilon(y) := \int_{\mathbb{R}^n} \chi_\varepsilon(x-y)\varphi(x)\, dx. \tag{6.19}$$

Then we have

$$(u_\varepsilon, \varphi) = (u, \varphi_\varepsilon). \tag{6.20}$$

REMARK 6.9. *Lemma 6.8 applies for any fixed ε. In fact, the precise form of the integrand is not important in the proof – the main point is that we can exchange pairing with a distribution with integration in a parameter. We will use the same idea later in these notes, referring to Lemma 6.8 (or to Proposition 7.4 below) for the scheme of the proof.*

PROOF. 1. If $u_\varepsilon \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, then (6.20) follows from Fubini's Theorem:

$$(u_\varepsilon, \varphi) = \int_{\mathbb{R}^{2n}} u(y)\chi_\varepsilon(x-y)\varphi(x)\,dxdy = (u, \varphi_\varepsilon). \tag{6.21}$$

What we need is a version of (6.21) which works when u is a distribution. Informally, the argument goes as follows: we have

$$\varphi_\varepsilon = \int_{\mathbb{R}^n} \chi_\varepsilon(x-\bullet)\varphi(x)\,dx \tag{6.22}$$

with the integral converging in the space $C^\infty_{\mathrm{c}}(\mathbb{R}^n)$. Pairing both sides with the distribution u and putting the pairing inside the integral, we get

$$(u, \varphi_\varepsilon) = \int_{\mathbb{R}^n} (u, \chi_\varepsilon(x-\bullet))\varphi(x)\,dx = \int_{\mathbb{R}^n} u_\varepsilon(x)\varphi(x)\,dx$$

which gives (6.20).

2. To make sense of the argument above, we need to show that pairing with u can be put under the integral sign in (6.22), preferably without relying on the general theory of integral with values in a topological vector space. A common way to do this is by using Riemann sums. Namely, for $\delta > 0$ define the Riemann sum for the integral (6.22)

$$\mathcal{R}_\delta := \delta^n \sum_{x\in\delta\mathbb{Z}^n} \chi_\varepsilon(x-\bullet)\varphi(x) \ \in\ C^\infty_{\mathrm{c}}(\mathbb{R}^n).$$

We have

$$\mathcal{R}_\delta \to \varphi_\varepsilon \quad \text{as } \delta \to 0+ \quad \text{in } C^\infty_{\mathrm{c}}(\mathbb{R}^n). \tag{6.23}$$

Indeed, the support condition is immediate since φ and χ_ε are compactly supported. Any derivative of $\mathcal{R}_\delta$ has the same form with χ_ε replaced by its derivative, so it suffices to show that $\mathcal{R}_\delta(y) \to \varphi_\varepsilon(y)$ uniformly in y. The latter follows in the same way as convergence of the usual Riemann sums to the integral, writing the Riemann sum as the integral of a step function and using that the function $(x, y) \mapsto \chi_\varepsilon(x-y)\varphi(y)$ is continuous and compactly supported, and thus uniformly continuous.

Since $u \in \mathcal{D}'(\mathbb{R}^n)$, by Proposition 2.6 we can pair (6.23) with u to get

$$(u, \mathcal{R}_\delta) \to (u, \varphi_\varepsilon) \quad \text{as } \delta \to 0+\,.$$

On the other hand, since u is a linear map $C^\infty_{\mathrm{c}}(\mathbb{R}^n) \to \mathbb{C}$, we compute

$$(u, \mathcal{R}_\delta) = \delta^n \sum_{x\in\delta\mathbb{Z}^n} (u, \chi_\varepsilon(x-\bullet))\varphi(x) = \delta^n \sum_{x\in\delta\mathbb{Z}^n} u_\varepsilon(x)\varphi(x).$$

This is a Riemann sum for the function $u_\varepsilon\varphi$, which converges as $\delta \to 0+$ to the integral

$$\int_{\mathbb{R}^n} u_\varepsilon(x)\varphi(x)\,dx = (u_\varepsilon, \varphi),$$

finishing the proof of (6.20). □

We can now give

Proof of Theorem 6.7. Fix $\varphi \in C_c^\infty(\mathbb{R}^n)$. We need to show that

$$(u_\varepsilon, \varphi) \to (u, \varphi) \quad \text{as } \varepsilon \to 0+\,. \tag{6.24}$$

By Lemma 6.8, we have $(u_\varepsilon, \varphi) = (u, \varphi_\varepsilon)$. Since u is a distribution, by Proposition 2.6 it suffices to show that

$$\varphi_\varepsilon \to \varphi \quad \text{in } C_c^\infty(\mathbb{R}^n) \quad \text{as } \varepsilon \to 0+\,. \tag{6.25}$$

Note that the statement (6.25) is similar to the convergence statement (1.33) from §1.3.2; indeed, φ_ε is the convolution of φ with the function $\chi_\varepsilon(-x)$.

We now show (6.25). The support condition is immediate from the definition (6.19) of φ_ε since φ is compactly supported, so it remains to show that for each multtiindex α we have uniformly in $y \in \mathbb{R}^n$

$$\partial_y^\alpha \varphi_\varepsilon(y) \to \partial_y^\alpha \varphi(y) \quad \text{as } \varepsilon \to 0+\,.$$

By a change of variables we have

$$\varphi_\varepsilon(y) = \int_{\mathbb{R}^n} \chi(w)\varphi(y + \varepsilon w)\,dw.$$

Since $\int_{\mathbb{R}^n} \chi(w)\,dw = 1$, we have

$$\begin{aligned}|\partial_y^\alpha \varphi_\varepsilon(y) - \partial_y^\alpha \varphi(y)| &\leq \int_{\mathbb{R}^n} \big|\chi(w)(\partial_y^\alpha \varphi(y + \varepsilon w) - \partial_y^\alpha \varphi(y))\big|\,dw \\ &\leq \|\chi\|_{L^1(\mathbb{R}^n)} \sup_{x \in B(y,\varepsilon)} |\partial_y^\alpha \varphi(x) - \partial_y^\alpha \varphi(y)|\end{aligned}$$

which goes to 0 as $\varepsilon \to 0+$ uniformly in y since the function $\partial_y^\alpha \varphi$ is uniformly continuous (see (1.21)). □

6.2.2. The case of a general open set. We now generalize Theorem 6.7 to distributions on an open subset of $\mathbb{R}^n$, proving the stronger statement that C_c^∞ (rather than C^∞) is dense.

Theorem 6.10. *Let $U \Subset \mathbb{R}^n$ and $u \in \mathcal{D}'(U)$. Then there exists a sequence*

$$f_k \in C_c^\infty(U), \qquad f_k \to u \quad \textit{in } \mathcal{D}'(U).$$

PROOF. 1. Using (1.14), take a sequence of compact sets exhausting U:

$$U = \bigcup_{k=1}^{\infty} K_k, \quad K_k \Subset K_{k+1}.$$

For each k, choose a cutoff function

$$\psi_k \in C_c^\infty(U), \quad \operatorname{supp}(1 - \psi_k) \cap K_k = \emptyset$$

and fix a number $\varepsilon_k > 0$ small enough so that

$$\operatorname{supp} \psi_k + B(0, \varepsilon_k) \ \Subset \ U.$$

We also require that $\varepsilon_k \to 0$ as $k \to \infty$.

Let χ_ε be the function from (6.17) and put

$$f_k := (\psi_k u) * \chi_{\varepsilon_k}.$$

Here $\psi_k u \in \mathcal{E}'(U)$ is extended by zero to an element of $\mathcal{E}'(\mathbb{R}^n)$, see Proposition 4.7. The convolution f_k lies in $C^\infty(\mathbb{R}^n)$ by Theorem 6.4; by (6.15) it is supported in $\operatorname{supp} \psi_k + B(0, \varepsilon_k) \Subset U$, so it restricts to an element of $C_c^\infty(U)$.

2. We now claim that $f_k \to u$ in $\mathcal{D}'(U)$. Take arbitrary $\varphi \in C_c^\infty(U)$ and extend it by 0 to an element of $C_c^\infty(\mathbb{R}^n)$. We need to show that

$$(f_k, \varphi) \to (u, \varphi) \quad \text{as } k \to \infty.$$

By Lemma 6.8 we have

$$(f_k, \varphi) = (u, \psi_k \varphi_{\varepsilon_k})$$

where $\varphi_{\varepsilon_k} \in C_c^\infty(\mathbb{R}^n)$ is defined by (6.19).

Since $u \in \mathcal{D}'(U)$, by Proposition 2.6 it suffices to show that

$$\psi_k \varphi_{\varepsilon_k} \to \varphi \quad \text{in } C_c^\infty(U). \tag{6.26}$$

We have

$$\operatorname{supp} \varphi_{\varepsilon_k} \ \subset \ \operatorname{supp} \varphi + B(0, \varepsilon_k).$$

Fix $\varepsilon_0 > 0$ small enough so that the compact set $K_{\varepsilon_0} := \operatorname{supp} \varphi + B(0, \varepsilon_0)$ be contained in U. Then for k large enough we have $\operatorname{supp} \varphi_{\varepsilon_k} \subset K_{\varepsilon_0}$, which implies that all the supports of $\psi_k \varphi_{\varepsilon_k}$ are contained in a k-independent compact subset of U. Moreover, if k is large enough then $\operatorname{supp} \varphi_{\varepsilon_k} \subset K_{\varepsilon_0} \subset K_k$ and thus $\psi_k \varphi_{\varepsilon_k} = \varphi_{\varepsilon_k}$. By (6.25) we have $\varphi_{\varepsilon_k} \to \varphi$ in $C_c^\infty(\mathbb{R}^n)$ as $k \to \infty$, which gives (6.26) and finishes the proof. □

6.3. Notes and exercises

Our presentation mostly follows [**Hör03**, §4.1]. An alternative presentation is available in [**FJ98**, §§5.1–5.2].

EXERCISE 6.1. (4 = 1 + 1 + 1 + 1 pts) *This exercise discusses the convolution property* (6.15).

(a) *Assume that* $X \subset \mathbb{R}^n$ *is closed and* $Y \subset \mathbb{R}^n$ *is compact. Show that the set*

$$X + Y := \{x + y \mid x \in X,\ y \in Y\} \ \subset\ \mathbb{R}^n \tag{6.27}$$

is closed.

(b) *Give an example of two closed sets* $X, Y \subset \mathbb{R}$ *such that* $X + Y$ *is not closed.*

(c) *Show the property* (6.15) *when* $u \in \mathcal{D}'(\mathbb{R}^n)$ *and* $\varphi \in C^\infty_{\mathrm{c}}(\mathbb{R}^n)$.

(d) *Give an example when the inclusion in* (6.15) *is not an equality.*

EXERCISE 6.2. (1 pt) *Let* $u \in \mathcal{D}'(\mathbb{R}^n)$, $\varphi, \psi \in C^\infty_{\mathrm{c}}(\mathbb{R}^n)$. *Show that*

$$(u * \varphi) * \psi = u * (\varphi * \psi).$$

(Hint: one way is to use density of C^∞_{c} *in* $\mathcal{D}'$*.)*

CHAPTER 7

Tensor products and distributional kernels

In this chapter we discuss two topics on distributions on a product space $U \times V$. We first define the *tensor product* $u \otimes v \in \mathcal{D}'(U \times V)$ of two distributions $u \in \mathcal{D}'(U)$, $v \in \mathcal{D}'(V)$. We next identify continuous linear operators $A : C_c^\infty(V) \to \mathcal{D}'(U)$ with their *Schwartz kernels* $\mathcal{K} \in \mathcal{D}'(U \times V)$. We finish with a discussion of the transpose of an operator and extending operators to distributions by duality; this section could have been put much earlier but knowing about Schwartz kernels gives another way to look at the transpose.

7.1. Tensor product of distributions

Let $U \Subset \mathbb{R}^n$, $V \Subset \mathbb{R}^m$. We use the letter x to denote a point in U and the letter y to denote a point in V. If $f \in L^1_{\mathrm{loc}}(U)$, $g \in L^1_{\mathrm{loc}}(V)$, then we define their tensor product $f \otimes g \in L^1_{\mathrm{loc}}(U \times V)$: as follows:

$$(f \otimes g)(x, y) = f(x)g(y). \tag{7.1}$$

We would like to extend this definition to distributions. For that we use the following consequence of Fubini's Theorem valid for any $f \in L^1_{\mathrm{loc}}(U)$, $g \in L^1_{\mathrm{loc}}(V)$, $\varphi \in C_c^\infty(U)$, $\psi \in C_c^\infty(V)$:

$$(f \otimes g, \varphi \otimes \psi) = \int_{U \times V} f(x)g(y)\varphi(x)\psi(y)\, dxdy = (f, \varphi)(g, \psi). \tag{7.2}$$

The next theorem uses (7.2) to define tensor product of distributions:

THEOREM 7.1. *Let $U \Subset \mathbb{R}^n$, $V \Subset \mathbb{R}^m$, $u \in \mathcal{D}'(U)$, and $v \in \mathcal{D}'(V)$. Then there exists unique $w \in \mathcal{D}'(U \times V)$ such that*

$$(w, \varphi \otimes \psi) = (u, \varphi)(v, \psi) \quad \text{for all } \varphi \in C_c^\infty(U),\ \psi \in C_c^\infty(V). \tag{7.3}$$

We call w the tensor product *of u and v and denote $u \otimes v := w$.*

PROOF. 1. We first show existence of w. If f, g are locally integrable functions, then for any $\beta \in C_c^\infty(U \times V)$ we have by Fubini's Theorem

$$(f \otimes g, \beta) = \int_U f(x)\Big(\int_V g(y)\beta(x, y)\, dy\Big)\, dx.$$

We use the iterated integral formula above to define w in general. Let $u \in \mathcal{D}'(U)$, $v \in \mathcal{D}'(V)$. Take any $\beta \in C_c^\infty(U \times V)$ and consider the function

$$(v(y), \beta(x,y)) = (v, \beta(x, \bullet)) \tag{7.4}$$

where $\beta(x,\bullet) \in C_c^\infty(V)$ is defined by $\beta(x,\bullet)(y) = \beta(x,y)$. By Proposition 6.3 the function defined in (7.4) lies in $C^\infty(U)$; it is also compactly supported (since β is compactly supported) and thus lies in $C_c^\infty(U)$. Thus we can apply u to that function and get the iterated distributional pairing, which we use as a definition of w:

$$(w, \beta) := \big(u(x), (v(y), \beta(x,y))\big) \quad \text{for all } \beta \in C_c^\infty(U \times V). \tag{7.5}$$

The map $w : C_c^\infty(U \times V) \to \mathbb{C}$ is linear. We next show that $w \in \mathcal{D}'(U \times V)$ is a distribution, by establishing seminorm bounds (2.1). Take arbitrary $K \Subset U \times V$. Then we have $K \subset K_U \times K_V$ for some $K_U \Subset U$, $K_V \Subset V$. Since u and v are both distributions, they satisfy the bounds (2.1): there exist C, N such that

$$|(u, \varphi)| \le C \|\varphi\|_{C^N} \quad \text{for all } \varphi \in C_c^\infty(U),\ \operatorname{supp} \varphi \subset K_U, \tag{7.6}$$

$$|(v, \psi)| \le C \|\psi\|_{C^N} \quad \text{for all } \psi \in C_c^\infty(V),\ \operatorname{supp} \psi \subset K_V. \tag{7.7}$$

Take arbitrary $\beta \in C_c^\infty(U \times V)$ such that $\operatorname{supp} \beta \subset K$. We estimate

$$\begin{aligned}
|(w, \beta)| &\le C \|(v, \beta(x, \bullet))\|_{C_x^N(U)} \\
&= C \max_{|\alpha| \le N} \sup_{x \in U} \big|(v(y), \partial_x^\alpha \beta(x,y))\big| \\
&\le C^2 \max_{|\alpha| \le N} \sup_{x \in U} \|\partial_x^\alpha \beta(x,y)\|_{C_y^N(V)} \le C^2 \|\beta\|_{C^{2N}(U \times V)}.
\end{aligned}$$

Here in the first line we use (7.6). In the second line we use the formula for the x-derivatives of $(v(y), \beta(x,y))$ from Proposition 6.3. In the third line we use (7.7). This gives a bound of the form (2.1) for w, showing that it is indeed a distribution.

Finally, the distribution w satisfies (7.3): indeed, if $\beta = \varphi \otimes \psi$ then $(v(y), \beta(x,y)) = \varphi(x)(v, \psi)$ and thus $(w, \beta) = (u, \varphi)(v, \psi)$.

We remark that we could have alternatively defined w by the formula

$$(w, \beta) := \big(v(y), (u(x), \beta(x,y))\big) \quad \text{for all } \beta \in C_c^\infty(U \times V) \tag{7.8}$$

and until we prove uniqueness it is not clear that this gives the same distribution w as (7.5).

2. We now show uniqueness of w; that is, if $w \in \mathcal{D}'(U \times V)$ satisfies

$$(w, \varphi \otimes \psi) = 0 \quad \text{for all } \varphi \in C_c^\infty(U),\ \psi \in C_c^\infty(V) \tag{7.9}$$

then $w = 0$.

We use the proof of Theorem 6.10, choosing the functions there in tensor product form. (To avoid a notational clash, we use the notation θ_k for the function called ψ_k

in the proof of Theorem 6.10.) Let K_k^U, K_k^V be families of compact sets exhausting U, V in the sense of (1.14). Put

$$\theta_k := \theta_k^U \otimes \theta_k^V \ \in\ C_c^\infty(U \times V)$$

where $\theta_k^U \in C_c^\infty(U)$ satisfies $\operatorname{supp}(1-\theta_k^U) \cap K_k^U = \emptyset$ and $\theta_k^V \in C_c^\infty(V)$ satisfies $\operatorname{supp}(1-\theta_k^V) \cap K_k^V = \emptyset$. Next, define the function

$$\chi := \chi^U \otimes \chi^V, \quad \chi^U \in C_c^\infty(B_{\mathbb{R}^n}(0, \tfrac{1}{2})),\ \chi^V \in C_c^\infty(B_{\mathbb{R}^m}(0, \tfrac{1}{2})), \quad \int_{\mathbb{R}^n} \chi^U = \int_{\mathbb{R}^m} \chi^V = 1,$$

and define $\chi_\varepsilon^U \in C_c^\infty(\mathbb{R}^n)$, $\chi_\varepsilon^V \in C_c^\infty(\mathbb{R}^m)$, and $\chi_\varepsilon \in C_c^\infty(\mathbb{R}^{n+m})$ by (6.17). As shown in Step 2 of the proof of Theorem 6.10, for a certain sequence $\varepsilon_k \to 0$ we have

$$(\theta_k w) * \chi_{\varepsilon_k} \to w \quad \text{in } \mathcal{D}'(U \times V).$$

For any $(x, y) \in U \times V$ we have

$$\begin{aligned} (\theta_k w) * \chi_{\varepsilon_k}(x, y) &= \big(\theta_k w, \chi_{\varepsilon_k}((x, y) - \bullet)\big) = (w, \varphi_{k,x} \otimes \psi_{k,y}) \\ \text{where} \quad \varphi_{k,x}(\tilde{x}) &= \theta_k^U(\tilde{x})\chi_{\varepsilon_k}^U(x - \tilde{x}), \quad \psi_{k,y}(\tilde{y}) = \theta_k^V(\tilde{y})\chi_{\varepsilon_k}^V(y - \tilde{y}). \end{aligned}$$

If w satisfies (7.9), then $(\theta_k w) * \chi_{\varepsilon_k} = 0$, which implies that $w = 0$. □

The formulas (7.5) and (7.8) are important, so we repeat them here for later use:

$$(u(x) \otimes v(y), \beta(x, y)) = \big(u(x), (v(y), \beta(x, y))\big), \tag{7.10}$$

$$(u(x) \otimes v(y), \beta(x, y)) = \big(v(y), (u(x), \beta(x, y))\big). \tag{7.11}$$

As an example, we compute the tensor product of two delta functions:

Proposition 7.2. *Let $u := \delta_0 \in \mathcal{D}'(\mathbb{R}^n)$, $v := \delta_0 \in \mathcal{D}'(\mathbb{R}^m)$. Then $u \otimes v = \delta_0 \in \mathcal{D}'(\mathbb{R}^{n+m})$.*

Proof. It suffices to note that for all $\varphi \in C_c^\infty(\mathbb{R}^n)$, $\psi \in C_c^\infty(\mathbb{R}^m)$ we have

$$(\delta_0, \varphi \otimes \psi) = \varphi(0)\psi(0) = (\delta_0, \varphi)(\delta_0, \psi).$$

□

We collect various properties of tensor products in

Proposition 7.3. *Let $U \Subset \mathbb{R}^n$, $V \Subset \mathbb{R}^m$, $W \Subset \mathbb{R}^\ell$, $u \in \mathcal{D}'(U)$, $v \in \mathcal{D}'(V)$, and $w \in \mathcal{D}'(W)$. Then:*

(1) $\operatorname{supp}(u \otimes v) = \operatorname{supp} u \times \operatorname{supp} v$;
(2) *if* $u_k \to u$ *in* $\mathcal{D}'(U)$ *and* $v_k \to v$ *in* $\mathcal{D}'(V)$, *then* $u_k \otimes v_k \to u \otimes v$ *in* $\mathcal{D}'(U \times V)$;
(3) $\partial_{x_j}(u \otimes v) = (\partial_{x_j} u) \otimes v$ *and* $\partial_{y_j}(u \otimes v) = u \otimes (\partial_{y_j} v)$;
(4) *if* $a \in C^\infty(U)$, $b \in C^\infty(V)$, *then* $(a \otimes b)(u \otimes v) = (au) \otimes (bv)$;
(5) $(u \otimes v) \otimes w = u \otimes (v \otimes w)$.

We leave the proof as an exercise below.

We finish with a statement about exchanging pairing with a fixed distribution and integration in a parameter. This is the integral counterpart of Proposition 6.3. Such a statement was used before in the proof of Lemma 6.8 but the proof below relies on existence and uniqueness of tensor product of distributions, whose proof in these notes in turn relies on Lemma 6.8.

PROPOSITION 7.4. *Let $U \Subset \mathbb{R}^n$, $V \Subset \mathbb{R}^m$, and assume that $v \in \mathcal{D}'(V)$ and $\psi \in C_c^\infty(U \times V)$. Then (in the notation of Proposition 6.3)*

$$\int_U (v, \psi(x, \bullet))\, dx = \Big(v, \int_U \psi(x, \bullet)\, dx\Big). \tag{7.12}$$

PROOF. Consider the constant function $1 \in \mathcal{D}'(U)$. Then the left-hand side of (7.12) is equal to

$$\big(1(x), (v(y), \psi(x, y))\big),$$

the right-hand side of (7.12) is equal to

$$\big(v(y), (1(x), \psi(x, y))\big),$$

and by (7.10), (7.11) both of these are equal to $(1 \otimes v, \psi)$. □

7.2. Distributional kernels

Let $U \Subset \mathbb{R}^n$, $V \Subset \mathbb{R}^m$. For any function $\mathcal{K}(x, y) \in L^1_{\mathrm{loc}}(U \times V)$, we can define the integral operator $A : L^\infty_c(V) \to L^1_{\mathrm{loc}}(U)$ by

$$A\varphi(x) = \int_V \mathcal{K}(x, y)\varphi(y)\, dy, \quad \varphi \in L^\infty_c(V), \quad x \in U. \tag{7.13}$$

We want to define an operator of the form (7.13) when $\mathcal{K}$ is a distribution. For all $\varphi \in C_c^\infty(V)$ and $\psi \in C_c^\infty(U)$ we compute by Fubini's Theorem

$$(A\varphi, \psi) = \int_{U\times V} \mathcal{K}(x, y)\psi(x)\varphi(y)\, dxdy = (\mathcal{K}, \psi \otimes \varphi). \tag{7.14}$$

We use the identity (7.14) as the definition of A when $\mathcal{K}$ is a distribution:

DEFINITION 7.5. *Let $\mathcal{K} \in \mathcal{D}'(U \times V)$. Define the linear operator $A : C_c^\infty(V) \to \mathcal{D}'(U)$ as follows:*

$$(A\varphi, \psi) = (\mathcal{K}, \psi \otimes \varphi) \quad \textit{for all } \varphi \in C_c^\infty(V),\ \psi \in C_c^\infty(U). \tag{7.15}$$

We call $\mathcal{K}$ the distributional kernel *or the* Schwartz kernel *of A.*

To check that $A\varphi$ is indeed a distribution for all $\varphi \in C_c^\infty(V)$, take any sequence $\psi_k \to 0$ in $C_c^\infty(U)$. Then $\psi_k \otimes \varphi \to 0$ in $C_c^\infty(U \otimes V)$ and thus $(A\varphi, \psi_k) \to 0$. Similarly if $\psi \in C_c^\infty(U)$ is fixed and $\varphi_k \to 0$ in $C_c^\infty(V)$, then $(A\varphi_k, \psi) \to 0$. This shows that A is a *sequentially continuous* operator $C_c^\infty(V) \to \mathcal{D}'(U)$, that is

$$\varphi_k \to 0 \quad \text{in } C_c^\infty(V) \quad \Longrightarrow \quad A\varphi_k \to 0 \quad \text{in } \mathcal{D}'(U). \tag{7.16}$$

7.2.1. The Schwartz Kernel Theorem. The next statement shows that *every* sequentially continuous operator $A : C_c^\infty(V) \to \mathcal{D}'(U)$ has the form (7.15) for a unique choice of the kernel $\mathcal{K}$. Thus we have a bijection

$$\text{Operators } C_c^\infty(V) \to \mathcal{D}'(U) \quad \simeq \quad \text{Distributions in } \mathcal{D}'(U \times V).$$

This is in contrast with integral operators on functions: for example the identity operator cannot be written in the form (7.13) for any function $\mathcal{K}$.

THEOREM 7.6 (Schwartz Kernel Theorem). *Let $U \Subset \mathbb{R}^n$, $V \Subset \mathbb{R}^m$, and $A : C_c^\infty(V) \to \mathcal{D}'(U)$ be a sequentially continuous linear operator. Then there exists unique $\mathcal{K} \in \mathcal{D}'(U \times V)$ such that A has the form* (7.15).

PROOF. We only give a sketch of the proof, sending the reader to [**Hör03**, Theorem 5.2.1] and [**FJ98**, Theorem 6.1.1] for details (we follow the latter for the sketch below). Uniqueness of $\mathcal{K}$ follows immediately from the uniqueness part of Theorem 7.1: if $\mathcal{K}$ satisfies (7.15) with $A = 0$ then $\mathcal{K} = 0$. So what we really need to show is existence.

1. Let $A : C_c^\infty(V) \to \mathcal{D}'(U)$ be sequentially continuous. We can reduce to the case when $U = (0,1)^n$, $V = (0,1)^m$ are rectangles and A is compactly supported in the sense that it is sequentially continuous $C^\infty(V) \to \mathcal{E}'(U)$. From Proposition 4.18 (which is a corollary of the Banach–Steinhaus Theorem in distributions) we see that

$$\varphi_k \to 0 \quad \text{in } C^\infty(V), \ \psi_k \to 0 \quad \text{in } C^\infty(U) \quad \Longrightarrow \quad (A\varphi_k, \psi_k) \to 0.$$

From here (arguing similarly to the proof of Proposition 2.6) we can derive the following norm bound: there exist C, N such that

$$|(A\varphi, \psi)| \le C\|\varphi\|_{C^N}\|\psi\|_{C^N} \quad \text{for all } \varphi \in C_c^\infty(V), \ \psi \in C_c^\infty(U). \tag{7.17}$$

2. We now construct the kernel $\mathcal{K}$ as a Fourier series. Namely, for $p \in \mathbb{Z}^n$, $q \in \mathbb{Z}^m$ define the complex number

$$c_{pq} := (Ae^{-2\pi i\langle y,q\rangle}, e^{-2\pi i\langle x,p\rangle}).$$

If A has the form (7.15) then c_{pq} are just the Fourier series coefficients of $\mathcal{K}$. Thus for general A we define

$$\mathcal{K}(x,y) := \sum_{p\in\mathbb{Z}^n, \ q\in\mathbb{Z}^m} c_{pq}e^{2\pi i\langle x,p\rangle}e^{2\pi i\langle y,q\rangle} \ \in \ \mathcal{D}'(U \times V). \tag{7.18}$$

The bound (7.17) implies that the sequence c_{pq} is polynomially bounded in terms of p, q, so the series (7.18) converges in $\mathcal{D}'(U\times V)$ similarly to Exercise 2.3.

It remains to show that A is given by (7.15) with $\mathcal{K}$ defined by (7.18). Take $\varphi\in C^\infty_{\mathrm{c}}(V)$, $\psi\in C^\infty_{\mathrm{c}}(U)$. Define the Fourier coefficients

$$a_p:=\int_U e^{-2\pi i\langle x,p\rangle}\psi(x)\,dx,\qquad b_q:=\int_V e^{-2\pi i\langle y,q\rangle}\varphi(y)\,dy.$$

Note that a_p,b_q are rapidly decaying (i.e. faster than any negative power of p,q) as $p,q\to\infty$. We now compute

$$\begin{aligned}(\mathcal{K},\psi\otimes\varphi)&=\sum_{p\in\mathbb{Z}^n,\ q\in\mathbb{Z}^m} c_{pq}a_{-p}b_{-q}\\ &=\sum_{p\in\mathbb{Z}^n,\ q\in\mathbb{Z}^m} a_pb_q\big(Ae^{2\pi i\langle y,q\rangle},e^{2\pi i\langle x,p\rangle}\big)\\ &=\Big(A\sum_{q\in\mathbb{Z}^m}b_qe^{2\pi i\langle y,q\rangle},\sum_{p\in\mathbb{Z}^n}a_pe^{2\pi i\langle x,p\rangle}\Big)=(A\varphi,\psi)\end{aligned}$$

proving (7.15). Here in the last line we use the bound (7.17) and the fact that the Fourier series

$$\varphi(y)=\sum_{q\in\mathbb{Z}^m}b_qe^{2\pi i\langle y,q\rangle},\qquad \psi(x)=\sum_{p\in\mathbb{Z}^n}a_pe^{2\pi i\langle x,p\rangle}$$

converge in $C^N(V)$ and $C^N(U)$ respectively. □

7.2.2. Examples and properties. As an example, we compute the Schwartz kernel of the identity operator:

PROPOSITION 7.7. *Let $U\Subset\mathbb{R}^n$. Then the Schwartz kernel of the identity operator*

$$A:C^\infty_{\mathrm{c}}(U)\to\mathcal{D}'(U),\quad A\varphi=\varphi$$

is given by the distribution $\delta_0(x-y)\in\mathcal{D}'(U\times U)$ defined as follows:

$$(\delta_0(x-y),\beta)=\int_U\beta(x,x)\,dx\quad\textit{for all }\beta\in C^\infty_{\mathrm{c}}(U\times U).$$

PROOF. Let $\varphi,\psi\in C^\infty_{\mathrm{c}}(U)$. Then

$$(A\varphi,\psi)=(\varphi,\psi)=\int_U\varphi(x)\psi(x)\,dx=(\delta_0(x-y),\psi(x)\varphi(y))$$

showing that (7.15) holds. □

REMARK 7.8. *Note that the support of $\delta_0(x-y)$ is the diagonal $\{(x,x)\mid x\in U\}$.*

We now discuss the relation between certain properties of the Schwartz kernel $\mathcal{K}$ and mapping properties of the corresponding operator A. The next lemma shows that operators with compactly supported Schwartz kernels are exactly those that extend to operators $C^\infty \to \mathcal{E}'$:

PROPOSITION 7.9. *Let $A : C_c^\infty(V) \to \mathcal{D}'(U)$ be a sequentially continuous operator with Schwartz kernel $\mathcal{K} \in \mathcal{D}'(U \times V)$. Then A extends to a sequentially continuous operator $\widetilde{A} : C^\infty(V) \to \mathcal{E}'(U)$ if and only if $\mathcal{K} \in \mathcal{E}'(U \times V)$.*

We leave the proof as an exercise below.

Another important class of operators is those that have smooth Schwartz kernels. It turns out that these correspond exactly to *smoothing operators* which will be important in later parts of the course:

PROPOSITION 7.10. *Let $A : C_c^\infty(V) \to \mathcal{D}'(U)$ be a sequentially continuous operator with Schwartz kernel $\mathcal{K} \in \mathcal{D}'(U \times V)$. Then A extends to a sequentially continuous operator $\widetilde{A} : \mathcal{E}'(V) \to C^\infty(U)$ if and only if $\mathcal{K} \in C^\infty(U \times V)$.*

PROOF. We only show the more useful direction that the smoothness of $\mathcal{K}$ implies that A is a smoothing operator. For the other direction see [**Hör03**, Theorem 5.2.6].

Let $\mathcal{K} \in C^\infty(U \times V)$. We let $\widetilde{A}$ be the integral operator (7.13) where integration is understood as distributional pairing:

$$\widetilde{A}u(x) = (u(y), \mathcal{K}(x, y)) \quad \text{for all } u \in \mathcal{E}'(V),\ x \in U.$$

By Proposition 6.3, we have $\widetilde{A}u \in C^\infty(U)$. Moreover, by (7.14) we see that $\widetilde{A}\varphi = A\varphi$ for all $\varphi \in C_c^\infty(V)$.

It remains to show that $\widetilde{A}$ is sequentially continuous. Assume that $u_k \to 0$ in $\mathcal{E}'(V)$. We need to show that $\widetilde{A}u_k \to 0$ in $C^\infty(U)$, that is for any $K_U \Subset U$ and any multiindex α we have

$$\sup_{x \in K_U} |\partial_x^\alpha \widetilde{A}u_k(x)| \to 0.$$

Arguing by contradiction we see that it suffices to show that for any sequence $x_k \to x_\infty \in K$ we have

$$\partial_x^\alpha \widetilde{A}u_k(x_k) \to 0. \tag{7.19}$$

By Proposition 6.3 we compute

$$\partial_x^\alpha \widetilde{A}u_k(x_k) = (u_k, \partial_x^\alpha \mathcal{K}(x_k, \bullet)).$$

We have $\partial_x^\alpha \mathcal{K}(x_k, \bullet) \to \partial_x^\alpha \mathcal{K}(x_\infty, \bullet)$ in $C^\infty(V)$. Since $u_k \to 0$ in $\mathcal{E}'(V)$, Proposition 4.18 shows (7.19) which finishes the proof. □

As an example we give the Schwartz kernel of a convolution operator with a smooth function (see §6.1). The proof is immediate from the definitions.

Proposition 7.11. *Assume that* $\varphi \in C^\infty(\mathbb{R}^n)$. *Define the operator*

$$A : \mathcal{E}'(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n), \quad Au = u * \varphi.$$

Then A *has the Schwartz kernel*

$$\mathcal{K}(x, y) = \varphi(x - y).$$

7.3. The transpose of an operator and defining operators by duality

We now study transposes of operators, which are useful in particular in defining various operations on distributions by duality (something we have already done in §3 without using the notion of transpose explicitly).

Definition 7.12. *Let* $U \Subset \mathbb{R}^n$, $V \Subset \mathbb{R}^m$, *and* $A : C_c^\infty(V) \to \mathcal{D}'(U)$ *be a sequentially continuous linear operator. Define its* transpose $A^t : C_c^\infty(U) \to \mathcal{D}'(V)$ *by the formula*

$$(A^t\psi, \varphi) = (A\varphi, \psi) \quad \text{for all } \varphi \in C_c^\infty(V),\ \psi \in C_c^\infty(U). \tag{7.20}$$

Note that (7.20) uniquely determines the operator A^t, which is sequentially continuous $C_c^\infty(U) \to \mathcal{D}'(V)$, and we have $(A^t)^t = A$.

There is an easy formula for the Schwartz kernel of the transpose operator: if A has Schwartz kernel $\mathcal{K} \in \mathcal{D}'(U \times V)$, then the Schwartz kernel $\mathcal{K}^t \in \mathcal{D}'(V \times U)$ of A^t is given by

$$\mathcal{K}^t(y, x) = \mathcal{K}(x, y), \tag{7.21}$$

in the sense that for each $\beta \in C_c^\infty(V \times U)$ we have

$$(\mathcal{K}^t, \beta) = (\mathcal{K}, \beta^t) \quad \text{where } \beta^t \in C_c^\infty(U \times V),\ \beta^t(x, y) = \beta(y, x).$$

Remark 7.13. *We can also consider the* adjoint operator *with respect to the sesquilinear pairing*

$$\langle u, \varphi \rangle_{L^2} := (u, \overline{\varphi}), \quad u \in \mathcal{D}'(U),\ \varphi \in C_c^\infty(U).$$

If $A : C_c^\infty(V) \to \mathcal{D}'(U)$, *then its adjoint* $A^* : C_c^\infty(U) \to \mathcal{D}'(V)$ *is given by*

$$\langle A^*\psi, \varphi \rangle_{L^2} = \langle A\varphi, \psi \rangle_{L^2} \quad \text{for all } \varphi \in C_c^\infty(V),\ \psi \in C_c^\infty(U). \tag{7.22}$$

The Schwartz kernel of the adjoint is given by

$$\mathcal{K}^*(y, x) = \overline{\mathcal{K}(x, y)}.$$

(Here the complex conjugate $\overline{u}$ *of a distribution* u *is defined by the identity* $(\overline{u}, \varphi) = \overline{(u, \overline{\varphi})}$.*)*

As an example, we compute the transpose of a partial derivative operator:

Proposition 7.14. *Let* $U \Subset \mathbb{R}^n$ *and take* $A := \partial_{x_j} : C_c^\infty(U) \to C_c^\infty(U)$. *Then* $A^t = -\partial_{x_j}$.

PROOF. What needs to be checked is that

$$(\partial_{x_j}\varphi,\psi)=-(\partial_{x_j}\psi,\varphi)\quad\text{for all }\varphi,\psi\in C_c^\infty(U)$$

and this follows from integration by parts, Theorem 1.17. □

The next theorem shows that if the transpose of an operator A maps smooth functions to smooth functions, then A can be extended to an operator on distributions. This conceptualizes the strategy used before in (3.2) and (3.6).

THEOREM 7.15. *Assume that* $U\Subset\mathbb{R}^n$, $V\Subset\mathbb{R}^m$, *and* $A:C_c^\infty(V)\to\mathcal{D}'(U)$ *is a sequentially continuous linear operator. Assume furthermore that the transpose* $A^t:C_c^\infty(U)\to\mathcal{D}'(V)$ *is a sequentially continuous operator*

$$A^t:C_c^\infty(U)\to C^\infty(V). \tag{7.23}$$

Then there exists unique sequentially continuous operator

$$\widetilde{A}:\mathcal{E}'(V)\to\mathcal{D}'(U)$$

such that $\widetilde{A}\varphi=A\varphi$ *for all* $\varphi\in C_c^\infty(V)$.

REMARK 7.16. *Once the theorem is proved, we will identify* A *and* $\widetilde{A}$, *saying that* A maps *spaces of distributions when strictly speaking* A extends *to spaces of distributions.*

REMARK 7.17. *If* A^t *has the stronger mapping property* $C_c^\infty(U)\to C_c^\infty(V)$, *then we can extend* A *to an operator* $\mathcal{D}'(V)\to\mathcal{D}'(U)$.

PROOF OF THEOREM 7.15. 1. We first show uniqueness, which follows from the density of C_c^∞ in $\mathcal{D}'$. Indeed, assume that $\widetilde{A}:\mathcal{E}'(V)\to\mathcal{D}'(U)$ is a sequentially continuous operator such that $\widetilde{A}\varphi=0$ for all $\varphi\in C_c^\infty(V)$. Take arbitrary $v\in\mathcal{E}'(V)$. By Theorem 6.10 (or rather its version for $\mathcal{E}'$), there exists a sequence

$$v_k\in C_c^\infty(V),\quad v_k\to v\quad\text{in }\mathcal{E}'(V).$$

By the sequential continuity of $\widetilde{A}$, we have $\widetilde{A}v_k\to\widetilde{A}v$ in $\mathcal{D}'(U)$. But $\widetilde{A}v_k=0$ for all k, so $\widetilde{A}v=0$. Since v was arbitrary, we get $\widetilde{A}=0$.

2. We next show existence of the extension $\widetilde{A}$. We define this extension by the simple formula

$$(\widetilde{A}v,\psi)=(v,A^t\psi)\quad\text{for all }v\in\mathcal{E}'(V),\ \psi\in C_c^\infty(U). \tag{7.24}$$

Here $A^t\psi\in C^\infty(V)$ by (7.23) and thus it can be paired with the distribution v. The rest of the proof is a routine verification:

- $\widetilde{A}v\in\mathcal{D}'(U)$ for all $v\in\mathcal{E}'(V)$. Indeed, assume that $\psi_k\to0$ in $C_c^\infty(U)$. Then from the sequential continuity of A^t between the spaces (7.23) we have $A^t\psi_k\to0$ in $C^\infty(V)$. It follows that $(\widetilde{A}v,\psi_k)\to0$, giving by Proposition 2.6 that $\widetilde{A}v\in\mathcal{D}'(U)$.

- $\widetilde{A}:\mathcal{E}'(V)\to\mathcal{D}'(U)$ is sequentially continuous. Indeed, assume that $v_k\to 0$ in $\mathcal{E}'(V)$. Then for each $\psi\in C_c^\infty(U)$ we have $(\widetilde{A}v_k,\psi)=(v_k,A^t\psi)\to 0$ and thus $\widetilde{A}v_k\to 0$ in $\mathcal{D}'(U)$.
- $\widetilde{A}\varphi = A\varphi$ for all $\varphi\in C_c^\infty(V)$. This follows immediately from the definition (7.24) of $\widetilde{A}$ and the definition (7.20) of the transpose A^t. □

7.4. Notes and exercises

We largely follow [**Hör03**, Chapter 5]. The presentation in [**FJ98**, §§4.1,4.3,6.1,2.8] is similar but with one nontrivial difference: to show uniqueness in Theorem 7.1 we follow [**Hör03**] (which is why we needed to define convolution of distributions with smooth functions before tensor products) and [**FJ98**] instead uses Fourier series.

Exercise 7.1. ($3=1+1+0.5+0.5$ pts) *Prove Proposition 7.3 (1)–(4).*

Exercise 7.2. (0.5 pts) *Prove Proposition 7.3 (5).*

Exercise 7.3. (2 pts) *Assume that $U\subset\mathbb{R}^n$, $V\subset\mathbb{R}^m$ are open, $0\in U$, and write elements of $\mathbb{R}^{n+m}$ as $(x,y)\in\mathbb{R}^n\times\mathbb{R}^m$. Show that the space of solutions $w\in\mathcal{D}'(U\times V)$ to the equations*

$$x_1 w=\ldots=x_n w=0$$

is given by distributions of the form $\delta_0\otimes v$ where $\delta_0\in\mathcal{D}'(U)$ is the delta distribution and $v\in\mathcal{D}'(V)$ is arbitrary.

Exercise 7.4. (1 pt) *Find the Schwartz kernels of the differentiation operators $\partial_{x_j}:C_c^\infty(U)\to C_c^\infty(U)$ and the multiplication operators $u\mapsto au$, where $a\in C^\infty(U)$.*

Exercise 7.5. (1 pt) *Let $A:C_c^\infty(V)\to\mathcal{D}'(U)$ be a sequentially continuous operator with Schwartz kernel $\mathcal{K}\in\mathcal{D}'(U\times V)$. Denote by $\partial_{x_j}:\mathcal{D}'(U)\to\mathcal{D}'(U)$, $\partial_{y_\ell}:C_c^\infty(V)\to C_c^\infty(V)$ the differentiation operators. Show that the composition $\partial_{x_j}A$ has Schwartz kernel $\partial_{x_j}\mathcal{K}$ and $A\partial_{y_\ell}$ has Schwartz kernel $-\partial_{y_\ell}\mathcal{K}$.*

Exercise 7.6. (1 pt) *Prove Proposition 7.9.*

CHAPTER 8

Convolution II

Armed with tensor products, we now introduce convolution of two distributions on $\mathbb{R}^n$ (under appropriate condition on their supports), which is itself a distribution on $\mathbb{R}^n$. This moves us one step closer to the formula (5.1) for a solution of a constant coefficient PDE.

8.1. The case of compact supports

We first consider the technically simpler case when both distributions have compact support. Assume first that $f, g \in L^1_c(\mathbb{R}^n)$ are functions. For any $\varphi \in C^\infty(\mathbb{R}^n)$ we compute using Fubini's Theorem

$$(f * g, \varphi) = \int_{\mathbb{R}^{2n}} f(x-y)g(y)\varphi(x)\,dxdy$$

$$= \int_{\mathbb{R}^{2n}} f(x)g(y)\varphi(x+y)\,dxdy = (f(x)\otimes g(y), \varphi(x+y)).$$

This motivates the following definition of convolution in terms of the distributional tensor product introduced in §7.1:

Definition 8.1. *Let $u, v \in \mathcal{E}'(\mathbb{R}^n)$. Define the convolution $u * v \in \mathcal{E}'(\mathbb{R}^n)$ as follows:*

$$(u * v, \varphi) = (u(x)\otimes v(y), \varphi(x+y)) \quad \textit{for all } \varphi \in C^\infty(\mathbb{R}^n). \tag{8.1}$$

It is immediate that $u * v$ is indeed a distribution: if $\varphi_k \to 0$ in $C^\infty(\mathbb{R}^n)$ then $\varphi_k(x+y) \to 0$ in $C^\infty(\mathbb{R}^{2n})$ and thus (as $u\otimes v \in \mathcal{E}'(\mathbb{R}^{2n})$) we have $(u * v, \varphi_k) \to 0$. The resulting operation is sequentially continuous:

Proposition 8.2. *Assume that $u_k \to u$ and $v_k \to v$ in $\mathcal{E}'(\mathbb{R}^n)$. Then*

$$u_k * v_k \to u * v \quad \textit{in } \mathcal{E}'(\mathbb{R}^n).$$

Proof.**S** Take arbitrary $\varphi \in C^\infty(\mathbb{R}^n)$. By part (2) of Proposition 7.3 (or rather its version for $\mathcal{E}'$) we have $u_k \otimes v_k \to u\otimes v$ in $\mathcal{E}'(\mathbb{R}^n)$. Thus $(u_k * v_k, \varphi) \to (u * v, \varphi)$. □

The next statement shows that if u is a distribution and v is a smooth function, then the convolution defined in §6.1 is the same as the convolution defined in the present section:

PROPOSITION 8.3. *Assume that* $u \in \mathcal{E}'(\mathbb{R}^n)$ *and* $v \in C_c^\infty(\mathbb{R}^n)$. *Let* $u \star v \in C_c^\infty(\mathbb{R}^n)$ *be as in Definition 6.1, namely*

$$u \star v(x) = (u, v(x - \bullet)) \quad \textit{for all } x \in \mathbb{R}^n.$$

Then $u \star v = u * v$ *where* $u * v \in \mathcal{E}'(\mathbb{R}^n)$ *is defined by (8.1).*

PROOF. First proof: Take arbitrary $\varphi \in C^\infty(\mathbb{R}^n)$. We need to show that

$$(u \star v, \varphi) = (u(x) \otimes v(y), \varphi(x + y)). \tag{8.2}$$

The left-hand side is equal to

$$(u \star v, \varphi) = \int_{\mathbb{R}^n} (u, v(x - \bullet))\varphi(x)\, dx.$$

Using Riemann sums similarly to the proof of Lemma 6.8 (or alternatively using a slight modification of Proposition 7.4), we pull the pairing with u out of the integral to get

$$(u \star v, \varphi) = (u, \psi) \quad \text{where } \psi(y) := \int_{\mathbb{R}^n} v(x - y)\varphi(x)\, dx, \quad \psi \in C^\infty(\mathbb{R}^n).$$

Making a change of variables, we see that

$$\psi(y) = \int_{\mathbb{R}^n} v(x)\varphi(x + y)\, dx,$$

thus (recalling the formula (7.10) for the distributional tensor product, or strictly speaking, its analog for $\mathcal{E}'$)

$$(u \star v, \varphi) = \big(u(y), (v(x), \varphi(x + y))\big) = (u(x) \otimes v(y), \varphi(x + y)).$$

Second proof: By Theorem 6.7 (or rather, its version for $\mathcal{E}'$) there exists a sequence $u_k \in C_c^\infty(\mathbb{R}^n)$ converging to u in $\mathcal{E}'(\mathbb{R}^n)$. We have $u_k \star v = u_k * v$ since both are given by the integral formula (1.27). By Propositions 6.5 and 8.2 we have

$$u_k \star v \to u \star v, \quad u_k * v \to u * v \quad \text{in } \mathcal{E}'(\mathbb{R}^n).$$

Thus $u \star v = u * v$. □

We collect basic properties of convolution of two compactly supported distributions in

Proposition 8.4. *Let* $u, v, w \in \mathcal{E}'(\mathbb{R}^n)$. *Then*

$$u * v = v * u, \tag{8.3}$$

$$u * (v * w) = (u * v) * w, \tag{8.4}$$

$$\operatorname{supp}(u * v) \subset \operatorname{supp} u + \operatorname{supp} v, \tag{8.5}$$

$$\partial^\alpha (u * v) = (\partial^\alpha u) * v = u * (\partial^\alpha v), \tag{8.6}$$

$$\delta_0 * u = u. \tag{8.7}$$

Proof.[S] All of the properties except (8.5) can be proved by approximating u, v, w by test functions in $C_c^\infty(\mathbb{R}^n)$ (Theorem 6.7) and using that these properties hold for test functions together with continuity of convolution (Proposition 8.2). In fact, (8.5) can also be shown this way if we pay attention to the supports of the approximating functions.

Below we provide more direct proofs for the sake of completeness. Let $\varphi \in C^\infty(\mathbb{R}^n)$ be arbitrary.

(8.3): Follows immediately from (8.1) and the fact that $\varphi(x + y) = \varphi(y + x)$.

(8.4): Using (7.10) we see that when paired with φ, both sides are equal to

$$(u(x) \otimes v(y) \otimes w(z), \varphi(x + y + z))$$

where $u \otimes v \otimes w = (u \otimes v) \otimes w = u \otimes (v \otimes w)$ is well-defined thanks to Proposition 7.3(5).

(8.5): We know (see Exercise 6.1(1)) that $\operatorname{supp} u + \operatorname{supp} v$ is closed. Next, if $\varphi \in C^\infty(\mathbb{R}^n)$ and $\operatorname{supp} \varphi \cap (\operatorname{supp} u + \operatorname{supp} v) = \emptyset$ then by Proposition 7.3(1) we have

$$\operatorname{supp}(u(x) \otimes v(y)) \cap \operatorname{supp}(\varphi(x + y)) = \emptyset$$

and thus by Proposition 4.9 we get $(u * v, \varphi) = 0$.

(8.6): It is enough to differentiate once. We compute using Proposition 7.3(3)

$$\begin{aligned}(\partial_{x_j}(u * v), \varphi) &= -(u * v, \partial_{x_j}\varphi) = -(u(x) \otimes v(y), (\partial_{x_j}\varphi)(x + y)) \\ &= -\big(u(x) \otimes v(y), \partial_{x_j}(\varphi(x + y))\big) = (\partial_{x_j}(u(x) \otimes v(y)), \varphi(x + y)) \\ &= ((\partial_{x_j} u)(x) \otimes v(y), \varphi(x + y)) = ((\partial_{x_j} u) * v, \varphi).\end{aligned}$$

(8.7): We compute using (7.10)

$$(\delta_0 * u, \varphi) = (\delta_0(x) \otimes u(y), \varphi(x + y)) = \big(\delta_0(x), (u(y), \varphi(x + y))\big) = (u, \varphi).$$

□

8.2. The case of properly summing supports

We now generalize the construction of §8.1 to cases when u, v are not necessarily compactly supported. This requires that the supports $\operatorname{supp} u, \operatorname{supp} v$ sum properly in the sense defined below. The significance of this condition is explained in the discussion following (8.10).

Definition 8.5. *Let* $V_1, V_2 \subset \mathbb{R}^n$ *be closed subsets. We say that* V_1, V_2 sum properly *if for each* $R > 0$ *there exists* $T > 0$ *such that for all* $x, y \in \mathbb{R}^n$

$$x \in V_1,\ y \in V_2,\ |x+y| \leq R \quad \Longrightarrow \quad |x|, |y| \leq T. \tag{8.8}$$

In other words, if $x \in V_1, y \in V_2$ *and* $|x|$ *and/or* $|y|$ *is large, then* $|x+y|$ *is also large.*

Some basic properties of properly summing sets are collected in

Proposition 8.6. *1. If* V_1, V_2 *sum properly then their sum*

$$V_1 + V_2 = \{x + y \mid x \in V_1,\ y \in V_2\}$$

is a closed subset of $\mathbb{R}^n$.

2. If one of V_1, V_2 *is compact, then* V_1, V_2 *sum properly.*

We leave the proof as an exercise below.

We now come back to the more general definition of convolution. Let $u, v \in \mathcal{D}'(\mathbb{R}^n)$ and assume that $\operatorname{supp} u, \operatorname{supp} v$ sum properly. We would like to define the convolution $u * v$ as an element of $\mathcal{D}'(\mathbb{R}^n)$. Take arbitrary $\varphi \in C^\infty_c(\mathbb{R}^n)$, then, following (8.1) we want to put

$$(u * v, \varphi) := (u(x) \otimes v(y), \varphi(x+y)). \tag{8.9}$$

Here $u \otimes v \in \mathcal{D}'(\mathbb{R}^{2n})$ and $\varphi(x+y) \in C^\infty(\mathbb{R}^{2n})$ need not be compactly supported. However, the intersection of their supports is compact:

$$\operatorname{supp}(u \otimes v) \cap \operatorname{supp}(\varphi(x+y)) \ \Subset\ \mathbb{R}^n. \tag{8.10}$$

Indeed, since $\varphi \in C^\infty_c(\mathbb{R}^n)$, there exists $R > 0$ such that $\operatorname{supp} \varphi \subset B(0,R)$. Let $(x,y) \in \operatorname{supp}(u \otimes v) \cap \operatorname{supp}(\varphi(x+y))$. By Proposition 7.3(1) we have $x \in \operatorname{supp} u$, $y \in \operatorname{supp} v$. Moreover, $x + y \in \operatorname{supp} \varphi$ so $|x+y| \leq R$. Since $\operatorname{supp} u, \operatorname{supp} v$ sum properly, we have $|x|, |y| \leq T$ for some fixed $T > 0$. Thus

$$\operatorname{supp}(u \otimes v) \cap \operatorname{supp}(\varphi(x+y)) \ \subset\ B(0,T) \times B(0,T)$$

is bounded and thus compact, giving (8.10).

Given the compact intersection of supports, we can make sense of the pairing (8.9) using the following general

Proposition 8.7. *Let* $U \Subset \mathbb{R}^n$, $u \in \mathcal{D}'(U)$, $\varphi \in C^\infty(U)$, *and assume that* $\operatorname{supp} u \cap \operatorname{supp} \varphi$ *is compact. Define* $(u, \varphi) \in \mathbb{C}$ *as follows (see Figure 8.1):*

$$(u, \varphi) := (u, \chi\varphi) \quad \textit{where } \chi \in C^\infty_c(U),\ \operatorname{supp}(1-\chi) \cap \operatorname{supp} u \cap \operatorname{supp} \varphi = \emptyset. \tag{8.11}$$

Then:

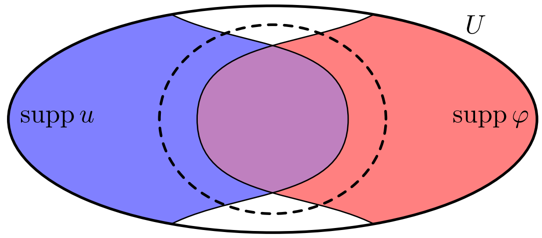

FIGURE 8.1. An illustration of (8.11). The support of u is in blue and the support of φ is in red. The dashed line denotes the support of χ.

(1) *the value of* (u, φ) *does not depend on the choice of the cutoff* χ*;*
(2) *if* $\varphi \in C_c^\infty(U)$ *or* $u \in \mathcal{D}'(U)$*, then* (u, φ) *equals to the distributional pairing defined before (in* (2.4) *and* §*4.2);*
(3) *the expression* (u, φ) *is bilinear in* u, φ*;*
(4) *if* $\operatorname{supp} u \cap \operatorname{supp} \varphi = \emptyset$*, then* $(u, \varphi) = 0$.

REMARK 8.8. *The formula* (8.11) *gives the* only *way to extend the distributional pairing which satisfies the above properties – see* [**Hör03**, Theorem 2.2.5].

PROOF.[S] (1): If $\widetilde{\chi}$ is another cutoff satisfying (8.11) then $\operatorname{supp} u \cap \operatorname{supp}((\chi - \widetilde{\chi})\varphi) = \emptyset$ and thus $(u, \chi\varphi) = (u, \widetilde{\chi}\varphi)$ by Proposition 4.2.

(2): if $\varphi \in C_c^\infty(U)$, then take χ such that $\chi\varphi = \varphi$. If $u \in \mathcal{E}'(U)$, then take χ such that $\operatorname{supp}(1 - \chi) \cap \operatorname{supp} u = \emptyset$ and use Proposition 4.9.

(3): we need to check the formula for $(a_1 u_1 + a_2 u_2, b_1\varphi_1 + b_2\varphi_2)$ where $a_j, b_j \in \mathbb{C}$, and it suffices to choose χ so that (8.11) holds for each u_j, φ_k.

(4): we may choose $\chi = 0$. □

We can now formally give

DEFINITION 8.9. *Let* $u, v \in \mathcal{D}'(\mathbb{R}^n)$ *and assume that* $\operatorname{supp} u, \operatorname{supp} v$ *sum properly. Define* $u * v \in \mathcal{D}'(\mathbb{R}^n)$ *by the formula* (8.9)*, where the pairing is defined using Proposition 8.7 and* (8.10).

It is straightforward to check that $u * v$ is indeed a distribution, and that for $u, v \in \mathcal{E}'(\mathbb{R}^n)$ it coincides with the convolution defined in §8.1. Moreover, if R, T satisfy (8.8) with $V_1 = \operatorname{supp} u$, $V_2 = \operatorname{supp} v$, and $\chi \in C_c^\infty(\mathbb{R}^n)$ satisfies $\operatorname{supp}(1 - \chi) \cap B(0, T) = \emptyset$, then

$$u * v|_{B^\circ(0,R)} = (\chi u) * (\chi v)|_{B^\circ(0,R)} \tag{8.12}$$

where the right-hand side of (8.12) features the convolution of two compactly supported distributions $\chi u, \chi v$.

Other properties of convolution are collected in

PROPOSITION 8.10.[S] *1. For any $u \in \mathcal{D}'(\mathbb{R}^n)$ we have $\delta_0 * u = u$.*

2. If $u, v \in \mathcal{D}'(\mathbb{R}^n)$ and $\operatorname{supp} u, \operatorname{supp} v$ sum properly, then

$$u * v = v * u, \tag{8.13}$$

$$\operatorname{supp}(u * v) \subset \operatorname{supp} u + \operatorname{supp} v, \tag{8.14}$$

$$\partial^\alpha(u * v) = (\partial^\alpha u) * v = u * (\partial^\alpha v). \tag{8.15}$$

3. If $u, v, w \in \mathcal{D}'(\mathbb{R}^n)$ and $\operatorname{supp} u, \operatorname{supp} v, \operatorname{supp} w$ sum properly (defined similarly to (8.8)), then $u * (v * w) = (u * v) * w$.

4. If $u \in \mathcal{D}'(\mathbb{R}^n)$, $\varphi \in C^\infty(\mathbb{R}^n)$, and either u or φ is compactly supported, then the convolution $u * \varphi$ from Definition 8.9 is the same as the convolution from Definition 6.1.

PROOF. This follows routinely from the properties of convolution of distributions in $\mathcal{E}'(\mathbb{R}^n)$ discussed in §8.1 and the formula (8.12). □

We finish with some examples illustrating the properly summing condition:

- By part 2 of Proposition 8.6, a compactly supported distribution in $\mathcal{E}'(\mathbb{R}^n)$ can be convolved with any distribution in $\mathcal{D}'(\mathbb{R}^n)$.
- The set $[0, \infty) \subset \mathbb{R}$ sums properly with itself, in fact we can take $T := R$ in (8.8). Thus we can for example define the convolution of two Heaviside functions $H * H$; one can compute $H * H(x) = x_+^1$. See Figure 8.2.
- On the other hand, the set $[0, \infty)$ does not sum properly with the set $(-\infty, 0]$. Thus we cannot define, for example, the convolution of the Heaviside function H with the function $\check{H}(x) := H(-x)$. Note that the usual definition (1.27) does not work either: we get

$$H * \check{H}(x) = \int_{\max(0,x)} dy = \infty.$$

8.3. Singular support and convolutions

In §4.1 we defined the support of a distribution $u \in \mathcal{D}'(U)$ as follows: a point x does not lie in $\operatorname{supp} u$ if it has a neighborhood V such that $u|_V = 0$. We now define *singular support* by replacing the requirement that $u|_V = 0$ by $u|_V$ being smooth:

DEFINITION 8.11. *Let $U \Subset \mathbb{R}^n$ and $u \in \mathcal{D}'(U)$. We say that a point $x \in U$ does **not** lie in $\operatorname{sing\,supp} u$ if there exists $V \Subset U$ containing x and such that $u|_V \in C^\infty(V)$.*

Here we recall that $C^\infty(V)$ is embedded into $\mathcal{D}'(V)$ by (2.2), so when we say $u|_V \in C^\infty(V)$ we strictly speaking mean that there exists $f \in C^\infty(V)$ such that

$$(u, \varphi) = \int_V f(x)\varphi(x)\,dx \quad \text{for all } \varphi \in C_c^\infty(V).$$

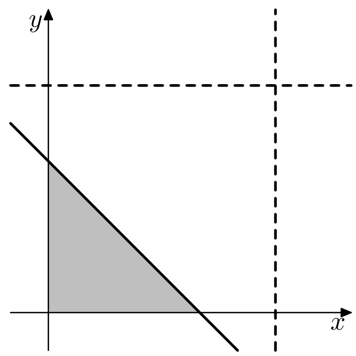

FIGURE 8.2. An illustration of (8.12) for $\operatorname{supp} u, \operatorname{supp} v \subset [0,\infty)$. The shaded region is the intersection $(\operatorname{supp} u \times \operatorname{supp} v) \cap \{(x,y) : |x+y| \le R\}$. The dashed lines denote the boundaries of the sets $\{\chi(x) = 1\}$ and $\{\chi(y) = 1\}$.

Similarly to Proposition 4.2 we have

PROPOSITION 8.12. *Let* $u \in \mathcal{D}'(U)$. *Then* $u|_{U \setminus \operatorname{sing\,supp} u} \in C^\infty(U \setminus \operatorname{sing\,supp} u)$. *In particular,* $\operatorname{sing\,supp} u = \emptyset$ *if and only if* $u \in C^\infty(U)$.

PROOF.[S] For each $x \in U \setminus \operatorname{sing\,supp} u$, there exists an open set $V_x \subset U \setminus \operatorname{sing\,supp} u$ containing x and a smooth function $v_x \in C^\infty(V_x)$ such that $u|_{V_x} = v_x$. The sets V_x cover $U \setminus \operatorname{sing\,supp} u$, and we have the compatibility conditions $v_x|_{V_x \cap V_y} = v_y|_{V_x \cap V_y}$ for all x, y. Thus there exists $v \in C^\infty(U \setminus \operatorname{sing\,supp} u)$ such that $v|_{V_x} = v_x$ for all x. By the uniqueness part of Theorem 2.13 we have $u|_{U \setminus \operatorname{sing\,supp} u} = v$. □

Some basic properties of singular support are collected in

PROPOSITION 8.13. *Assume that* $U \Subset \mathbb{R}^n$, $u \in \mathcal{D}'(U)$, *and* $a \in C^\infty(U)$. *Then*

$$\operatorname{sing\,supp}(\partial_{x_j} u) \subset \operatorname{sing\,supp} u, \tag{8.16}$$

$$\operatorname{sing\,supp}(au) \subset \operatorname{sing\,supp} u, \tag{8.17}$$

$$\operatorname{sing\,supp} u \subset \operatorname{supp} u. \tag{8.18}$$

The proofs are immediate.

A somewhat harder to establish property, used crucially in the proof of elliptic regularity in §9.2 below, is the behavior of singular support under convolution:

PROPOSITION 8.14. *Assume that* $u, v \in \mathcal{D}'(\mathbb{R}^n)$ *and* $\operatorname{supp} u, \operatorname{supp} v$ *sum properly. Then*

$$\operatorname{sing\,supp}(u * v) \ \subset \ \operatorname{sing\,supp} u + \operatorname{sing\,supp} v. \tag{8.19}$$

REMARK 8.15. *Note that* (8.19) *is nontrivial even if one of the sets* $\operatorname{sing\,supp} u$, $\operatorname{sing\,supp} v$ *is empty: in this case it states that the convolution of a smooth function with a distribution is smooth. This special case is used as a step in the proof below.*

PROOF. 1. We first assume that $u, v \in \mathcal{E}'(\mathbb{R}^n)$. We write u as the sum of two pieces: one whose *support* is in a small neighborhood of the *singular support* of u and another one which is smooth. Namely, fix $\varepsilon > 0$ and take a cutoff function

$$\psi_u \in C_c^\infty(\mathbb{R}^n), \quad \operatorname{supp} \psi_u \subset \operatorname{sing\,supp} u + B(0,\varepsilon), \quad \operatorname{supp}(1-\psi_u) \cap \operatorname{sing\,supp} u = \emptyset.$$

Then we write $u = \psi_u u + (1-\psi_u)u$ where

$$\operatorname{supp}(\psi_u u) \subset \operatorname{sing\,supp} u + B(0,\varepsilon), \quad (1-\psi_u)u \in C_c^\infty(\mathbb{R}^n).$$

In the same way we write $v = \psi_v v + (1-\psi_v)v$ where

$$\operatorname{supp}(\psi_v v) \subset \operatorname{sing\,supp} v + B(0,\varepsilon), \quad (1-\psi_v)v \in C_c^\infty(\mathbb{R}^n).$$

We now decompose

$$u * v = (\psi_u u) * (\psi_v v) + (\psi_u u) * ((1-\psi_v)v) + ((1-\psi_u)u) * v.$$

We have by (8.18) and (8.5)

$$\begin{aligned}\operatorname{sing\,supp}\big((\psi_u u) * (\psi_v v)\big) &\subset \operatorname{supp}(\psi_u u) + \operatorname{supp}(\psi_v v)\\ &\subset \operatorname{sing\,supp} u + \operatorname{sing\,supp} v + B(0,2\varepsilon)\end{aligned}$$

where we used that $B(0,\varepsilon)+B(0,\varepsilon) \subset B(0,2\varepsilon)$. On the other hand, $(\psi_u u)*((1-\psi_v)v)$ and $((1-\psi_u)u)*v$ are convolutions of a distribution in $\mathcal{E}'(\mathbb{R}^n)$ and a function in $C_c^\infty(\mathbb{R}^n)$, thus by Proposition 8.3 they lie in $C_c^\infty(\mathbb{R}^n)$. It follows that

$$\operatorname{sing\,supp}(u * v) \ \subset \ \operatorname{sing\,supp} u + \operatorname{sing\,supp} v + B(0,2\varepsilon).$$

Since this is true for any $\varepsilon > 0$ and $\operatorname{sing\,supp} u + \operatorname{sing\,supp} v$ is closed, we get (8.19).

2. We now consider the general case, when $u, v \in \mathcal{D}'(\mathbb{R}^n)$ and $\operatorname{supp} u, \operatorname{supp} v$ sum properly. It suffices to show that for each $R > 0$, we have

$$(\operatorname{sing\,supp} u * v) \cap B^\circ(0,R) \ \subset \ \operatorname{sing\,supp} u + \operatorname{sing\,supp} v. \tag{8.20}$$

The left-hand side of (8.20) is the singular support of $u * v|_{B^\circ(0,R)}$, and by (8.12)

$$u * v|_{B^\circ(0,R)} = (\chi u) * (\chi v)|_{B^\circ(0,R)}$$

for a correct choice of the cutoff $\chi \in C_c^\infty(\mathbb{R}^n)$. Applying Step 1 of the present proof to $\chi u, \chi v \in \mathcal{E}'(\mathbb{R}^n)$, we have

$$\begin{aligned}(\operatorname{sing\,supp} u * v) \cap B^\circ(0,R) \ &\subset \ \operatorname{sing\,supp}(\chi u) + \operatorname{sing\,supp}(\chi v)\\ &\subset \ \operatorname{sing\,supp} u + \operatorname{sing\,supp} v\end{aligned}$$

giving (8.20). □

8.4. Notes and exercises

Our presentation mostly follows [**FJ98**, §5.1–5.3 and Lemma 8.6.1]. The presentation in [**Hör03**, §4.2] is different because it comes before the definition of the tensor product of distributions.

EXERCISE 8.1. ($1 = 0.5 + 0.5$ pts) *Prove Proposition 8.6.*

EXERCISE 8.2. (1 pt) *Assume that* $\operatorname{Re} a, \operatorname{Re} b > 0$. *Show that* $x_+^{a-1} * x_+^{b-1} = B(a,b)x_+^{a+b-1}$ *where* B *denotes the beta function. (You can use the standard integral formula for convolution, no need to do things distributionally here. Note: using analytic continuation one can show that the same formula actually holds for all* $a, b \in \mathbb{C}$*, but you don't have to do this.)*

EXERCISE 8.3. (1 pt) *Denote elements in* $\mathbb{R}^n$ *(where* $n \geq 2$*) by* $x = (x_1, x')$ *where* $x' \in \mathbb{R}^{n-1}$. *Define the set* $\Omega := \{x \colon x_1 \geq |x'|\}$. *Show that* $\Omega + \Omega = \Omega$. *Show also that* Ω *sums properly with the set* $\{x_1 \geq 0\}$. *Does the set* $\{x_1 \geq 0\}$ *sum properly with itself?*

CHAPTER 9

Fundamental solutions and elliptic regularity

In this chapter we show the formula (5.1) for a solution to a constant coefficient partial differential equation. We give several basic examples and then prove the first version of Elliptic Regularity, for constant coefficient operators which have fundamental solutions with singular support at the origin.

9.1. Fundamental solutions

9.1.1. Basic properties. We first give the general definition of a linear differential operator (with smooth coefficients):

DEFINITION 9.1. *Let* $U \Subset \mathbb{R}^n$ *and* $m \in \mathbb{N}_0$. *A* differential operator *of order* m *on* U *is an expression of the form*

$$P = \sum_{|\alpha| \le m} a_\alpha(x) \partial_x^\alpha \tag{9.1}$$

where $a_\alpha \in C^\infty(U)$ *are called the* coefficients *of* P. *Denote by*

$$\mathrm{Diff}^m(U) \tag{9.2}$$

the space of all differential operators on U. *For* $P \in \mathrm{Diff}^m(U)$ *we say that* P *has* constant coefficients *if each of the functions* a_α *is constant.*

Here are some basic properties of differential operators:

- each $P \in \mathrm{Diff}^m(U)$ maps each of the spaces $C_c^\infty(U)$, $C^\infty(U)$, $\mathcal{E}'(U)$, $\mathcal{D}'(U)$ to itself;
- if $P \in \mathrm{Diff}^m(U)$, $Q \in \mathrm{Diff}^\ell(U)$, then their composition PQ is a differential operator in $\mathrm{Diff}^{m+\ell}(U)$;
- if $P \in \mathrm{Diff}^m(U)$, then the transpose P^t (see Definition 7.12) also lies in $\mathrm{Diff}^m(U)$, in fact if P is given by (9.1) then P^t is given by the formula

$$P^t u = \sum_{|\alpha| \le m} (-1)^{|\alpha|} \partial_x^\alpha (a_\alpha u) \quad \text{for all } u \in \mathcal{D}'(U) \tag{9.3}$$

  as we can see from Proposition 7.14.

- if $P \in \mathrm{Diff}^m(U)$ and $u \in \mathcal{D}'(U)$ then

$$\operatorname{supp}(Pu) \subset \operatorname{supp} u, \tag{9.4}$$

$$\operatorname{sing\,supp}(Pu) \subset \operatorname{sing\,supp} u. \tag{9.5}$$

In this chapter we study differential operators with constant coefficients. For such operators a key object is a *fundamental solution*:

Definition 9.2. *Let P be a differential operator with constant coefficients. A distribution $E \in \mathcal{D}'(\mathbb{R}^n)$ is called a* fundamental solution *of P if it solves the differential equation*

$$PE = \delta_0. \tag{9.6}$$

Remark 9.3. *For any $u \in \mathcal{D}'(\mathbb{R}^n)$ and $\varphi \in C_c^\infty(\mathbb{R}^n)$ we have*

$$(Pu, \varphi) = (u, P^t\varphi)$$

where P^t is the transpose of P. Thus E is a fundamental solution of P if and only if

$$(E, P^t\varphi) = \varphi(0) \quad \text{for all } \varphi \in C_c^\infty(\mathbb{R}^n).$$

Fundamental solutions are important because they give a way of describing (some) solutions of the more general equation $Pu = f$ where f is a distribution. To state this we use the notion of convolution of distributions from §8.2.

Theorem 9.4. *Let P be a differential operator with constant coefficients and E be a fundamental solution of P. Then:*

1. *if $u \in \mathcal{D}'(\mathbb{R}^n)$ and $\operatorname{supp} u, \operatorname{supp} E$ sum properly then $u = E * (Pu)$;*
2. *if $f \in \mathcal{D}'(\mathbb{R}^n)$ and $\operatorname{supp} f, \operatorname{supp} E$ sum properly then $P(E * f) = f$.*

Proof. From (8.15) we see that for any $v, w \in \mathcal{D}'(\mathbb{R}^n)$ such that $\operatorname{supp} v, \operatorname{supp} w$ sum properly we have

$$P(v * w) = (Pv) * w = v * (Pw). \tag{9.7}$$

To show part (1) of the theorem, we apply this statement to u and E, getting

$$E * (Pu) = (PE) * u = \delta_0 * u = u.$$

For part (2) of the theorem, we apply (9.7) to f and E, getting

$$P(E * f) = (PE) * f = \delta_0 * f = f.$$

$\square$

Remark 9.5. *The proper sum condition always holds if one of the sets is compact (see Proposition 8.6). Thus we have in particular:*

1. *if $u \in \mathcal{E}'(\mathbb{R}^n)$ solves the equation $Pu = f$ then $u = E * f$ (one can think of this as a* uniqueness *statement for the equation $Pu = f$);*

(2) *if $f \in \mathcal{E}'(\mathbb{R}^n)$ and we define $u := E * f$, then u solves the equation $Pu = f$ (one can think of this as an* existence *statement).*

However, these statements come with a very serious restriction that u or f be compactly supported. As an example, the Laplace equation $\Delta u = 0$ has no nontrivial compactly supported solutions but it does have plenty of non-compactly supported ones (e.g. $u \equiv 1$). This also shows that in part (1) of Theorem 9.4 it is important that $\operatorname{supp} E$ *sums properly with* $\operatorname{supp} u$*, not just with* $\operatorname{supp}(Pu)$*. See §9.1.3 below for why this is needed on a simple example.*

9.1.2. Examples of fundamental solutions. We now give a few examples of fundamental solutions for important constant coefficient operators. We start with the Laplace operator:

PROPOSITION 9.6. *A fundamental solution of the Laplace operator $\Delta = \partial_{x_1}^2 + \cdots + \partial_{x_n}^2$ on $\mathbb{R}^n$ is given by the locally integrable function*

$$E(x) = \begin{cases} \frac{1}{2}|x|, & n = 1, \\ \frac{1}{2\pi}\log|x|, & n = 2, \\ -c_n|x|^{2-n}, & n \geq 3. \end{cases} \tag{9.8}$$

Here $c_n = \frac{1}{(n-2)\operatorname{vol}(\mathbb{S}^{n-1})}$ and $\operatorname{vol}(\mathbb{S}^{n-1})$ is the area of the unit sphere in $\mathbb{R}^n$.

REMARK 9.7. *Note that, except for $n = 2$, the function E is homogeneous of degree $2 - n$. Thus ΔE is homogeneous of degree $-n$, which matches the degree of homogeneity of the delta function. This, and the fact that E is invariant under rotations (i.e. orthogonal changes of variables), explains why we would expect a fundamental solution to have the form (9.8).*

PROOF. We just consider the case $n = 2$, with the case of general n proved similarly. By Remark 9.3, and since Δ is its own transpose, it suffices to show that for each $\varphi \in C_c^\infty(\mathbb{R}^2)$ we have

$$\int_{\mathbb{R}^2} E(x)\Delta\varphi(x)\,dx = \varphi(0). \tag{9.9}$$

This is done similarly to Exercise 1.1. For $x \in \mathbb{R}^2 \setminus \{0\}$ we compute

$$\partial_{x_j} E(x) = \frac{x_j}{2\pi|x|^2}, \quad \partial_{x_j}^2 E(x) = \frac{|x|^2 - 2x_j^2}{2\pi|x|^4}, \quad \Delta E(x) = 0.$$

Since $E \in L^1_{\text{loc}}(\mathbb{R}^2)$ we have

$$\int_{\mathbb{R}^2} E(x)\Delta\varphi(x) = \lim_{\varepsilon\to 0+} I_\varepsilon \quad \text{where } I_\varepsilon := \int_{\mathbb{R}^2\setminus B(0,\varepsilon)} E(x)\Delta\varphi(x)\,dx.$$

Integrating by parts twice (using (1.39)) and using that $\Delta E = 0$ on $\mathbb{R}^2 \setminus \{0\}$ and $\varphi \in C^\infty_c(\mathbb{R}^2)$ we write I_ε as a surface integral:

$$I_\varepsilon = \int_{\partial B(0,\varepsilon)} E(x)(\vec{n}(x)\cdot\nabla\varphi(x))\,dS(x) - \int_{\partial B(0,\varepsilon)} \varphi(x)(\vec{n}(x)\cdot\nabla E(x))\,dS(x). \tag{9.10}$$

Here dS is the length measure on the circle $\partial B(0,\varepsilon)$ and $\vec{n}(x)$ is the unit normal on the circle which points inside the circle, i.e. outside of the region $\mathbb{R}^2\setminus B(0,\varepsilon)$. We have for $x \in \partial B(0,\varepsilon)$

$$\vec{n}(x) = -\frac{x}{|x|},\quad \nabla E(x) = \frac{x}{2\pi|x|^2},\quad \vec{n}(x)\cdot\nabla E(x) = -\frac{1}{2\pi\varepsilon}.$$

Now, the first term on the right-hand side of (9.10) is $\mathcal{O}(\varepsilon\log(1/\varepsilon))$ which goes to 0 as $\varepsilon \to 0+$. The second term is

$$\frac{1}{2\pi\varepsilon}\int_{\partial B(0,\varepsilon)}\varphi(x)\,dS(x) \to \varphi(0)\quad \text{as } \varepsilon \to 0+\,.$$

Thus we obtain (9.9) which finishes the proof. □

One can similarly obtain a fundamental solution for the Cauchy–Riemann operator $\partial_{\bar z} := \frac{1}{2}(\partial_x + i\partial_y)$ on $\mathbb{R}^2_{x,y}$. We leave the proof as an exercise below.

PROPOSITION 9.8. *A fundamental solution of $\partial_{\bar z}$ is given by the locally integrable function*

$$E(x,y) = \frac{1}{\pi(x+iy)}.$$

We next consider the heat operator $\partial_t - \Delta_x$ on $\mathbb{R}^{n+1} = \mathbb{R}_t \times \mathbb{R}^n_x$. The proof is again left as an exercise below.

PROPOSITION 9.9. *A fundamental solution of $\partial_t - \Delta_x$ is given by the locally integrable function*

$$E(t,x) = \begin{cases} (4\pi t)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4t}}, & t > 0,\\ 0, & t \le 0.\end{cases} \tag{9.11}$$

REMARK 9.10. *One can show that* $\operatorname{sing\,supp} E = \{0\}$, *that is E is smooth on $\mathbb{R}^{n+1}\setminus\{0\}$, similarly to the bump function (1.26).*

We now discuss the wave operator $\partial_t^2 - \Delta_x$ on $\mathbb{R}_t \times \mathbb{R}^n_x$. The situation is more complicated here since in general fundamental solutions are not locally integrable function. For now we just consider the case $n = 1$ (in §10.2 below we handle the case $n = 3$):

PROPOSITION 9.11. *A fundamental solution for the operator $\partial_t^2 - \partial_x^2$ on $\mathbb{R}^2_{t,x}$ is given by the locally integrable function*

$$E(t,x) = \begin{cases} \frac{1}{2}, & t > |x|,\\ 0, & t < |x|\end{cases} \tag{9.12}$$

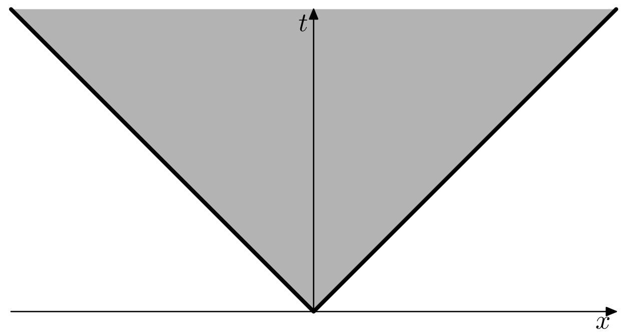

FIGURE 9.1. The support (shaded) and the singular support (the bold lines) of the fundamental solution (9.12).

PROOF. We change variables to

$$y_1 = t + x, \quad y_2 = t - x$$

which transforms the wave operator to a multiple of $\partial_{y_1}\partial_{y_2}$ and the function E to a multiple of $H(y_1)H(y_2)$, where H is the Heaviside function.

Since we have not introduced pullbacks of distributions yet, we argue on the side of test functions. Take arbitrary $\varphi \in C_c^\infty(\mathbb{R}^2)$. Denote $\Box = \partial_t^2 - \partial_x^2$ and define $\psi \in C_c^\infty(\mathbb{R}^2)$ by putting $\varphi(t,x) = \psi(t+x, t-x)$. We compute

$$\begin{aligned}\int_{\mathbb{R}^2} E(t,x)\Box\varphi(t,x)\,dxdt &= 4\int_{\mathbb{R}^2} E(t,x)\partial_{y_1}\partial_{y_2}\psi(t+x,t-x)\,dxdt\\ &= \int_{\mathbb{R}^2} H(y_1)H(y_2)\partial_{y_1}\partial_{y_2}\psi(y_1,y_2)\,dy_1dy_2\\ &= \big(\partial_{y_1}\partial_{y_2}(H(y_1)\otimes H(y_2)), \psi(y_1,y_2)\big)\\ &= \psi(0,0) = \varphi(0,0).\end{aligned}$$

which shows that $\Box E = \delta_0$ by Remark 9.3, since $\Box$ is its own transpose. In the last line above we used that

$$\partial_{y_1}\partial_{y_2}(H(y_1)\otimes H(y_2)) = \delta_0(y_1)\otimes\delta_0(y_2) = \delta_0(y_1,y_2)$$

which follows from the properties of tensor product (see Propositions 7.2 and 7.3) and the fact that $H' = \delta_0$ (see (3.4)). $\Box$

REMARK 9.12. *Note that* $\operatorname{supp} E = \{t \geq |x|\}$ *and* $\operatorname{sing\,supp} E = \{t = |x|\}$, *see Figure 9.1.*

We finish this section with the following general

THEOREM 9.13 (Malgrange–Ehrenpreis Theorem). *Let $P \neq 0$ be a differential operator with constant coefficients. Then P has a fundamental solution.*

We do not give the proof here, sending a curious reader to [**Hör03**, Theorem 7.3.10] or [**FJ98**, Theorem 10.4.1].

9.1.3. A negative example$^{\mathbf{X}}$. We now give a simple example illustrating why in part 1 of Theorem 9.4 it is important that the supports of E and u sum properly, going through the construction of distributional convolution in §8 in this particular case. This section is optional for reading.

Our example is as follows: on $\mathbb{R}$ we have

$$Pu = 0 \quad \text{where } P = \partial_x, \quad u \equiv 1,$$

and a fundamental solution of P is given by the Heaviside function:

$$E(x) = H(x).$$

Clearly we do not have $u = E * Pu$, even though the convolution $E * Pu$ makes perfect sense (and equals 0). There is no contradiction with Theorem 9.4 since the supports of E, u do not sum properly. But what if we were to repeat the proof of that theorem while ignoring the support issue?

Looking at the proof of Theorem 9.4, we see that what fails is the identity

$$(\partial_x H) * 1 = H * (\partial_x 1),$$

and properties of convolution do not apply here since $\operatorname{supp} H, \operatorname{supp} 1$ do not sum properly, so the convolution $H * 1$ cannot be defined. Let us look a bit more into the proof of this property of convolution to see what goes wrong. Take arbitrary $\varphi \in C_{\mathrm{c}}^{\infty}(\mathbb{R})$, then if we ignore the support issue then

$$\begin{aligned}((\partial_x H) * 1, \varphi) &= (\partial_x H(x) \otimes 1(y), \varphi(x+y)) = -(H(x) \otimes 1(y), \varphi'(x+y)),\\ (H * (\partial_x 1), \varphi) &= (H(x) \otimes \partial_y 1(y), \varphi(x+y)) = -(H(x) \otimes 1(y), \varphi'(x+y)),\end{aligned}$$

which indicates that the two expressions are equal. However, in the second equality in each line above we used the definition of distributional derivative which does not apply since $H(x) \otimes 1(y)$ cannot be paired with $\varphi(x+y)$ as their supports do not have compact intersection. More concretely, we could try to write

$$(H(x) \otimes 1(y), \varphi'(x+y)) = \int_{\{x>0\}} \varphi'(x+y)\, dxdy$$

and compute it by Fubini's Theorem in two ways (which corresponds to the iterated tensor product formulas (7.10) and (7.11)) as

$$\int_{\{x>0\}} \varphi'(x+y)\, dxdy = \int_0^\infty \Big(\int_{\mathbb{R}} \varphi'(x+y)\, dy \Big)\, dx = \int_{\mathbb{R}} \Big(\int_0^\infty \varphi'(x+y)\, dx \Big)\, dy$$

But the function $\varphi'(x+y)$ is not integrable on $\{x > 0\}$, so Fubini's Theorem does not apply. While the two iterated integrals above both converge, their values are different: the first one is equal to 0 and the second one is equal to $-\int_{\mathbb{R}} \varphi(y)\, dy$.

9.2. Elliptic regularity I

We now give the first version of elliptic regularity, which is one of the main results in this course. Further versions will be proved in §§12.2,14 below. Recall the notion of singular support defined in §8.3.

THEOREM 9.14 (Elliptic Regularity I). *Assume that P is a differential operator with constant coefficients on $\mathbb{R}^n$ and that there exists a fundamental solution E of P such that*

$$\operatorname{sing\,supp} E = \{0\}. \tag{9.13}$$

Let $U \Subset \mathbb{R}^n$ and $u \in \mathcal{D}'(U)$. Then

$$\operatorname{sing\,supp} u = \operatorname{sing\,supp}(Pu). \tag{9.14}$$

In particular, we have

$$Pu \in C^\infty(U) \quad \Longrightarrow \quad u \in C^\infty(U). \tag{9.15}$$

REMARK 9.15. *Theorem 9.14 is not a completely satisfactory result since to apply it we need to find a fundamental solution of P with singular support at the origin. Still, looking at the examples in §9.1.2 we get elliptic regularity for the Laplace operator, the Cauchy–Riemann operator, and the heat operator. On the other hand, for the wave operator elliptic regularity fails; in the case of 1 spatial dimension this follows from Proposition 9.11 since the fundamental solution E does not satisfy* $\operatorname{sing\,supp} E \subset \operatorname{sing\,supp} \Box E$. *Also, strictly speaking Theorem 9.14 should be called* hypoelliptic regularity *since it applies to some operators which are not elliptic, such as the heat operator – see §12.2.*

PROOF. We have $\operatorname{sing\,supp}(Pu) \subset \operatorname{sing\,supp} u$ by (8.16), so we need to show that $\operatorname{sing\,supp} u \subset \operatorname{sing\,supp}(Pu)$.

1. Fix arbitrary $x_0 \in U$ such that $x_0 \notin \operatorname{sing\,supp}(Pu)$; we need to show that $x_0 \notin \operatorname{sing\,supp} u$. Fix a cutoff function

$$\chi \in C^\infty_{\mathrm{c}}(U), \quad x_0 \notin \operatorname{supp}(1-\chi).$$

Consider the product

$$\chi u \ \in \ \mathcal{E}'(U)$$

and extend it by 0 (using Proposition 4.7) to an element of $\mathcal{E}'(\mathbb{R}^n)$, which we still denote by χu.

By part 1 of Theorem 9.4 we have

$$\chi u = E * (P\chi u).$$

Since $\operatorname{sing\,supp} E = \{0\}$, by Proposition 8.14 we have

$$\operatorname{sing\,supp}(\chi u) \subset \operatorname{sing\,supp}(P\chi u).$$

It thus suffices to show that

$$x_0 \notin \operatorname{sing\,supp}(P\chi u). \tag{9.16}$$

Indeed, then $x_0 \notin \operatorname{sing\,supp}(\chi u)$ and thus, as $\chi = 1$ near x_0, we have $x_0 \notin \operatorname{sing\,supp} u$ as needed.

2. We compute

$$P\chi u = [P, \chi]u + \chi P u.$$

Here $[P, \chi] = P\chi - \chi P$ is the commutator of P with the multiplication operator by χ. It is a differential operator with variable coefficients, and (as $\chi = 1$ near x_0 and $[P, 1] = 0$) these coefficients are supported away from x_0. Thus $x_0 \notin \operatorname{supp}([P, \chi]u)$ and thus $x_0 \notin \operatorname{sing\,supp}([P, \chi]u)$. Since $x_0 \notin \operatorname{sing\,supp}(Pu)$, we also have $x_0 \notin \operatorname{sing\,supp}(\chi P u)$. Adding these together we get $x_0 \notin \operatorname{sing\,supp}(P\chi u)$, giving (9.16) and finishing the proof. □

9.3. Notes and exercises

Our presentation largely follows [**Hör03**, §4.4]; see also [**FJ98**, §5.4] for an alternative presentation of the material of §9.1. For a detailed introduction to the general theory of differential operators with constant coefficients see [**Hör05**].

EXERCISE 9.1. (1 pt) *Prove Proposition 9.8.*

EXERCISE 9.2. (1.5 pts) *Prove Proposition 9.9. (Hint: first check that* $(\partial_t - \Delta_x)E = 0$ *for* $t > 0$*. Then compute the integral in Remark 9.3 as an iterated integral* $dxdt$*, integrate by parts in the integral* dx*, and use the Fundamental Theorem of Calculus in* t *to write the integral in Remark 9.3 as a limit as* $t \to 0+$*. Finally compute this limit by a change of variables* $x = 2\sqrt{t}y$ *and the Dominated Convergence Theorem, using also the value of the Gaussian integral.)*

EXERCISE 9.3. (0.5 pt) *Using the fact that the Heaviside function is a fundamental solution for* ∂_x*, show that for* $u \in \mathcal{D}'(\mathbb{R})$*, if* $\operatorname{supp} u \subset [a, \infty)$ *and* $\operatorname{supp}(\partial_x u) \subset [b, \infty)$ *for some* $a \leq b$*, then* $\operatorname{supp} u \subset [b, \infty)$*.*

EXERCISE 9.4. (2.5 = 1 + 0.5 + 1 + 0.5 + 0.5 pts)

This exercise studies solutions to the initial value problem for the wave operator on $\mathbb{R}^2$*,* $\Box := \partial_t^2 - \partial_x^2$*. Assume that*

$$\Box u = f, \quad u(0, x) = g_0(x), \quad \partial_t u(0, x) = g_1(x).$$

Here $u \in C^2(\mathbb{R}^2)$ *is the solution,* $f \in C^0(\mathbb{R}^2)$ *is the forcing term, and* $g_0 \in C^2(\mathbb{R}), g_1 \in C^1(\mathbb{R})$ *are the initial data.*

(a) *Define* $v(t,x) = H(t)u(t,x) \in L^1_{\mathrm{loc}}(\mathbb{R}^2)$ *where* H *is the Heaviside function. Show that, with derivatives in the sense of distributions,*

$$\Box v = \delta_0'(t) \otimes g_0(x) + \delta_0(t) \otimes g_1(x) + H(t)f.$$

(b) *Using that* $\operatorname{supp} v \subset \{t \geq 0\}$ *show that* $v = E * \Box v$ *where* E *is defined in (9.12).*

(c) *Assume that* $w \in \mathcal{D}'(\mathbb{R}^2)$ *and* $\operatorname{supp} w \subset \{t \geq 0\}$. *Show that for each* $\varphi \in C^\infty_{\mathrm{c}}(\mathbb{R}^2)$ *we have*

$$(E * w, \varphi) = (w, \psi)$$

for some $\psi \in C^\infty_{\mathrm{c}}(\mathbb{R}^2)$ *such that*

$$\psi(t,x) = \frac{1}{2}\int_{|y|<s} \varphi(t+s, x+y)\, ds dy, \quad t \geq 0.$$

(d) *Assume that* $f = 0$ *and* $\operatorname{supp} g_0, \operatorname{supp} g_1 \subset [-R, R]$. *Show that*

$$\operatorname{supp} u \cap \{t > 0\} \ \subset\ \{|x| \leq t + R\}.$$

(This is called 'finite speed of propagation'.)

(e) *Assume that* $g_0 = g_1 = 0$ *and* $\operatorname{supp} f \subset \{t > 0\}$. *Show that singularities propagate at unit speed: namely, if* $(t,x) \in \operatorname{sing\,supp} u$ *and* $t > 0$, *then we have* $(t,x) = (s,y) + (\tau, -\tau)$ *or* $(t,x) = (s,y) + (\tau,\tau)$ *for some* $\tau \geq 0$ *and* $(s,y) \in \operatorname{sing\,supp} f$.

EXERCISE 9.5. (1.5 pts) *Using the previous exercise, show d'Alembert's formula: for* $t > 0$

$$\begin{aligned} u(t,x) = {} & \frac{1}{2}(g_0(x+t) + g_0(x-t)) + \frac{1}{2}\int_{x-t}^{x+t} g_1(s)\, ds \\ & + \frac{1}{2}\int_0^t \int_{x-(t-\tau)}^{x+(t-\tau)} f(\tau, s)\, ds d\tau. \end{aligned} \tag{9.17}$$

(This would need a fair amount of computation.)

CHAPTER 10

Pullbacks by smooth maps

In this chapter we define the composition of a distribution with a C^∞ map, under appropriate conditions on the map. As an application, we find a fundamental solution of the wave operator in 3+1 dimensions.

10.1. Defining pullback

10.1.1. Pullback of functions. We first review the classical concept of pullback of smooth functions. Assume that $U \Subset \mathbb{R}^n$, $V \Subset \mathbb{R}^m$, and $\Phi : U \to V$ is a C^∞ map. For $f \in C^\infty(V)$, define the *pullback* of f by Φ as

$$\Phi^* f := f \circ \Phi \ \in \ C^\infty(U). \tag{10.1}$$

This gives a linear sequentially continuous operator

$$\Phi^* : C^\infty(V) \to C^\infty(U). \tag{10.2}$$

The pullback operator acts on locally integrable functions:

$$\Phi^* : L^1_{\mathrm{loc}}(V) \to L^1_{\mathrm{loc}}(U) \tag{10.3}$$

provided thats Φ satisfies the following condition: for each $K \Subset U$ there exists a constant C_K so that

$$\mathrm{vol}(K \cap \Phi^{-1}(\Omega)) \leq C_K \, \mathrm{vol}(\Omega) \quad \text{for all measurable } \Omega \subset V. \tag{10.4}$$

A basic example of when this condition fails is the following map:

$$\Phi : \mathbb{R} \to \mathbb{R}, \quad \Phi(x) = 0 \quad \text{for all } x. \tag{10.5}$$

In this case pullback does not make sense on locally integrable functions already because of identification of functions which are equal almost everywhere: there exist functions f on $\mathbb{R}$ which satisfy $f = 0$ almost everywhere but $\Phi^* f$ is not equal to 0 almost everywhere (e.g. take f to be the indicator function of the set $\{0\}$).

Note that pullback is contravariant, namely if we have two C^∞ maps

$$U \xrightarrow{\Phi_2} V \xrightarrow{\Phi_1} W,$$

then the pullback by $\Phi_1 \circ \Phi_2 : U \to W$ satisfies

$$(\Phi_1 \circ \Phi_2)^* = \Phi_2^* \Phi_1^*. \tag{10.6}$$

10.1.2. Pullback on distributions. The counterexample (10.5) shows that we do not expect to be able to define the pullback of an arbitrary distribution by an arbitrary smooth map Φ. Instead we restrict to Φ satisfying the following condition:

DEFINITION 10.1. *Let $U \Subset \mathbb{R}^n$, $V \Subset \mathbb{R}^m$. A C^∞ map $\Phi : U \to V$ is called a* submersion *if*

$$d\Phi(x) : \mathbb{R}^n \to \mathbb{R}^m \quad \text{is surjective for all } x \in U.$$

Note that we necessarily have $n \geq m$. Moreover, any submersion satisfies the condition (10.4) (for example, one can see this by following the proof of Theorem 10.2 below).

The main result of this section is

THEOREM 10.2. *Assume that $U \Subset \mathbb{R}^n$, $V \Subset \mathbb{R}^m$, and $\Phi : U \to V$ is a C^∞ submersion. Then there exists a sequentially continuous linear operator*

$$\Phi^* : \mathcal{D}'(V) \to \mathcal{D}'(U) \tag{10.7}$$

such that for any $f \in L^1_{\text{loc}}(V)$, $\Phi^* f \in L^1_{\text{loc}}(U)$ is the classical pullback defined in (10.1).

REMARK 10.3. *Since $C^\infty(V)$ is dense in $\mathcal{D}'(V)$ (by Theorem 6.10), such an extension of Φ^* to distributions is unique. Since $C^\infty(V)$ is also dense in $L^1_{\text{loc}}(V)$ (by Theorem 1.14), it suffices to construct Φ^* such that $\Phi^* f = f \circ \Phi$ for any $f \in C^\infty(V)$ (rather than for any $f \in L^1_{\text{loc}}(V)$).*

REMARK 10.4.**X** *The requirement that Φ be a submersion is almost necessary to define Φ^* with the properties in Theorem 10.2 – see* [**Hör03**, Theorem 6.1.1] *and the paragraph following it.*

We prove Theorem 10.2 in steps, treating first two special cases and then writing a general submersion as a composition of those cases.

10.1.3. Case 1: diffeomorphism. We start with the case when $n = m$ and $\Phi : U \to V$ is a C^∞ *diffeomorphism*, that is Φ is bijective and the inverse $\Phi^{-1} : V \to U$ is a C^∞ map. (By the Inverse Mapping Theorem, this is equivalent to Φ being bijective and the differential $d\Phi(x)$ being an invertible linear map for each $x \in U$.) We will use the following standard result from multivariable calculus/Lebesgue integral theory:

THEOREM 10.5 (Change of variables formula).**R** *Assume that $\Phi : U \to V$ is a C^1 diffeomorphism. Define the* Jacobian

$$\mathcal{J}_\Phi \in C^0(U), \quad \mathcal{J}_\Phi(x) = |\det d\Phi(x)|.$$

Let $f : V \to \mathbb{C}$ be a measurable function. Then $f \in L^1(V)$ if and only if the function $(\Phi^* f)\mathcal{J}_\Phi$ lies in $L^1(U)$, and in this case

$$\int_V f(y)\, dy = \int_U f(\Phi(x))\mathcal{J}_\Phi(x)\, dx. \tag{10.8}$$

For the proof, see for example [**Rud87**, Theorem 8.26] or [**Str11**, Theorem 5.2.2].

We now define the pullback operator $\Phi^* : \mathcal{D}'(V) \to \mathcal{D}'(U)$, we use the general extension by duality procedure of Theorem 7.15. All that we need is to show that the transpose operator $(\Phi^*)^t$ is sequentially continuous on test functions,

$$(\Phi^*)^t : C_c^\infty(U) \to C_c^\infty(V). \tag{10.9}$$

We take arbitrary $\varphi \in C^\infty(V)$, $\psi \in C_c^\infty(U)$ and compute using Theorem 10.5

$$((\Phi^*)^t\psi, \varphi) = (\Phi^*\varphi, \psi) = \int_U \varphi(\Phi(x))\psi(x)\,dx = \int_V \varphi(y)\psi(\Phi^{-1}(y))\mathcal{J}_{\Phi^{-1}}(y)\,dy.$$

It follows that

$$(\Phi^*)^t\psi(y) = \psi(\Phi^{-1}(y))\mathcal{J}_{\Phi^{-1}}(y), \quad y \in V, \quad \psi \in C_c^\infty(U).$$

This gives the mapping property (10.9) and shows that the operator Φ^* extends to distributions by the formula

$$(\Phi^* v, \psi) = (v, (\Phi^*)^t\psi) \quad \text{for all } v \in \mathcal{D}'(V),\ \psi \in C_c^\infty(U).$$

We discuss two important examples. One is the pullback of a delta function:

Proposition 10.6. *Assume that $\Phi : U \to V$ is a C^∞ diffeomorphism and $y_0 \in V$. Then*

$$\Phi^*\delta_{y_0} = \mathcal{J}_{\Phi^{-1}}(y_0)\delta_{\Phi^{-1}(y_0)}. \tag{10.10}$$

Proof.**S** We compute for any $\psi \in C_c^\infty(U)$

$$(\Phi^*\delta_{y_0}, \psi) = (\delta_{y_0}, (\Phi^*)^t\psi) = (\Phi^*)^t\psi(y_0) = \mathcal{J}_{\Phi^{-1}}(y_0)\psi(\Phi^{-1}(y_0))$$

which gives the needed identity. □

Another one is the relation to homogeneous distributions defined in §5.1.2. Define the diffeomorphism $\lambda_t : \mathbb{R}^n \to \mathbb{R}^n$, $\lambda_t(x) = tx$, $t > 0$, so that $\Lambda_t = \lambda_t^*$.

Proposition 10.7. *Let $u \in \mathcal{D}'(\mathbb{R}^n)$. Then u is homogeneous of degree $a \in \mathbb{C}$ if and only if $\lambda_t^* u = t^a u$ for all $t > 0$.*

Proof.**S** We compute for any $\psi \in C_c^\infty(\mathbb{R}^n)$

$$(\lambda_t^* u, \psi) = (u, t^{-n}\Lambda_t^{-1}\psi).$$

Thus $(\lambda_t^* u, \psi) = t^a(u, \psi)$ if and only if $(u, \Lambda_t^{-1}\psi) = t^{n+a}(u, \psi)$. □

10.1.4. Case 2: projection. We now consider the case of a *projection map*

$$\Phi : U \to V, \quad \Phi(x', x'') = x'$$

where $n \geq m$, we write elements of $\mathbb{R}^n$ as $x = (x', x'')$ with $x' \in \mathbb{R}^m$, $x'' \in \mathbb{R}^{n-m}$, and $U \Subset \mathbb{R}^n$, $V \Subset \mathbb{R}^m$ satisfy $U \subset V \times \mathbb{R}^{n-m}$.

If $f \in L^1_{\text{loc}}(V)$, then we can write the pullback $\Phi^* f \in L^1_{\text{loc}}(U)$ as the restriction of a tensor product: $(\Phi^* f)(x', x'') = f(x')$, thus $\Phi^* f = (f \otimes 1)|_U$ where 1 is treated as a constant function on $\mathbb{R}^{n-m}$, so that $f \otimes 1 \in L^1_{\text{loc}}(V \times \mathbb{R}^{n-m})$.

Same definition works for distributions: for $v \in \mathcal{D}'(V)$ put

$$\Phi^* v := (v \otimes 1)|_U \ \in \ \mathcal{D}'(U) \tag{10.11}$$

and this defines an operator satisfying the conditions in Theorem 10.2.

10.1.5. The general case. We now give the proof of Theorem 10.2 in the case when $U \Subset \mathbb{R}^n$, $V \Subset \mathbb{R}^m$, and $\Phi : U \to V$ is an arbitrary C^∞ submersion. The following lemma shows that locally Φ is the composition of a diffeomorphism and a projection:

Lemma 10.8. *Assume that $\Phi : U \to V$ is a C^∞ submersion. Fix $x_0 \in U$. Then there exist open sets $U_{x_0} \Subset U$, $W_{x_0} \Subset V \times \mathbb{R}^{n-m}$, $x_0 \in U_{x_0}$, and a C^∞ diffeomorphism $\varkappa_{x_0} : U_{x_0} \to W_{x_0}$ such that*

$$\Phi(x) = y' \quad \textit{for all } x \in U_{x_0} \textit{ where } \quad (y', y'') := \varkappa_{x_0}(x). \tag{10.12}$$

Proof. Since Φ is a submersion, the linear map $d\Phi(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ is surjective. Thus there exists a linear map $\Psi : \mathbb{R}^n \to \mathbb{R}^{n-m}$ such that the linear map

$$v \in \mathbb{R}^n \ \mapsto \ (d\Phi(x_0)v, \Psi(v))$$

is invertible. For $x \in U$, define

$$\varkappa_{x_0}(x) = (\Phi(x), \Psi(x)) \ \in V \times \mathbb{R}^{n-m},$$

then $d\varkappa_{x_0}(x_0)$ is invertible. By the Inverse Mapping Theorem there exist open neighborhoods U_{x_0}, W_{x_0} of $x_0, \varkappa_{x_0}(x_0)$ such that $\varkappa_{x_0} : U_{x_0} \to W_{x_0}$ is a diffeomorphism. From the definition of $\varkappa_{x_0}$ we see that it satisfies (10.12). □

Coming back to the proof of Theorem 10.2, take arbitrary $x_0 \in U$ and let $U_{x_0}, W_{x_0}, \varkappa_{x_0}$ be given by Lemma 10.8. Then we can write

$$\Phi|_{U_{x_0}} = \pi_{x_0} \circ \varkappa_{x_0}$$

where $\pi_{x_0} : W_{x_0} \to V$ is defined by $\pi_{x_0}(y', y'') = y'$. Define the pullback operator

$$(\Phi|_{U_{x_0}})^* := \varkappa_{x_0}^* \pi_{x_0}^* : \mathcal{D}'(V) \to \mathcal{D}'(U_{x_0}) \tag{10.13}$$

where $\varkappa_{x_0}^* : \mathcal{D}'(W_{x_0}) \to \mathcal{D}'(U_{x_0})$ is defined in §10.1.3 and $\pi_{x_0}^* : \mathcal{D}'(V) \to \mathcal{D}'(W_{x_0})$ is defined in §10.1.4. Then $(\Phi|_{U_{x_0}})^*$ satisfies the conditions of Theorem 10.2 for the map $\Phi|_{U_{x_0}}$.

We now glue the different operators $(\Phi|_{U_{x_0}})^*$ together to get the global operator Φ^*. If $x_0, x_1 \in U$ then the pullback operators $(\Phi|_{U_{x_0}})^*$ and $(\Phi|_{U_{x_1}})^*$ agree on $U_{x_0} \cap U_{x_1}$, that is for all $v \in \mathcal{D}'(V)$ we have

$$(\Phi|_{U_{x_0}})^* v|_{U_{x_0} \cap U_{x_1}} = (\Phi|_{U_{x_1}})^* v|_{U_{x_0} \cap U_{x_1}}. \tag{10.14}$$

Indeed, this is immediate for $v \in C^\infty(V)$ (as both sides are equal to the classical pullback $(v \circ \Phi)|_{U_{x_0} \cap U_{x_1}}$) and follows for general v since $C^\infty(V)$ is dense in $\mathcal{D}'(V)$ by Theorem 6.10.

Applying the sheaf property of distributions (Theorem 2.13) for the covering $U = \bigcup_{x_0 \in U} U_{x_0}$, we see that for each $v \in \mathcal{D}'(V)$ there exists unique $\Phi^* v \in \mathcal{D}'(U)$ such that

$$(\Phi^* v)|_{U_{x_0}} = (\Phi|_{U_{x_0}})^* v \quad \text{for all } x_0 \in U. \tag{10.15}$$

It it straightforward to check that this defines an operator $\Phi^* : \mathcal{D}'(V) \to \mathcal{D}'(U)$ satisfying the conditions of Theorem 10.2, finishing its proof.

Remark 10.9. *Recalling the constructions in §§10.1.3–10.1.4, we get the following concrete expression for* $(\Phi|_{U_{x_0}})^*$*: for all* $v \in \mathcal{D}'(V)$ *and* $\psi \in C_c^\infty(U_{x_0})$

$$\begin{aligned} ((\Phi|_{U_{x_0}})^* v, \psi) &= (v \otimes 1, (\varkappa_{x_0}^*)^T \psi) = (v, \widetilde{\psi}) \\ \text{where} \quad \widetilde{\psi}(y') &= \int_{\mathbb{R}^{n-m}} \mathcal{J}_{\varkappa_0^{-1}}(y', y'') \psi(\varkappa_{x_0}^{-1}(y', y''))\, dy''. \end{aligned}$$

10.1.6. Properties of pullback. We now discuss properties of the pullback operator on distributions. We start with

Proposition 10.10 (Chain Rule). *Assume that* $U \Subset \mathbb{R}^n$, $V \Subset \mathbb{R}^m$, $\Phi : U \to V$ *is a* C^∞ *submersion, and* $v \in \mathcal{D}'(V)$. *Denote* $\Phi(x) = (\Phi_1(x), \dots, \Phi_m(x))$ *where* $\Phi_1, \dots, \Phi_m : U \to \mathbb{R}$. *Then we have for all* $j = 1, \dots, n$

$$\partial_{x_j}(\Phi^* v) = \sum_{k=1}^m (\partial_{x_j} \Phi_k)\, \Phi^*(\partial_{y_k} v). \tag{10.16}$$

Remark 10.11. *Writing* $v(\Phi(x))$ *in place of* $\Phi^* v(x)$*, we see that* (10.16) *takes the more familiar form*

$$\partial_{x_j}\big(v(\Phi_1(x), \dots, \Phi_m(x))\big) = \sum_{k=1}^m \partial_{x_j} \Phi_k(x) \partial_{y_k} v(\Phi_1(x), \dots, \Phi_m(x)).$$

Proof. This follows from the usual Chain Rule when $v \in C^\infty(V)$, and is true in general since $C^\infty(V)$ is dense in $\mathcal{D}'(V)$ (by Theorem 6.10). ☐

As an application of Proposition 10.10 we compute the pullback of the delta function $\delta_0 \in \mathcal{D}'(\mathbb{R})$ by a submersion Φ, which produces a delta function on the hypersurface $\Phi^{-1}(0)$. See Proposition 13.2 for a review of the concept of embedded submanifold (more specifically, a hypersurface) used below. The surface measure used below coincides with the Riemannian volume density induced by the restriction of the Euclidean metric to Σ (see §13.1.7).

PROPOSITION 10.12. *Let $U \Subset \mathbb{R}^n$ and $\Phi : U \to \mathbb{R}$ be a submersion (that is, the gradient $d\Phi$ is nonzero everywhere); in particular, then $\Sigma := \Phi^{-1}(0) \subset U$ is a hypersurface. Define the distribution $\delta_\Sigma \in \mathcal{D}'(U)$ by integration with respect to the surface measure dS on Σ:*

$$(\delta_\Sigma, \varphi) := \int_\Sigma \varphi(x)\, dS(x) \quad \text{for all } \varphi \in C_c^\infty(U). \tag{10.17}$$

Then

$$\Phi^* \delta_0(x) = \frac{1}{|d\Phi(x)|} \delta_\Sigma(x). \tag{10.18}$$

PROOF. Denote by $H \in L^1_{\mathrm{loc}}(\mathbb{R})$ the Heaviside function. Then $\Phi^* H = \mathbf{1}_\Omega$ is the indicator function of the set

$$\Omega := \{x \in U \mid \Phi(x) \geq 0\}.$$

By Proposition 10.10 and since $H' = \delta_0$ by (3.4) we have for each $j = 1, \dots, n$

$$\partial_{x_j}(\Phi^* H) = (\partial_{x_j}\Phi)\, \Phi^* \delta_0. \tag{10.19}$$

Now, by the Divergence Theorem (see (1.38), where we effectively have $\partial\Omega = \Sigma$ since $\operatorname{supp}\varphi \Subset U$) we compute for each $\varphi \in C_c^\infty(U)$

$$(\partial_{x_j}(\Phi^* H), \varphi) = -\int_\Omega \partial_{x_j}\varphi\, dx = -\int_\Sigma \varphi \nu_j\, dS(x).$$

Here $\vec{\nu}(x) = (\nu_1(x), \dots, \nu_n(x))$ is the normal vector to Σ at $x \in \Sigma$ pointing outside of Ω. We have

$$\vec{\nu}(x) = -\frac{d\Phi(x)}{|d\Phi(x)|},$$

thus we get the identity

$$\partial_{x_j}(\Phi^* H) = \frac{\partial_{x_j}\Phi}{|d\Phi|} \delta_\Sigma. \tag{10.20}$$

Together (10.19) and (10.20) show that for each j

$$(\partial_{x_j}\Phi)\left(\Phi^*\delta_0 - \frac{1}{|d\Phi|\delta_\Sigma}\right) = 0$$

which gives (10.18) since $d\Phi \neq 0$ everywhere on U. $\square$

We finally state several more properties of pullbacks. The proofs are left as exercises below.

Proposition 10.13. *Assume that* $U \Subset \mathbb{R}^n$, $V \Subset \mathbb{R}^m$, $\Phi : U \to V$ *is a submersion,* $v \in \mathcal{D}'(V)$, *and* $a \in C^\infty(V)$. *Then*

$$\Phi^*(av) = (\Phi^* a)(\Phi^* v), \tag{10.21}$$

$$\operatorname{supp}(\Phi^* v) = \Phi^{-1}(\operatorname{supp} v), \tag{10.22}$$

$$\operatorname{sing\,supp}(\Phi^* v) = \Phi^{-1}(\operatorname{sing\,supp} v). \tag{10.23}$$

Moreover, the contravariant property (10.6) *holds on distributions.*

10.2. Application to the wave equation

10.2.1. Construction of a fundamental solution. We now come back to the question of constructing a fundamental solution for the wave operator, started in Proposition 9.11. The theorem below gives the existence of what is known as *advanced*, or *future*, fundamental solution:

Theorem 10.14. *The operator* $\Box = \partial_t^2 - \Delta_x$ *on* $\mathbb{R}_t \times \mathbb{R}^n_x$ *has a fundamental solution* $E \in \mathcal{D}'(\mathbb{R}^{n+1})$ *with the following properties:*

$$\operatorname{supp} E \subset \{(t,x) \colon t \geq |x|\}, \tag{10.24}$$

$$\operatorname{sing\,supp} E = \{(t,x) \colon t = |x|\}. \tag{10.25}$$

We only prove Theorem 10.14 for $n = 3$, that is in the case of 3 spatial dimensions. See [**Hör03**, Theorem 6.2.3] for the case of general n. We break the proof into several steps.

1. We first construct the restriction $\widetilde{E}_+ = E_+|_{\mathbb{R}^4 \setminus \{0\}}$ as a pullback of the delta function on $\mathbb{R}$. (For other values of n, one has to instead pull back the homogeneous distribution χ^a_+ defined in (5.19) with $a := \frac{1-n}{2}$.) Consider the function

$$\Phi : \mathbb{R}^4 \setminus \{0\} \to \mathbb{R}, \quad \Phi(t,x) = t^2 - |x|^2.$$

In special relativity, this is known as the *interval* between (t,x) and the origin. Just like the function $|x|$ featured in the fundamental solution of the Laplace operator (Proposition 9.6) is invariant under rotations, the function Φ is invariant under the Lorentz group $\mathrm{O}(1,3)$ which also leaves the wave equation invariant.

The map Φ is a submersion, so we may define the pullback

$$\widetilde{E} := \Phi^* \delta_0 \ \in \ \mathcal{D}'(\mathbb{R}^4 \setminus \{0\}).$$

We claim that

$$\Box \widetilde{E} = 0. \tag{10.26}$$

This follows from the Chain Rule (Proposition 10.10): we compute for any $v \in \mathcal{D}'(\mathbb{R})$ and $j = 1, 2, 3$

$$\begin{aligned}
\partial_t(\Phi^* v) &= 2t\Phi^* v', \\
\partial_t^2(\Phi^* v) &= 2\Phi^* v' + 4t^2 \Phi^* v'', \\
\partial_{x_j}(\Phi^* v) &= -2x_j \Phi^* v', \\
\partial_{x_j}^2(\Phi^* v) &= -2\Phi^* v' + 4x_j^2 \Phi^* v''.
\end{aligned}$$

This gives

$$\Box(\Phi^* v) = 8\Phi^* v' + 4\Phi\, \Phi^* v'' = \Phi^* w \quad \text{where } w(s) := 8v'(s) + 4sv''(s).$$

The distribution $\delta_0 \in \mathcal{D}'(\mathbb{R})$ is homogeneous of degree -1, so δ_0' is homogeneous of degree -2. Then Euler's equation (see part 2 of Proposition 5.5; of course this can also be checked directly in this case e.g. by differentiating the identity $s\delta_0(s) = 0$ twice) shows that $s\delta_0''(s) = -2\delta_0'(s)$. Thus if $v = \delta_0$ above, then $w = 0$, and we obtain (10.26).

By Proposition 10.12, we see that

$$\widetilde{E} = \frac{1}{|d\Phi|}\delta_{\mathcal{C}}$$

where the *light cone* $\mathcal{C} := \Phi^{-1}(0)$ consists of two parts:

$$\mathcal{C} = \mathcal{C}_+ \sqcup \mathcal{C}_-, \quad \mathcal{C}_\pm := \{(t, x) \in \mathbb{R}^4 \setminus \{0\} \colon \pm t > 0\}.$$

We now define

$$\widetilde{E}_\pm := \frac{1}{|d\Phi|}\delta_{\mathcal{C}_\pm} \;\in\; \mathcal{D}'(\mathbb{R}^4 \setminus \{0\}),$$

so that

$$\widetilde{E} = \widetilde{E}_+ + \widetilde{E}_-, \quad \operatorname{supp} \widetilde{E}_\pm \subset \mathcal{C}_\pm.$$

By (10.26) we have $\Box\widetilde{E}_+ + \Box\widetilde{E}_- = 0$, but the supports of $\Box\widetilde{E}_\pm$ do not intersect each other, thus

$$\Box\widetilde{E}_\pm = 0. \tag{10.27}$$

2. We now extend $\widetilde{E}_+$ through the origin. This can be done using homogeneity: since δ_0 is homogeneous of degree -1 and Φ is homogeneous of degree 2, one can check that $\widetilde{E}_+$ is homogeneous of degree -2 and thus by Theorem 5.6 there exists a unique extension of $\widetilde{E}_+$ to a distribution $E_+ \in \mathcal{D}'(\mathbb{R}^4)$ which is homogeneous of degree -2.

However, we can also argue directly by obtaining a more explicit formula for $\widetilde{E}_+$. Let us parametrize $\mathcal{C}_+$ by $x \in \mathbb{R}^3 \setminus \{0\}$ as the graph $t = |x|$, then the surface element dS is given by

$$dS = \sqrt{1 + \left|\frac{d|x|}{dx}\right|^2}\, dx = \sqrt{2}\, dx.$$

Next, $d\Phi(t,x) = 2(t,-x)$, so $|d\Phi| = 2\sqrt{2}|x|$ on $\mathcal{C}_+$. Thus we have for each $\varphi \in C_c^\infty(\mathbb{R}^4 \setminus \{0\})$

$$(\widetilde{E}_+, \varphi) = \int_{\mathcal{C}_+} \frac{\varphi}{|d\Phi|}\, dS = \int_{\mathbb{R}^3\setminus\{0\}} \frac{\varphi(|x|,x)}{2|x|}\, dx.$$

The integral above still converges if $\operatorname{supp}\varphi$ is allowed to contain the origin, so we define the distribution $E_+ \in \mathcal{D}'(\mathbb{R}^4)$ by

$$(E_+, \varphi) := \int_{\mathbb{R}^3} \frac{\varphi(|x|,x)}{2|x|}\, dx \quad \text{for all } \varphi \in C_c^\infty(\mathbb{R}^4). \tag{10.28}$$

We have $E_+|_{\mathbb{R}^4\setminus\{0\}} = \widetilde{E}_+$. Moreover, recalling Definition 5.3 we see that E_+ is homogeneous of degree -2. We also have

$$\operatorname{supp} E_+ = \operatorname{sing\,supp} E_+ = \{(t,x) \in \mathbb{R}^4 \colon t = |x|\}. \tag{10.29}$$

3. We finally compute $\Box E_+$, which is a distribution in $\mathcal{D}'(\mathbb{R}^4)$. By (10.27) we have $\Box E_+|_{\mathbb{R}^4\setminus\{0\}} = \Box\widetilde{E}_+ = 0$. Thus $\operatorname{supp}\Box E_+ \subset \{0\}$. By Theorem 4.19, $\Box E_+$ is a linear combination of δ_0 and its derivatives. On the other hand, by part 1 of Proposition 5.5 we see that $\Box E_+$ is homogeneous of degree -4, and $\partial^\alpha \delta_0$ is homogeneous of degree $-4-|\alpha|$. Arguing similarly to (5.8) we see that $\Box E_+$ is a multiple of δ_0:

$$\Box E_+ = c\delta_0 \quad \text{for some } c \in \mathbb{C}. \tag{10.30}$$

To compute the constant c, we pair $\Box E_+$ with a function of the form $\psi(t)$ where $\psi \in C_c^\infty(\mathbb{R})$ satisfies $\psi(0) = 1$. This is possible by Proposition 8.7 since the intersection of $(\operatorname{supp}\psi) \times \mathbb{R}^3$ with $\operatorname{supp} E_+$ is compact. We have

$$\begin{aligned} c &= (\Box E_+, \psi(t)) = (E_+, \Box\psi(t)) = (E_+, \psi''(t)) = \int_{\mathbb{R}^3} \frac{\psi''(|x|)}{2|x|}\, dx \\ &= 2\pi \int_0^\infty r\psi''(r)\, dr = 2\pi. \end{aligned}$$

where in the last line we used spherical coordinates and then integrated by parts.

It remains to put

$$E := \frac{1}{2\pi} E_+ \ \in\ \mathcal{D}'(\mathbb{R}^4)$$

to obtain a fundamental solution of $\Box$ satisfying the conditions of Theorem 10.14. Recalling (10.28) we obtain the following explicit expression for E:

$$(E, \varphi) = \int_{\mathbb{R}^3} \frac{\varphi(|x|,x)}{4\pi|x|}\, dx \quad \text{for all } \varphi \in C_c^\infty(\mathbb{R}^4). \tag{10.31}$$

10.2.2. The Cauchy problem. We now use the fundamental solution E of $\Box$ on $\mathbb{R} \times \mathbb{R}^n_x$ from Theorem 10.14 to obtain a few partial results on the *forward Cauchy problem*

$$\begin{aligned} \Box u(t,x) &= f(t,x), \quad t \geq 0, \\ u(0,x) &= g_0(x), \\ \partial_t u(0,x) &= g_1(x). \end{aligned} \tag{10.32}$$

See [**Hör03**, Theorem 6.2.4] for a more comprehensive treatment.

Assume that u is a classical solution to (10.32):

$$u \in C^2([0,\infty)_t \times \mathbb{R}^n_x), \quad g_0 \in C^2(\mathbb{R}^n), \quad g_1 \in C^1(\mathbb{R}^n).$$

Using the Heaviside function H, define

$$v(t,x) := H(t)u(t,x), \quad v \in L^1_{\mathrm{loc}}(\mathbb{R}^4).$$

Arguing in the same way as for Exercise 9.4(a) we compute

$$\Box v = \delta_0'(t) \otimes g_0(x) + \delta_0(t) \otimes g_1(x) + H(t)f. \tag{10.33}$$

By Exercise 8.3, $\operatorname{supp} E$ and $\operatorname{supp} v \subset \{t \geq 0\}$ sum properly, thus part 1 of Theorem 9.4 gives

$$v = E * \Box v. \tag{10.34}$$

This gives uniqueness for the Cauchy problem (10.32): if $f = 0$ and $g_0 = g_1 = 0$ then $\Box v = 0$ and thus $v = 0$, implying that $u = 0$.

Arguing similarly to Exercise 9.4(d,e) we obtain *finite speed of propagation*:

$$\begin{aligned} \operatorname{supp} u \subset \{(t,x) \mid \exists y \in \operatorname{supp} g_0 \cup \operatorname{supp} g_1,\ |x-y| \leq t\} \\ \cup \{(t,x) \mid \exists (s,y) \in \operatorname{supp} f,\ |x-y| \leq t-s\} \end{aligned} \tag{10.35}$$

and a weak version of *propagation of singularities*: if $g_0 = g_1 = 0$ and $\operatorname{supp} f \subset \{t > 0\}$ then

$$\operatorname{sing\,supp} u \subset \{(t,x) \mid \exists (s,y) \in \operatorname{sing\,supp} f,\ |x-y| = t-s\}. \tag{10.36}$$

10.3. Notes and exercises

Our presentation largely follows [**Hör03**, §§6.1–6.2] and [**FJ98**, Chapter 7].

Exercise 10.1. (2.5 = 0.5 + 1 + 1 pts) *Let $\Phi : U \to V$ be a submersion and $v \in \mathcal{D}'(V)$.*

(a) *Show that if $\widetilde{U} \Subset U$, $\widetilde{V} \Subset V$, and $\Phi(\widetilde{U}) \subset \widetilde{V}$, then $(\Phi^* v)|_{\widetilde{U}} = (\Phi|_{\widetilde{U}})^*(v|_{\widetilde{V}})$, where $\Phi|_{\widetilde{U}} : \widetilde{U} \to \widetilde{V}$.*

(b) *Show that if* Φ *is surjective (that is,* $\Phi(U) = V$*), then*

$$\Phi^* v = 0 \quad \Longrightarrow \quad v = 0, \tag{10.37}$$

$$\Phi^* v \in C^\infty(U) \quad \Longrightarrow \quad v \in C^\infty(V). \tag{10.38}$$

(You might need to review the construction of the extension Φ^* *in §§10.1.3–10.1.5.)*

(c) *Prove* (10.22) *and* (10.23). *(You might find the previous two parts of the exercise useful.)*

Exercise 10.2. (0.5 pt) *Prove the properties* (10.6) *and* (10.21) *for pullback on distributions.*

Exercise 10.3. (1 pt) *Let* $\Phi : \mathbb{R} \to \mathbb{R}$ *be given by* $\Phi(x) = x^2$. *Show that the pullback operator* $\Phi^* : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ *does not extend to a sequentially continuous operator* $\mathcal{D}'(\mathbb{R}) \to \mathcal{D}'(\mathbb{R})$. *(Hint: let* $\chi \in C_c^\infty(\mathbb{R})$ *be equal to 1 near 0, put* $\chi_\varepsilon(x) := \varepsilon^{-1}\chi(x/\varepsilon)$*, and look at the limit of* $(\Phi^*\chi_\varepsilon, \chi)$*.)*

Exercise 10.4. (2 = 1 + 1 pts) *Compute the transposes* $(\Phi^*)^t : C_c^\infty(U) \to \mathcal{D}'(V)$ *of pullbacks by the following two maps* $\Phi : U \to V$. *In each case decide whether* $(\Phi^*)^t$ *maps* $C_c^\infty(U)$ *to* $C_c^\infty(V)$ *(which would allow to extend* Φ^* *to distributions):*

(a) $\Phi : \mathbb{R}^2 \to \mathbb{R}$, $\Phi(x_1, x_2) = x_1$;

(b) $\Phi : \mathbb{R} \to \mathbb{R}^2$, $\Phi(x_1) = (x_1, 0)$.

Exercise 10.5. (2 = 1 + 1 pts) *Assume that* $W \subset \mathbb{R}^n$ *is open and* $F : W \to \mathbb{R}^m$ *is a* C^∞ *map. Define the submersion* $\Phi : W \times \mathbb{R}^m \to \mathbb{R}^m$ *by* $\Phi(x, y) = y - F(x)$.

(a) *Show that for each* $v \in \mathcal{D}'(\mathbb{R}^m)$ *the distribution* $\Phi^* v \in \mathcal{D}'(W \times \mathbb{R}^m)$ *is given by*

$$(\Phi^* v, \varphi) = \left(v(y), \int_W \varphi(x, y + F(x))\, dx \right) \quad \text{for all} \quad \varphi \in C_c^\infty(W \times \mathbb{R}^m). \tag{10.39}$$

(b) *Show that the Schwartz kernel of the pullback operator* $F^* : C^\infty(\mathbb{R}^m) \to C^\infty(W)$ *is given by* $\mathcal{K}(x, y) = \delta_0(y - F(x))$ *where* $\delta_0(y - F(x))$ *is defined as* $\Phi^*\delta_0$. *(In the special case when* F *is the identity map we see that the Schwartz kernel of the identity operator is given by* $\delta_0(y - x) = \delta_0(x - y)$*.)*

Exercise 10.6. (1 pt) *Check that the distribution* E *given in* (10.31) *satisfies* $\Box E = \delta_0$ *directly, without appealing to the classification of distributions supported at the origin. To do this, introduce the spherical coordinates* $x = r\theta$ *where* $\theta \in \mathbb{S}^2$. *You may use the formula*

$$\Delta_x = \partial_r^2 + \frac{2}{r}\partial_r + \frac{1}{r^2}\Delta_\theta$$

where $\Delta_\theta : C^\infty(\mathbb{S}^2) \to C^\infty(\mathbb{S}^2)$ *is the Laplace–Beltrami operator for the standard metric on the 2-sphere. You may also use that* $\Delta_\theta f$ *integrates to 0 on* $\mathbb{S}^2$ *for all* $f \in C^\infty(\mathbb{S}^2)$. *After getting rid of* Δ_θ*, you might find it useful to write everything in terms of the function* $\psi(u, v, \theta) = \varphi(u + v, (u - v)\theta)$ *where* $\varphi \in C_c^\infty(\mathbb{R}^4)$ *and* $u, v \in \mathbb{R}$, $\theta \in \mathbb{S}^2$.

EXERCISE 10.7. $(1 = 0.5 + 0.5 \text{ pt})$ *Let* $E \in \mathcal{D}'(\mathbb{R}^4)$ *be defined in* (10.31).

(a) *Assume that* $w \in \mathcal{D}'(\mathbb{R}^4)$ *and* $\operatorname{supp} w \subset \{t \geq 0\}$. *Show that for each* $\varphi \in C_c^\infty(\mathbb{R}^4)$ *we have*

$$(E * w, \varphi) = (w, \psi)$$

for some $\psi \in C_c^\infty(\mathbb{R}^4)$ *such that*

$$\psi(t,x) = \frac{1}{4\pi} \int_{\mathbb{R}^3} \frac{\varphi(t + |y|, x + y)}{|y|} \, dy, \quad t \geq 0.$$

(b) *Using part (a) and* (10.33), *show the following version of Kirchhoff's formula: if* $u \in C^2(\{t \geq 0\})$ *is the solution to*

$$\Box u(t,x) = 0, \quad u(0,x) = 0, \quad \partial_t u(0,x) = g_1(x),$$

then we have for all $t \geq 0$ *and* $x \in \mathbb{R}^3$

$$u(t,x) = \frac{t}{4\pi} \int_{\mathbb{S}^2} g_1(x + t\theta) \, dS(\theta).$$

That is, the value of the solution at time t *and space* x *is equal to* t *times the average of the initial data* g_1 *over the sphere of radius* t *centered at* x.

CHAPTER 11

Fourier transform I

In this chapter we define the Fourier transform on distributions, which is a powerful tool in the study of PDEs in particular because it turns constant coefficient differential operators into multiplication operators.

11.1. Fourier transform on Schwartz functions

11.1.1. Fourier transform on L^1. We start by defining Fourier transform on functions. For two vectors $x, \xi \in \mathbb{R}^n$, denote by $x \cdot \xi$ their usual Euclidean inner product, that is

$$x \cdot \xi := \sum_{j=1}^{n} x_j \xi_j.$$

DEFINITION 11.1. *Let $f \in L^1(\mathbb{R}^n)$. Define the Fourier transform*

$$\widehat{f} = \mathcal{F}(f) \ \in \ L^\infty(\mathbb{R}^n)$$

by the formula

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\,dx. \tag{11.1}$$

Note that $\widehat{f}(0)$ is the integral of f.

It is immediate to see that $\mathcal{F} : L^1(\mathbb{R}^n) \to L^\infty(\mathbb{R}^n)$ is a bounded linear operator, in fact we have from its definition

$$\|\widehat{f}\|_{L^\infty(\mathbb{R}^n)} \le \|f\|_{L^1(\mathbb{R}^n)} \quad \text{for all } f \in L^1(\mathbb{R}^n). \tag{11.2}$$

Moreover, $\widehat{f}$ is a continuous function:

PROPOSITION 11.2. *Assume that $f \in L^1(\mathbb{R}^n)$. Then $\widehat{f} \in C^0(\mathbb{R}^n)$.*

PROOF. We have for any $\xi \in \mathbb{R}^n$

$$\widehat{f}(\eta) = \int_{\mathbb{R}^n} e^{-ix\cdot\eta} f(x)\,dx \ \to \ \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\,dx = \widehat{f}(\xi) \quad \text{as } \eta \to \xi$$

by the Dominated Convergence Theorem, since $|e^{-ix\cdot\eta} f(x)| = |f(x)|$, $f \in L^1(\mathbb{R}^n)$, and $e^{-ix\cdot\eta} \to e^{-ix\cdot\xi}$ as $\eta \to \xi$ for all $x \in \mathbb{R}^n$. □

11.1.2. Schwartz functions. The operator $\mathcal{F} : L^1 \to L^\infty$ is very far from invertible. It is highly desirable to have spaces on which the Fourier transform is an isomorphism. One of such spaces is given by *Schwartz functions*:

DEFINITION 11.3. *We say that $\varphi \in C^\infty(\mathbb{R}^n)$ is a* Schwartz function *if for all multiindices α, β we have*

$$\sup_{x\in\mathbb{R}^n} |x^\alpha \partial_x^\beta \varphi(x)| < \infty. \tag{11.3}$$

Denote by $\mathscr{S}(\mathbb{R}^n)$ the space of all Schwartz functions on $\mathbb{R}^n$. For a sequence $\varphi_j \in \mathscr{S}(\mathbb{R}^n)$, we say it converges to $\varphi \in \mathscr{S}(\mathbb{R}^n)$ in $\mathscr{S}(\mathbb{R}^n)$ if for all α, β we have

$$\sup_{x\in\mathbb{R}^n} |x^\alpha \partial_x^\beta (\varphi_j - \varphi)| \to 0 \quad \text{as } j \to \infty.$$

REMARK 11.4. *We think of $\mathscr{S}(\mathbb{R}^n)$ as a space of test functions which is well suited to study the Fourier transform. We sometimes call elements of this space* rapidly decreasing *functions, since for $\varphi \in \mathscr{S}(\mathbb{R}^n)$ every derivative $\partial_x^\beta \varphi$ is $\mathcal{O}((1+|x|)^{-N})$ for all N.*

REMARK 11.5. *We have the inclusions*

$$C_c^\infty(\mathbb{R}^n) \ \subset \ \mathscr{S}(\mathbb{R}^n) \ \subset \ C^\infty(\mathbb{R}^n). \tag{11.4}$$

Correspondingly, convergence of sequences in $C_c^\infty(\mathbb{R}^n)$ is stronger than in $\mathscr{S}(\mathbb{R}^n)$, which in turn is stronger than in $C^\infty(\mathbb{R}^n)$. The space $C_c^\infty(\mathbb{R}^n)$ is dense in $\mathscr{S}(\mathbb{R}^n)$, see Exercise 11.1 below.

Note however that unlike $C_c^\infty(U)$ and $C^\infty(U)$, which are defined for any $U \Subset \mathbb{R}^n$, the Schwartz space is only defined for functions on the entire $\mathbb{R}^n$.

A family of seminorms on $\mathscr{S}(\mathbb{R}^n)$ is given by

$$\|\varphi\|_{N,M} := \max_{|\alpha|\le N,\, |\beta|\le M} \sup_{\mathbb{R}^n} |x^\alpha \partial_x^\beta \varphi|, \quad N, M \in \mathbb{N}_0. \tag{11.5}$$

We have $\varphi_n \to 0$ in $\mathscr{S}'(\mathbb{R}^n)$ if and only if $\|\varphi_n\|_{N,M} \to 0$ for all N, M. In fact, it is enough to require that $\|\varphi_n\|_{N,N} \to 0$ for all N. The collection of seminorms $\| \bullet \|_{N,N}$ makes $\mathscr{S}(\mathbb{R}^n)$ into a Fréchet space similarly to §4.3.1.

From the definition of the above seminorms we see immediately that the multiplication operators x_j and the differentiation operators ∂_{x_j} are sequentially continuous $\mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$, in fact for all all N, M there exists a constant C such that for all $\varphi \in C_c^\infty(\mathbb{R}^n)$ we have

$$\|x_j \varphi\|_{N,M} \le C\|\varphi\|_{N+1,M}, \qquad \|\partial_{x_j}\varphi\|_{N,M} \le C\|\varphi\|_{N,M+1}. \tag{11.6}$$

The Schwartz space $\mathscr{S}(\mathbb{R}^n)$ is contained in $L^1(\mathbb{R}^n)$. In fact, if $\varphi \in \mathscr{S}(\mathbb{R}^n)$, then $|\varphi(x)| \le C_n \|\varphi\|_{n+1,0}(1+|x|)^{-n-1}$ for some constant C_n depending only on n, so

$$\|\varphi\|_{L^1(\mathbb{R}^n)} \le C_n \|\varphi\|_{n+1,0}. \tag{11.7}$$

11.1.3. Fourier transform acts on Schwartz functions. For any $\varphi \in \mathscr{S}(\mathbb{R}^n)$, we can use (11.1) to define the Fourier transform $\widehat{\varphi} \in L^\infty(\mathbb{R}^n)$, and (11.2) and (11.7) together show that

$$\|\widehat{\varphi}\|_{L^\infty(\mathbb{R}^n)} \leq C_n \|\varphi\|_{n+1,0}. \tag{11.8}$$

A remarkable property of the space $\mathscr{S}(\mathbb{R}^n)$ is that the Fourier transform of a Schwartz function is again a Schwartz function:

THEOREM 11.6. *For each $\varphi \in \mathscr{S}(\mathbb{R}^n)$, $\mathcal{F}(\varphi) = \widehat{\varphi}$ also lies in $\mathscr{S}(\mathbb{R}^n)$. Moreover, the operator $\mathcal{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is sequentially continuous.*

The proof of Theorem 11.6 relies on the fact that Fourier transform intertwines differentiation and multiplication. To state it we introduce the modified differentiation operators

$$D_{x_j} := -i\partial_{x_j}. \tag{11.9}$$

For a multiindex α, we have

$$D_x^\alpha = D_{x_1}^{\alpha_1} \dots D_{x_n}^{\alpha_n} = (-i)^{|\alpha|}\partial_x^\alpha.$$

PROPOSITION 11.7. *Let $\varphi \in \mathscr{S}(\mathbb{R}^n)$. Then $\widehat{\varphi} \in C^1(\mathbb{R}^n)$ and*

$$\widehat{D_{x_j}\varphi}(\xi) = \xi_j \widehat{\varphi}(\xi), \tag{11.10}$$

$$\widehat{x_j\varphi}(\xi) = -D_{\xi_j}\widehat{\varphi}(\xi). \tag{11.11}$$

PROOF. 1. To show (11.10), we integrate by parts:

$$\begin{aligned}\widehat{D_{x_j}\varphi}(\xi) &= \int_{\mathbb{R}^n} e^{-ix\cdot\xi} D_{x_j}\varphi(x)\,dx = -\int_{\mathbb{R}^n} (D_{x_j} e^{-ix\cdot\xi})\varphi(x)\,dx \\ &= \int_{\mathbb{R}^n} \xi_j e^{-ix\cdot\xi}\varphi(x)\,dx = \xi_j\widehat{\varphi}(\xi).\end{aligned}$$

Here to justify integration by parts, we can first integrate on the ball $B(0,R)$ and then let $R \to \infty$; the boundary terms will go to 0 since φ is rapidly decreasing.

2. To show (11.11), we differentiate under the integral sign:

$$-D_{\xi_j}\widehat{\varphi}(\xi) = \int_{\mathbb{R}^n} (-D_{\xi_j} e^{-ix\cdot\xi})\varphi(x)\,dx = \int_{\mathbb{R}^n} x_j e^{-ix\cdot\xi}\varphi(x)\,dx = \widehat{x_j\varphi}(\xi). \tag{11.12}$$

To justify differentiation under the integral sign, denote by e_j the jth coordinate vector on $\mathbb{R}^n$ and write for any $\xi \in \mathbb{R}^n$ and $t \in \mathbb{R} \setminus \{0\}$

$$\frac{\widehat{\varphi}(\xi + te_j) - \widehat{\varphi}(\xi)}{t} = \int_{\mathbb{R}^n} \frac{e^{-ix\cdot(\xi+te_j)} - e^{-ix\cdot\xi}}{t}\varphi(x)\,dx. \tag{11.13}$$

Applying the inequality $|e^{i\alpha} - 1| \leq |\alpha|$ with $\alpha := -tx_j$, we see that

$$\Big|\frac{e^{-ix\cdot(\xi+te_j)} - e^{-ix\cdot\xi}}{t}\varphi(x)\Big| = \Big|\frac{e^{-itx_j} - 1}{t}\varphi(x)\Big| \leq |x_j\varphi(x)|.$$

Since φ is a Schwartz function, we have $x_j\varphi \in L^1(\mathbb{R}^n)$. Thus we can pass to the limit $t \to 0$ under the integral in (11.13), which means that we can differentiate under the integral in (11.12). □

REMARK 11.8. *The above proof shows that* (11.11) *holds for all* $\varphi \in L^1(\mathbb{R}^n)$ *such that* $x_j\varphi \in L^1(\mathbb{R}^n)$. *Together with Proposition 11.2 this implies that for each* $k \in \mathbb{N}_0$

$$(1+|x|)^k f(x) \in L^1(\mathbb{R}^n) \quad \Longrightarrow \quad \widehat{f} \in C^k(\mathbb{R}^n). \tag{11.14}$$

We can now give

PROOF OF THEOREM 11.6. Assume that $\varphi \in \mathscr{S}(\mathbb{R}^n)$. Applying Proposition 11.7 iteratively and using that the operators x_j, ∂_{x_j} map $\mathscr{S}(\mathbb{R}^n)$ to itself, we see that $\widehat{\varphi} \in C^\infty(\mathbb{R}^n)$ and for any multiindices α, β

$$\xi^\alpha D_\xi^\beta \widehat{\varphi} = (-1)^{|\beta|} \widehat{D_x^\alpha x^\beta \varphi}. \tag{11.15}$$

Here $D_x^\alpha x^\beta \varphi \in \mathscr{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$ and thus (11.15) is a bounded function on $\mathbb{R}^n$. Since α, β are chosen arbitrary, we see that $\widehat{\varphi} \in \mathscr{S}(\mathbb{R}^n)$. The continuity of Fourier transform on $\mathscr{S}(\mathbb{R}^n)$ follows from the estimates (where C depends only on N, M, n)

$$\|\widehat{\varphi}\|_{N,M} \leq C\|\varphi\|_{M+n+1,N}$$

which are an immediate corollary of (11.15), (11.2), (11.7), and (11.6). □

11.1.4. Properties of Fourier transform. We now give some properties of the Fourier transform. We first show that Fourier transform is its own transpose in the sense of §7.3. Similarly to (2.3) we use the notation (f, g) to denote the integral of fg, where f, g are functions on $\mathbb{R}^n$ and $fg \in L^1(\mathbb{R}^n)$.

PROPOSITION 11.9. *Let* $f, g \in L^1(\mathbb{R}^n)$. *Then*

$$(\widehat{f}, g) = (f, \widehat{g}). \tag{11.16}$$

PROOF.$^{\mathbf{S}}$ By Fubini's theorem both sides are equal to

$$\int_{\mathbb{R}^{2n}} e^{-ix\cdot\xi} f(x) g(\xi)\, dx d\xi.$$

□

We next give the relation between Fourier transform, convolution, and multiplication. Note that if $f, g \in L^1(\mathbb{R}^n)$, then by Fubini's Theorem the convolution $f * g$ (defined by the integral (1.27) which converges for almost every x) is in $L^1(\mathbb{R}^n)$.

PROPOSITION 11.10. *Let* $f, g \in L^1(\mathbb{R}^n)$. *Then*

$$\widehat{f * g}(\xi) = \widehat{f}(\xi)\widehat{g}(\xi). \tag{11.17}$$

PROOF.[S] By Fubini's theorem and the change of variables $x = y + z$ we have

$$\begin{aligned}\widehat{f * g}(\xi) &= \int_{\mathbb{R}^{2n}} e^{-ix\cdot\xi} f(y)g(x-y)\,dxdy \\ &= \int_{\mathbb{R}^{2n}} e^{-iy\cdot\xi}e^{-iz\cdot\xi} f(y)g(z)\,dydz = \widehat{f}(\xi)\widehat{g}(\xi).\end{aligned}$$

□

We give three more properties. The proofs are left as exercises below.

PROPOSITION 11.11. *Assume that $f \in L^1(\mathbb{R}^n)$ and $g \in L^1(\mathbb{R}^m)$. Then*

$$\widehat{f \otimes g} = \widehat{f} \otimes \widehat{g}. \tag{11.18}$$

PROPOSITION 11.12. *Assume that $A : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear map and $f \in L^1(\mathbb{R}^n)$. Then*

$$\widehat{A^* f}(\xi) = |\det A|^{-1} \widehat{f}(A^{-T}\xi) \tag{11.19}$$

where A^{-T} denotes the inverse of the transpose of A.

PROPOSITION 11.13. *Let $f \in L^1(\mathbb{R}^n)$ and $\overline{f}$ be the complex conjugate of f, i.e. $\overline{f}(x) = \overline{f(x)}$. Then*

$$\mathcal{F}(\overline{f})(\xi) = \overline{(\mathcal{F}f)(-\xi)}. \tag{11.20}$$

We finally compute the Fourier transform of the *Gaussian function*

$$G(x) = e^{-\frac{|x|^2}{2}}, \quad x \in \mathbb{R}^n. \tag{11.21}$$

Note that $G \in \mathscr{S}(\mathbb{R}^n)$, since for any α, β the function $x^\alpha \partial_x^\beta G$ is the product of G with a polynomial and thus is bounded on $\mathbb{R}^n$.

PROPOSITION 11.14. *If G is given by (11.21), then*

$$\widehat{G}(\xi) = (2\pi)^{\frac{n}{2}} e^{-\frac{|\xi|^2}{2}}, \tag{11.22}$$

that is $\widehat{G} = (2\pi)^{\frac{n}{2}} G$.

PROOF. It suffices to consider the case of dimension $n = 1$. Indeed, if G_n is the Gaussian in $\mathbb{R}^n$, then $G_{n+m} = G_n \otimes G_m$, so the formula for $\widehat{G}_{n+m}$ can be deduced from the ones for $\widehat{G}_n$ and $\widehat{G}_m$ by Proposition 11.11.

First proof: The function $G(x)$ satisfies the linear first order ODE

$$\partial_x G(x) = -xG(x). \tag{11.23}$$

Taking the Fourier transform of both sides and using (11.10) and (11.11), we see that $\widehat{G}$ satisfies the same ODE (11.23). From standard theory of linear ODEs we see that

$$\widehat{G}(\xi) = cG(\xi)$$

for some constant $c \in \mathbb{R}$. To compute c, we take $\xi = 0$, then

$$c = \widehat{G}(0) = \int_{\mathbb{R}} e^{-\frac{x^2}{2}}\, dx = \sqrt{2\pi}$$

is the Gaussian integral.

Second proof: We write down the integral for $\widehat{G}(\xi)$ and complete the square:

$$\widehat{G}(\xi) = \int_{\mathbb{R}} e^{-\frac{x^2}{2} - ix\xi}\, dx = \int_{\mathbb{R}} e^{-\frac{(x+i\xi)^2}{2} - \frac{\xi^2}{2}}\, dx.$$

We write this as a complex integral:

$$\widehat{G}(\xi) = e^{-\frac{\xi^2}{2}} \int_{\mathbb{R}+i\xi} e^{-\frac{z^2}{2}}\, dz.$$

Since $F(z) := e^{-\frac{z^2}{2}}$ is holomorphic in $z \in \mathbb{C}$ and satisfies $F(x + i\xi) \to 0$ as $|x| \to \infty$ locally uniformly in ξ, we can deform the contour above from $\mathbb{R} + i\xi$ back to $\mathbb{R}$ (more precisely, use Cauchy's integral theorem on the boundary of the domain $[-R, R]+[0, \xi]$ and let $R \to \infty$) and get

$$\widehat{G}(\xi) = e^{-\frac{\xi^2}{2}} \int_{\mathbb{R}} e^{-\frac{x^2}{2}}\, dx = \sqrt{2\pi} e^{-\frac{\xi^2}{2}}$$

where we again used the Gaussian integral. □

11.1.5. Fourier inversion formula. We are now ready to prove one of the most magical properties of the Fourier transform, which is a formula for its inverse:

THEOREM 11.15. *Assume that $\varphi \in \mathscr{S}(\mathbb{R}^n)$. Then we have for all $x \in \mathbb{R}^n$,*

$$\varphi(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{ix\cdot\xi} \widehat{\varphi}(\xi)\, d\xi. \tag{11.24}$$

REMARK 11.16. *It follows from Theorem 11.15 that the operator $\mathcal{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is invertible and its inverse is given by the formula*

$$\mathcal{F}^{-1}\psi(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{ix\cdot\xi} \psi(\xi)\, d\xi. \tag{11.25}$$

Note that $\mathcal{F}^{-1}$ maps $\mathscr{S}(\mathbb{R}^n)$ to itself by Theorem 11.6, since $\mathcal{F}^{-1}\psi(x) = (2\pi)^{-n}\widehat{\psi}(-x)$.

REMARK 11.17.$^{\mathbf{X}}$ *An interpretation of Theorem 11.15 is as follows: $\widehat{\varphi}(\xi)$ is the L^2 inner product $\langle \varphi, e_\xi \rangle_{L^2(\mathbb{R}^n)}$ between φ and the complex exponential wave at frequency ξ defined as $e_\xi(x) = e^{ix\cdot\xi}$. We can think of $\widehat{\varphi}(\xi)$ as the (complex) amplitude of the function φ at frequency ξ. Now the inversion formula (11.24) can be rewritten as*

$$\varphi = (2\pi)^{-n} \int_{\mathbb{R}^n} \widehat{\varphi}(\xi) e_\xi\, d\xi.$$

We can interpret this as φ being reconstructed from the basic waves e_ξ as an integral (which is analogous to a linear combination) with $\widehat{\varphi}(\xi)$ giving the coefficients. In a *way this formula is similar to writing an element of a Hilbert space in terms of its coefficients in an orthonormal basis. Of course this is only a heuristic – the functions* e_ξ *do not lie in* L^2 *and the argument above does not explain the factor* $(2\pi)^{-n}$.

Proof of Theorem 11.15. We can write the right-hand side of (11.24) as an iterated integral:

$$(2\pi)^{-n}\int_{\mathbb{R}^n} e^{ix\cdot\xi}\Big(\int_{\mathbb{R}^n} e^{-iy\cdot\xi}\varphi(y)\,dy\Big)\,d\xi.$$

However, Fubini's theorem does not apply here since the function $e^{i(x-y)\cdot\xi}\varphi(y)$ is not integrable on $\mathbb{R}^{2n}$.

To fix this issue, we regularize the integral using the Gaussian G defined in (11.21), which is a useful function because we have previously computed its Fourier transform in (11.22). Since $\widehat{\varphi}\in\mathscr{S}(\mathbb{R}^n)\subset L^1(\mathbb{R}^n)$ and $G(0)=1$, by the Dominated Convergence Theorem we see that the right-hand side of (11.24) is equal to

$$\begin{aligned}\lim_{\varepsilon\to0+}(2\pi)^{-n}\int_{\mathbb{R}^n} e^{ix\cdot\xi}G(\varepsilon\xi)\widehat{\varphi}(\xi)\,d\xi &= \lim_{\varepsilon\to0+}(2\pi)^{-n}\int_{\mathbb{R}^{2n}} e^{i(x-y)\cdot\xi}G(\varepsilon\xi)\varphi(y)\,dyd\xi\\ &= \lim_{\varepsilon\to0+}(2\pi\varepsilon)^{-n}\int_{\mathbb{R}^{2n}} e^{\frac{i(x-y)\cdot\eta}{\varepsilon}}G(\eta)\varphi(y)\,dyd\eta\\ &= \lim_{\varepsilon\to0+}(2\pi\varepsilon)^{-n}\int_{\mathbb{R}^n}\widehat{G}\Big(\frac{y-x}{\varepsilon}\Big)\varphi(y)\,dy\\ &= \lim_{\varepsilon\to0+}(2\pi)^{-n}\int_{\mathbb{R}^n}\widehat{G}(w)\varphi(x+\varepsilon w)\,dw.\end{aligned}\tag{11.26}$$

Here in the first line we use Fubini's theorem (which applies now since $G\in L^1(\mathbb{R}^n)$). In the second line we make the change of variables $\xi=\eta/\varepsilon$. In the third line we use Fubini's theorem again to integrate out η, and in the last line we make the change of variables $y=x+\varepsilon w$.

Since $\widehat{G}\in L^1(\mathbb{R}^n)$, we can use the Dominated Convergence Theorem, the explicit formula (11.22) for $\widehat{G}$, and the Gaussian integral to compute the limit on the last line of (11.26) as

$$(2\pi)^{-n}\int_{\mathbb{R}^n}\widehat{G}(w)\varphi(x)\,dw=\varphi(x)$$

which finishes the proof. □

As an application, we obtain the formula for the Fourier transform of a product. Note that for two Schwartz functions, their product and convolution are still Schwartz functions (see Exercise 11.8 for the latter).

Proposition 11.18. *Assume that* $\varphi,\psi\in\mathscr{S}(\mathbb{R}^n)$. *Then*

$$\widehat{\varphi\psi}=(2\pi)^{-n}\widehat{\varphi}*\widehat{\psi}.\tag{11.27}$$

PROOF. Similarly to (11.17) we have for each $f, g \in L^1(\mathbb{R}^n)$

$$\mathcal{F}^{-1}(f * g) = (2\pi)^n \mathcal{F}^{-1}(f)\mathcal{F}^{-1}(g).$$

It remains to apply this with $f := \widehat{\varphi}$, $g := \widehat{\psi}$. □

11.2. Fourier transform on tempered distributions

In this section we extend the Fourier transform from $L^1(\mathbb{R}^n)$ to the much larger space of tempered distributions.

11.2.1. Tempered distributions. We first define tempered distributions as the dual space to $\mathscr{S}(\mathbb{R}^n)$, similarly to the spaces $\mathcal{D}'$ and $\mathcal{E}'$:

DEFINITION 11.19. *Let $u : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ be a linear functional. We say that u is a* tempered distribution *if for each sequence $\varphi_k \to 0$ in $\mathscr{S}(\mathbb{R}^n)$ we have $(u, \varphi_k) \to 0$. Denote by $\mathscr{S}'(\mathbb{R}^n)$ the space of all tempered distributions on $\mathbb{R}^n$.*

For a sequence $u_k \in \mathscr{S}(\mathbb{R}^n)$, we say that $u_k \to u$ in $\mathscr{S}'(\mathbb{R}^n)$ if $(u_k, \varphi) \to (u, \varphi)$ for all $\varphi \in \mathscr{S}(\mathbb{R}^n)$.

REMARK 11.20.$^{\mathbf{S}}$ *Similarly to Proposition 2.6, a linear functional u on $\mathscr{S}(\mathbb{R}^n)$ lies in $\mathscr{S}'(\mathbb{R}^n)$ if and only if there exist C, N, M such that*

$$|(u, \varphi)| \leq C\|\varphi\|_{N,M} \quad \text{for all } \varphi \in \mathscr{S}(\mathbb{R}^n) \tag{11.28}$$

where the seminorm $\|\bullet\|_{N,M}$ was defined in (11.5).

REMARK 11.21.$^{\mathbf{S}}$ *There is a natural version of the Banach–Steinhaus Theorem for the space $\mathscr{S}'(\mathbb{R}^n)$ which is proved in the same way as Theorem 4.14. In particular, we have the following analog of Proposition 4.18:*

$$u_k \to u \quad \text{in } \mathscr{S}'(\mathbb{R}^n), \quad \varphi_k \to \varphi \quad \text{in } \mathscr{S}(\mathbb{R}^n) \quad \Longrightarrow \quad (u_k, \varphi_k) \to (u, \varphi). \tag{11.29}$$

The space $\mathscr{S}'(\mathbb{R}^n)$ is fairly large, in particular for any N we have

$$(1 + |x|)^N L^1(\mathbb{R}^n) \ \subset \ \mathscr{S}'(\mathbb{R}^n).$$

More precisely, if $f : \mathbb{R}^n \to \mathbb{C}$ is a function such that $(1 + |x|)^{-N} f \in L^1(\mathbb{R}^n)$ for some N, then we treat f as an element of $\mathscr{S}'(\mathbb{R}^n)$ by defining the pairing (f, φ) for $\varphi \in \mathscr{S}(\mathbb{R}^n)$ as the integral (2.3). In particular, the space $L^p(\mathbb{R}^n)$ embeds into $\mathscr{S}'(\mathbb{R}^n)$ for any $p \in [1, \infty]$ and any polynomial function lies in $\mathscr{S}'(\mathbb{R}^n)$.

Similarly to the inclusion $\mathcal{E}' \subset \mathcal{D}'$ (see (4.2)) we have the inclusions

$$\mathcal{E}'(\mathbb{R}^n) \ \subset \ \mathscr{S}'(\mathbb{R}^n) \ \subset \ \mathcal{D}'(\mathbb{R}^n), \tag{11.30}$$

since $C^\infty_c(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n) \subset C^\infty(\mathbb{R}^n)$, $C^\infty_c(\mathbb{R}^n)$ is dense in $\mathscr{S}(\mathbb{R}^n)$ (see Exercise 11.1), and $\mathscr{S}(\mathbb{R}^n)$ contains $C^\infty_c(\mathbb{R}^n)$ which is dense in $C^\infty(\mathbb{R}^n)$. Moreover, $C^\infty_c(\mathbb{R}^n)$ is dense in $\mathscr{S}'(\mathbb{R}^n)$ (see Exercise 11.2).

We now briefly discuss how previously defined operations on distributions act on the space $\mathscr{S}'(\mathbb{R}^n)$, viewed as a subspace of $\mathcal{D}'(\mathbb{R}^n)$. All the operations below are sequentially continuous on the indicated spaces. The first property is straightforward to verify and the rest are assigned as exercises below.

(1) If $u \in \mathscr{S}'(\mathbb{R}^n)$, then its distributional derivative $\partial_{x_j} u \in \mathcal{D}'(\mathbb{R}^n)$ also lies in $\mathscr{S}'(\mathbb{R}^n)$ and satisfies

$$(\partial_{x_j} u, \varphi) = -(u, \partial_{x_j} \varphi) \quad \text{for all } \varphi \in \mathscr{S}(\mathbb{R}^n).$$

(2) If $u \in \mathscr{S}'(\mathbb{R}^n)$ and $a \in C^\infty(\mathbb{R}^n)$ has polynomially bounded derivatives, i.e. for each α there exists N such that $\partial_x^\alpha a(x) = \mathcal{O}((1+|x|)^N)$, then $au \in \mathcal{D}'(\mathbb{R}^n)$ lies in $\mathscr{S}'(\mathbb{R}^n)$. In particular, this applies if a is a polynomial or $a \in \mathscr{S}(\mathbb{R}^n)$.

(3) If $u \in \mathscr{S}'(\mathbb{R}^n)$ and $v \in \mathscr{S}'(\mathbb{R}^m)$ then $u \otimes v \in \mathcal{D}'(\mathbb{R}^{n+m})$ lies in $\mathscr{S}'(\mathbb{R}^{n+m})$.

(4) If $A : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear map and $u \in \mathscr{S}'(\mathbb{R}^n)$ then $A^* u \in \mathcal{D}'(\mathbb{R}^n)$ lies in $\mathscr{S}'(\mathbb{R}^n)$.

(5) If $u \in \mathscr{S}'(\mathbb{R}^n)$ and $\varphi \in \mathscr{S}(\mathbb{R}^n)$, then the convolution

$$u * \varphi(x) = (u, \varphi(x - \bullet)), \quad x \in \mathbb{R}^n \tag{11.31}$$

is a smooth function on $\mathbb{R}^n$ with polynomially bounded derivatives, and thus in particular lies in $\mathscr{S}'(\mathbb{R}^n)$.

(6) If $u \in \mathscr{S}'(\mathbb{R}^n)$ and $v \in \mathcal{E}'(\mathbb{R}^n)$, then the convolution $u * v \in \mathcal{D}'(\mathbb{R}^n)$ lies in $\mathscr{S}'(\mathbb{R}^n)$.

11.2.2. Extending Fourier transform to tempered distributions. We now define Fourier transform of tempered distributions. As with many other operations before, we use duality. Recall from Proposition 11.9 that for all $f, g \in L^1(\mathbb{R}^n)$ we have

$$(\widehat{f}, g) = (f, \widehat{g}). \tag{11.32}$$

This motivates the following

Definition 11.22. *Let $u \in \mathscr{S}'(\mathbb{R}^n)$. Define the Fourier transform $\mathcal{F}u = \widehat{u} \in \mathscr{S}'(\mathbb{R}^n)$ by the formula*

$$(\widehat{u}, \varphi) := (u, \widehat{\varphi}) \quad \textit{for all } \varphi \in \mathscr{S}(\mathbb{R}^n). \tag{11.33}$$

Since $\mathcal{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is sequentially continuous by Theorem 11.6, we see that $\widehat{u}$ is indeed a tempered distribution. Moreover, the operator $\mathcal{F} : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ defined in (11.33) is sequentially continuous. By (11.32), if $u \in L^1(\mathbb{R}^n)$ then the distribution $\widehat{u}$ defined in (11.33) agrees with the classical Fourier transform of u defined in (11.1).

Having defined the Fourier transform on $\mathscr{S}'(\mathbb{R}^n)$, we give two important examples:

PROPOSITION 11.23. *On $\mathbb{R}^n$, we have*

$$\widehat{\delta_0} = 1, \tag{11.34}$$

$$\widehat{1} = (2\pi)^n \delta_0. \tag{11.35}$$

PROOF. (11.34): We compute for each $\varphi \in \mathscr{S}(\mathbb{R}^n)$

$$(\widehat{\delta_0}, \varphi) = (\delta_0, \widehat{\varphi}) = \widehat{\varphi}(0) = \int_{\mathbb{R}^n} \varphi(x)\,dx = (1, \varphi).$$

(11.35): We compute for each $\varphi \in \mathscr{S}(\mathbb{R}^n)$

$$(\widehat{1}, \varphi) = (1, \widehat{\varphi}) = \int_{\mathbb{R}^n} \widehat{\varphi}(\xi)\,d\xi = (2\pi)^n \varphi(0) = (2\pi)^n(\delta_0, \varphi)$$

where we use Fourier Inversion Formula (Theorem 11.15) with $x := 0$. □

REMARK 11.24. *In PDE papers,* (11.35) *is often written as*

$$\int_{\mathbb{R}^n} e^{-ix\cdot\xi}\,dx = (2\pi)^n \delta_0(\xi)$$

despite the fact that the integral does not converge. (One could actually make sense of an integral here by repeated integration by parts, see for example [**Hör03**, §7.8]*.) If we formally substitute $\xi = 0$, we obtain the nonsensical statement*

$$\int_{\mathbb{R}^n} dx = (2\pi)^n \delta_0(0).$$

We now discuss some properties of Fourier transform on tempered distributions. We start with the Fourier inversion formula. The inverse Fourier transform operator $\mathcal{F}^{-1} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ defined in (11.25) extends to a sequentially continuous operator $\mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$, since we have $\mathcal{F}^{-1}u(x) = (2\pi)^{-n}\,\widehat{u}(-x)$. Since $\mathscr{S}(\mathbb{R}^n)$ is dense in $\mathscr{S}'(\mathbb{R}^n)$, the operators $\mathcal{F}$ and $\mathcal{F}^{-1}$ are still inverses of each other when acting on $\mathscr{S}'(\mathbb{R}^n)$.

A similar argument using the density of $\mathscr{S}$ in $\mathscr{S}'$ shows that the identities (11.10) and (11.11) hold on $\mathscr{S}'$. Iterating these, we see that for all $u \in \mathscr{S}'(\mathbb{R}^n)$ and α

$$\widehat{D_x^\alpha u}(\xi) = \xi^\alpha \widehat{u}(\xi), \tag{11.36}$$

$$\widehat{x^\alpha u}(\xi) = (-1)^{|\alpha|} D_\xi^\alpha \widehat{u}(\xi). \tag{11.37}$$

Arguing again by density, we also get the distributional analogues of the formulas (11.18) and (11.19): if $u \in \mathscr{S}'(\mathbb{R}^n)$ and $v \in \mathscr{S}'(\mathbb{R}^m)$ then

$$\widehat{u \otimes v} = \widehat{u} \otimes \widehat{v}, \tag{11.38}$$

and if $u \in \mathscr{S}'(\mathbb{R}^n)$ and $A : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear map, then

$$\widehat{A^* u} = |\det A|^{-1} (A^{-T})^* \widehat{u}. \tag{11.39}$$

We also give a version of the convolution formula:

Proposition 11.25. *Assume that* $u \in \mathscr{S}'(\mathbb{R}^n)$ *and* $\varphi \in \mathscr{S}(\mathbb{R}^n)$. *Then*

$$\widehat{u * \varphi} = \widehat{u}\,\widehat{\varphi}, \tag{11.40}$$

$$\widehat{u\varphi} = (2\pi)^{-n}\,\widehat{u} * \widehat{\varphi}. \tag{11.41}$$

Proof. (11.40): We first review why both sides of the formula make sense. Since $u \in \mathscr{S}'(\mathbb{R}^n)$ and $\varphi \in \mathscr{S}(\mathbb{R}^n)$, we can define their convolution $u*\varphi \in \mathscr{S}'(\mathbb{R}^n)$ by (11.31). The product $\widehat{u}\widehat{\varphi}$ lies in $\mathscr{S}'(\mathbb{R}^n)$ (see Exercise 11.3).

We now argue by density. All the operations used are sequentially continuous on appropriate spaces, so if $u_k \to u$ in $\mathscr{S}'(\mathbb{R}^n)$, then

$$\widehat{u_k * \varphi} \to \widehat{u * \varphi}, \quad \widehat{u_k}\,\widehat{\varphi} \to \widehat{u}\,\widehat{\varphi} \quad \text{in } \mathscr{S}'(\mathbb{R}^n).$$

Since $\mathscr{S}(\mathbb{R}^n)$ is dense in $\mathscr{S}'(\mathbb{R}^n)$, we can choose $u_k \in \mathscr{S}(\mathbb{R}^n)$ converging in $\mathscr{S}'(\mathbb{R}^n)$ to any given $u \in \mathscr{S}'(\mathbb{R}^n)$. The formula (11.40) holds for u_k, φ by (11.17) and it remains to pass to the limit.

(11.41): This is proved in the same way, using the identity (11.27). □

11.2.3. Fourier transform of compactly supported distributions. We previously saw that each $u \in \mathcal{E}'(\mathbb{R}^n)$ also lies in $\mathscr{S}'(\mathbb{R}^n)$. The Fourier transform of u is a smooth function given by a simple formula:

Proposition 11.26. *Let* $u \in \mathcal{E}'(\mathbb{R}^n)$. *Then* $\widehat{u} \in C^\infty(\mathbb{R}^n)$ *has polynomially bounded derivatives and*

$$\widehat{u}(\xi) = (u(x), e^{-ix\cdot\xi}) \quad \text{for all } \xi \in \mathbb{R}^n. \tag{11.42}$$

Here we can pair u *with* $e_{-\xi}(x) := e^{-ix\cdot\xi}$ *since* $u \in \mathcal{E}'(\mathbb{R}^n)$ *and* $e_{-\xi} \in C^\infty(\mathbb{R}^n)$.

Proof. 1. Define $v(\xi) := (u(x), e^{-ix\cdot\xi})$ following (11.42). By Proposition 6.3 we see that $v \in C^\infty(\mathbb{R}^n)$ and for all α

$$\partial_\xi^\alpha v(\xi) = (u(x), \partial_\xi^\alpha e^{-ix\cdot\xi}) = (-i)^{|\alpha|}(u(x), x^\alpha e^{-ix\cdot\xi}).$$

Since $u \in \mathcal{E}'(\mathbb{R}^n)$, by Proposition 4.12 there exist $K \Subset \mathbb{R}^n$ and constants C, N such that for all ξ

$$|\partial_\xi^\alpha v(\xi)| \leq C\|x^\alpha e_{-\xi}\|_{C^N(\mathbb{R}^n,K)} = C \max_{|\beta|\leq N} \sup_{x\in K} |\partial_x^\beta (x^\alpha e^{-ix\cdot\xi})|.$$

Thus v has polynomially bounded derivatives, more precisely for each α there exists C_α such that

$$|\partial_\xi^\alpha v(\xi)| \leq C_\alpha (1 + |\xi|)^N \quad \text{for all } \xi \in \mathbb{R}^n. \tag{11.43}$$

2. It remains to show that $\widehat{u} = v$. One way to see is this by approximation: this is true when $u \in C_c^\infty(\mathbb{R}^n)$, both $\widehat{u}$ and v depend continuously on $u \in \mathcal{E}'(\mathbb{R}^n)$, and $C_c^\infty(\mathbb{R}^n)$ is dense in $\mathcal{E}'(\mathbb{R}^n)$ similarly to Theorem 6.7.

We give here a more direct way. Fix $\varphi \in \mathscr{S}(\mathbb{R}^n)$, then we need to show that

$$(u, \widehat{\varphi}) = (v, \varphi). \tag{11.44}$$

This is proved as follows:

$$\begin{aligned}(u, \widehat{\varphi}) &= \Big(u(x), \int_{\mathbb{R}^n} e^{-ix\cdot\xi}\varphi(\xi)\,d\xi\Big)\\ &= \int_{\mathbb{R}^n} (u(x), e^{-ix\cdot\xi}\varphi(\xi))\,d\xi\\ &= \int_{\mathbb{R}^n} v(\xi)\varphi(\xi)\,d\xi = (v, \varphi).\end{aligned}$$

Here in the second line we use that the integral $\int_{\mathbb{R}^n} e^{-ix\cdot\xi}\varphi(\xi)\,d\xi$ converges in $C^\infty(\mathbb{R}^n)$ in the x variable, so one can exchange the integral with the pairing with $u \in \mathcal{E}'(\mathbb{R}^n)$ similarly to the proof of Lemma 6.8. □

REMARK 11.27.$^{\mathbf{X}}$ *The function $\widehat{u}$ is in fact real analytic – see Theorem 11.31 below.*

If $u \in \mathscr{S}'(\mathbb{R}^n)$ and $v \in \mathcal{E}'(\mathbb{R}^n)$, then the convolution $u*v$ lies in $\mathscr{S}'(\mathbb{R}^n)$ by item (6) at the end of §11.2.1. By Exercise 11.3, the product $\widehat{u}\,\widehat{v}$ lies in $\mathscr{S}'(\mathbb{R}^n)$. The two are related by a convolution formula:

PROPOSITION 11.28. *Assume that $u \in \mathscr{S}'(\mathbb{R}^n)$ and $v \in \mathcal{E}'(\mathbb{R}^n)$. Then*

$$\widehat{u * v} = \widehat{u}\,\widehat{v}. \tag{11.45}$$

PROOF. One possibility is to argue using the density of $C_c^\infty(\mathbb{R}^n)$ in both $\mathscr{S}'(\mathbb{R}^n)$ and $\mathcal{E}'(\mathbb{R}^n)$. Here we present a more direct proof.

1. We first consider the case when both u, v lie in $\mathcal{E}'(\mathbb{R}^n)$. In this case $\widehat{u}, \widehat{v}$ are smooth functions and the proof is simple: denoting $e_\xi(x) = e^{ix\cdot\xi}$ we have (recalling the definition of convolution in §8.1)

$$\begin{aligned}\widehat{u * v}(\xi) &= (u(x) \otimes v(y), e_{-\xi}(x+y))\\ &= (u(x) \otimes v(y), e_{-\xi}(x) \otimes e_{-\xi}(y))\\ &= (u(x), e_{-\xi}(x))(v(y), e_{-\xi}(y)) = \widehat{u}(\xi)\widehat{v}(\xi).\end{aligned}$$

2.$^{\mathbf{X}}$ We now consider the general case when $u \in \mathscr{S}'(\mathbb{R}^n)$ and $v \in \mathcal{E}'(\mathbb{R}^n)$. Using the definition of convolution in §8.2 we have for each $\varphi \in \mathscr{S}(\mathbb{R}^n)$

$$(\widehat{u * v}, \varphi) = (u * v, \widehat{\varphi}) = \big(u(x), (v(y), \widehat{\varphi}(x+y))\big)$$

where $(v(y), \widehat{\varphi}(x+y)) \in \mathscr{S}(\mathbb{R}^n)$ (see Exercise 11.9). We now compute

$$
\begin{aligned}
(v(y), \widehat{\varphi}(x+y)) &= \left(v(y), \int_{\mathbb{R}^n} e^{-ix\cdot\xi} e^{-iy\cdot\xi} \varphi(\xi)\, d\xi\right) \\
&= \int_{\mathbb{R}^n} (v(y), e^{-iy\cdot\xi}) e^{-ix\cdot\xi} \varphi(\xi)\, d\xi \\
&= \int_{\mathbb{R}^n} \widehat{v}(\xi) e^{-ix\cdot\xi} \varphi(\xi)\, d\xi = \widehat{\widehat{v}\,\varphi}(x).
\end{aligned}
$$

Here in the second line we use that the integral $\int_{\mathbb{R}^n} e^{-ix\cdot\xi} e^{-iy\cdot\xi}\varphi(\xi)\, d\xi$ converges in $C^\infty(\mathbb{R}^n)$ in the y variable, so one can exchange the integral with the pairing with $v \in \mathcal{E}'(\mathbb{R}^n)$ similarly to the proof of Lemma 6.8. We now have

$$
(\widehat{u * v}, \varphi) = (u, \widehat{\widehat{v}\,\varphi}) = (\widehat{u}, \widehat{v}\,\varphi) = (\widehat{u}\,\widehat{v}, \varphi),
$$

giving (11.45). □

11.2.4. Fourier transform on L^2. The next theorem shows that the Fourier transform acts as a unitary operator on the space $L^2(\mathbb{R}^n) \subset \mathscr{S}'(\mathbb{R}^n)$ (up to a constant). It is key in the Fourier transform characterization of Sobolev spaces and gives one more reason why the space L^2 is the best for many applications to PDEs.

THEOREM 11.29. *Assume that $f \in L^2(\mathbb{R}^n)$. Then the Fourier transform $\widehat{f}$, defined by (11.33), also lies in $L^2(\mathbb{R}^n)$ and we have*

$$
\|\widehat{f}\|_{L^2(\mathbb{R}^n)} = (2\pi)^{\frac{n}{2}} \|f\|_{L^2(\mathbb{R}^n)}. \tag{11.46}
$$

Similarly we have $\mathcal{F}^{-1}(f) \in L^2(\mathbb{R}^n)$, so $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is an invertible linear operator.

PROOF. 1. We first show the identity

$$
\|\widehat{\varphi}\|_{L^2(\mathbb{R}^n)} = (2\pi)^{\frac{n}{2}} \|\varphi\|_{L^2(\mathbb{R}^n)} \quad \text{for all } \varphi \in \mathscr{S}(\mathbb{R}^n). \tag{11.47}
$$

To do this, we apply (11.27) to the function φ and its complex conjugate $\overline{\varphi}$:

$$
\widehat{|\varphi|^2} = \widehat{\varphi\overline{\varphi}} = (2\pi)^{-n} \widehat{\varphi} * \widehat{\overline{\varphi}}.
$$

Evaluating both sides at 0 and using (11.20) we get

$$
\|\varphi\|^2_{L^2(\mathbb{R}^n)} = \widehat{|\varphi|^2}(0) = (2\pi)^{-n} \int_{\mathbb{R}^n} \widehat{\varphi}(\xi) \overline{\widehat{\varphi}(\xi)}\, d\xi = (2\pi)^{-n} \|\widehat{\varphi}\|^2_{L^2(\mathbb{R}^n)}
$$

which gives (11.47).

2. Now take arbitrary $f \in L^2(\mathbb{R}^n)$. By Theorem 1.14, there exists a sequence $\varphi_k \in \mathscr{S}(\mathbb{R}^n)$ converging to f in $L^2(\mathbb{R}^n)$. In particular, φ_k is a Cauchy sequence in $L^2(\mathbb{R}^n)$. By (11.47) we have

$$
\|\widehat{\varphi}_k - \widehat{\varphi}_\ell\|_{L^2(\mathbb{R}^n)} = (2\pi)^{\frac{n}{2}} \|\varphi_k - \varphi_\ell\|_{L^2(\mathbb{R}^n)}.
$$

Thus $\widehat{\varphi}_k$ is also a Cauchy sequence in $L^2(\mathbb{R}^n)$. Since $L^2(\mathbb{R}^n)$ is a complete space, the sequence $\widehat{\varphi}_k$ converges in $L^2(\mathbb{R}^n)$ to some $g \in L^2(\mathbb{R}^n)$, with

$$\|g\|_{L^2(\mathbb{R}^n)} = \lim_{k\to\infty} \|\widehat{\varphi}_k\|_{L^2(\mathbb{R}^n)} = (2\pi)^{\frac{n}{2}} \lim_{k\to\infty} \|\varphi_k\|_{L^2(\mathbb{R}^n)} = (2\pi)^{\frac{n}{2}} \|f\|_{L^2(\mathbb{R}^n)}.$$

Since the convergence in $L^2(\mathbb{R}^n)$ is stronger than in $\mathscr{S}'(\mathbb{R}^n)$, and the Fourier transform is sequentially continuous on $\mathscr{S}'(\mathbb{R}^n)$, we see that

$$\widehat{\varphi}_k \to \widehat{f}, \quad \widehat{\varphi}_k \to g \quad \text{in } \mathscr{S}'(\mathbb{R}^n).$$

Thus $\widehat{f} = g$. □

REMARK 11.30. *Following the proof of Theorem 11.29 we also see that for all $f, g \in L^2(\mathbb{R}^n)$*

$$\langle \widehat{f}, \widehat{g} \rangle_{L^2(\mathbb{R}^n)} = (2\pi)^n \langle f, g \rangle_{L^2(\mathbb{R}^n)}, \tag{11.48}$$

$$(f, g) = (2\pi)^{-n} \int_{\mathbb{R}^n} \widehat{f}(\xi)\widehat{g}(-\xi)\, d\xi. \tag{11.49}$$

11.2.5. Paley–Wiener theorem$^{\mathbf{X}}$. We saw in Proposition 11.26 that when $u \in \mathcal{E}'(\mathbb{R}^n)$, the Fourier transform $\widehat{u}$ is a smooth function. In fact, this function is real analytic and one can characterize the space $\mathcal{E}'(\mathbb{R}^n)$ in terms of the properties of the holomorphic extension of $\widehat{u}$:

THEOREM 11.31 (Paley–Wiener theorem). *Let $R \geq 0$. Then:*

1. If $u \in \mathcal{E}'(\mathbb{R}^n)$ and $\operatorname{supp} u \subset B(0, R)$ then $\widehat{u} \in C^\infty(\mathbb{R}^n)$ extends to a holomorphic function $U : \mathbb{C}^n \to \mathbb{C}$ and there exist constants C, N such that

$$|U(\zeta)| \leq C(1 + |\zeta|)^N e^{R|\operatorname{Im}\zeta|} \quad \textit{for all } \zeta \in \mathbb{C}^n. \tag{11.50}$$

2. Conversely, if U is a holomorphic function on $\mathbb{C}^n$ satisfying the bound (11.50) for some C, N then there exists $u \in \mathcal{E}'(\mathbb{R}^n)$ such that $\operatorname{supp} u \subset B(0, R)$ and $U|_{\mathbb{R}^n} = \widehat{u}$.

For the proof, see [**Hör03**, Theorem 7.3.1] or [**FJ98**, Theorem 10.2.2]. Here we just give some informal explanations:

- Recall from Proposition 11.26 that $\widehat{u}(\xi) = (u(x), e^{-ix\cdot\xi})$. We define the extension of $\widehat{u}$ by $U(\zeta) := (u(x), e^{-ix\cdot\zeta})$ for $\zeta \in \mathbb{C}^n$ and U is holomorphic. The bound (11.50) can be verified by following the proof of Proposition 11.26.
- For part 2, let us consider the case when (11.50) is replaced by the following stronger estimate: for each N there exists C_N such that

$$|U(\zeta)| \leq C_N(1 + |\zeta|)^{-N} e^{R|\operatorname{Im}\zeta|} \quad \text{for all } \zeta \in \mathbb{C}^n. \tag{11.51}$$

  The function $U|_{\mathbb{R}^n}$ is Schwartz (as follows from (11.51) and Cauchy estimates for derivatives of holomorphic functions), thus there exists $u \in \mathscr{S}(\mathbb{R}^n)$ such that $\widehat{u} = U|_{\mathbb{R}^n}$.

- It remains to show that $\operatorname{supp} u \subset B(0, R)$. Fix $x \in \mathbb{R}^n$ with $|x| > R$. Then there exists $\eta \in \mathbb{R}^n$ such that $|\eta| = 1$ and $x \cdot \eta > R$ (e.g. take $\eta = x/|x|$). By the Fourier inversion formula (Theorem 11.15) we write

$$u(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{ix\cdot\xi} U(\xi)\, d\xi.$$

Since U is holomorphic, using the estimate (11.51) we can deform the contour of integration to get

$$u(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{ix\cdot(\xi+it\eta)} U(\xi + it\eta)\, d\xi \quad \text{for all } t \in \mathbb{R}.$$

Using the estimate (11.51) with $N = n + 1$ we get

$$|u(x)| \le (2\pi)^{-n} \int_{\mathbb{R}^n} e^{-tx\cdot\eta} |U(\xi + it\eta)|\, d\xi \le C e^{-tx\cdot\eta + R|t|}.$$

Letting $t \to \infty$ we see that $u(x) = 0$ as needed.

11.2.6. Poisson summation formula$^{\mathbf{X}}$. We finally state a formula for the Fourier transform of the delta function on a periodic lattice:

THEOREM 11.32 (Poisson summation formula). *Define $u \in \mathscr{S}'(\mathbb{R}^n)$ by*

$$u := \sum_{k\in\mathbb{Z}^n} \delta_k,$$

that is for each $\varphi \in \mathscr{S}(\mathbb{R}^n)$ we have

$$(u, \varphi) = \sum_{k\in\mathbb{Z}^n} \varphi(k).$$

Then

$$\widehat{u} = (2\pi)^n \sum_{\ell\in\mathbb{Z}^n} \delta_{2\pi\ell}. \tag{11.52}$$

Equivalently, we have

$$\sum_{k\in\mathbb{Z}^n} e^{ik\cdot x} = (2\pi)^n \sum_{\ell\in\mathbb{Z}^n} \delta_{2\pi\ell}(x). \tag{11.53}$$

Here the series converge in $\mathscr{S}'(\mathbb{R}^n)$, that is for all $\varphi \in \mathscr{S}(\mathbb{R}^n)$

$$\sum_{k\in\mathbb{Z}^n} \widehat{\varphi}(k) = (2\pi)^n \sum_{\ell\in\mathbb{Z}^n} \varphi(2\pi\ell). \tag{11.54}$$

For the proof, see [**Hör03**, Theorem 7.2.1] or [**FJ98**, Theorem 8.5.1]. Here we just give some informal explanations for how Theorem 11.32 is related to Fourier series:

- Denote by $\mathbb{T}^n := \mathbb{R}^n/2\pi\mathbb{Z}^n$ the n-torus. Each $\psi \in C^\infty(\mathbb{T}^n)$ is the sum of its Fourier series

$$\psi(x) = \sum_{k\in\mathbb{Z}^n} \psi_k e^{ik\cdot x} \quad \text{where } \psi_k := (2\pi)^{-n} \int_{\mathbb{T}^n} e^{-ik\cdot x}\psi(x)\,dx. \tag{11.55}$$

We can think of ψ as a $2\pi\mathbb{Z}^n$-periodic function in $C^\infty(\mathbb{R}^n)$. The integral in (11.55) can then be computed by

$$\psi_k = (2\pi)^{-n} \int_{[0,2\pi]^n} e^{-ik\cdot x}\psi(x)\,dx. \tag{11.56}$$

- An alternative to (11.56), which is better for applications to distributions, is as follows. Fix $\chi \in C^\infty_c(\mathbb{R}^n)$ whose translates form a partition of unity:

$$\sum_{\ell\in\mathbb{Z}^n} \chi(x - 2\pi\ell) = 1 \quad \text{for all } x \in \mathbb{R}^n.$$

Then we can write

$$\psi_k = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{-ik\cdot x}\chi(x)\psi(x)\,dx. \tag{11.57}$$

We have $\psi_k = (2\pi)^{-n}\widehat{\chi\psi}(k)$, so (11.54) for the function $\varphi := \chi\psi$ is the same as (11.55) with $x = 0$.
- Here is a way to derive (11.54) (and thus Theorem 11.32) from (11.55). Take arbitrary $\varphi \in \mathscr{S}(\mathbb{R}^n)$ and define $\psi \in C^\infty(\mathbb{T}^n)$ as the symmetrization of φ:

$$\psi(x) := \sum_{\ell\in\mathbb{Z}^n} \varphi(x + 2\pi\ell) \quad \text{for all } x \in \mathbb{R}^n.$$

Using (11.57) we compute the Fourier coefficients of ψ:

$$\begin{aligned}\psi_k &= (2\pi)^{-n} \int_{\mathbb{R}^n} \sum_{\ell\in\mathbb{Z}^n} e^{-ik\cdot x}\chi(x)\varphi(x + 2\pi\ell)\,dx \\ &= (2\pi)^{-n} \sum_{\ell\in\mathbb{Z}^n} \int_{\mathbb{R}^n} e^{-ik\cdot x}\chi(x - 2\pi\ell)\varphi(x)\,dx = (2\pi)^{-n}\widehat{\varphi}(k).\end{aligned}$$

Using (11.55) at $x = 0$, we get

$$(2\pi)^{-n} \sum_{k\in\mathbb{Z}^n} \widehat{\varphi}(k) = \psi(0) = \sum_{\ell\in\mathbb{Z}^n} \varphi(2\pi\ell),$$

giving (11.54).
- Another way to interpret Poisson summation formula in terms of Fourier series is to show that (11.55) actually holds for any $2\pi\mathbb{Z}^n$-periodic distribution $\psi \in \mathcal{D}'(\mathbb{R}^n)$, with the series converging in $\mathcal{D}'(\mathbb{R}^n)$ and ψ_k defined by (11.57) with

the integral replaced by distributional pairing. If we now take ψ to be the delta function at $0 \in \mathbb{T}^n$, that is

$$\psi = \sum_{\ell \in \mathbb{Z}^n} \delta_{2\pi\ell},$$

then $\psi_k = (2\pi)^{-n}$ for all k, so (11.55) implies (11.53).

11.3. Notes and exercises

Our presentation follows [**Hör03**, §§7.1–7.3] and [**FJ98**, §§8.1–8.5,9.2,10.2]. Our proof of Theorem 11.15 follows a direct route by regularizing the double integral; there is an alternative proof by using the intertwining relations of Proposition 11.7, see [**Hör03**, Theorem 7.1.5].

Exercise 11.1. (1 pt) *This exercise shows that $C_c^\infty(\mathbb{R}^n)$ is dense in $\mathscr{S}(\mathbb{R}^n)$. Let $\varphi \in \mathscr{S}(\mathbb{R}^n)$ and $\psi \in C_c^\infty(\mathbb{R}^n)$ satisfy $\psi(0) = 1$. Put $\varphi_\varepsilon(x) := \psi(\varepsilon x)\varphi(x)$ for $\varepsilon > 0$. Show that $\varphi_\varepsilon \to \varphi$ in $\mathscr{S}(\mathbb{R}^n)$ as $\varepsilon \to 0+$.*

Exercise 11.2. (1 pt) *Show that $C_c^\infty(\mathbb{R}^n)$ is dense in $\mathscr{S}'(\mathbb{R}^n)$. (Hint: show that for an appropriate choice of $\psi, \chi \in C_c^\infty(\mathbb{R}^n)$ and each $u \in \mathscr{S}'(\mathbb{R}^n)$, we have $(\psi_\varepsilon u) * \chi_\varepsilon \to u$ in $\mathscr{S}'(\mathbb{R}^n)$ as $\varepsilon \to 0+$ where $\psi_\varepsilon(x) := \psi(\varepsilon x)$, $\chi_\varepsilon(x) := \varepsilon^{-n}\chi(x/\varepsilon)$. To do that, you can follow part of the proof of Theorem 6.10. You can use without proof that Lemma 6.8 applies when $u \in \mathcal{E}'(\mathbb{R}^n)$ and $\varphi \in C^\infty(\mathbb{R}^n)$.)*

Exercise 11.3. ($1 = 0.5 + 0.5$ pt) *Assume that $a \in C^\infty(\mathbb{R}^n)$ has polynomially bounded derivatives, i.e. for each α there exists N such that $\partial_x^\alpha a(x) = \mathcal{O}((1 + |x|)^N)$. Show that:*

(a) *multiplication by a is a sequentially continuous operator $\mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$;*

(b) *if $u \in \mathscr{S}'(\mathbb{R}^n)$, then the product $au \in \mathcal{D}'(\mathbb{R}^n)$ lies in $\mathscr{S}'(\mathbb{R}^n)$ and the map $u \mapsto au$ is sequentially continuous on $\mathscr{S}'(\mathbb{R}^n)$.*

Exercise 11.4. *(0.5 pt) Prove Proposition 11.13.*

Exercise 11.5. ($1.5 = 0.5{+}1$ pts) *This exercise shows the relation between Fourier transform and tensor product.*

(a) *Show that if $\varphi \in \mathscr{S}(\mathbb{R}^n)$, $\psi \in \mathscr{S}(\mathbb{R}^m)$ then $\varphi \otimes \psi \in \mathscr{S}(\mathbb{R}^{n+m})$ and prove Proposition 11.11.*

(b) *Show that if $u \in \mathscr{S}'(\mathbb{R}^n)$, $v \in \mathscr{S}'(\mathbb{R}^m)$ then the distributional tensor product $u \otimes v \in \mathcal{D}'(\mathbb{R}^{n+m})$ (defined in §7.1) lies in $\mathscr{S}'(\mathbb{R}^{n+m})$ and $\widehat{u \otimes v} = \widehat{u} \otimes \widehat{v}$.*

Exercise 11.6. ($1.5 = 0.5{+}0.5{+}0.5$ pts) *This exercise shows the relation between Fourier transform and pullback by an invertible linear map $A : \mathbb{R}^n \to \mathbb{R}^n$.*

(a) *Show then A^* is a sequentially continuous operator $\mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ and prove Proposition 11.12.*

(b) *Show that if $u \in \mathscr{S}'(\mathbb{R}^n)$, then the distributional pullback $A^*u \in \mathcal{D}'(\mathbb{R}^n)$ (defined in §10.1.3) lies in $\mathscr{S}'(\mathbb{R}^n)$ and the Fourier transform formula (11.19) holds.*

(c) *Assume that $u \in \mathscr{S}'(\mathbb{R}^n)$ is homogeneous of degree $a \in \mathbb{C}$. Show that $\widehat{u}$ is homogeneous and compute its degree of homogeneity. You may use Proposition 10.7.*

EXERCISE 11.7. (1 = 0.5 + 0.5 pt) *For $w \in \mathbb{R}^n$, define the following operators on $C^\infty(\mathbb{R}^n)$:*

$$\tau_w f(x) = f(x - w), \quad \sigma_w f(x) = e^{ix \cdot w} f(x).$$

(a) *Show that τ_w, σ_w define sequentially continuous operators on $\mathscr{S}(\mathbb{R}^n)$. Use this to show that for $u \in \mathscr{S}'(\mathbb{R}^n)$, the distributional pullback and product $\tau_w u, \sigma_w u \in \mathcal{D}'(\mathbb{R}^n)$ lie in $\mathscr{S}'(\mathbb{R}^n)$.*

(b) *Show that for each $u \in \mathscr{S}'(\mathbb{R}^n)$*

$$\widehat{\tau_w u} = \sigma_{-w} \widehat{u}, \quad \widehat{\sigma_w u} = \tau_w \widehat{u}.$$

EXERCISE 11.8. (2 = 1 + 1 pts) *This exercise studies the properties of convolution on Schwartz functions and tempered distributions.*

(a) *Assume that $\varphi, \psi \in \mathscr{S}(\mathbb{R}^n)$. Show that the convolution $\varphi * \psi$, defined by (1.27), lies in $\mathscr{S}(\mathbb{R}^n)$. (Hint: you can use the Leibniz Rule for convolutions, which states that $x_j(\varphi * \psi) = (x_j\varphi) * \psi + \varphi * (x_j\psi)$.)*

(b) *Assume that $u \in \mathscr{S}'(\mathbb{R}^n)$ and $\varphi \in \mathscr{S}(\mathbb{R}^n)$. Show that the convolution $u * \varphi$, defined by (11.31), is a smooth function on $\mathbb{R}^n$ with polynomially bounded derivatives.*

EXERCISE 11.9. (1 pt) *Assume that $u \in \mathscr{S}'(\mathbb{R}^n)$ and $v \in \mathcal{E}'(\mathbb{R}^n)$. Show that the convolution $u * v \in \mathcal{D}'(\mathbb{R}^n)$, defined in §8.2, lies in $\mathscr{S}'(\mathbb{R}^n)$. (Hint: use (8.9), (7.10), and show that for each $\varphi \in \mathscr{S}(\mathbb{R}^n)$ the function $x \mapsto (v(y), \varphi(x + y))$ lies in $\mathscr{S}(\mathbb{R}^n)$.)*

EXERCISE 11.10. (1 pt) *Let $\varphi \in \mathscr{S}(\mathbb{R}^n)$. For $t > 0$ and $x \in \mathbb{R}^n$, define*

$$u(t, x) = (4\pi t)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} \varphi(y)\, dy. \tag{11.58}$$

Using Proposition 11.10, show that u solves the heat equation $\partial_t u = \Delta_x u$ in $(0, \infty)_t \times \mathbb{R}^n$ and that $u(t, \bullet) \to \varphi$ as $t \to 0+$ in $\mathscr{S}(\mathbb{R}^n)$. (We can think of u as the convolution of $\delta_0(t) \otimes \varphi(x)$ with the fundamental solution to the heat equation given in (9.11), but for this problem it is useful to apply Proposition 11.10 in the x variable only for fixed t. You don't need to rigorously justify being able to exchange ∂_t with taking the Fourier transform in the x variable.)

EXERCISE 11.11. (3 = 1 + 1 + 1 pts) *This exercise gives a method to compute Fourier transforms of certain distributions using analytic continuation.*

(a) *Assume that* $u \in \mathscr{S}'(\mathbb{R})$ *and* $\operatorname{supp} u \subset [a, \infty)$ *for some* $a \in \mathbb{R}$. *Take a cutoff* $\chi \in C^\infty(\mathbb{R})$ *such that* $\chi = 1$ *near* $[a, \infty)$ *and* $\operatorname{supp} \chi \subset [a-1, \infty)$, *and define the function*

$$F(\eta) := (u(x), \chi(x) e^{-ix\eta}), \quad \eta \in \mathbb{C}, \quad \operatorname{Im} \eta < 0.$$

Explain why $F(\eta)$ *is well-defined and independent of* χ *and show that it is holomorphic in* $\{\operatorname{Im} \eta < 0\}$.

(b) *Show that* $F(\xi - i\varepsilon) \to \widehat{u}(\xi)$ *in* $\mathscr{S}'(\mathbb{R})$ *as* $\varepsilon \to 0+$. *(Hint:* $F(\xi - i\varepsilon)$ *is the Fourier transform of* $e^{-\varepsilon x} u(x)$ *but you should justify your arguments carefully.)*

(c) *Assume that* $a \in \mathbb{C}$ *and* $\operatorname{Re} a > -1$. *Show that the Fourier transform of* x_+^a *is given by* $e^{-i\pi(a+1)/2}\Gamma(a+1)(\xi - i0)^{-a-1}$ *where* Γ *is the Euler Gamma function and* $(\xi - i0)^{-a-1}$ *was defined in Exercise 5.4. In particular, compute the Fourier transform of the Heaviside function. (Hint: use parts (a)–(b), computing* $F(\eta)$ *for* $\eta = -is$, $s > 0$ *and then arguing by analytic continuation in* η. *The result actually holds for all* $a \in \mathbb{C}$ *by analytic continuation in* a.*)*

CHAPTER 12

Fourier transform II

In this chapter we explore applications of Fourier transform. We first define Sobolev spaces and establish their fundamental properties. We next proving the second version of Elliptic Regularity, applying to elliptic constant coefficient differential operators.

12.1. Sobolev spaces

12.1.1. A simple case. The Sobolev space $H^s(\mathbb{R}^n)$ is the subspace of $\mathscr{S}'(\mathbb{R}^n)$ whose elements are thought of as 'having derivatives up to order s lying in L^2'. This is an informal definition since s can be any real number (integer or non-integer, positive or negative). The easiest formal definition of these spaces for us is on the Fourier transform side. To prepare for this, we consider first the simplest case when s is a nonnegative integer:

Proposition 12.1. *Let $u \in \mathscr{S}'(\mathbb{R}^n)$ and $k \geq 0$ be an integer. Then the following are equivalent:*

(1) *u has derivatives up to order k in $L^2(\mathbb{R}^n)$, that is*

$$\partial_x^\alpha u \in L^2(\mathbb{R}^n) \quad \text{for all } \alpha,\ |\alpha| \leq k. \tag{12.1}$$

(Here as before, $\partial_x^\alpha u$ is defined in the sense of distributions.)

(2) *the Fourier transform $\widehat{u}$ (defined a priori as an element of $\mathscr{S}'(\mathbb{R}^n)$) is a locally integrable function such that*

$$(1+|\xi|)^k\, \widehat{u}(\xi) \in L^2(\mathbb{R}^n). \tag{12.2}$$

Proof. Since the Fourier transform on $\mathscr{S}'(\mathbb{R}^n)$ maps $L^2(\mathbb{R}^n)$ onto itself (by Theorem 11.29), either condition (1) or (2) above implies that $\widehat{u} \in L^2(\mathbb{R}^n)$.

Since the Fourier transform also intertwines differentiation with multiplication (by (11.36)), we have for any multiindex α

$$\partial_x^\alpha u \in L^2(\mathbb{R}^n) \iff \widehat{\partial_x^\alpha u} \in L^2(\mathbb{R}^n) \iff \xi^\alpha \widehat{u}(\xi) \in L^2(\mathbb{R}^n).$$

Thus (12.1) is equivalent to the statement

$$\xi^\alpha \widehat{u}(\xi) \in L^2(\mathbb{R}^n) \quad \text{for all } \alpha,\ |\alpha| \leq k$$

which is equivalent to (12.2) since

$$C^{-1}(1+|\xi|)^k \le 1 + \sum_{j=1}^n |\xi_j|^k \le \sum_{|\alpha|\le k} |\xi^\alpha| \le C(1+|\xi|)^k$$

for some constant C depending only on n, k. □

REMARK 12.2. *For ξ bounded, the condition (12.2) does not depend on k, it just states that $\widehat{u}(\xi)$ lies in L^2 (locally). The difference for different k is only in the asymptotic behavior of $\widehat{u}(\xi)$ as $|\xi| \to \infty$. This is a basic example of the general principle that* regularity *of a distribution u is related to the* decay *of its Fourier transform $\widehat{u}(\xi)$ as $|\xi| \to \infty$.*

12.1.2. General definition and basic properties. We now define general Sobolev spaces following (12.2). For convenience we define the function

$$\langle \xi \rangle := \sqrt{1+|\xi|^2}, \quad \xi \in \mathbb{R}^n \tag{12.3}$$

which is smooth on $\mathbb{R}^n$ and satisfies for some constant C depending only on n

$$C^{-1}(1+|\xi|) \le \langle \xi \rangle \le C(1+|\xi|).$$

DEFINITION 12.3. *Let $s \in \mathbb{R}$. Define the* Sobolev space of order s

$$H^s(\mathbb{R}^n) := \{u \in \mathscr{S}'(\mathbb{R}^n) \colon \langle \xi \rangle^s \widehat{u}(\xi) \in L^2(\mathbb{R}^n)\}. \tag{12.4}$$

Here are some basic properties of Sobolev spaces:

(1) Each $H^s(\mathbb{R}^n)$ is a Hilbert space, with the norm customarily defined by
$$\|u\|_{H^s(\mathbb{R}^n)} := \|\langle \xi \rangle^s \widehat{u}(\xi)\|_{L^2(\mathbb{R}^n)}, \quad u \in H^s(\mathbb{R}^n). \tag{12.5}$$
This follows from the fact that $H^s(\mathbb{R}^n)$ is isometric to the weighted L^2 space $\langle \xi \rangle^{-s} L^2(\mathbb{R}^n)$.

(2) We have the containment $H^s(\mathbb{R}^n) \subset H^t(\mathbb{R}^n)$ whenever $s \ge t$.

(3) If $s = k$ is a nonnegative integer, then $H^k(\mathbb{R}^n)$ consists of all functions $u \in L^2(\mathbb{R}^n)$ satisfying the equivalent conditions of Proposition 12.1, and the norm (12.5) is equivalent to the alternative Hilbert norm
$$\Big(\sum_{|\alpha|\le k} \|\partial_x^\alpha u\|^2_{L^2(\mathbb{R}^n)}\Big)^{\frac{1}{2}}.$$
In particular, $H^0(\mathbb{R}^n) = L^2(\mathbb{R}^n)$.

(4) We have the containments
$$\mathscr{S}(\mathbb{R}^n) \;\subset\; H^s(\mathbb{R}^n) \;\subset\; \mathscr{S}'(\mathbb{R}^n) \tag{12.6}$$
and convergence in $\mathscr{S}$ is stronger than convergence in H^s, which in turn is stronger than convergence in $\mathscr{S}'$. Moreover, $\mathscr{S}(\mathbb{R}^n)$ is dense in $H^s(\mathbb{R}^n)$, since

the space $\mathscr{S}(\mathbb{R}^n)$ is dense in the weighted L^2 space $\langle\xi\rangle^{-s}L^2(\mathbb{R}^n)$ (by Theorem 1.14) and the Fourier transform maps $\mathscr{S}(\mathbb{R}^n)$ onto itself. In fact, $C_c^\infty(\mathbb{R}^n)$ is dense in $H^s(\mathbb{R}^n)$ as well, since any Schwartz function can be approximated by elements of $C_c^\infty(\mathbb{R}^n)$ in H^s norm.

(5) The differential operator ∂_{x_j} on $\mathbb{R}^n$ restricts to a bounded operator

$$\partial_{x_j} : H^{s+1}(\mathbb{R}^n) \to H^s(\mathbb{R}^n). \tag{12.7}$$

Indeed, if $u \in H^{s+1}(\mathbb{R}^n)$, then by (11.36) we have

$$\|\partial_{x_j} u\|_{H^s(\mathbb{R}^n)} = \|\langle\xi\rangle^s \xi_j \widehat{u}(\xi)\|_{L^2(\mathbb{R}^n)} \le \|\langle\xi\rangle^{s+1}\widehat{u}(\xi)\|_{L^2(\mathbb{R}^n)} = \|u\|_{H^{s+1}(\mathbb{R}^n)}.$$

12.1.3. Characterization of Sobolev spaces. The definition (12.4) is convenient because it immediately works for all values of s. However, it is useful to have a characterization of Sobolev spaces which does not feature the Fourier transform. (We will in particular use it in the proof of Proposition 12.15 below.)

We start with a characterization of H^{s+1} in terms of H^s which generalizes Proposition 12.1:

PROPOSITION 12.4. *For any $s \in \mathbb{R}$ we have*

$$H^{s+1}(\mathbb{R}^n) = \{u \in H^s(\mathbb{R}^n) \colon \partial_{x_j} u \in H^s(\mathbb{R}^n),\ j = 1, \dots, n\} \tag{12.8}$$

with the corresponding norm equivalence: there exists a constant C such that for all $u \in H^{s+1}(\mathbb{R}^n)$

$$C^{-1}\|u\|_{H^{s+1}(\mathbb{R}^n)} \le \|u\|_{H^s(\mathbb{R}^n)} + \sum_{j=1}^n \|\partial_{x_j} u\|_{H^s(\mathbb{R}^n)} \le C\|u\|_{H^{s+1}(\mathbb{R}^n)}. \tag{12.9}$$

PROOF.**S** If $u \in H^{s+1}(\mathbb{R}^n)$, then $\partial_{x_j} u \in H^s(\mathbb{R}^n)$ by (12.7). On the other hand, if $u \in H^s(\mathbb{R}^n)$, then we estimate

$$\begin{aligned}\|u\|_{H^{s+1}(\mathbb{R}^n)} = \|\langle\xi\rangle^{s+1}\widehat{u}(\xi)\|_{L^2(\mathbb{R}^n)} &\le \|\langle\xi\rangle^s\widehat{u}(\xi)\|_{L^2(\mathbb{R}^n)} + \sum_{j=1}^n \|\langle\xi\rangle^s\xi_j\widehat{u}(\xi)\|_{L^2(\mathbb{R}^n)} \\ &= \|u\|_{H^s(\mathbb{R}^n)} + \sum_{j=1}^n \|\partial_{x_j} u\|_{H^s(\mathbb{R}^n)}.\end{aligned}$$

□

We next consider the special case $s \in (0, 1)$ and characterize H^s in terms of convergence of a double integral (reminding one of Hölder continuity but with sup-norm replaced by square-integral):

PROPOSITION 12.5. *Fix $s \in (0,1)$. Assume that $u \in L^2(\mathbb{R}^n)$. Define the integral*

$$I_s(u) := \int_{\mathbb{R}^{2n}} \frac{|u(x)-u(y)|^2}{|x-y|^{n+2s}}\,dxdy \;\in\; [0,\infty]. \tag{12.10}$$

Then $u \in H^s(\mathbb{R}^n)$ if and only if $I_s(u) < \infty$. Moreover, we have the norm equivalence

$$C^{-1}\|u\|_{H^s(\mathbb{R}^n)} \le \|u\|_{L^2(\mathbb{R}^n)} + \sqrt{I_s(u)} \le C\|u\|_{H^s(\mathbb{R}^n)}. \tag{12.11}$$

REMARK 12.6. *If $u \in L^2(\mathbb{R}^n)$ and $s > 0$, then by Fubini's Theorem we have for any $\varepsilon > 0$*

$$\begin{aligned}\int_{\{|x-y|\ge\varepsilon\}} \frac{|u(x)-u(y)|^2}{|x-y|^{n+2s}}\,dxdy &\le 4\int_{\{|x-y|\ge\varepsilon\}} \frac{|u(x)|^2}{|x-y|^{n+2s}}\,dxdy \\ &\le C_\varepsilon \|u\|^2_{L^2(\mathbb{R}^n)}.\end{aligned} \tag{12.12}$$

Therefore, the convergence of $I_s(u)$ is a question about the neighborhood of the diagonal $\{|x-y| \le \varepsilon\}$.

PROOF. 1. Making the change of variables $y = x + w$, we write

$$I_s(u) = \int_{\mathbb{R}^{2n}} \frac{|u(x+w)-u(x)|^2}{|w|^{n+2s}}\,dxdw.$$

We first compute the integral

$$\int_{\mathbb{R}^n} |u(x+w)-u(x)|^2\,dx = \|\tau_{-w}u - u\|^2_{L^2(\mathbb{R}^n)} \tag{12.13}$$

where $\tau_{-w}u(x) := u(x+w)$. By Exercise 11.7, we have

$$\mathcal{F}(\tau_{-w}u - u)(\xi) = (e^{iw\cdot\xi} - 1)\widehat{u}(\xi),$$

thus by Theorem 11.29 the integral (12.13) is equal to

$$(2\pi)^{-n}\int_{\mathbb{R}^n} |e^{iw\cdot\xi}-1|^2\,|\widehat{u}(\xi)|^2\,d\xi.$$

It follows that

$$I_s(u) = (2\pi)^{-n}\int_{\mathbb{R}^{2n}} \frac{|e^{iw\cdot\xi}-1|^2}{|w|^{n+2s}}|\widehat{u}(\xi)|^2\,dwd\xi. \tag{12.14}$$

2. We now integrate out w. For $\xi \in \mathbb{R}^n$, define

$$F(\xi) := (2\pi)^{-n}\int_{\mathbb{R}^n} \frac{|e^{iw\cdot\xi}-1|^2}{|w|^{n+2s}}\,dw, \tag{12.15}$$

so that

$$I_s(u) = \int_{\mathbb{R}^n} F(\xi)|\widehat{u}(\xi)|^2\,d\xi. \tag{12.16}$$

The integral (12.15) converges: on $\{|w| \ge 1\}$ this follows from the fact that $s > 0$ and on $\{|w| \le 1\}$ this follows from the bound $|e^{iw\cdot\xi} - 1| \le |\xi|\cdot|w|$ and the fact that $s < 1$.

Moreover, this integral only depends on $|\xi|$ (since $|w|$ is invariant under orthogonal transformations) and we have for all $t > 0$, making the change of variables $w = v/t$

$$\begin{aligned} F(t\xi) &= (2\pi)^{-n} \int_{\mathbb{R}^n} \frac{|e^{itw\cdot\xi} - 1|^2}{|w|^{n+2s}}\, dw \\ &= (2\pi)^{-n} t^{2s} \int_{\mathbb{R}^n} \frac{|e^{iv\cdot\xi} - 1|^2}{|v|^{n+2s}}\, dv = t^{2s} F(\xi). \end{aligned}$$

It follows that $F(\xi) = c|\xi|^{2s}$ for some constant $c > 0$. Thus, recalling (12.16), we have

$$I_s(u) = c \int_{\mathbb{R}^n} |\xi|^{2s} \cdot |\widehat{u}(\xi)|^2\, d\xi.$$

Since $\widehat{u} \in L^2(\mathbb{R}^n)$, we see that

$$I_s(u) < \infty \quad\iff\quad |\xi|^s \widehat{u}(\xi) \in L^2(\mathbb{R}^n) \quad\iff\quad \langle\xi\rangle^s \widehat{u}(\xi) \in L^2(\mathbb{R}^n) \quad\iff\quad u \in H^s(\mathbb{R}^n).$$

The bound (12.12) follows directly from the proof. □

Together Propositions 12.4 and 12.5 (and the fact that $H^0 = L^2$) characterize the spaces $H^s(\mathbb{R}^n)$ for $s \geq 0$. To handle the case $s < 0$, we use the following proposition, whose proof is left as an exercise below.

Proposition 12.7. *Fix $s \in \mathbb{R}$. Then the space $H^{-s}(\mathbb{R}^n)$ is dual to $H^s(\mathbb{R}^n)$ in the following sense:*

1. There exists a unique bilinear pairing

$$u \in H^s(\mathbb{R}^n),\ v \in H^{-s}(\mathbb{R}^n) \quad\mapsto\quad (u, v) \in \mathbb{C} \tag{12.17}$$

which coincides with the usual pairing (2.3) when $u, v \in \mathscr{S}(\mathbb{R}^n)$ and is continuous in the sense that whenever $u_k \to u$ in $H^s(\mathbb{R}^n)$ and $v_k \to v$ in $H^{-s}(\mathbb{R}^n)$ we have $(u_k, v_k) \to (u, v)$.

2. Let $v \in \mathscr{S}'(\mathbb{R}^n)$. Then $v \in H^{-s}(\mathbb{R}^n)$ if and only if there exists a constant C_v such that

$$|(v, \varphi)| \leq C_v \|\varphi\|_{H^s(\mathbb{R}^n)} \quad \text{for all } \varphi \in \mathscr{S}(\mathbb{R}^n). \tag{12.18}$$

Moreover, there exists a constant C depending only on n, s such that for each $v \in H^{-s}(\mathbb{R}^n)$ we have

$$C^{-1}\|v\|_{H^{-s}(\mathbb{R}^n)} \ \leq\ C_v \ \leq\ C\|v\|_{H^{-s}(\mathbb{R}^n)} \tag{12.19}$$

where C_v is the smallest constant such that the inequality (12.18) holds.

Remark 12.8. *Since $H^s(\mathbb{R}^n)$, Riesz Representation Theorem (Theorem 1.4) shows that any bounded linear functional $F : H^s(\mathbb{R}^n) \to \mathbb{C}$ has the form*

$$F(v) = \langle w, v\rangle_{H^s(\mathbb{R}^n)} \quad \text{for some } w \in H^s(\mathbb{R}^n). \tag{12.20}$$

On the other hand, Proposition 12.7 shows that

$$F(v) = (u, v) \quad \text{for some } u \in H^{-s}(\mathbb{R}^n). \tag{12.21}$$

There is no contradiction between (12.20) and (12.21) since the former features the inner product $\langle \bullet, \bullet \rangle_{H^s}$ and the latter uses the standard pairing $(\bullet, \bullet)$ which is related to the L^2 inner product. A more proper way to explain Proposition 12.7 is to say that it shows that H^{-s} is dual to H^s with respect to the L^2 pairing.

12.1.4. Multiplication by Schwartz functions. Since we defined Sobolev spaces using Fourier transform, it is not immediately clear that they are invariant under multiplication by smooth functions (except when the order is a nonnegative integer, where one can use Proposition 12.1 and the Leibniz rule). The next proposition shows that Sobolev spaces are invariant under multiplication by Schwartz functions a. Since we will typically use local Sobolev spaces, the restriction that a decays rapidly at infinity will not be too strong; in fact, we will typically the statement below for $a \in C_c^\infty(\mathbb{R}^n)$.

PROPOSITION 12.9. *Assume that $s \in \mathbb{R}$ and $a \in \mathscr{S}(\mathbb{R}^n)$. Then there exists a constant $C_{s,a}$ such that for each $u \in H^s(\mathbb{R}^n)$, the product au also lies in $H^s(\mathbb{R}^n)$ and*

$$\|au\|_{H^s(\mathbb{R}^n)} \le C_{s,a}\|u\|_{H^s(\mathbb{R}^n)}. \tag{12.22}$$

The proof will write the Fourier transform of au in terms of the convolution of the Fourier transforms of a and u. It will use the following

LEMMA 12.10 (Young's convolution inequality, special case). *Assume that $f \in L^2(\mathbb{R}^n)$ and $g \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, and define $f * g$ by (1.27). Then $f * g \in L^2(\mathbb{R}^n)$ and*

$$\|f * g\|_{L^2(\mathbb{R}^n)} \le \|f\|_{L^2(\mathbb{R}^n)}\|g\|_{L^1(\mathbb{R}^n)}. \tag{12.23}$$

REMARK 12.11. *The requirement that $g \in L^2(\mathbb{R}^n)$ is just to make the integral (1.27) converge at every point; it is not necessary but we do not want to do the extra work to remove it here.*

PROOF. Take any $\xi \in \mathbb{R}^n$. We estimate

$$\begin{aligned} |f * g(\xi)|^2 &= \left| \int_{\mathbb{R}^n} f(\xi - \eta) g(\eta)\, d\eta \right|^2 \\ &\le \left(\int_{\mathbb{R}^n} \Big(|f(\xi-\eta)| \cdot \sqrt{|g(\eta)|} \Big) \sqrt{|g(\eta)|}\, d\eta \right)^2 \\ &\le \left(\int_{\mathbb{R}^n} |f(\xi-\eta)|^2 |g(\eta)|\, d\eta \right) \left(\int_{\mathbb{R}^n} |g(\eta)|\, d\eta \right) \end{aligned}$$

Here in the last line we use the Cauchy–Schwartz inequality. Integrating in ξ we get

$$\begin{aligned} \|f * g\|^2_{L^2(\mathbb{R}^n)} &\le \left(\int_{\mathbb{R}^{2n}} |f(\xi-\eta)|^2 |g(\eta)|\, d\xi d\eta \right) \|g\|_{L^1(\mathbb{R}^n)} \\ &= \|f\|^2_{L^2(\mathbb{R}^n)} \|g\|^2_{L^1(\mathbb{R}^n)}. \end{aligned}$$

□

We now give

Proof of Proposition 12.9. 1. By (11.41) we have $\widehat{au} = (2\pi)^{-n}\,\widehat{u} * \widehat{a}$. Since $u \in H^s(\mathbb{R}^n)$, the Fourier transform $\widehat{u}$ is a function and, recalling (11.31), we have

$$(2\pi)^n \widehat{au}(\xi) = \int_{\mathbb{R}^n} \widehat{a}(\xi - \eta)\widehat{u}(\eta)\,d\eta.$$

Define the functions on $\mathbb{R}^n$

$$v(\eta) := \langle \eta \rangle^s \widehat{u}(\eta), \quad w(\xi) := (2\pi)^n \langle \xi \rangle^s \widehat{au}(\xi),$$

then

$$w(\xi) = \int_{\mathbb{R}^n} \frac{\langle \xi \rangle^s}{\langle \eta \rangle^s} \widehat{a}(\xi - \eta) v(\eta)\,d\eta \tag{12.24}$$

and we need to show that if $v \in L^2(\mathbb{R}^n)$ then $w \in L^2(\mathbb{R}^n)$ and

$$\|w\|_{L^2(\mathbb{R}^n)} \leq C_{s,a} \|v\|_{L^2(\mathbb{R}^n)}. \tag{12.25}$$

2. Since $a \in \mathscr{S}(\mathbb{R}^n)$, we have $\widehat{a} \in \mathscr{S}(\mathbb{R}^n)$. Looking at (12.24), we see that $\widehat{a}(\xi - \eta)$ is small unless $\xi - \eta$ is bounded, and when $\xi - \eta$ is bounded the ratio $\langle \xi \rangle^s / \langle \eta \rangle^s$ is bounded. This observation motivates the use of the following inequality (where C_s is a constant depending only on s):

$$\frac{\langle \xi \rangle^s}{\langle \eta \rangle^s} \leq C_s \langle \xi - \eta \rangle^{|s|}. \tag{12.26}$$

To show (12.26), we recall the definition (12.3), which implies the inequality

$$\langle \xi \rangle^2 = 1 + |\eta + (\xi - \eta)|^2 \leq 1 + 2|\eta|^2 + 2|\xi - \eta|^2 \leq 2\langle \eta \rangle^2 \langle \xi - \eta \rangle^2.$$

Switching the roles of ξ and η, we also get the inequality $\langle \eta \rangle^2 \leq 2\langle \xi \rangle^2 \langle \xi - \eta \rangle^2$. Taking these inequalities to the power $|s|$, we get (12.26).

Recalling (12.24) we see that

$$|w(\xi)| \leq C_s \int_{\mathbb{R}^n} \langle \xi - \eta \rangle^{|s|} |\widehat{a}(\xi - \eta)| \cdot |v(\eta)|\,d\eta. \tag{12.27}$$

The right-hand side of (12.27) is the convolution of $|v| \in L^2(\mathbb{R}^n)$ with the function $\langle \zeta \rangle^{|s|} |\widehat{a}(\zeta)| \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. By Lemma 12.10 we have

$$\|w\|_{L^2(\mathbb{R}^n)} \leq C_s \|\langle \zeta \rangle^{|s|} \widehat{a}(\zeta)\|_{L^1(\mathbb{R}^n)} \cdot \|v\|_{L^2(\mathbb{R}^n)} \tag{12.28}$$

giving (12.25) and finishing the proof. □

12.1.5. Further properties. We defined Sobolev spaces on the whole $\mathbb{R}^n$. One can localize these to obtain spaces of locally Sobolev distributions and compactly supported Sobolev distributions on any open subset of $\mathbb{R}^n$:

Definition 12.12. *Let $U \Subset \mathbb{R}^n$ and $s \in \mathbb{R}$. Define the spaces of* locally H^s distributions *and* compactly supported H^s distributions

$$H^s_{\mathrm{loc}}(U) \subset \mathcal{D}'(U), \qquad H^s_{\mathrm{c}}(U) \subset \mathcal{E}'(U)$$

as follows:

- *for $u \in \mathcal{E}'(U)$, we say that $u \in H^s_{\mathrm{c}}(U)$ if the extension of u by zero to an element of $\mathcal{E}'(\mathbb{R}^n)$ (see Proposition 4.7) lies in $H^s(\mathbb{R}^n)$;*
- *for $u \in \mathcal{D}'(U)$, we say that $u \in H^s_{\mathrm{loc}}(U)$ if for each $\chi \in C^\infty_{\mathrm{c}}(U)$ we have $\chi u \in H^s_{\mathrm{c}}(U)$.*

Note that Proposition 12.9 implies that

$$H^s_{\mathrm{c}}(U) = H^s_{\mathrm{loc}}(U) \cap \mathcal{E}'(U). \tag{12.29}$$

We define convergence of sequences in the newly introduced spaces as follows:

- we say that $u_k \to u$ in $H^s_{\mathrm{c}}(U)$ if there exists $K \Subset U$ such that $\operatorname{supp} u_k \subset K$ for all k, and $\|u_k - u\|_{H^s(\mathbb{R}^n)} \to 0$ where we identify $u_k - u \in \mathcal{E}'(U)$ with its extension by zero to the entire $\mathbb{R}^n$;
- we say that $u_k \to u$ in $H^s_{\mathrm{loc}}(U)$ if for each $\chi \in C^\infty_{\mathrm{c}}(U)$ we have $\|\chi(u_k - u)\|_{H^s(\mathbb{R}^n)} \to 0$, where we again identify $\chi(u_k - u) \in \mathcal{E}'(U)$ with its extension by zero to $\mathbb{R}^n$.

We list below some properties of the spaces H^s_{c}, H^s_{loc}. We leave the proof as an exercise below.

Proposition 12.13.[S] *Let $U \Subset \mathbb{R}^n$ and $s \in \mathbb{R}$. Then:*

1. *for any $a \in C^\infty(U)$, multiplication by a is a sequentially continuous operator $H^s_{\mathrm{c}}(U) \to H^s_{\mathrm{c}}(U)$ and $H^s_{\mathrm{loc}}(U) \to H^s_{\mathrm{loc}}(U)$;*
2. *for any $a \in C^\infty_{\mathrm{c}}(U)$, multiplication by a is a sequentially continuous operator $H^s_{\mathrm{loc}}(U) \to H^s_{\mathrm{c}}(U)$;*
3. *the differentiation operator ∂_{x_j} is sequentially continuous $H^{s+1}_{\mathrm{c}}(U) \to H^s_{\mathrm{c}}(U)$ and $H^{s+1}_{\mathrm{loc}}(U) \to H^s_{\mathrm{loc}}(U)$;*
4. *the space $C^\infty_{\mathrm{c}}(U)$ is dense in $H^s_{\mathrm{c}}(U)$ and in $H^s_{\mathrm{loc}}(U)$.*

Similarly to Proposition 12.7, the spaces $H^s_{\mathrm{loc}}(U)$ and $H^{-s}_{\mathrm{c}}(U)$ are dual to each other. We again leave the proof as an exercise below.

PROPOSITION 12.14.[S] *Let $U \Subset \mathbb{R}^n$ and $s \in \mathbb{R}$.*

1. There exists a unique sequentially continuous bilinear pairing

$$u \in H^s_{\mathrm{loc}}(U),\ v \in H^{-s}_{\mathrm{c}}(U) \quad \mapsto \quad (u, v) \in \mathbb{C} \tag{12.30}$$

which coincides with the usual pairing (2.3) when $u \in C^\infty(U)$, $v \in C^\infty_{\mathrm{c}}(U)$.

2. For $u \in \mathcal{D}'(U)$ we have $u \in H^s_{\mathrm{loc}}(U)$ if and only if $(u, \varphi_k) \to 0$ for any sequence $\varphi_k \in C^\infty_{\mathrm{c}}(U)$ converging to 0 in $H^{-s}_{\mathrm{c}}(U)$.

3. For $v \in \mathcal{E}'(U)$ we have $v \in H^{-s}_{\mathrm{c}}(U)$ if and only if $(v, \psi_k) \to 0$ for any sequence $\psi_k \in C^\infty(U)$ converging to 0 in $H^s_{\mathrm{loc}}(U)$.

A more difficult property to establish (but still left as an exercise below) is invariance of Sobolev spaces under pullback by diffeomorphisms:

PROPOSITION 12.15. *Assume that $U, V \Subset \mathbb{R}^n$ and $\Phi : U \to V$ is a C^∞ diffeomorphism. Fix $s \in \mathbb{R}$. Then the pullback operator Φ^* (defined on distributions in §10.1.3) is a sequentially continuous operator $H^s_{\mathrm{c}}(V) \to H^s_{\mathrm{c}}(U)$ and $H^s_{\mathrm{loc}}(V) \to H^s_{\mathrm{loc}}(U)$.*

We finish this section with one case of *Sobolev embedding*, which allows us to convert Sobolev regularity (at a loss in the number of derivatives) to classical C^k regularity. The proof is left as an exercise below.

THEOREM 12.16. *Assume that $s \in \mathbb{R}$, $k \in \mathbb{N}_0$, and $s > \frac{n}{2} + k$. Then for any $U \Subset \mathbb{R}^n$*

$$H^s_{\mathrm{loc}}(U) \ \subset \ C^k(U) \tag{12.31}$$

and if $u_j \to 0$ in $H^s_{\mathrm{loc}}(U)$ then $u_j \to 0$ in $C^k(U)$ (i.e. uniformly on compact subsets with k derivatives).

12.2. Elliptic regularity II

In this section we present the second version of elliptic regularity. For the first version, see §9.2 above. The conclusion is the same as for the first version, but the assumption is different, featuring the coefficients of the operator rather than requiring existence of a fundamental solution with a particular property.

12.2.1. Symbols of operators. We start by making the definitions needed to state the theorem. Let P be a constant coefficient differential operator of order $m \in \mathbb{N}_0$ on $\mathbb{R}^n$ (see Definition 9.1). We write it in the form

$$P = \sum_{|\alpha| \le m} a_\alpha D^\alpha_x, \quad D^\alpha_x = (-i)^{|\alpha|}\partial^\alpha_x \tag{12.32}$$

for some constants $a_\alpha \in \mathbb{C}$.

Definition 12.17. *Let P be given by (12.32). Define the* full symbol *of P as the following polynomial on $\mathbb{R}^n$:*

$$p(\xi) = \sum_{|\alpha|\leq m} a_\alpha \xi^\alpha. \tag{12.33}$$

Define the principal symbol *as consisting of order m terms in the full symbol:*

$$p_0(\xi) = \sum_{|\alpha|=m} a_\alpha \xi^\alpha. \tag{12.34}$$

We say that P is an elliptic *differential operator if*

$$p_0(\xi) \neq 0 \quad \text{for all } \xi \in \mathbb{R}^n \setminus \{0\}. \tag{12.35}$$

The presence of powers of i in (12.32) compared to (9.1) is convenient because of the relation to the Fourier transform. More precisely, from (11.36) we get that the Fourier transform conjugates P to multiplication by the full symbol p:

$$\widehat{Pu}(\xi) = p(\xi)\,\widehat{u}(\xi) \quad \text{for all } u \in \mathscr{S}'(\mathbb{R}^n). \tag{12.36}$$

As an example, we compute the full and principal symbols of the Laplace operator, the Cauchy–Riemann operator, the heat operator, and the wave operator (see §9.1.2):

$$P = \Delta \quad \Longrightarrow \quad p(\xi) = p_0(\xi) = -|\xi|^2, \tag{12.37}$$

$$P = \tfrac{1}{2}(\partial_x + i\partial_y) \quad \Longrightarrow \quad p(\xi,\eta) = p_0(\xi,\eta) = \tfrac{i}{2}(\xi + i\eta), \tag{12.38}$$

$$P = \partial_t - \Delta_x \quad \Longrightarrow \quad p(\tau,\xi) = i\tau + |\xi|^2, \quad p_0(\tau,\xi) = |\xi|^2, \tag{12.39}$$

$$P = \partial_t^2 - \Delta_x \quad \Longrightarrow \quad p(\tau,\xi) = p_0(\tau,\xi) = -\tau^2 + |\xi|^2. \tag{12.40}$$

12.2.2. Statement of elliptic regularity. We are now ready to state

Theorem 12.18 (Elliptic regularity II). *Assume that P is an order m constant coefficient differential operator on $\mathbb{R}^n$ which is elliptic in the sense of Definition 12.17. Then for any $U \Subset \mathbb{R}^n$ and $u \in \mathcal{D}'(U)$ we have*

$$\operatorname{sing\,supp} u = \operatorname{sing\,supp}(Pu). \tag{12.41}$$

Remark 12.19. *Looking at (12.37)–(12.40), we see that the Laplace operator Δ and the Cauchy–Riemann operator $\frac{1}{2}(\partial_x + i\partial_y)$ are elliptic, and the heat operator $\partial_t - \Delta_x$ and the wave operator $\partial_t^2 - \Delta_x$ are not elliptic. The ellipticity condition is sufficient but not necessary for (12.41) to hold, since the heat operator satisfies the assumptions of Elliptic Regularity I (Theorem 9.14).*

Following the proof of Theorem 12.18 below we obtain the following analog in Sobolev spaces. The proof is left as an exercise below.

THEOREM 12.20. *Under the assumptions of Theorem 12.18 we have for each $s \in \mathbb{R}$*

$$Pu \in H^s_{\mathrm{loc}}(U) \quad \Longrightarrow \quad u \in H^{s+m}_{\mathrm{loc}}(U) \tag{12.42}$$

where m is the order of the elliptic operator P.

As an example, if $\Delta u \in L^2_{\mathrm{loc}}(U)$ then $u \in H^2_{\mathrm{loc}}(U)$.

12.2.3. Kohn–Nirenberg symbols. The proof of Theorem 12.18 uses *Kohn–Nirenberg symbols*, which also play an important role in the proof of Elliptic Regularity III in §14 below. Here we introduce these symbols and study their basic properties.

DEFINITION 12.21. *Let $m \in \mathbb{R}$ and $a \in C^\infty(\mathbb{R}^n)$. We say that a is a Kohn–Nirenberg symbol of order m if for each multiindex α there exists a constant C_α such that*

$$|\partial_\xi^\alpha a(\xi)| \leq C_\alpha \langle \xi \rangle^{m-|\alpha|} \quad \text{for all } \xi \in \mathbb{R}^n. \tag{12.43}$$

Denote by $S^m(\mathbb{R}^n)$ the space of all Kohn–Nirenberg symbols of order m.

The condition (12.43) can be interpreted as follows: $a(\xi) = \mathcal{O}(\langle \xi \rangle^m)$ and each differentiation makes a one order smaller as $|\xi| \to \infty$.

From the definition and using the Leibniz rule one can check that

$$a \in S^m(\mathbb{R}^n),\ b \in S^\ell(\mathbb{R}^n) \quad \Longrightarrow \quad ab \in S^{m+\ell}(\mathbb{R}^n), \tag{12.44}$$

$$a \in S^m(\mathbb{R}^n) \quad \Longrightarrow \quad \partial_{x_j} a \in S^{m-1}(\mathbb{R}^n). \tag{12.45}$$

A fundamental example of a Kohn–Nirenberg symbol is the symbol of a differential operator:

PROPOSITION 12.22. *Assume that $m \in \mathbb{N}_0$ and $p(\xi)$ is a polynomial of degree m in ξ. Then $p \in S^m(\mathbb{R}^n)$.*

PROOF.[S] The derivative $\partial^\alpha p$ is a polynomial of degree $m - |\alpha|$ (and is equal to 0 if $|\alpha| > m$) which implies the bounds (12.43). □

We next prove two properties of the class S^m which will be used in the next subsection to prove Theorem 12.18. The first one is that inverses of elliptic Kohn–Nirenberg symbols are also Kohn–Nirenberg symbols:

PROPOSITION 12.23. *Assume that $p \in S^m(\mathbb{R}^n)$ and there exists a constant $c > 0$ such that*

$$|p(\xi)| \geq c\langle \xi \rangle^m \quad \text{for all } \xi \in \mathbb{R}^n. \tag{12.46}$$

Define $q := 1/p \in C^\infty(\mathbb{R}^n)$. Then $q \in S^{-m}(\mathbb{R}^n)$.

Proof. By induction on $|\alpha|$, we see that for any multiindex α the derivative $\partial_\xi^\alpha q$ is a linear combination with constant coefficients of expressions of the form

$$\frac{\partial_\xi^{\alpha_1} p(\xi)\cdots\partial_\xi^{\alpha_k} p(\xi)}{p(\xi)^{k+1}} \tag{12.47}$$

where $|\alpha_1|,\dots,|\alpha_k|\geq 1$ and $\alpha_1+\dots+\alpha_k=\alpha$. Since p satisfies the bounds (12.43) and (12.46), we see that (12.47) is

$$\mathcal{O}\left(\frac{\langle\xi\rangle^{m-|\alpha_1|}\dots\langle\xi\rangle^{m-|\alpha_k|}}{|p(\xi)|^{k+1}}\right)=\mathcal{O}(\langle\xi\rangle^{-m-|\alpha|})$$

which shows that $q\in S^{-m}(\mathbb{R}^n)$. □

Remark 12.24. *The same proof shows that if (12.46) holds for all $|\xi|\geq T$ and some fixed T, and $q\in C^\infty(\mathbb{R}^n)$ satisfies $p(\xi)q(\xi)=1$ for all $|\xi|\geq T$, then $q\in S^{-m}(\mathbb{R}^n)$.*

The second property concerns the Fourier transform of a symbol. Note that each $a\in S^m(\mathbb{R}^n)$ has polynomially bounded derivatives and in particular lies in $\mathscr{S}'(\mathbb{R}^n)$.

Proposition 12.25. *Assume that $a\in S^m(\mathbb{R}^n)$ for some m, and let $\widehat{a}\in\mathscr{S}'(\mathbb{R}^n)$ be the Fourier transform of a. Then*

$$\operatorname{sing\,supp}\widehat{a}\ \subset\ \{0\}. \tag{12.48}$$

Remark 12.26. *An example is when a is a polynomial, then $\widehat{a}$ is a derivative of the delta function δ_0 by (11.35) and (11.37).*

Proof. 1. For any multiindex α, we have by (11.36)

$$x^\alpha\widehat{a}(x)=\widehat{D_\xi^\alpha a}(x). \tag{12.49}$$

By (12.43) we have $D_\xi^\alpha a(\xi)=\mathcal{O}(\langle\xi\rangle^{m-|\alpha|})$. Thus

$$D_\xi^\alpha a\in L^1(\mathbb{R}^n)\quad\text{when } |\alpha|>m+n.$$

By Proposition 11.2 we see that

$$x^\alpha\widehat{a}(x)\in C^0(\mathbb{R}^n)\quad\text{when } |\alpha|>m+n.$$

In particular, if we take $N\in\mathbb{N}_0$ large enough so that $2N>m+n$ then $|x|^{2N}\widehat{a}(x)\in C^0(\mathbb{R}^n)$ which implies that

$$\widehat{a}|_{\mathbb{R}^n\setminus\{0\}}\in C^0(\mathbb{R}^n\setminus\{0\}).$$

2. A modification of the above argument shows that a is in C^∞ away from the origin. Namely, fix $k\in\mathbb{N}_0$ and choose $N\in\mathbb{N}_0$ large enough so that $2N>k+m+n$. Then for each α with $|\alpha|=2N$ we have

$$D_\xi^\alpha a\in\langle\xi\rangle^{-k}L^1(\mathbb{R}^n).$$

By (11.14) and (12.49) we see that

$$x^\alpha \widehat{a}(x) \in C^k(\mathbb{R}^n).$$

This shows that $|x|^{2N}\widehat{a}(x) \in C^k(\mathbb{R}^n)$, implying that $a|_{\mathbb{R}^n\setminus\{0\}}$ lies in C^k. Since this is true for all k, we see that $a|_{\mathbb{R}^n\setminus\{0\}} \in C^\infty(\mathbb{R}^n \setminus \{0\})$. Thus $\operatorname{sing\,supp} \widehat{a} \subset \{0\}$. □

12.2.4. Proof of elliptic regularity. We now prove Theorem 12.18.

1. We first construct an *elliptic parametrix*, which is a distribution $E \in \mathcal{D}'(\mathbb{R}^n)$ such that

$$R := \delta_0 - PE \in C^\infty(\mathbb{R}^n), \quad \operatorname{sing\,supp} E \subset \{0\}. \tag{12.50}$$

We can think of E as a fundamental solution of P modulo smooth functions: instead of $PE = \delta_0$ we require that $\delta_0 - PE$ be smooth.

Let p be the full symbol of P and p_0 be its principal symbol (see Definition 12.17). Since P is elliptic, the restriction of p_0 to the unit sphere $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ is a nonvanishing continuous function, so there exists a constant $c > 0$ such that $|p_0(\xi)| \geq c$ for all $\xi \in \mathbb{S}^{n-1}$. Since the function p_0 is homogeneous of degree m, we have

$$|p_0(\xi)| \geq c|\xi|^m \quad \text{for all } \xi \in \mathbb{R}^n. \tag{12.51}$$

The difference $p - p_0$ is a polynomial of degree $m-1$, so $p(\xi) = p_0(\xi) + \mathcal{O}(\langle\xi\rangle^{m-1})$. Therefore there exists $T > 0$ such that

$$|p(\xi)| \geq \tfrac{c}{2}\langle\xi\rangle^m \quad \text{for all } \xi,\ |\xi| \geq T.$$

Fix a function

$$q \in C^\infty(\mathbb{R}^n), \quad q(\xi) = \frac{1}{p(\xi)} \quad \text{for } |\xi| \geq T. \tag{12.52}$$

For example, we can put $q := (1-\chi)/p$ where $\chi \in C^\infty_{\mathrm{c}}(B^\circ(0,T))$ satisfies $\chi = 1$ near $p^{-1}(0)$.

By Proposition 12.23 and Remark 12.24 the function q is a Kohn–Nirenberg symbol:

$$q \in S^{-m}(\mathbb{R}^n).$$

We now define the distribution $E \in \mathscr{S}'(\mathbb{R}^n)$ as the inverse Fourier transform of q:

$$E := \mathcal{F}^{-1}(q), \qquad \widehat{E} = q. \tag{12.53}$$

By Proposition 12.48 (which applies to the inverse Fourier transform since $E(x) = (2\pi)^{-n}\widehat{q}(-x)$) we have $\operatorname{sing\,supp} E \subset \{0\}$.

It remains to show that $R := \delta_0 - PE \in C^\infty(\mathbb{R}^n)$. By (12.36) and (11.34) we compute the Fourier transform

$$\widehat{R}(\xi) = 1 - p(\xi)q(\xi).$$

Recalling (12.52) we see that $\widehat{R}$ is a smooth compactly supported function, and thus in particular in $\mathscr{S}(\mathbb{R}^n)$. Since the inverse Fourier transform maps $\mathscr{S}(\mathbb{R}^n)$ to itself, we see that $R \in \mathscr{S}(\mathbb{R}^n)$, so in particular it lies in $C^\infty(\mathbb{R}^n)$.

2. We now argue similarly to the proof of Theorem 9.14. Fix arbitrary $x_0 \in U \setminus \operatorname{sing\,supp} Pu$ and take a cutoff function

$$\chi \in C_c^\infty(U), \quad x_0 \notin \operatorname{supp}(1-\chi).$$

Treating χu as an element of $\mathcal{E}'(\mathbb{R}^n)$ and using (9.7) we see that

$$\chi u = \delta_0 * (\chi u) = (PE + R) * (\chi u) = E * (P\chi u) + R * (\chi u).$$

Since $\operatorname{sing\,supp} E \subset \{0\}$, by Proposition 8.14 we have $\operatorname{sing\,supp} E*(P\chi u) \subset \operatorname{sing\,supp}(P\chi u)$. Since $R \in C^\infty(\mathbb{R}^n)$, by Theorem 6.4 we have $R * (\chi u) \in C^\infty(\mathbb{R}^n)$. Therefore

$$\operatorname{sing\,supp}(\chi u) \subset \operatorname{sing\,supp}(P\chi u).$$

Arguing as in the proof of Theorem 9.14 we see that $x_0 \notin \operatorname{sing\,supp}(P\chi u)$, thus $x_0 \notin \operatorname{sing\,supp}(\chi u)$ which implies that $x_0 \notin \operatorname{sing\,supp} u$. Since x_0 was arbitrary this shows that $\operatorname{sing\,supp} u \subset \operatorname{sing\,supp} Pu$ and finishes the proof.

12.3. Notes and exercises

Our presentation mostly follows [**Hör03**, §7.9] and [**FJ98**, §§8.6,9.3]. The book of Sobolev [**Sob91**], first published in 1950, is a nice introduction to Sobolev spaces and their applications for anyone interested in the history of their development before Schwartz.

A simple explanations for how Sobolev spaces appear in the study of hyperbolic equations is as follows: if u solves the wave equation $(\partial_t^2 - \Delta_x)u = 0$, then the energy

$$E_u(t) = \frac{1}{2}\int_{\mathbb{R}^n} |\partial_t u(t,x)|^2 + |\partial_x u(t,x)|^2\,dx$$

is a conserved quantity. However, this energy controls the H^1 norm of u, rather than the C^2 norm which would be needed to make sense of u as a classical solution.

The theory of Sobolev spaces extends considerably past what is presented here. In particular, one can define Sobolev spaces based on L^p rather than L^2, as well as Sobolev spaces on domains with boundary. The latter are important in solving boundary value problems for elliptic equations (such as the Poisson equation $\Delta u = f$) and the Hilbert theory of these spaces underlies the finite element method of solving such equations numerically. See [**Tay11a, Eva10**] for more information.

EXERCISE 12.1. $(1 = 0.5 + 0.5$ pt) *For the distributions below, find out for which s they lie in $H^s(\mathbb{R}^n)$:*

(a) δ_0;

(b) *the indicator function of the some interval $[a, b] \subset \mathbb{R}$ (here $n = 1$).*

EXERCISE 12.2. (0.5 pt) *Let $u \in \mathcal{E}'(\mathbb{R}^n)$. Show that there exists $s \in \mathbb{R}$ such that $u \in H^s(\mathbb{R}^n)$. (Hint: use Proposition 11.26.)*

EXERCISE 12.3. (1.5 = 0.5 + 1 pts) *Prove Proposition 12.7. (Hint: use* (11.49) *and Exercise 2.1.)*

EXERCISE 12.4. (1 pt) *Prove Proposition 12.13. (Hint: to show density of $C^\infty_c(U)$ in $H^s_{\text{loc}}(U)$, take arbitrary $u \in H^s_{\text{loc}}(U)$, consider a sequence of functions $\chi_k \in C^\infty_c(U)$ defined in* (4.4) *and take $\varphi_k \in C^\infty_c(U)$ such that $\|\chi_k u - \varphi_k\|_{H^s(\mathbb{R}^n)} \leq 1/k$. Then show that $\varphi_k \to u$ in $H^s_{\text{loc}}(U)$.)*

EXERCISE 12.5. (1 pt) *Prove Proposition 12.14.*

EXERCISE 12.6. (2.5 pts) *Prove Proposition 12.15. (Hint: use the results of §12.1.3, considering first the case $s = 0$, then $0 < s < 1$, then using these to treat the case of general $s \geq 0$, and finally using duality to treat the case $s < 0$.)*

EXERCISE 12.7. (1 pt) *Prove Theorem 12.16. (Hint: use* (11.14)*.)*

EXERCISE 12.8. (1 = 0.5 + 0.5 pts) *This exercise extends the previous one by comparing Sobolev spaces with Hölder spaces. Assume that $0 < \gamma < 1$. Define the Hölder space $C^\gamma(\mathbb{R}^n) \subset C^0(\mathbb{R}^n)$ consisting of all functions f such that for each $K \Subset \mathbb{R}^n$ there exists a constant C_K such that for all $x, y \in K$ we have $|f(x) - f(y)| \leq C_K|x-y|^\gamma$. Denote by $C^\gamma_c(\mathbb{R}^n)$ the set of compactly supported functions in $C^\gamma(\mathbb{R}^n)$.*

(a) *Show that $C^\gamma_c(\mathbb{R}^n) \subset H^s(\mathbb{R}^n)$ for each $s < \gamma$. (Hint: use Proposition 12.5.)*

(b) *Show that $H^s(\mathbb{R}^n) \subset C^\gamma(\mathbb{R}^n)$ for each $s > \gamma + \frac{n}{2}$. (Hint: write each $u \in H^s(\mathbb{R}^n)$ in terms of $\widehat{u}$ using the Fourier inversion formula, and use the inequality $|e^{ix\cdot\xi} - e^{iy\cdot\xi}| = |e^{i(x-y)\cdot\xi} - 1| \leq C_\gamma|x-y|^\gamma|\xi|^\gamma$.)*

EXERCISE 12.9. (1.5 pts) *This exercise forms the basis for the theorem about restricting elements of Sobolev spaces to hypersurfaces, which is important for the study of boundary value problems. We write elements of $\mathbb{R}^n$ as (x_1, x') where $x' \in \mathbb{R}^{n-1}$, and consider the restriction operator to $\{x_1 = 0\}$,*

$$T : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^{n-1}), \quad T\varphi(x') = \varphi(0, x').$$

Show that when $s > \frac{1}{2}$, there exists a constant C such that we have the bound

$$\|T\varphi\|_{H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})} \leq C\|\varphi\|_{H^s(\mathbb{R}^n)} \quad \text{for all} \quad \varphi \in \mathscr{S}(\mathbb{R}^n).$$

Thus by Continuous Linear Extension T extends to a bounded operator $H^s(\mathbb{R}^n) \to H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})$. (Hint: use Fourier Inversion Formula to write the Fourier transform of $T\varphi$ in terms of the integral of $\widehat{\varphi}$ in the ξ_1 variable. Next, if $v \in L^2(\mathbb{R}^n)$, then we can use Cauchy–Schwartz to estimate $\int_{\mathbb{R}} \langle\xi\rangle^{-s} v(\xi_1, \xi')\, d\xi_1$ in terms of the L^2 norms of the *functions* $\xi_1 \mapsto (1+|\xi_1|^2+|\xi'|^2)^{-s/2}$ *and* $\xi_1 \mapsto v(\xi_1,\xi')$*. It remains to show that the first of these norms is bounded by* $C\langle\xi'\rangle^{\frac{1}{2}-s}$*.)*

EXERCISE 12.10. (1 pt) *Prove Theorem 12.20. (Hint: show first that if* E *is defined in* (12.53)*, then for any* $v \in H^s_{\mathrm{c}}(\mathbb{R}^n)$ *we have* $E * v \in H^{s+m}(\mathbb{R}^n)$*. You might want to choose arbitrary* $\psi \in C^\infty_{\mathrm{c}}(U)$ *and show that* $\psi u \in H^{s+m}$ *by taking* $\chi \in C^\infty_{\mathrm{c}}(U)$ *in the proof of Theorem 12.18 such that* $\operatorname{supp}(1-\chi)\cap\operatorname{supp}\psi = \emptyset$*. You can freely use anything in the proof of Theorem 12.18.)*

CHAPTER 13

Manifolds and differential operators

In this chapter we discuss manifolds, distributions on manifolds, and differential operators. One of the advantages of manifolds for us is the existence of compact manifolds, which are the setting of several of the most interesting applications of the material of this course (see Theorems 15.13, 16.1, 17.15 below).

A lot of definitions and proofs can be transferred from open subsets of $\mathbb{R}^n$ to a manifold (often via pushforwards by charts), and we try to give the list of statements that are true and the new details of the proofs compared to the case of open subsets of $\mathbb{R}^n$, but skip the more technical details which can hopefully be worked out by a dedicated reader and would potentially add many more pages to this chapter without making it any easier to read.

13.1. Manifolds$^{\mathbf{R}}$

In this section we briefly review some concepts from the theory of smooth manifolds. We skip a lot of definitions and almost all the proofs, referring the reader to [**Lee13**, Chapters 1–3,10,11,13,16] for details. For a more gentle introduction to some of the topics below, see alternatively [**Spi65**]. On the other end of the spectrum, [**Hör03**, §§6.3–6.4] provides a very fast introduction to the theory of manifolds.

13.1.1. Basics. A *manifold* is informally thought of as a space which is locally diffeomorphic to $\mathbb{R}^n$. More precisely, for us an n-dimensional *manifold* is:

- a Hausdorff topological space $\mathcal{M}$ which is second countable (there exists a countable basis of the topology of $\mathcal{M}$),
- and a collection of homeomorphisms $U \to V$ where $U \Subset \mathcal{M}$, $V \Subset \mathbb{R}^n$, which we call *charts*,

so that the following properties hold:

- the domains U of the given charts cover the entire $\mathcal{M}$;
- if $\varkappa : U \to V$ is a chart, then for any nonempty $W \Subset U$ the restriction $\varkappa|_W : W \to \varkappa(W)$ is also a chart;
- if $\varkappa_1 : U \to V_1$ and $\varkappa_2 : U \to V_2$ are two charts, then the *transition map* $\varkappa_2 \circ \varkappa_1^{-1}$ is a diffeomorphism $V_1 \to V_2$.

Remark 13.1.**X** *One can show that every manifold has metrizable topology. The second countability and Hausdorff property above are for correctness sake, we will not be using them directly in these notes.*

We denote the manifold above by just $\mathcal{M}$, suppressing the smooth structure (i.e. the choice of the collection of charts on $\mathcal{M}$) in the notation. It is actually better to define the manifold as having a complete atlas, *which is a collection of charts which includes any chart compatible with all the charts in it (in the sense of smoothness of transition maps).*

We think of charts as local coordinate systems: if $\varkappa : U \to V$ is a chart, then $\varkappa(x)$ is the coordinate vector of a point $x \in U$. The inverse $\varkappa^{-1} : V \to U$ is called the *parametrization map.*

13.1.2. Examples. A fundamental example of an n-dimensional manifold is $\mathbb{R}^n$ itself, with charts given e.g. by identity maps $I : U \to U$ for all nonempty $U \subset \mathbb{R}^n$. An open subset of a manifold is a manifold itself, so any open subset of $\mathbb{R}^n$ is an n-dimensional manifold. (We purposely avoid the question about whether the empty set is a manifold.)

A more nontrivial example is given by

Proposition 13.2. *Assume that $\mathcal{U} \Subset \mathbb{R}^N$ and $F : \mathcal{U} \to \mathbb{R}^m$ is a C^∞ map, with $N \geq m$. Fix $y_0 \in \mathbb{R}^m$ and assume that for each $x \in F^{-1}(y_0)$, the differential $dF(x)$ is a surjective linear map. Then $F^{-1}(y_0)$ is an $N - m$ dimensional manifold.*

The proof of Proposition 13.2 uses the Inverse Mapping Theorem. Arguing the same way as in the proof of Lemma 10.8 we see that for each $x_0 \in F^{-1}(y_0)$ we can find a local system of coordinates y on $\mathbb{R}^N$ near x_0 in which $F(y', y'') = y'$ where $y' \in \mathbb{R}^m, y'' \in \mathbb{R}^{N-m}$. Then $F^{-1}(y_0)$ is an open subset of the $N - m$ dimensional affine subspace $\{y' = y_0\} \subset \mathbb{R}^N$ and a chart on $F^{-1}(y_0)$ near x_0 is given by the map $x \mapsto y''$.

An important example of a manifold constructed this way is the sphere

$$\mathbb{S}^n := \{x \in \mathbb{R}^{n+1} \colon |x| = 1\}. \tag{13.1}$$

Another commonly used manifold is the torus

$$\mathbb{T}^n := \underbrace{\mathbb{S}^1 \times \cdots \times \mathbb{S}^1}_{n \text{ times}} \tag{13.2}$$

which is also often thought of as the quotient $\mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n$.

13.1.3. Functions and maps. For a manifold $\mathcal{M}$ and a function $f : \mathcal{M} \to \mathbb{C}$, we say that f is *smooth*, and write

$$f \in C^\infty(\mathcal{M})$$

if for each chart $\varkappa : U \to V$, the *pushforward* of f by $\varkappa$,

$$\varkappa_* f := f \circ \varkappa^{-1} \ : \ V \to \mathbb{C} \tag{13.3}$$

lies in $C^\infty(V)$. The pushforward here is just the pullback by $\varkappa^{-1}$, but it will be notationally convenient for us to write $\varkappa_*$ rather than $(\varkappa^{-1})^*$.

We define $C_c^\infty(\mathcal{M})$ to be the space of compactly supported functions in $C^\infty(\mathcal{M})$, with support of f defined to be the closure of $\{x \in \mathcal{M} \mid f(x) \neq 0\}$. If $\mathcal{N} \Subset \mathcal{M}$, then we have the natural restriction operator $C^\infty(\mathcal{M}) \to C^\infty(\mathcal{N})$ and the extension by zero operator $C_c^\infty(\mathcal{N}) \to C_c^\infty(\mathcal{M})$. The partition of unity Theorem 1.15 still applies to manifolds.

The notions of convergence in C_c^∞ (from Definition 2.5) and in C^∞ (from Definition 4.4) make sense on a manifold:

DEFINITION 13.3. *We say that a sequence $u_k \in C^\infty(\mathcal{M})$ converges to u in $C^\infty(\mathcal{M})$ if for each chart $\varkappa : U \to V$ we have $\varkappa_*(u_k - u) \to 0$ in $C^\infty(V)$ in the sense of Definition 4.4. A sequence $u_k \in C_c^\infty(\mathcal{M})$ converges to u in $C_c^\infty(\mathcal{M})$ if $u_k \to u$ in $C^\infty(\mathcal{M})$ and there exists $K \Subset \mathcal{M}$ such that $\operatorname{supp} u_k \subset K$ for all k.*

Similarly to the spaces C^∞ and C_c^∞, we can define the spaces $L^p_{\mathrm{loc}}(\mathcal{M})$, $L^p_c(\mathcal{M})$. The key observation, just like with smooth functions, is that pullbacks by diffeomorphisms preserve the spaces L^p_{loc} on open subsets of $\mathbb{R}^n$, so it does not matter what chart $\varkappa : U \to V$ we choose to determine whether $f \in L^p_{\mathrm{loc}}(U)$ for $U \Subset \mathcal{M}$.

More generally one can define smooth maps between two manifolds, $\Phi : \mathcal{M} \to \mathcal{N}$. Such a smooth map is called a *diffeomorphism* if the inverse Φ^{-1} is also a smooth map.

13.1.4. Tangent bundle and vector fields. If $\mathcal{M}$ is an n-dimensional manifold and $x \in \mathcal{M}$, then the *tangent space* $T_x\mathcal{M}$ is an n-dimensional (real) vector space. Elements of $T_x\mathcal{M}$ are called *tangent vectors* to $\mathcal{M}$ at x. There are several ways to define it (derivations at x on the space of smooth functions, or equivalence classes of paths through x) but neither is particularly fast to describe so we will just have to refer to [**Lee13**, Chapter 3] for a proper definition.

In the example $\mathcal{M} = F^{-1}(y_0)$ given by Proposition 13.2, the tangent space at $x \in \mathcal{M}$ is the following $N - m$-dimensional subspace of $\mathbb{R}^N$:

$$T_x F^{-1}(y_0) = \{v \in \mathbb{R}^N \mid dF(x)v = 0\}. \tag{13.4}$$

If $\Phi : \mathcal{M} \to \mathcal{N}$ is a smooth map between two manifolds, then its differential is a linear map of tangent spaces:

$$d\Phi(x) : T_x\mathcal{M} \to T_{\Phi(x)}\mathcal{N}, \quad x \in \mathcal{M},$$

and we have a version of the Chain Rule: if $\mathcal{M} \xrightarrow{\Phi_1} \mathcal{N} \xrightarrow{\Phi_2} \mathcal{L}$ are smooth maps then

$$d(\Phi_2 \circ \Phi_1)(x) = d\Phi_2(\Phi_1(x))d\Phi_1(x), \quad x \in \mathcal{M}.$$

If $V \Subset \mathbb{R}^n$, then the tangent space to V at each point is just $\mathbb{R}^n$. Thus if $\varkappa : U \to V$, $U \Subset \mathcal{M}$, is a chart, then we have the linear isomorphisms

$$d\varkappa(x) : T_x\mathcal{M} \to \mathbb{R}^n, \quad x \in \mathcal{U}. \tag{13.5}$$

The *tangent bundle* of $\mathcal{M}$ is the set of all tangent vectors:

$$T\mathcal{M} := \{(x, v) \mid x \in \mathcal{M},\ v \in T_x\mathcal{M}\} \tag{13.6}$$

and it is a $2n$-dimensional smooth manifold. More precisely, any chart $\varkappa : U \to V$ on $\mathcal{M}$ induces the following chart on $T\mathcal{M}$:

$$(x, v) \mapsto (\varkappa(x), d\varkappa(x)v) \in \mathbb{R}^{2n}, \quad x \in U,\ v \in T_x\mathcal{M}.$$

If $\mathcal{M} = F^{-1}(y_0)$ is the example of Proposition 13.2, then the tangent bundle of $\mathcal{M}$ is

$$T\mathcal{M} = \{(x, v) \in \mathcal{U} \times \mathbb{R}^N \mid F(x) = y_0,\ dF(x)v = 0\}$$

which is again a manifold of the type given by Proposition 13.2. In particular, if $\mathcal{M}$ is the sphere defined in (13.1) then

$$T\mathbb{S}^n = \{(x, v) \in \mathbb{R}^{2n+2} \colon |x| = 1,\ x \cdot v = 0\},$$

so for example

$$T\mathbb{S}^2 = \{(x_1, x_2, x_3, v_1, v_2, v_3) \in \mathbb{R}^6 \colon x_1^2 + x_2^2 + x_3^2 = 1,\ x_1v_1 + x_2v_2 + x_3v_3 = 0\}.$$

A C^∞ *vector field* on $\mathcal{M}$ is a map

$$X : x \in \mathcal{M} \ \mapsto\ X(x) \in T_x\mathcal{M}$$

so that the map $x \in \mathcal{M} \mapsto (x, X(x)) \in T\mathcal{M}$ is smooth. Denote by $C^\infty(\mathcal{M}; T\mathcal{M})$ the space of all C^∞ vector fields on $\mathcal{M}$.

If $\varkappa : U \to V$ is a chart on $\mathcal{M}$, then the *pushforward* of a vector field $X \in C^\infty(\mathcal{M}; T\mathcal{M})$ by $\varkappa$ is the vector field $\varkappa_*X$ on V defined by

$$\varkappa_*X(\varkappa(x)) = d\varkappa(x)X(x), \quad x \in U. \tag{13.7}$$

A vector field on $V \Subset \mathbb{R}^n$ is just a smooth map $V \to \mathbb{R}^n$, so we can write

$$\varkappa_*X = \sum_{j=1}^{n} X_j(x)\partial_{x_j} \tag{13.8}$$

where $\partial_{x_1}, \dots, \partial_{x_n}$ denotes the canonical basis of $\mathbb{R}^n$ and $X_j \in C^\infty(V; \mathbb{R})$.

REMARK 13.4. *Note that we should have used a different letter instead of x in (13.8); indeed, x in (13.7) is a point in $U \Subset \mathcal{M}$ and in (13.8) it is a point in $V \Subset \mathbb{R}^n$. We denote both by the same letter, which is a common abuse of notation in differential geometry. In fact, we often suppress the pushforward $\varkappa_*$ in the notation and just say that* in the chart $\varkappa$, we have $X = \sum_{j=1}^n X_j(x)\partial_{x_j}$. *This takes some time to get used to, but it saves a lot of time and ink later. Same applies to forms, densities, and Riemannian metrics studied below.*

A vector field $X \in C^\infty(\mathcal{M}; T\mathcal{M})$ defines an operator $X : C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$ as follows:

$$Xf(x) = df(x)X(x) \quad \text{for all } f \in C^\infty(\mathcal{M}),\ x \in \mathcal{M}. \tag{13.9}$$

If $\varkappa : U \to V$ is a chart in which $X = \sum_{j=1}^n X_j(x)\partial_{x_j}$ for some $X_j \in C^\infty(V)$ then in this chart

$$Xf(x) = \sum_{j=1}^n X_j(x)\partial_{x_j} f(x). \tag{13.10}$$

Here we push forward both Xf and f to V by $\varkappa$, so strictly speaking (13.10) should be stated as

$$\varkappa_* X = \sum_{j=1}^n X_j(x)\,\partial_{x_j} \quad \Longrightarrow \quad \varkappa_*(Xf)(x) = \sum_{j=1}^n X_j(x)\partial_{x_j}(\varkappa_* f)(x). \tag{13.11}$$

13.1.5. Cotangent bundle and 1-forms. If $\mathcal{M}$ is a manifold and $x \in \mathcal{M}$, the *cotangent space* $T_x^*\mathcal{M}$ is the dual of the tangent space, that is the space of linear maps $T_x\mathcal{M} \to \mathbb{R}$. Similarly to the tangent spaces, one can put cotangent spaces together to form the *cotangent bundle*

$$T^*\mathcal{M} = \{(x, \xi) \mid x \in \mathcal{M},\ \xi \in T_x^*\mathcal{M}\}. \tag{13.12}$$

If $\Phi : \mathcal{M} \to \mathcal{N}$ is a smooth map, then we can define the transpose differential

$$d\Phi(x)^T : T^*_{\Phi(x)}\mathcal{N} \to T^*_x\mathcal{M}, \quad x \in \mathcal{M} \tag{13.13}$$

by the formula

$$d\Phi(x)^T\eta(v) = \eta(d\Phi(x)v) \quad \text{for all } \eta \in T^*_{\Phi(x)}\mathcal{N},\ v \in T_x\mathcal{M}.$$

If Φ is a diffeomorphism, then we can define the inverse-transpose

$$d\Phi(x)^{-T} := (d\Phi(x)^T)^{-1} : T^*_x\mathcal{M} \to T^*_{\Phi(x)}\mathcal{N}. \tag{13.14}$$

Similarly to vector fields, we define *1-forms* on $\mathcal{M}$ as maps

$$\omega : x \in \mathcal{M} \ \mapsto \ \omega(x) \in T^*_x\mathcal{M}$$

such that the map $x \in \mathcal{M} \mapsto (x, \omega(x)) \in T^*\mathcal{M}$ is smooth. Denote by $C^\infty(\mathcal{M}; T^*\mathcal{M})$ the space of all 1-forms on $\mathcal{M}$.

If $\varkappa : U \to V$ is a chart, we define the pushforward of a 1-form $\omega \in C^\infty(\mathcal{M}; T^*\mathcal{M})$ by $\varkappa$ to be the following 1-form on V:

$$\varkappa_*\omega(\varkappa(x)) = d\varkappa(x)^{-T}\omega(x), \quad x \in U. \tag{13.15}$$

A 1-form on $V \Subset \mathbb{R}^n$ is the same as a smooth map ω from V to the dual of $\mathbb{R}^n$, which is canonically identified with $\mathbb{R}^n$, so we can write

$$\varkappa_*\omega = \sum_{j=1}^n \omega_j(x)\, dx_j \tag{13.16}$$

where $dx_j : \mathbb{R}^n \to \mathbb{R}$ is the j-th coordinate map on $\mathbb{R}^n$ and $\omega_j \in C^\infty(V; \mathbb{R})$.

If $f : \mathcal{M} \to \mathbb{R}$ is a smooth function, then the differential df is naturally a 1-form, since for each $x \in \mathcal{M}$, $df(x)$ is a linear map from $T_x\mathcal{M}$ to $\mathbb{R}$. In any chart $\varkappa : U \to V$ this 1-form is given by

$$df = \sum_{j=1}^n (\partial_{x_j} f)\, dx_j. \tag{13.17}$$

For a vector field $X \in C^\infty(\mathcal{M}; T\mathcal{M})$ and a 1-form $\omega \in C^\infty(\mathcal{M}; T^*\mathcal{M})$, we can define the pairing $\omega(X) \in C^\infty(\mathcal{M}; \mathbb{R})$ by

$$(\omega(X))(x) = \omega(x)(X(x)), \quad x \in \mathcal{M},$$

so that $Xf = df(X)$.

13.1.6. Riemannian metrics. A *Riemannian metric* g on a smooth manifold $\mathcal{M}$ is a smooth choice of a (positive definite) inner product on tangent spaces to $\mathcal{M}$. That is, for each $x \in \mathcal{M}$, $g(x)$ is an inner product on $T_x\mathcal{M}$ (we also denote this inner product by $\langle \bullet, \bullet \rangle_{g(x)}$) and the norm-squared

$$|v|^2_{g(x)} := \langle v, v \rangle_{g(x)}, \quad (x, v) \in T\mathcal{M}$$

is a smooth function on $T\mathcal{M}$.

A Riemannian metric g defines also an inner product on cotangent spaces, so that for each $(x, \xi) \in T^*\mathcal{M}$ the corresponding norm is

$$|\xi|_{g(x)} = \max\{\xi(v) \colon v \in T_x\mathcal{M},\ |v|_{g(x)} = 1\}.$$

If $\Phi : \mathcal{M} \to \mathcal{N}$ is a diffeomorphism, and h is a metric on $\mathcal{N}$, then the pullback metric Φ^*h on $\mathcal{M}$ is defined by

$$\langle v, w \rangle_{\Phi^*h(x)} = \langle d\Phi(x)v, d\Phi(x)w \rangle_{h(\Phi(x))}, \quad x \in \mathcal{M},\ v, w \in T_x\mathcal{M}.$$

If g is a metric on $\mathcal{M}$, we say that $\Phi : (\mathcal{M}, g) \to (\mathcal{N}, h)$ is an *isometry* if $\Phi^*h = g$.

If $\varkappa : U \to V$ is a chart, we define the pushforward of a metric g on $\mathcal{M}$ to be the pullback of g by $\varkappa^{-1}$:

$$\varkappa_* g = (\varkappa^{-1})^* g$$

which is a metric on V. A metric on $V \Subset \mathbb{R}^n$ is the same as a smooth map from V to the space of positive definite matrices, so we can write

$$\varkappa_* g = \sum_{j,k=1}^{n} g_{jk}(x)\, dx_j dx_k \tag{13.18}$$

where $G(x) = (g_{jk}(x))_{j,k=1}^n$ is a real symmetric positive definite $n \times n$ matrix depending smoothly on $x \in V$ and for $v, w \in \mathbb{R}^n$

$$\langle v, w \rangle_{\varkappa_* g(x)} = \sum_{j,k=1}^{n} g_{jk}(x) v_j w_k.$$

Note that if $\xi = \sum_{j=1}^n \xi_j\, dx_j$, $\eta = \sum_{k=1}^n \eta_k\, dx_k$ are two vectors in the dual space to $\mathbb{R}^n$, then their inner product with respect to $\varkappa_* g$ is

$$\langle \xi, \eta \rangle_{\varkappa_* g(x)} = \sum_{j,k=1}^{n} g^{jk}(x) \xi_j \eta_k$$

where

$$G^{-1}(x) = (g^{jk}(x))_{j,k=1}^n \tag{13.19}$$

is the inverse of the matrix $G(x)$, which is again a positive definite matrix.

If $\mathcal{M} = \Phi^{-1}(y_0)$ is the example from Proposition 13.2, then a Riemannian metric on $\mathcal{M}$ can be defined by restricting the Euclidean inner product on $\mathbb{R}^N$ to tangent spaces of $\mathcal{M}$. For example, in the case of the sphere $\mathbb{S}^{n-1}$ defined in (13.1) this produces the standard metric known as the *round metric* on the sphere.

As an example of computation in coordinates, if g is the round metric on $\mathbb{S}^2$, and we consider the spherical coordinate chart $\varkappa : U \to V$ with

$$\begin{gathered} U = \mathbb{S}^2 \setminus \{ y \in \mathbb{R}^3 \mid y_1 \geq 0,\ y_2 = 0 \}, \quad V = (0, \pi)_\theta \times (0, 2\pi)_\varphi, \\ \varkappa^{-1}(\theta, \varphi) = (\sin\theta \cos\varphi,\ \sin\theta \sin\varphi,\ \cos\theta) \end{gathered} \tag{13.20}$$

then the pushforward $\varkappa_* g$ has the form

$$\varkappa_* g = d\theta^2 + \sin^2\theta\, d\varphi^2. \tag{13.21}$$

13.1.7. Integration of densities. Let $\mathcal{M}$ be a manifold and $a \in L^1_{\mathrm{c}}(\mathcal{M})$. We would like to define the integral of a on $\mathcal{M}$ but this is not possible: the resulting definition cannot be invariant under diffeomorphisms of manifolds since the change of variables formula (Theorem 10.5) features multiplication by the Jacobian. To fix this problem, we introduce a different kind of object on $\mathcal{M}$, called *density*, which can be integrated in a coordinate independent way.

We start with a bit of linear algebra:

DEFINITION 13.5. *Let $\mathcal{V}$ be an n-dimensional real vector space. A* density *on $\mathcal{V}$ is a map $\omega : \mathcal{V}^n \to \mathbb{C}$ such that for any linear map $A : \mathcal{V} \to \mathcal{V}$ and any vectors $v_1, \dots, v_n \in \mathcal{V}$ we have*

$$\omega(Av_1, \dots, Av_n) = |\det A|\, \omega(v_1, \dots, v_n). \tag{13.22}$$

Denote by $\mathrm{Den}(\mathcal{V})$ the set of all densities on $\mathcal{V}$. A density ω is called positive *if $\omega(v_1, \dots, v_n) > 0$ whenever $v_1, \dots, v_n$ form a basis of $\mathcal{V}$.*

It is immediate from the definition that $\mathrm{Den}(\mathcal{V})$ is a vector space. It is at most one-dimensional since (13.22) implies that ω is determined by $\omega(e_1, \dots, e_n)$ for any fixed choice of basis $e_1, \dots, e_n$ of $\mathcal{V}$. On $\mathbb{R}^n$ we have the canonical positive density $|dx| = dx_1 \dots dx_n$ defined as follows:

$$|dx|(v_1, \dots, v_n) = \big|\det[v_1 \dots v_n]\big| \quad \text{for all } v_1, \dots, v_n \in \mathbb{R}^n \tag{13.23}$$

where $[v_1 \dots v_n]$ is the matrix with columns $v_1, \dots, v_n$. This shows that for general $\mathcal{V}$ the space of densities is nontrivial and thus one-dimensional.

Coming back to the manifold $\mathcal{M}$, define for each $x \in \mathcal{M}$

$$|\Omega|_x := \mathrm{Den}(T_x\mathcal{M}).$$

In particular, an element of $|\Omega|_x$ is a map from $(T_x\mathcal{M})^n$ to $\mathbb{R}$. A (rough) *density* on $\mathcal{M}$ is then a map

$$\omega : x \in \mathcal{M} \ \mapsto \ \omega(x) \in |\Omega|_x.$$

If $\Phi : \mathcal{M} \to \mathcal{N}$ is a diffeomorphism, and ω is a density on $\mathcal{N}$, then the pullback $\Phi^*\omega$ is the density on $\mathcal{M}$ defined as follows: for all $x \in \mathcal{M}$ and $v_1, \dots, v_n \in T_x\mathcal{M}$

$$\Phi^*\omega(x)(v_1, \dots, v_n) = \omega(\Phi(x))(d\Phi(x)v_1, \dots, d\Phi(x)v_n). \tag{13.24}$$

If $\varkappa : U \to V$ is a chart on $\mathcal{M}$, then the *pushforward* of a density ω on $\mathcal{M}$ by $\varkappa$ is just the pullback of ω by $\varkappa^{-1}$:

$$\varkappa_*\omega := (\varkappa^{-1})^*\omega$$

which is a density on V. We can write

$$\varkappa_*\omega = \omega(x)|dx| \tag{13.25}$$

where $\omega : V \to \mathbb{C}$ and $|dx|$ is the standard density defined in (13.23).

We say that a density ω on $\mathcal{M}$ is *smooth* if for any chart $\varkappa : U \to V$ the function $\omega(x)$ from (13.25) lies in $C^\infty(V)$. Denote by $C^\infty(\mathcal{M}; |\Omega|)$ the space of all smooth densities on $\mathcal{M}$. Similarly one can define the spaces $C^\infty_{\mathrm{c}}(\mathcal{M}; |\Omega|)$, $L^p_{\mathrm{loc}}(\mathcal{M}; |\Omega|)$, and $L^p_{\mathrm{c}}(\mathcal{M}; |\Omega|)$.

We now describe how to integrate densities on manifolds. If $\varkappa : U \to V$ is a chart on $\mathcal{M}$ and $\omega \in L^1_{\mathrm{c}}(\mathcal{M}; |\Omega|)$ is supported inside U, then we define

$$\int_{\mathcal{M}} \omega := \int_V \omega(x)\, dx \tag{13.26}$$

where the right-hand side is the integral with respect to Lebesgue measure and $\omega(x)$ is defined in (13.25). Using the change of variables formula we see that this integral does not depend on the choice of the chart; we leave the verification of this as an exercise below.

For general $\omega \in L^1_{\mathrm{c}}(\mathcal{M}; |\Omega|)$ we take a partition of unity $1 = \chi_1 + \cdots + \chi_N$ near $\operatorname{supp}\omega$, where each χ_ℓ is supported in the domain of a single chart, and define

$$\int_{\mathcal{M}} \omega := \sum_{\ell=1}^{N} \int_{\mathcal{M}} \chi_\ell\, \omega \tag{13.27}$$

where the integrals on the right-hand side are defined by (13.26). The resulting integral is independent of the choice of parition of unity (something we again leave as an exercise below). Moreover, we have the following invariance under pullback: if $\Phi : \mathcal{M} \to \mathcal{N}$ is a diffeomorphism and $\omega \in L^1_{\mathrm{c}}(\mathcal{N}; |\Omega|)$ then

$$\int_{\mathcal{M}} \Phi^*\omega = \int_{\mathcal{N}} \omega. \tag{13.28}$$

Given a Riemannian metric g on $\mathcal{M}$, then we can define the *Riemannian volume density* $d\operatorname{vol}_g \in C^\infty(\mathcal{M}; |\Omega|)$ as follows: for $x \in \mathcal{M}$ and $v_1, \dots, v_n \in T_x\mathcal{M}$ we put

$$d\operatorname{vol}_g(x)(v_1, \dots, v_n) = \sqrt{|\det B|} \quad \text{where } B = (b_{jk})_{j,k=1}^n, \quad b_{jk} = \langle v_j, v_k\rangle_{g(x)}. \tag{13.29}$$

(We leave the fact that $d\operatorname{vol}_g$ is indeed a density as an exercise below.) In any chart $\varkappa : U \to V$ we have (using the notation (13.18))

$$\varkappa_* d\operatorname{vol}_g = \sqrt{|\det G(x)|}\, |dx|, \quad G(x) = (g_{jk}(x))_{j,k=1}^n. \tag{13.30}$$

Since each manifold has a Riemannian metric, and since $d\operatorname{vol}_g$ is positive, we see that each manifold has a positive C^∞ density. Denoting one such density by ω_0, we can identify densities ω on $\mathcal{M}$ with functions f by the formula $\omega = f\omega_0$. Under this identification, the integral (13.27) of ω is just the integral of f with respect to the measure on $\mathcal{M}$ induced by the density ω_0.

As an example of (13.30), in the spherical coordinate chart on $\mathbb{S}^2$ given by (13.20) the Riemannian volume density for the round metric g is given by

$$\varkappa_* d\operatorname{vol}_g = \sin\theta\, d\theta d\varphi$$

which corresponds to the integration in spherical coordinates formula from multivariable calculus.

13.1.8. Vector bundles. We finally introduce general vector bundles, several particular cases of which (the tangent bundle, the cotangent bundle, and the bundle of densities) have already appeared above.

The idea of a vector bundle over a manifold $\mathcal{M}$ is to fix a vector space $\mathscr{E}(x)$ for each point $x \in \mathcal{M}$, in a way which in some sense is smooth in x. More precisely, a *smooth m-dimensional real vector bundle* over an n-dimensional manifold $\mathcal{M}$ is a

- smooth $n+m$-dimensional manifold $\mathscr{E}$, called the *total space* of the bundle,
- a surjective smooth map $\pi : \mathscr{E} \to \mathcal{M}$, with each preimage

$$\mathscr{E}(x) := \pi^{-1}(x), \quad x \in \mathcal{M}$$

  called the *fiber* of $\mathscr{E}$ at x,
- a structure of a real m-dimensional vector space on each fiber $\mathscr{E}(x)$,
- and a collection of diffeomorphisms (called *trivializations* of $\mathscr{E}$)

$$\Theta : \pi^{-1}(U) \to U \times \mathbb{R}^m$$

  where $U \Subset \mathcal{M}$, such that for each $x \in U$, Θ maps the fiber $\mathscr{E}(x)$ to $\{x\} \times \mathbb{R}^m$, and this restricted map is a linear isomorphism with respect to the vector space structure fixed on $\mathscr{E}(x)$ and the standard vector space structure on $\{x\} \times \mathbb{R}^m \simeq \mathbb{R}^m$,
- so that the domains U of trivializations cover the whole $\mathcal{M}$ and the restriction of a trivialization Θ to $\pi^{-1}(W)$ for any $W \Subset U$ is again a trivialization.

We have transition maps between different trivializations: if $\Theta_1, \Theta_2 : \pi^{-1}(U) \to U \times \mathbb{R}^m$ are trivializations then

$$\Theta_2 \circ \Theta_1^{-1}(x, w) = (x, \mathcal{A}(x)w) \quad \text{for all } x \in U,\ w \in \mathbb{R}^m \tag{13.31}$$

where $\mathcal{A}(x) : \mathbb{R}^m \to \mathbb{R}^m$ is a family of linear isomorphisms depending smoothly on $x \in U$.

A basic example of a vector bundle is the *trivial bundle*

$$\mathscr{E} := \mathcal{M} \times \mathbb{R}^m, \quad \pi(x, v) = x,$$

with trivializations given by identity maps. More interesting examples are given by

- the n-dimensional tangent bundle $\mathscr{E} = T\mathcal{M}$, where for each chart $\varkappa : U \to V$ on $\mathcal{M}$ we have a trivialization

$$\Theta_\varkappa(x, v) = (x, d\varkappa(x)v) \in U \times \mathbb{R}^n, \quad x \in U,\ v \in T_x\mathcal{M}, \tag{13.32}$$

- the n-dimensional cotangent bundle $\mathscr{E} = T^*\mathcal{M}$, where for each chart $\varkappa : U \to V$ on $\mathcal{M}$ we have a trivialization

$$\Theta_\varkappa(x, \xi) = (x, d\varkappa(x)^{-T}\xi) \in U \times \mathbb{R}^n, \quad x \in U,\ \xi \in T_x^*\mathcal{M}, \tag{13.33}$$

  with $d\varkappa(x)^{-T}$ defined in (13.14),

- and the 1-dimensional bundle of (real) densities $\mathscr{E} = \{(x, \omega) \mid x \in \mathcal{M},\ \omega \in \mathrm{Den}(T_x\mathcal{M})\}$, where for each chart $\varkappa : U \to V$ on $\mathcal{M}$ we have a trivialization

$$\Theta_\varkappa(x, \omega) = (x, \omega(d\varkappa(x)^{-1}e_1, \dots, d\varkappa(x)^{-1}e_n)), \quad x \in U,\ \omega \in \mathrm{Den}(T_x\mathcal{M}) \tag{13.34}$$

where $e_1, \dots, e_n$ is the canonical basis of $\mathbb{R}^n$.

If $\mathscr{E}$ is a vector bundle over $\mathcal{M}$, then a *smooth section* of $\mathscr{E}$ is a map

$$\beta : x \in \mathcal{M} \ \mapsto\ \beta(x) \in \mathscr{E}(x)$$

such that $x \mapsto (x, \beta(x))$ is a smooth map $\mathcal{M} \to \mathscr{E}$. Denote by $C^\infty(\mathcal{M}; \mathscr{E})$ the space of all smooth sections of $\mathscr{E}$. Note that sections of the tangent bundle, the cotangent bundle, and the density bundle are respectively vector fields, 1-forms, and densities. On the other hand, sections of the trivial bundle $\mathcal{M} \times \mathbb{R}^m$ are just smooth functions $\mathcal{M} \to \mathbb{R}^m$.

An equivalent characterization of the map β being smooth is the following: for any trivialization $\Theta : \pi^{-1}(U) \to U \times \mathbb{R}^m$ the map $\beta_\Theta : U \to \mathbb{R}^m$ defined by

$$\Theta(\beta(x)) = (x, \beta_\Theta(x)) \quad \text{for all } x \in U \tag{13.35}$$

is smooth. We call β_Θ the *representation* of β in the trivialization Θ. Using these representations, we can define the space of locally L^p sections $L^p_{\mathrm{loc}}(\mathcal{M}; \mathscr{E})$; here we use that the transition maps (13.31) preserve the space of locally L^p maps $\mathcal{M} \to \mathbb{R}^m$. Restricting to compactly supported sections we get the spaces $C^\infty_{\mathrm{c}}(\mathcal{M}; \mathscr{E})$ and $L^p_{\mathrm{c}}(\mathcal{M}; \mathscr{E})$.

Building on Definition 13.3 we give

DEFINITION 13.6.$^{\mathbf{S}}$ *We say that $\beta_k \in C^\infty(\mathcal{M}; \mathscr{E})$ converges to β in $C^\infty(\mathcal{M}; \mathscr{E})$ if for each trivialization Θ of $\mathscr{E}$, with the representations $\beta_{k,\Theta}, \beta_\Theta : U \to \mathbb{R}^m$ defined by (13.35), we have $\beta_{k,\Theta} \to \beta_\Theta$ in $C^\infty(U; \mathbb{R}^m)$. We say that $\beta_k \in C^\infty_{\mathrm{c}}(\mathcal{M}; \mathscr{E})$ converges to β in $C^\infty_{\mathrm{c}}(\mathcal{M}; \mathscr{E})$ if $\beta_k \to \beta$ in $C^\infty(\mathcal{M}; \mathscr{E})$ and there exists $K \Subset \mathcal{M}$ such that* $\operatorname{supp} \beta_k \subset K$ *for all k.*

If $\mathscr{E}, \mathscr{F}$ are two vector bundles over the same manifold $\mathcal{M}$, then a *bundle homomorphism* is a smooth map

$$B : \mathscr{E} \to \mathscr{F} \tag{13.36}$$

such that for each $x \in \mathcal{M}$, B maps the fiber $\mathscr{E}(x)$ to the fiber $\mathscr{F}(x)$, and the corresponding map is linear (with respect to the vector space structures on $\mathscr{E}(x)$, $\mathscr{F}(x)$). We can think of bundle homomorphisms as sections of the homomorphism bundle $\mathrm{Hom}(\mathscr{E} \to \mathscr{F})$ over $\mathcal{M}$ defined by

$$\mathrm{Hom}(\mathscr{E} \to \mathscr{F})(x) = \{A : \mathscr{E}(x) \to \mathscr{F}(x) \text{ linear map}\}.$$

Thus we denote the space of all bundle homomorphisms $\mathscr{E} \to \mathscr{F}$ by

$$C^\infty(\mathcal{M}; \mathrm{Hom}(\mathscr{E} \to \mathscr{F})).$$

For a section $\omega \in C^\infty(\mathcal{M};\mathscr{E})$, we can apply B to ω to yield a section $B\omega \in C^\infty(\mathcal{M};\mathscr{F})$. The resulting operators $B : C^\infty(\mathcal{M};\mathscr{E}) \to C^\infty(\mathcal{M};\mathscr{F})$ are generalizations of multiplication by smooth functions to sections of vector bundles.

13.2. Distributions on a manifold

13.2.1. Basic properties. As discussed in (2.6), if $U \Subset \mathbb{R}^n$ then a distribution $u \in \mathcal{D}'(U)$ is determined by specifying the 'integrals' $\int_U u\varphi\,dx$ for all test functions $\varphi \in C^\infty_{\mathrm{c}}(U)$. If $\mathcal{M}$ is a manifold, then there is no canonical way to integrate functions on $\mathcal{M}$; instead, as explained in §13.1.7 we can integrate densities. The product of a function and a density is a density, so if u is a function on $\mathcal{M}$ and ω is a density on $\mathcal{M}$, then the integral $\int_{\mathcal{M}} u\omega$ makes invariant sense. Thus on manifolds, we should revise (2.6) to make a distribution u on U answer the question

$$\text{For any test density } \omega \in C^\infty_{\mathrm{c}}(\mathcal{M};|\Omega|), \text{ what is the integral } \int_{\mathcal{M}} u(x)\omega(x)\ ? \tag{13.37}$$

This means that the space $\mathcal{D}'(\mathcal{M})$ of distributions on $\mathcal{M}$ should be defined as the dual space to $C^\infty_{\mathrm{c}}(\mathcal{M};|\Omega|)$ and the space $\mathcal{E}'(\mathcal{M})$ of compactly supported distributions should be defined as the dual to $C^\infty(\mathcal{M};|\Omega|)$:

DEFINITION 13.7. *Let $\mathcal{M}$ be a manifold and $|\Omega|$ be the bundle of densities on $\mathcal{M}$. A linear functional $u : C^\infty_{\mathrm{c}}(\mathcal{M};|\Omega|) \to \mathbb{C}$ is called a* distribution *on $\mathcal{M}$ if for each ω_k converging to 0 in $C^\infty_{\mathrm{c}}(\mathcal{M};|\Omega|)$ in the sense of Definition 13.6 we have $(u,\omega_k) \to 0$. Denote by $\mathcal{D}'(\mathcal{M})$ the space of distributions on $\mathcal{M}$.*

We similarly define the class of distributions $\mathcal{E}'(\mathcal{M})$ as the space of sequentially continuous linear functionals on $C^\infty(\mathcal{M};|\Omega|)$.

As in Proposition 2.3 and the discussion following it, we denote by (u,ω) the result of applying a distribution $u \in \mathcal{D}'(\mathcal{M})$ to a density $\omega \in C^\infty_{\mathrm{c}}(\mathcal{M};|\Omega|)$ and we embed $L^1_{\mathrm{loc}}(\mathcal{M})$ into $\mathcal{D}'(\mathcal{M})$ by putting

$$(f,\omega) := \int_{\mathcal{M}} f\omega \quad \text{for all } f \in L^1_{\mathrm{loc}}(\mathcal{M}),\ \omega \in C^\infty_{\mathrm{c}}(\mathcal{M};|\Omega|). \tag{13.38}$$

A lot of the fundamental theory of distributions that we established before works on manifolds, with essentially the same proofs. This includes:

- the notion of weak convergence of distributions (see Definitions 2.7 and 4.8);
- restriction of distributions to open subsets and the sheaf property (see §2.3), since partitions of unity still exist on manifolds;
- multiplication of distributions by functions in $C^\infty(\mathcal{M})$ (see §3.2.1);
- the notion of support $\operatorname{supp} u \subset \mathcal{M}$ of a distribution $u \in \mathcal{D}'(\mathcal{M})$ and the identification of $\mathcal{E}'(\mathcal{M})$ with the space of compactly supported elements of $\mathcal{D}'(\mathcal{M})$ (see §§4.1–4.2);

- the notion of singular support $\operatorname{sing\,supp} u \subset \mathcal{M}$ of a distribution $u \in \mathcal{D}'(\mathcal{M})$ (see §8.3);
- Banach–Steinhaus theorems (see §4.3) where to define the Fréchet space structure on $C^\infty(\mathcal{M}; |\Omega|)$ we take a countable partition of unity $1 = \sum_\ell \chi_\ell$, with each χ_ℓ compactly supported in the domain U_ℓ of some chart $\varkappa_\ell : U_\ell \to V_\ell$, and taking for $\omega \in C^\infty(\mathcal{M}; |\Omega|)$ the seminorms $\|(\varkappa_\ell^{-1})^*(\chi_\ell \omega)\|_{C^N}$ for all ℓ, N.

We can also differentiate distributions, extending to them the action of vector fields on smooth functions defined in (13.9). Let $X \in C^\infty(\mathcal{M}; T\mathcal{M})$ be a vector field, considered as an operator on $C^\infty(\mathcal{M})$. Then there exists the transpose operator X^t on $C^\infty_c(\mathcal{M}; |\Omega|)$ such that in terms of the pairing defined by (13.38)

$$(Xf, \omega) = (f, X^t\omega) \quad \text{for all } f \in C^\infty(\mathcal{M}),\ \omega \in C^\infty_c(\mathcal{M}; |\Omega|), \tag{13.39}$$

and moreover $\operatorname{supp}(X^t\omega) \subset \operatorname{supp} \omega$.

To show the existence of X^t, take a chart $\varkappa : U \to V$, in which $\varkappa_* X = \sum_{j=1}^n X_j(x)\partial_{x_j}$. Recalling (13.26) we see that (13.39) holds for any $f \in C^\infty(\mathcal{M})$, $\omega \in C^\infty_c(U; |\Omega|)$ if and only if

$$\int_V \varkappa_*(Xf)\varkappa_*\omega = \int_V \varkappa_* f\, \varkappa_*(X^t\omega)$$

which by Theorem 1.17 gives the following formula for $\varkappa_*(X^t\omega)$:

$$\varkappa_*(X^t\omega) = -\sum_{j=1}^n \partial_{x_j}(X_j(x)\omega(x))|dx| \quad \text{where } \varkappa_*\omega = \omega(x)|dx|. \tag{13.40}$$

One can now check for any $\omega \in C^\infty_c(\mathcal{M}; |\Omega|)$, the formula (13.40) defines the same density $X^t\omega|_U$ for any choice of chart $\varkappa : U \to V$, and use the sheaf property for $C^\infty(\mathcal{M}; |\Omega|)$ to piece (13.40) together to a global density $X^t\omega \in C^\infty_c(\mathcal{M}; |\Omega|)$.

Now we can define the result of applying a vector field X to a distribution as follows:

$$(Xu, \omega) = (u, X^t\omega) \quad \text{for all } u \in \mathcal{D}'(\mathcal{M}),\ \omega \in C^\infty_c(\mathcal{M}; |\Omega|). \tag{13.41}$$

This gives a sequentially continuous operator $X : \mathcal{D}'(\mathcal{M}) \to \mathcal{D}'(\mathcal{M})$. The Leibniz Rule (Proposition 3.4) takes the form

$$X(au) = (Xa)u + a(Xu) \quad \text{for all } a \in C^\infty(\mathcal{M}),\ u \in \mathcal{D}'(\mathcal{M}).$$

13.2.2. Pushforwards by charts and further properties. We previously defined pushforwards of functions, vector fields, 1-forms, Riemannian metrics, and densities by charts, which allowed us to locally view these as corresponding objects on open subsets of $\mathbb{R}^n$. We now do the same with distributions.

Let $\mathcal{M}$ be a manifold and $\varkappa : U \to V$ be a chart. For $u \in \mathcal{D}'(\mathcal{M})$, define the pushforward $\varkappa_* u \in \mathcal{D}'(V)$ of u by $\varkappa$ as follows:

$$(\varkappa_* u, \varphi) = (u, \varkappa^*(\varphi|dx|)) \quad \text{for all } \varphi \in C_c^\infty(V) \tag{13.42}$$

where the pullback $\varkappa^*(\varphi|dx|)$ is a density in $C_c^\infty(U; |\Omega|) \subset C_c^\infty(\mathcal{M}; |\Omega|)$. If $u \in L^1_{\mathrm{loc}}(\mathcal{M})$, then the pushforward $\varkappa_* u$ as a distribution coincides with the usual pushforward $\varkappa_* u = u \circ \varkappa^{-1} \in L^1_{\mathrm{loc}}(V)$, as follows from (13.26) and (13.38).

If $\varkappa : U \to V$ is a chart and $W \Subset U$, then the pushforward of u by the restricted chart $\varkappa|_W$ is equal to $\varkappa_* u|_{\varkappa(W)}$. Moreover, if $\varkappa_1 : U \to V_1$, $\varkappa_2 : U \to V_2$ are two charts, then

$$\varkappa_{1*} u = (\varkappa_2 \circ \varkappa_1^{-1})^* \varkappa_{2*} u \tag{13.43}$$

where $\varkappa_2 \circ \varkappa_1^{-1} : V_1 \to V_2$ is the transition diffeomorphism and we use the notion of the pullback of a distribution from §10.1.3. Conversely, using the sheaf property for distributions we see that if for each chart $\varkappa : U \to V$ we are given a distribution $u_\varkappa \in \mathcal{D}'(V)$ and the compatibility conditions above are satisfied, then there exists unique $u \in \mathcal{D}'(\mathcal{M})$ such that $u_\varkappa = \varkappa_* u$ for all $\varkappa$.

Using pushforwards and previously proved results on distributions on open subsets of $\mathbb{R}^n$, we can establish the following properties of distributions on manifolds:

- The space $C_c^\infty(\mathcal{M})$ is dense in $\mathcal{D}'(\mathcal{M})$ and in $\mathcal{E}'(\mathcal{M})$. To show this, fix a countable partition of unity

$$1 = \sum_{\ell=1}^\infty \chi_\ell, \quad \chi_\ell \in C_c^\infty(U_\ell) \tag{13.44}$$

such that each $U_\ell \Subset \mathcal{M}$ is the domain of a chart $\varkappa_\ell : U_\ell \to V_\ell$, and the partition (13.44) is locally finite in the sense that any $K \Subset \mathcal{M}$ intersects only finitely many of the sets U_ℓ. (See for example [**Lee13**, Theorem 2.23] for a proof of existence of such a partition.)

Take arbitrary $u \in \mathcal{D}'(\mathcal{M})$. For each ℓ there exists a sequence $\varphi_{k\ell} \in C_c^\infty(U_\ell)$ such that $\varphi_{k\ell} \to \chi_\ell u$ as $k \to \infty$ in $\mathcal{D}'(\mathcal{M})$. Indeed, it suffices to use Theorem 6.10 to construct a sequence of functions in $C_c^\infty(V_\ell)$ converging to $\varkappa_{\ell*}(\chi_\ell u)$ in $\mathcal{E}'(V_\ell)$, and pull these functions back to U_ℓ by $\varkappa_\ell$. Now put

$$\varphi_k := \sum_{\ell \le k} \varphi_{k\ell} \ \in \ C_c^\infty(\mathcal{M}) \tag{13.45}$$

Take arbitrary $\omega \in C_c^\infty(\mathcal{M}; |\Omega|)$. Then there exists $\ell_0 > 0$ such that for all $\ell > \ell_0$ and k we have $U_\ell \cap \operatorname{supp} \omega = \emptyset$. We have for all $k \ge \ell_0$

$$(u, \omega) = \sum_{\ell \le \ell_0} (\chi_\ell u, \omega), \qquad (\varphi_k, \omega) = \sum_{\ell \le \ell_0} (\varphi_{k\ell}, \omega),$$

which implies that $(\varphi_k, \omega) \to (u, \omega)$ as $k \to \infty$ and thus $\varphi_k \to u$ in $\mathcal{D}'(\mathcal{M})$. The same argument shows density of $C_c^\infty(\mathcal{M})$ in $\mathcal{E}'(\mathcal{M})$.

- If $\mathcal{M}, \mathcal{N}$ are two manifolds and $u \in \mathcal{D}'(\mathcal{M})$, $v \in \mathcal{D}'(\mathcal{N})$, then one can define the tensor product $u \otimes v \in \mathcal{D}'(\mathcal{M} \times \mathcal{N})$ satisfying the conditions of Theorem 7.1. Here the tensor product of a (smooth) density on $\mathcal{M}$ and a density on $\mathcal{N}$ is a density on $\mathcal{M} \times \mathcal{N}$. Note that if $\varkappa' : U' \to V'$ is a chart on $\mathcal{M}$ and $\varkappa'' : U'' \to V''$ is a chart on $\mathcal{N}$, then the pushforward of $u \otimes v$ by the chart $\varkappa' \times \varkappa'' : U' \times U'' \to V' \times V''$ is equal to the tensor product $(\varkappa'_* u) \otimes (\varkappa''_* v)$.
- If $\mathcal{M}, \mathcal{N}$ are two manifolds, then sequentially continuous operators $C_c^\infty(\mathcal{N}) \to \mathcal{D}'(\mathcal{M})$ are in 1-to-1 correspondence with the corresponding Schwartz kernels on $\mathcal{M} \times \mathcal{N}$, but the presence of densities makes this more cumbersome to state. For example, if $\mathcal{K} \in \mathcal{D}'(\mathcal{M} \times \mathcal{N})$, then (7.15) gives the corresponding operator which maps $C_c^\infty(\mathcal{N}; |\Omega|) \to \mathcal{D}'(\mathcal{M})$.
- If $\Phi : \mathcal{M} \to \mathcal{N}$ is a smooth submersion, then the pullback operator $\Phi^* : C^\infty(\mathcal{N}) \to C^\infty(\mathcal{M})$ extends to an operator $\mathcal{D}'(\mathcal{N}) \to \mathcal{D}'(\mathcal{M})$, which can be seen by following the construction in §10.1.
- On the other hand, the concepts of homogeneity, convolution, tensor product, constant coefficient differential operator, or fundamental solution do not extend to general manifolds.

If $\mathscr{E}$ is a vector bundle over $\mathcal{M}$, then we can define the space of distributions on $\mathcal{M}$ with values in $\mathscr{E}$, denoted by $\mathcal{D}'(\mathcal{M}; \mathscr{E})$, as the dual to the space of smooth compactly supported sections $C_c^\infty(\mathcal{M}; \mathrm{Hom}(\mathscr{E} \to |\Omega|))$, with $|\Omega|$ denoting the bundle of densities over $\mathcal{M}$. Indeed, for each $f \in L^1_{\mathrm{loc}}(\mathcal{M}; \mathscr{E})$ and $\omega \in C_c^\infty(\mathcal{M}; \mathrm{Hom}(\mathscr{E} \to |\Omega|))$ the product ωf lies in $L^1_c(\mathcal{M}; |\Omega|)$ and thus can be integrated in an invariant way, yielding the pairing (f, ω).

A representation of $u \in \mathcal{D}'(\mathcal{M}; \mathscr{E})$ in a trivialization $\Theta : \pi^{-1}(U) \to U \times \mathbb{R}^m$ (see (13.35)) is a distribution on U with values in $\mathbb{R}^m$, which is the same as an element of the direct sum of m copies of $\mathcal{D}'(U)$. We leave it to the reader to fill in the technical details of the construction of such local representations.

A particular case is when $\mathscr{E} = |\Omega|$, with the space $\mathcal{D}'(\mathcal{M}; |\Omega|)$ being the dual to the space of (scalar) functions $C_c^\infty(\mathcal{M})$. (Note that since $|\Omega|$ is a one-dimensional vector bundle, the homomorphism bundle $\mathrm{Hom}(|\Omega| \to |\Omega|)$ is canonically isomorphic to the trivial bundle $\mathcal{M} \times \mathbb{R}$.) This space of density-valued distributions includes in particular the delta function δ_y for any $y \in \mathcal{M}$, defined by

$$(\delta_y, \varphi) = \varphi(y) \quad \text{for all } \varphi \in C_c^\infty(\mathcal{M}). \tag{13.46}$$

13.2.3. Sobolev spaces. We now introduce Sobolev spaces on a manifold:

DEFINITION 13.8. *Assume that* $\mathcal{M}$ *is a manifold and* $s \in \mathbb{R}$*. Define the spaces*

$$H^s_{\mathrm{loc}}(\mathcal{M}) \subset \mathcal{D}'(\mathcal{M}), \quad H^s_{\mathrm{c}}(\mathcal{M}) := H^s_{\mathrm{loc}}(\mathcal{M}) \cap \mathcal{E}'(\mathcal{M})$$

as follows: a distribution $u \in \mathcal{D}'(\mathcal{M})$ *lies in* $H^s_{\mathrm{loc}}(\mathcal{M})$ *if and only if for any chart* $\varkappa : U \to V$ *the pushforward* $\varkappa_* u \in \mathcal{D}'(V)$ *lies in the local Sobolev space* $H^s_{\mathrm{loc}}(V)$ *(see Definition 12.12).*

This definition makes sense since for any two charts $\varkappa_1 : U \to V_1$, $\varkappa_2 : U \to V_2$, we have $\varkappa_{1*} u \in H^s_{\mathrm{loc}}(V_1)$ if and only if $\varkappa_{2*} u \in H^s_{\mathrm{loc}}(V_2)$, as follows from Proposition 12.15 and (13.43). See Exercise 13.3 below for more information.

We say that a sequence $u_k \in H^s_{\mathrm{loc}}(\mathcal{M})$ converges to u in $H^s_{\mathrm{loc}}(\mathcal{M})$ if for any chart $\varkappa : U \to V$ the pushforward $\varkappa_* u_k$ converges to $\varkappa_* u$ in $H^s_{\mathrm{loc}}(V)$. Convergence in $H^s_{\mathrm{c}}(\mathcal{M})$ is defined by requiring convergence in $H^s_{\mathrm{loc}}(\mathcal{M})$ and all the supports being contained in a fixed compact subset of $\mathcal{M}$.

An important special case is when $\mathcal{M}$ is a compact manifold. Then $H^s_{\mathrm{loc}}(\mathcal{M}) = H^s_{\mathrm{c}}(\mathcal{M})$, and we denote this space by just $H^s(\mathcal{M})$. One can make $H^s(\mathcal{M})$ naturally into a Hilbert space, convergence in which corresponds to H^s_{loc} (or equivalently H^s_{c}) convergence. See Exercise 13.4 below for details.

Coming back to the case of general $\mathcal{M}$, the properties in §12.1.5 still hold on manifolds, in particular:

- if $a \in C^\infty(\mathcal{M})$, then multiplication by a is a continuous operator on $H^s_{\mathrm{loc}}(\mathcal{M})$ and on $H^s_{\mathrm{c}}(\mathcal{M})$;
- if X is a smooth vector field on $\mathcal{M}$, then it defines a continuous operator $H^{s+1}_{\mathrm{loc}}(\mathcal{M}) \to H^s_{\mathrm{loc}}(\mathcal{M})$ and $H^{s+1}_{\mathrm{c}}(\mathcal{M}) \to H^s_{\mathrm{c}}(\mathcal{M})$;
- the space $C^\infty_{\mathrm{c}}(\mathcal{M})$ is dense in both $H^s_{\mathrm{loc}}(\mathcal{M})$ and $H^s_{\mathrm{c}}(\mathcal{M})$ (for the case of $H^s_{\mathrm{loc}}(\mathcal{M})$ one can argue similarly to (13.45)).

One can also define Sobolev spaces $H^s(\mathcal{M}; \mathscr{E})$ inside the spaces of distributions with values in a vector bundle $\mathscr{E}$. We leave to the reader to work out the technical details, noting that the transition maps (13.31) preserve the spaces $H^s_{\mathrm{loc}}(U; \mathbb{R}^m) \simeq \oplus^m H^s_{\mathrm{loc}}(U)$ since multiplication by smooth functions preserves $H^s_{\mathrm{loc}}(U)$.

The spaces $H^s_{\mathrm{loc}}(\mathcal{M}; \mathscr{E})$ and $H^{-s}_{\mathrm{c}}(\mathcal{M}; \mathrm{Hom}(\mathscr{E} \to |\Omega|))$ are dual to each other in the same sense as Proposition 12.14. In particular, $H^s_{\mathrm{loc}}(\mathcal{M})$ is dual to $H^{-s}_{\mathrm{c}}(\mathcal{M}; |\Omega|)$.

13.3. Differential operators

We now introduce the notion of the principal symbol of a general differential operator and study its basic properties.

13.3.1. The case of $\mathbb{R}^n$. We start with the case of $U \Subset \mathbb{R}^n$. Recall from §9.1.1 the definition of the algebra of differential operators $\mathrm{Diff}^m(U)$, where $m \in \mathbb{N}_0$ is the order of the operator. The key object associated to a differential operator is its *principal symbol*, defined as follows:

DEFINITION 13.9. *Let $U \Subset \mathbb{R}^n$ and consider an operator $P \in \mathrm{Diff}^m(U)$ given by*

$$P = \sum_{|\alpha| \leq m} a_\alpha(x) D_x^\alpha, \quad a_\alpha \in C^\infty(U), \quad D_x^\alpha := (-i)^{|\alpha|} \partial_x^\alpha. \tag{13.47}$$

Define the principal symbol of P to be the function

$$p \in C^\infty(U \times \mathbb{R}^n), \quad p(x,\xi) = \sum_{|\alpha|=m} a_\alpha(x) \xi^\alpha, \quad x \in U, \ \xi \in \mathbb{R}^n. \tag{13.48}$$

We use the notation $\sigma_m(P) := p$.

Here are a few basic properties of the principal symbol:

- $\sigma_m(P)$ is a homogeneous polynomial of degree m in $\xi \in \mathbb{R}^n$ with coefficients that are smooth functions of $x \in U$;
- if P has constant coefficients then the definition of principal symbol that we just gave agrees with the one given before in (12.34) (note that there we denoted the principal symbol by p_0 rather than p);
- $\sigma_m(P) = 0$ if and only if $P \in \mathrm{Diff}^{m-1}(U)$.

We will sometimes suppress the subscript m and just denote the principal symbol of P by $\sigma(P)$, when the order of P is understood from the context.

The principal symbol has several important algebraic properties. We do not use them in the course but list them in the following proposition. The proof is left as an exercise below.

PROPOSITION 13.10. *Let $U \Subset \mathbb{R}^n$, $A \in \mathrm{Diff}^m(U)$, and $B \in \mathrm{Diff}^\ell(U)$. Then:*

1. (Product Rule) The principal symbol of the composition $AB \in \mathrm{Diff}^{m+\ell}(U)$ is

$$\sigma_{m+\ell}(AB) = \sigma_m(A)\sigma_\ell(B). \tag{13.49}$$

2. (Commutator Rule) The principal symbol of the commutator $[A,B] := AB - BA$ is

$$\sigma_{m+\ell-1}([A,B]) = -i\{\sigma_m(A), \sigma_\ell(B)\} \tag{13.50}$$

where $[A,B] \in \mathrm{Diff}^{m+\ell-1}(U)$ since $\sigma_{m+\ell}([A,B]) = 0$ by (13.49), and the Poisson bracket $\{\bullet,\bullet\}$ is defined by

$$\{a,b\} = \sum_{j=1}^n (\partial_{\xi_j} a)(\partial_{x_j} b) - (\partial_{x_j} a)(\partial_{\xi_j} b), \quad a, b \in C^\infty(U \times \mathbb{R}^n). \tag{13.51}$$

3. (Adjoint Rule) The principal symbol of the adjoint $A^* \in \mathrm{Diff}^m(U)$ *(see (7.22)) is given by*

$$\sigma_m(A^*)(x,\xi) = \overline{\sigma_m(A)(x,\xi)}. \tag{13.52}$$

In preparation for defining the principal symbol of a differential operator on a manifold, we need to understand how principal symbols change under conjugation by diffeomorphisms. To do this, we first prove the following preliminary statement, which is interesting in its own right:

LEMMA 13.11. *Assume that* $U \Subset \mathbb{R}^n$, $P \in \mathrm{Diff}^m(U)$, *and we are given two functions*

$$\varphi \in C^\infty(U;\mathbb{R}), \quad b \in C^\infty(U;\mathbb{C}).$$

(One commonly calls φ *the* phase *and* b *the* amplitude.*) Denote* $p := \sigma_m(P)$. *Then we have for all* $\lambda \in \mathbb{R}$

$$P\big(e^{i\lambda\varphi(x)}b(x)\big) = e^{i\lambda\varphi(x)}\big(p(x,d\varphi(x))b(x)\lambda^m + r(x,\lambda)\big) \tag{13.53}$$

where $r(x,\lambda)$ *is a polynomial of degree* $m-1$ *in* λ *with coefficients in* $C^\infty(U)$ *and* $d\varphi(x)$ *denotes the vector* $(\partial_{x_1}\varphi(x),\dots,\partial_{x_n}\varphi(x))$.

REMARK 13.12. *An important special case is when* $b \equiv 1$ *and* $\varphi(x) = x\cdot\xi$ *for some fixed* $\xi \in \mathbb{R}^n$. *In this case we recover the full symbol of* P*:*

$$P(e^{i\lambda x\cdot\xi}) = e^{i\lambda x\cdot\xi}\sum_{|\alpha|\le m} a_\alpha(x)\xi^\alpha\lambda^{|\alpha|}$$

where $a_\alpha \in C^\infty(U)$ *are the coefficients of* P, *see (13.47).*

PROOF. Consider the conjugated operator $e^{-i\lambda\varphi}Pe^{i\lambda\varphi}$ defined by

$$e^{-i\lambda\varphi}Pe^{i\lambda\varphi}f(x) = e^{-i\lambda\varphi(x)}P(e^{i\lambda\varphi(x)}f(x)) \quad \text{for all } f \in C^\infty(U).$$

We compute

$$e^{-i\lambda\varphi}D_{x_j}e^{i\lambda\varphi} = D_{x_j} + \lambda\partial_{x_j}\varphi.$$

Therefore for a general differentiation operator $D_x^\alpha = D_{x_1}^{\alpha_1}\dots D_{x_n}^{\alpha_n}$ we have

$$e^{-i\lambda\varphi}D_x^\alpha e^{i\lambda\varphi} = (D_x + \lambda d\varphi)^\alpha = (D_{x_1}+\lambda\partial_{x_1}\varphi)^{\alpha_1}\dots(D_{x_n}+\lambda\partial_{x_n}\varphi)^{\alpha_n}.$$

It follows that, using the formula (13.47) for P,

$$P(e^{i\lambda\varphi(x)}b(x)) = e^{i\lambda\varphi(x)}\sum_{|\alpha|\le m} a_\alpha(x)(D_x+\lambda d\varphi)^\alpha b(x). \tag{13.54}$$

The sum on the right-hand side is a polynomial of degree m in λ with coefficients in $C^\infty(U)$. The coefficient of λ^m in this polynomial is equal to

$$\sum_{|\alpha|=m} a_\alpha(x)d\varphi(x)^\alpha b(x) = p(x,d\varphi(x))b(x),$$

which gives (13.53). $\square$

REMARK 13.13. *An a concrete example, if $P = \Delta$ is the Laplacian, then*

$$e^{-i\lambda\varphi(x)}P(e^{i\lambda\varphi(x)}a(x)) = -|d\varphi(x)|^2 a(x)\lambda^2 + (2d\varphi(x)\cdot da(x) + (\Delta\varphi(x))a(x))i\lambda + \Delta a(x).$$

The principal part as $\lambda \to \infty$ is given by (13.53); the other parts are harder to understand (though they come up in several advanced topics such as Carleman estimates or Witten Laplacians).

We now give the promised formula for how the principal symbol behaves under changes of variables:

PROPOSITION 13.14. *Let $U, V \Subset \mathbb{R}^n$ and $\Phi : U \to V$ be a diffeomorphism. Assume that $P \in \operatorname{Diff}^m(V)$ and define the* pullback of P by Φ *as the operator $\Phi^*P : C^\infty(U) \to C^\infty(U)$ given by*

$$(\Phi^*P)(u) = \Phi^*(P(\Phi^{-*}u)) \quad \text{for all } u \in C^\infty(U) \tag{13.55}$$

where Φ^{-*} is the pullback operator by Φ^{-1}. Then $\Phi^*P \in \operatorname{Diff}^m(U)$ and

$$\sigma_m(\Phi^*P)(x,\xi) = \sigma_m(P)(\Phi(x), d\Phi(x)^{-T}\xi) \quad \text{for all } x \in U,\ \xi \in \mathbb{R}^n \tag{13.56}$$

where $d\Phi(x)^{-T}$ denotes the inverse of the transpose of $d\Phi(x) : \mathbb{R}^n \to \mathbb{R}^n$.

PROOF. The fact that $\Phi^*P \in \operatorname{Diff}^m(U)$ follows from the Chain Rule. One can also use the Chain Rule to get (13.56), but we instead give a proof relying on Lemma 13.11.

Take arbitrary $\varphi \in C^\infty(U;\mathbb{R})$. By Lemma 13.11, we have as $\lambda \to \infty$

$$(\Phi^*P)(e^{i\lambda\varphi(x)}) = e^{i\lambda\varphi(x)}\big(\sigma_m(\Phi^*P)(x, d\varphi(x))\lambda^m + \mathcal{O}(\lambda^{m-1})\big). \tag{13.57}$$

Denote $\psi := \varphi \circ \Phi^{-1} \in C^\infty(V;\mathbb{R})$. Then, denoting $p := \sigma_m(P)$, we have

$$\begin{aligned}(\Phi^*P)(e^{i\lambda\varphi(x)}) &= \Phi^*(P(e^{i\lambda\psi(y)})) \\ &= \Phi^*\big(e^{i\lambda\psi(y)}\big(p(y, d\psi(y))\lambda^m + \mathcal{O}(\lambda^{m-1})\big)\big) \\ &= e^{i\lambda\varphi(x)}\big(p(\Phi(x), d\psi(\Phi(x)))\lambda^m + \mathcal{O}(\lambda^{m-1})\big).\end{aligned} \tag{13.58}$$

Here in the first equality we used (13.55) and in the second equality we again used Lemma 13.11. By the Chain Rule we have $d\psi(\Phi(x)) = d\Phi(x)^{-T}d\varphi(x)$. Comparing (13.57) and (13.58) we get

$$\sigma_m(\Phi^*P)(x, d\varphi(x)) = p(\Phi(x), d\Phi(x)^{-T}d\varphi(x)).$$

Since φ can be chosen arbitrary, we get (13.56). $\square$

13.3.2. The case of manifolds and examples. We now introduce differential operators on manifolds. As in §§13.1.4–13.1.7, we use pushfowards by charts.

Definition 13.15. *Let $\mathcal{M}$ be a manifold. We say that an operator $P : C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$ is a* differential operator *of order $m \in \mathbb{N}_0$ if for any chart $\varkappa : U \to V$, there exists a differential operator $\varkappa_* P \in \operatorname{Diff}^m(V)$, called the* pushforward of P by $\varkappa$, *such that*

$$\varkappa_*(Pu) = (\varkappa_* P)(\varkappa_* u) \quad \text{for all } u \in C^\infty(\mathcal{M}). \tag{13.59}$$

Denote by $\operatorname{Diff}^m(\mathcal{M})$ the space of all differential operators of order m on $\mathcal{M}$.

Each $P \in \operatorname{Diff}^m(\mathcal{M})$ is sequentially continuous on $C^\infty(\mathcal{M})$ and on $C^\infty_c(\mathcal{M})$, and satisfies the locality property

$$\operatorname{supp}(Pu) \subset \operatorname{supp} u \quad \text{for all } u \in C^\infty(\mathcal{M}).$$

Next, we define the principal symbol of a differential operator on a manifold, which is a function on its cotangent bundle:

Proposition 13.16. *Assume that $\mathcal{M}$ is a manifold and $P \in \operatorname{Diff}^m(\mathcal{M})$. Then there exists unique function $p \in C^\infty(T^*\mathcal{M})$ such that for each chart $\varkappa : U \to V$ we have*

$$p(x,\xi) = \sigma_m(\varkappa_* P)(\varkappa(x), d\varkappa(x)^{-T}\xi) \quad \text{for all } x \in U,\ \xi \in T^*_x\mathcal{M}. \tag{13.60}$$

where $\sigma_m(\varkappa_* P)$ is defined in (13.48). We call p the principal symbol *of P and denote $\sigma_m(P) := p$.*

The proof of Proposition 13.16, left as an exercise below, relies on Proposition 13.14, which shows that it is natural to consider the principal symbol of a differential operator on $V \Subset \mathbb{R}^n$ as a function on the cotangent bundle of V (which is canonically identified with $V \times \mathbb{R}^n$).

Remark 13.17.**X** *Proposition 13.10 can be extended to differential operators on manifolds. The Product Rule and the Commutator Rule there are the same as for open subsets of $\mathbb{R}^n$, though some work is needed to see why the Poisson bracket makes invariant sense on functions on $T^*\mathcal{M}$ (for this one typically uses the canonical symplectic form on $\mathcal{M}$, but we do not develop this here). The Adjoint Rule also holds but one has to be careful since the adjoint of an operator $A : C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$ is an operator on densities, $A^* : C^\infty(\mathcal{M}; |\Omega|) \to C^\infty(\mathcal{M}; |\Omega|)$; see §13.3.3 below for the definition of such an operator.*

We now give several important examples of differential operators on manifolds and compute their principal symbols:

- Let $a \in C^\infty(\mathcal{M})$. Then the multiplication operator $Pu = au$ lies in $\mathrm{Diff}^0(\mathcal{M})$ and we have
$$\sigma_0(P)(x,\xi) = a(x) \quad \text{for all } (x,\xi) \in T^*\mathcal{M}. \tag{13.61}$$
- Let $X \in C^\infty(\mathcal{M}; T\mathcal{M})$ be a vector field and consider it as an operator using (13.9). Then the operator $P := -iX$ lies in $\mathrm{Diff}^1(\mathcal{M})$ and we have
$$\sigma_1(P)(x,\xi) = \xi(X(x)) \quad \text{for all } (x,\xi) \in T^*\mathcal{M}. \tag{13.62}$$
This gives a simple explanation for why the principal symbol of a differential operator should be a function on the cotangent bundle (rather than, say, the tangent bundle): the vector $X(x)$ defines a linear function on the cotangent space $T_x^*\mathcal{M}$.
- Let g be a Riemannian metric on $\mathcal{M}$. The *Laplace–Beltrami operator* Δ_g is characterized by the identity
$$\begin{gathered} -\int_{\mathcal{M}} (\Delta_g f)\varphi \, d\,\mathrm{vol}_g = \int_{\mathcal{M}} \langle df(x), d\varphi(x)\rangle_{g(x)} \, d\,\mathrm{vol}_g(x) \\ \text{for all } f \in C^\infty(\mathcal{M}),\ \varphi \in C_{\mathrm{c}}^\infty(\mathcal{M}). \end{gathered} \tag{13.63}$$
If $\varkappa : U \to V$ is a chart and $\varkappa_* g = \sum_{j,k=1}^n g_{jk}(x)\, dx_j dx_k$, then the pushforward $\varkappa_*\Delta_g$ has the form
$$(\varkappa_*\Delta_g) f(x) = \frac{1}{\sqrt{|\det G(x)|}} \sum_{j,k=1}^n \partial_{x_j}\big(\sqrt{|\det G(x)|}\, g^{jk}(x) \partial_{x_k} f(x)\big) \tag{13.64}$$
where $G(x) = (g_{jk}(x))_{j,k=1}^n$ and $G^{-1}(x) = (g^{jk}(x))_{j,k=1}^n$. The operator Δ_g lies in $\mathrm{Diff}^2(\mathcal{M})$ and its principal symbol is given by
$$\sigma_2(\Delta_g)(x,\xi) = -\langle \xi, \xi\rangle_{g(x)} \quad \text{for all } (x,\xi) \in T^*\mathcal{M}. \tag{13.65}$$

13.3.3. Vector bundles. We finally discuss differential operators acting on sections of vector bundles over manifolds. We start with the trivial bundles. Let $\ell, \ell' \in \mathbb{N}$. An operator $\mathbf{P} : C^\infty(\mathcal{M}; \mathbb{C}^{\ell'}) \to C^\infty(\mathcal{M}; \mathbb{C}^\ell)$ is the same as a matrix of operators
$$\big(P_{kk'} : C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})\big)_{1 \le k \le \ell,\, 1 \le k' \le \ell'},$$
with the action of $\mathbf{P}$ given by
$$\mathbf{P}(u_1, \dots, u_{\ell'}) = (v_1, \dots, v_\ell) \quad \text{with } v_k = \sum_{k'=1}^{\ell'} P_{kk'} u_{k'}. \tag{13.66}$$
If each $P_{kk'}$ is a differential operator in $\mathrm{Diff}^m(\mathcal{M})$, then we say that $\mathbf{P}$ is a differential operator of order m, and write
$$\mathbf{P} \in \mathrm{Diff}^m(\mathcal{M}; \mathbb{C}^{\ell'} \to \mathbb{C}^\ell).$$

Now, let $\mathscr{E}, \mathscr{F}$ be two vector bundles of dimensions ℓ', ℓ over a manifold $\mathcal{M}$. To keep the theory consistent we should complexify $\mathscr{E}, \mathscr{F}$ (i.e. consider the bundle $\mathscr{E}_{\mathbb{C}}$ with $\mathscr{E}_{\mathbb{C}}(x) = \mathbb{C} \otimes_{\mathbb{R}} \mathscr{E}(x)$ for all x) but we suppress this complexification in the notation.

Definition 13.18. *We say that a sequentially continuous operator* $\mathbf{P} : C^\infty(\mathcal{M}; \mathscr{E}) \to C^\infty(\mathcal{M}; \mathscr{F})$ *is a* differential operator *of order* $m \in \mathbb{N}_0$, *and write*

$$\mathbf{P} \in \mathrm{Diff}^m(\mathcal{M}; \mathscr{E} \to \mathscr{F})$$

if for any trivializations $\Theta_{\mathscr{E}} : \pi_{\mathscr{E}}^{-1}(U) \to U \times \mathbb{C}^{\ell'}$, $\Theta_{\mathscr{F}} : \pi_{\mathscr{F}}^{-1}(U) \to U \times \mathbb{C}^{\ell}$, *where* $U \Subset \mathcal{M}$ *and* $\pi_{\mathscr{E}} : \mathscr{E} \to \mathcal{M}$, $\pi_{\mathscr{F}} : \mathscr{F} \to \mathcal{M}$ *are the projection maps, there exists a differential operator* $\widetilde{\mathbf{P}} \in \mathrm{Diff}^m(\mathcal{M}; \mathbb{C}^{\ell'} \to \mathbb{C}^{\ell})$ *which is the* representation of $\mathbf{P}$ in the trivializations $\Theta_{\mathscr{E}}$, $\Theta_{\mathscr{F}}$ *in the following sense:*

$$(\mathbf{P}\beta)_{\Theta_{\mathscr{F}}} = \widetilde{\mathbf{P}}\beta_{\Theta_{\mathscr{E}}} \quad \textit{for all } \beta \in C^\infty(\mathcal{M}; \mathscr{E}) \tag{13.67}$$

where $\beta_{\Theta_{\mathscr{E}}} \in C^\infty(U; \mathbb{C}^{\ell'})$, $(\mathbf{P}\beta)_{\Theta_{\mathscr{F}}} \in C^\infty(U; \mathbb{C}^{\ell})$ *are defined in* (13.35).

Note that $\mathbf{P}$ is sequentially continuous on the spaces C^∞ and C^∞_{c} and $\mathrm{supp}(\mathbf{P}\beta) \subset \mathrm{supp}\,\beta$ for all $\beta \in C^\infty(\mathcal{M}; \mathscr{E})$. Moreover, $\mathrm{Diff}^0(\mathcal{M}; \mathscr{E} \to \mathscr{F})$ is just the space of bundle homomorphisms $\mathscr{E} \to \mathscr{F}$ defined in (13.35) above.

An important example of a differential operator on bundles is the first order differential operator (see (13.17))

$$d : C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M}; T^*\mathcal{M}). \tag{13.68}$$

We now define the principal symbol of a differential operator on vector bundles. In case of trivial bundles, the principal symbol of an operator $\mathbf{P} \in \mathrm{Diff}^m(\mathcal{M}; \mathbb{C}^{\ell'} \to \mathbb{C}^{\ell})$ given by (13.66), is the matrix of the principal symbols of the operators $P_{kk'}$. It is useful to think of this matrix as a linear map of the fibers $\mathbb{C}^{\ell'}, \mathbb{C}^{\ell}$: if $(x, \xi) \in T^*\mathcal{M}$ then $\sigma_m(\mathbf{P})(x, \xi) : \mathbb{C}^{\ell'} \to \mathbb{C}^{\ell}$ is the linear map given by

$$\sigma_m(\mathbf{P})(x, \xi)(w_1, \dots, w_{\ell'}) = \left(\sum_{k'=1}^{\ell'} \sigma_m(P_{kk'})(x, \xi) w_{k'} \right)_{k=1}^{\ell}. \tag{13.69}$$

Next, if $\mathscr{E}, \mathscr{F}$ are two vector bundles over $\mathcal{M}$ and $\mathbf{P} \in \mathrm{Diff}^m(\mathcal{M}; \mathscr{E} \to \mathscr{F})$ then we define the principal symbol $\sigma_m(\mathbf{P})$ as follows: for each $(x, \xi) \in T^*\mathcal{M}$, the value of $\sigma_m(\mathbf{P})$ at (x, ξ) is a linear map

$$\sigma_m(\mathbf{P})(x, \xi) : \mathscr{E}(x) \to \mathscr{F}(x) \tag{13.70}$$

such that for any trivializations $\Theta_{\mathscr{E}}, \Theta_{\mathscr{F}}$ and with $\widetilde{\mathbf{P}}$ given by (13.67) we have for all $x \in U$, $\xi \in T_x^*\mathcal{M}$, and $\mathbf{w} \in \mathscr{E}(x)$

$$\Theta_{\mathscr{F}}(\sigma_m(\mathbf{P})(x, \xi)\mathbf{w}) = (x, \sigma_m(\widetilde{\mathbf{P}})(x, \xi)\widetilde{\mathbf{w}}) \quad \text{where } \Theta_{\mathscr{E}}(\mathbf{w}) = (x, \widetilde{\mathbf{w}}). \tag{13.71}$$

This definition does not depend on the choice of trivializations, as one can show that the symbol defined in (13.69) is equivariant under the transition maps (13.31); we skip the details. The resulting symbol $\sigma_m(\mathbf{P})$ is a section of the bundle $\pi^*\operatorname{Hom}(\mathscr{E}\to\mathscr{F})$ over $T^*\mathcal{M}$, which is the pullback of the homomorphism bundle $\operatorname{Hom}(\mathscr{E}\to\mathscr{F})$ by the projection map $\pi: T^*\mathcal{M}\to\mathcal{M}$; more precisely,

$$\pi^*\operatorname{Hom}(\mathscr{E}\to\mathscr{F})(x,\xi)=\operatorname{Hom}(\mathscr{E}\to\mathscr{F})(x)\quad\text{for all }(x,\xi)\in T^*\mathcal{M}.\tag{13.72}$$

As an example, if d is the differential operator from (13.68) then we compute the principal symbol of $-id$:

$$\sigma_1(-id)(x,\xi)w=w\xi\quad\text{for all }(x,\xi)\in T^*\mathcal{M},\ w\in\mathbb{C}.$$

which follows from (13.17).

One application of differential operators on vector bundles is extension of differential operators to distributions by duality. Namely, if $A\in\operatorname{Diff}^m(\mathcal{M})$, then we can extend it to a sequentially continuous operator on $\mathcal{D}'(\mathcal{M})$ by the formula

$$(Au,\omega)=(u,A^t\omega)\quad\text{for all }u\in\mathcal{D}'(\mathcal{M}),\ \omega\in C^\infty_{\mathrm{c}}(\mathcal{M};|\Omega|)\tag{13.73}$$

where the transpose operator A^t lies in $\operatorname{Diff}^m(\mathcal{M};|\Omega|\to|\Omega|)$. We omit the details since they are quite similar to what was done for vector fields in (13.41).

13.4. Notes and exercises

The material in §13.1 can be found in most differential geometry textbooks such as [**Lee13**]. The presentation in §13.2 partially follows [**Hör03**, §6.3], and the presentation in §13.3 partially follows the first half-page of [**Hör03**, §8.3].

Exercise 13.1. ($1=0.5+0.5$ pt) *This exercise proves coordinate invariance of the integral of a density, introduced in §13.1.7. Let $\mathcal{M}$ be a manifold.*

(a) *Assume that $\varkappa_1:U_1\to V_1$ and $\varkappa_2:U_2\to V_2$ are charts on $\mathcal{M}$ and $\omega\in L^1_{\mathrm{c}}(\mathcal{M};|\Omega|)$ is supported inside $U_1\cap U_2$. Show that the integrals $\int_{\mathcal{M}}\omega$ defined by (13.26) using the charts $\varkappa_1$ and $\varkappa_2$ are equal to each other. (Hint: use Theorem 10.5 for the transition map between $\varkappa_1$ and $\varkappa_2$.)*

(b) *Assume that $\omega\in L^1_{\mathrm{c}}(\mathcal{M};|\Omega|)$. Show that the integral $\int_{\mathcal{M}}\omega$ defined in (13.27) does not depend on the choice of partition of unity.*

Exercise 13.2. *(1 pt) Let $(\mathcal{M},g)$ be a Riemannian manifold. Show that the expression $d\operatorname{vol}_g$ defined in (13.29) is a density, that is for each $x\in\mathcal{M}$ the function $d\operatorname{vol}_g(x):(T_x\mathcal{M})^n\to\mathbb{R}$ satisfies (13.22). (Hint: reduce to the case when $v_1,\dots,v_n$ is a basis of $T_x\mathcal{M}$ use the matrix of the linear map A in this basis.)*

Exercise 13.3. (1 pt) *This exercise shows in particular that Sobolev spaces on manifolds are nontrivial, by constructing elements of these spaces from charts. Let $\mathcal{M}$ be a manifold, $\varkappa:U\to V$ be a chart on $\mathcal{M}$, $s\in\mathbb{R}$, and $v\in H^s_{\mathrm{c}}(V)$. Take the pullback $\varkappa^*v\in\mathcal{E}'(U)$ and extend it by zero to an element of $\mathcal{E}'(\mathcal{M})$. Show that $\varkappa^*v\in H^s_{\mathrm{c}}(\mathcal{M})$,* *and if $v_k \to 0$ in $H^s_{\mathrm{c}}(V)$ then $\varkappa^* v_k \to 0$ in $H^s_{\mathrm{c}}(\mathcal{M})$. (You may freely use properties of pullback of distributions on manifolds.)*

EXERCISE 13.4. (2 = 1+1 pts) *Let $\mathcal{M}$ be a compact manifold. Fix a finite collection of charts $\varkappa_\ell : U_\ell \to V_\ell$, $\ell = 1, \dots, N$, such that $\mathcal{M} = \bigcup_{\ell=1}^N U_\ell$, and a partition of unity*

$$1 = \sum_{\ell=1}^N \chi_\ell, \quad \chi_\ell \in C^\infty_{\mathrm{c}}(U_\ell).$$

Let $s \in \mathbb{R}$ and denote $H^s(\mathcal{M}) := H^s_{\mathrm{loc}}(\mathcal{M}) = H^s_{\mathrm{c}}(\mathcal{M})$. For $u \in H^s(\mathcal{M})$, define the norm $\|u\|_{H^s(\mathcal{M})}$ as follows:

$$\|u\|^2_{H^s(\mathcal{M})} = \sum_{\ell=1}^N \|\varkappa_{\ell *}(\chi_\ell u)\|^2_{H^s(\mathbb{R}^n)} \tag{13.74}$$

where each $\varkappa_{\ell *}(\chi_\ell u)$ is a distribution in $H^s_{\mathrm{c}}(V_\ell)$ and thus in $H^s(\mathbb{R}^n)$. (It is easy to see that $\|\bullet\|$ is a norm on $H^s(\mathcal{M})$ induced by an inner product – you do not need to show this explicitly.)

***(a)** Show that for each sequence $u_k \in H^s(\mathcal{M})$, we have $\|u_k\|_{H^s(\mathcal{M})} \to 0$ if and only if $u_k \to 0$ in $H^s_{\mathrm{loc}}(\mathcal{M})$ as defined in §13.2.3. (Hint: use Exercise 13.4, the decomposition $u = \sum_{\ell=1}^N \chi_\ell u$, and the fact that $\chi_\ell u$ is the extension by 0 of the pullback $\varkappa_\ell^* \varkappa_{\ell *}(\chi_\ell u)$.) This in particular implies that a different choice of the charts $\varkappa_\ell$ and the cutoff functions χ_ℓ yields an equivalent norm (13.74).*

***(b)** Show that $H^s(\mathcal{M})$ with the norm (13.74) is complete and thus a Hilbert space. (Hint: let u_k be a Cauchy sequence in $H^s(\mathcal{M})$. Use completeness of $H^s(\mathbb{R}^n)$ to show that for each ℓ, we have $\chi_\ell u_k \to v_\ell$ as $k \to \infty$ in $H^s_{\mathrm{loc}}(\mathcal{M})$ for some $v_\ell \in H^s_{\mathrm{c}}(U_\ell)$. Conclude that $u_k \to \sum_{\ell=1}^N v_\ell$ in $H^s_{\mathrm{loc}}(\mathcal{M})$.)*

EXERCISE 13.5. (2 = 1 + 1 pts) *Prove parts 1 and 3 of Proposition 13.10. (For part 3, you can use (9.3) and the relation between transpose and adjoint.)*

EXERCISE 13.6. (1 pt) *Prove part 2 of Proposition 13.10.*

EXERCISE 13.7. (0.5 pt) *Prove Proposition 13.16.*

EXERCISE 13.8. (2.5 = 0.5+1+1 pts) *Let $\mathbb{S}^n$ be the n-sphere defined in (13.1), with $n \geq 2$, endowed with the round metric g (i.e. the one coming from the ambient space $\mathbb{R}^{n+1}$). In this exercise you compute the eigenvalues of the operator $-\Delta_g$, namely the numbers $\lambda \in \mathbb{R}$ such that there exist nonzero $u \in C^\infty(\mathbb{S}^n; \mathbb{R})$ solving the eigenfunction equation*

$$-\Delta_g u = \lambda u.$$

***(a)** Show that each eigenvalue λ has to satisfy $\lambda \geq 0$. (Hint: compute the integral $\int_{\mathbb{S}^n} (\Delta_g u) u \, d\,\mathrm{vol}_g$ using the defining property of the Laplace–Beltrami operator.)*

(b) *Let $a \geq 0$. Denote by Δ_0 the usual Laplace operator on $\mathbb{R}^{n+1}$. Show that the equation*

$$\Delta_0 v = 0 \quad \textit{on} \quad \mathbb{R}^{n+1} \setminus \{0\} \tag{13.75}$$

has a nonzero solution $v \in C^\infty(\mathbb{R}^{n+1} \setminus \{0\})$ which is homogeneous of degree a if and only if a is a (nonnegative) integer. (Hint: show that v is a locally integrable function on $\mathbb{R}^{n+1}$ and defines a tempered distribution in $\mathscr{S}'(\mathbb{R}^{n+1})$, which we denote $\tilde{v}$. Arguing similarly to the proof of (10.30), show that $\Delta_0 \tilde{v} = 0$. Now pass to the Fourier transform of $\tilde{v}$ and show that it is supported at a single point; deduce from here that $\tilde{v}$ is a polynomial.)

(c) *The pullback of the operator Δ_0 by the polar coordinate diffeomorphism*

$$\Phi : (0, \infty) \times \mathbb{S}^n \to \mathbb{R}^{n+1} \setminus \{0\}, \quad \Phi(r, \theta) := r\theta$$

is equal to the operator $\partial_r^2 + \frac{n}{r}\partial_r + \frac{1}{r^2}\Delta_g$, with the spherical Laplacian Δ_g acting in the θ variable. (This can be checked by noting that this operator has to be the Laplace–Beltrami operator of the pullback by Φ of the Euclidean metric, but you don't need to do this computation here.) Using this, show that the eigenvalues of $-\Delta_g$ are given by $k(k+n-1)$ where k runs over nonnegative integers. (Hint: if u is an eigenfunction of $-\Delta_g$ then define $v(r\theta) = r^a u(\theta)$ in polar coordinates for a right choice of a so that $\Delta_0 v = 0$.) The eigenfunctions of $-\Delta_g$ are called spherical harmonics.

CHAPTER 14

Elliptic operators with variable coefficients

In this chapter we prove the third (and last) version of elliptic regularity, for elliptic differential operators on manifolds. (See §§9.2,12.2 for the previous versions.) To state it, we make the following

Definition 14.1. *Let $\mathcal{M}$ be a manifold and $P \in \mathrm{Diff}^m(\mathcal{M})$ be a differential operator. Denote by $p := \sigma_m(P)$ the principal symbol of P (see Proposition 13.16). We say that P is an* elliptic *differential operator if*

$$p(x, \xi) \neq 0 \quad \text{for all } (x, \xi) \in T^*\mathcal{M},\ \xi \neq 0. \tag{14.1}$$

Note that for constant coefficient differential operators on $\mathbb{R}^n$, this definition of ellipticity coincides with the one given in (12.35) above. An example of an elliptic differential operator is the Laplace–Beltrami operator $\Delta_g \in \mathrm{Diff}^2(\mathcal{M})$ associated to a Riemannian metric, see §13.3.2.

We can now state the main result of this chapter:

Theorem 14.2 (Elliptic Regularity III). *Let $\mathcal{M}$ be a manifold and $P \in \mathrm{Diff}^m(\mathcal{M})$ be an elliptic differential operator. Then for each $u \in \mathcal{D}'(\mathcal{M})$ we have*

$$\operatorname{sing\,supp} u = \operatorname{sing\,supp}(Pu). \tag{14.2}$$

A version for vector bundles is given by Theorem 14.23 below. One can replace C^∞ regularity by Sobolev regularity, see Theorem 15.1 below.

14.1. Pseudodifferential operators

Our proof of Theorem 14.2 relies on the construction of an *elliptic parametrix*, which is a generalization to the variable coefficient case of the convolution operator with the distribution E used in the proof of Theorem 12.18. This elliptic parametrix will be a *pseudodifferential operator*, and in this section we take some time to introduce pseudodifferential operators and establish some of their properties. For our purposes it will be enough to study pseudodifferential operators on open subsets of $\mathbb{R}^n$.

14.1.1. Kohn–Nirenberg symbols revisited. We start by revisiting the Kohn–Nirenberg symbols introduced in §12.2.3, allowing for dependence on x in addition to ξ:

DEFINITION 14.3. *Let $U \Subset \mathbb{R}^n$. A function $a \in C^\infty(U \times \mathbb{R}^n)$ is called a* Kohn–Nirenberg symbol *of order $m \in \mathbb{R}$, if for each $K \Subset U$ and multiindices α, β there exists a constant $C_{\alpha\beta K}$ such that*

$$|\partial_x^\alpha \partial_\xi^\beta a(x,\xi)| \leq C_{\alpha\beta K} \langle \xi \rangle^{m-|\beta|} \quad \textit{for all } (x,\xi) \in K \times \mathbb{R}^n. \tag{14.3}$$

Denote by $S^m(U \times \mathbb{R}^n)$ the space of all Kohn–Nirenberg symbols of order m.

REMARK 14.4. *The bounds (14.3) can be interpreted as follows: we have $a(x,\xi) = \mathcal{O}(\langle \xi \rangle^m)$, each differentiation in x keeps the bound the same but each differentiation in ξ makes a one order smaller, and the constants in the bounds are locally uniform in $x \in U$.*

We have $S^\ell(U \times \mathbb{R}^n) \subset S^m(U \times \mathbb{R}^n)$ whenever $\ell \leq m$. We will also use the residual class (whose elements are called *rapidly decaying* symbols)

$$S^{-\infty}(U \times \mathbb{R}^n) := \bigcap_{m \in \mathbb{R}} S^m(U \times \mathbb{R}^n) \tag{14.4}$$

which can be characterized as follows: a function $a \in C^\infty(U \times \mathbb{R}^n)$ lies in $S^{-\infty}(U \times \mathbb{R}^n)$ if and only if for each $K \Subset U$, multiindices α, β, and $N \in \mathbb{N}$ there exists a constant $C_{\alpha\beta KN}$ such that

$$|\partial_x^\alpha \partial_\xi^\beta a(x,\xi)| \leq C_{\alpha\beta KN} \langle \xi \rangle^{-N} \quad \text{for all } (x,\xi) \in K \times \mathbb{R}^n. \tag{14.5}$$

Informally, this means that $a(x,\xi) = \mathcal{O}(\langle \xi \rangle^{-\infty})$ locally uniformly in x with all derivatives.

Similarly to (12.44) and (12.45) we have

$$a \in S^m(U \times \mathbb{R}^n),\ b \in S^\ell(U \times \mathbb{R}^n) \quad \Longrightarrow \quad ab \in S^{m+\ell}(U \times \mathbb{R}^n), \tag{14.6}$$

$$a \in S^m(U \times \mathbb{R}^n) \quad \Longrightarrow \quad \partial_{x_j} a \in S^m(U \times \mathbb{R}^n),\ \partial_{\xi_j} a \in S^{m-1}(U \times \mathbb{R}^n). \tag{14.7}$$

We also have the following generalizations of Propositions 12.22 and 12.23:

PROPOSITION 14.5. *Assume that $p(x,\xi) = \sum_{|\alpha| \leq m} a_\alpha(x) \xi^\alpha$ is a polynomial of degree m in $\xi \in \mathbb{R}^n$ with coefficients $a_\alpha \in C^\infty(U)$. Then $p \in S^m(U \times \mathbb{R}^n)$.*

PROOF.$^{\mathbf{S}}$ The derivative $\partial_x^\alpha \partial_\xi^\beta p$ is a polynomial of degree $m - |\beta|$ in ξ with coefficients smooth in x (and it is equal to 0 if $|\beta| > m$), which gives the bounds (14.3). □

PROPOSITION 14.6. *Assume that $p \in S^m(U \times \mathbb{R}^n)$ and for each $K \Subset U$ there exists a constant $c_K > 0$ such that*

$$|p(x,\xi)| \geq c_K |\xi|^m \quad \textit{for all } x \in K,\ \xi \in \mathbb{R}^n,\ |\xi| \geq 1. \tag{14.8}$$

Let $q \in C^\infty(U \times \mathbb{R}^n)$ be such that $q(x,\xi) = 1/p(x,\xi)$ for all $x \in U$, $|\xi| \geq 1$. Then $q \in S^{-m}(U \times \mathbb{R}^n)$.

PROOF.[S] By induction in $|\alpha|+|\beta|$ we see that for all multiindices α, β and all $x \in U$, $\xi \in \mathbb{R}^n$ with $|\xi| \geq 1$, $\partial_x^\alpha \partial_\xi^\beta q(x,\xi)$ is a linear combination with constant coefficients of expressions of the form

$$\frac{(\partial_x^{\alpha_1}\partial_\xi^{\beta_1} p(x,\xi)) \cdots (\partial_x^{\alpha_k}\partial_\xi^{\beta_k} p(x,\xi))}{p(x,\xi)^{k+1}} \tag{14.9}$$

where $|\alpha_1|+|\beta_1|, \dots, |\alpha_k|+|\beta_k| \geq 1$ and $\alpha_1+\dots+\alpha_k=\alpha$, $\beta_1+\dots+\beta_k=\beta$. Using the bounds (14.3) and (14.8) we see that for each $K \Subset U$ there exists a constant C_K so that (14.9) is bounded in absolute value by $C_K|\xi|^{-m-|\beta|}$ for all $x \in K$, $|\xi| \geq 1$. This gives the bounds (14.3) for q, showing that it lies in $S^{-m}(U \times \mathbb{R}^n)$. ☐

14.1.2. Asymptotic sums and Borel's Theorem. In preparation for the construction of elliptic parametrix in §14.2.1 below we now introduce asymptotic sums of Kohn–Nirenberg symbols:

DEFINITION 14.7. *Assume that $U \Subset \mathbb{R}^n$, $m \in \mathbb{R}$, and we are given symbols*

$$a \in S^m(U \times \mathbb{R}^n); \qquad a_k \in S^{m-k}(U \times \mathbb{R}^n),\ k \in \mathbb{N}_0.$$

We say that a is asymptotic to $\sum_{k=0}^\infty a_k$, and write

$$a \sim \sum_{k=0}^\infty a_k$$

if for each $N \in \mathbb{N}_0$ we have

$$a - \sum_{k=0}^{N-1} a_k \ \in\ S^{m-N}(U \times \mathbb{R}^n). \tag{14.10}$$

REMARK 14.8. *It is important to distinguish between asymptotic sums and convergent series. As we see in Theorem 14.9 below,* any *sequence of symbols in the right classes has an asymptotic sum. The corresponding series $\sum_{k=0}^\infty a_k(x,\xi)$ may converge for all (x,ξ). This is similar to the difference between Taylor formula and Taylor series. See Exercise 14.2 below for a version of Theorem 14.9 for Taylor expansions of functions of one variable.*

The main result about asymptotic sums is that they always exist:

THEOREM 14.9 (Borel's Theorem). *Given any sequence $a_k \in S^{m-k}(U \times \mathbb{R}^n)$, $k \in \mathbb{N}_0$, there exists $a \in S^m(U \times \mathbb{R}^n)$ such that $a \sim \sum_{k=0}^\infty a_k$ in the sense of Definition 14.7. Moreover, any two such symbols a differ by an element of $S^{-\infty}(U \times \mathbb{R}^n)$.*

PROOF. 1. Fix a cutoff function

$$\chi \in C_c^\infty(\mathbb{R}^n), \quad \chi = 1 \quad \text{on } B(0,1).$$

We have

$$1-\chi(\varepsilon\xi)\to 0 \quad \text{in } S^1(\mathbb{R}^n) \quad \text{as } \varepsilon\to 0+, \tag{14.11}$$

where convergence is understood in the sense of the seminorms coming from (14.3). (The convergence (14.11) holds in S^δ for all $\delta>0$, but not in S^0.) To see (14.11), we first observe that $|1-\chi(\xi)|\leq C|\xi|$ for all ξ and thus

$$\sup_{\xi\in\mathbb{R}^n}\langle\xi\rangle^{-1}|(1-\chi(\varepsilon\xi))|\leq C\sup_{\xi\in\mathbb{R}^n}\varepsilon\langle\xi\rangle^{-1}|\xi|\leq C\varepsilon\to 0.$$

Next, for any multiindex β with $|\beta|\geq 1$ we have

$$\begin{aligned}\sup_{\xi\in\mathbb{R}^n}\langle\xi\rangle^{|\beta|-1}\big|\partial_\xi^\beta(1-\chi(\varepsilon\xi))\big| &= \varepsilon^{|\beta|}\sup_{\xi\in\mathbb{R}^n}\langle\xi\rangle^{|\beta|-1}\big|(\partial_\xi^\beta\chi)(\varepsilon\xi)\big|\\ &\leq \varepsilon^{|\beta|}\sup_{|\xi|\leq R/\varepsilon}\langle\xi\rangle^{|\beta|-1}\leq C\varepsilon\to 0\end{aligned}$$

where we fixed $R>0$ such that $\operatorname{supp}\chi\subset B(0,R)$. This shows (14.11).

Using the Leibniz Rule similarly to (14.6), we see that (14.11) and the fact that $a_k\in S^{m-k}(U\times\mathbb{R}^n)$ implies that for each k

$$(1-\chi(\varepsilon\xi))a_k(x,\xi)\to 0 \quad \text{in } S^{m-k+1}(U\times\mathbb{R}^n). \tag{14.12}$$

2. Take a sequence of compact sets $K_k\Subset K_{k+1}$ exhausting U in the sense of (1.14). Using (14.12), choose $\varepsilon_k>0$ small enough so that

$$\begin{gathered}|\partial_x^\alpha\partial_\xi^\beta b_k(x,\xi)|\leq 2^{-k}\langle\xi\rangle^{m-k+1-|\beta|} \quad \text{for all } |\alpha|,|\beta|\leq k,\ (x,\xi)\in K_k\times\mathbb{R}^n\\ \text{where } b_k(x,\xi):=(1-\chi(\varepsilon_k\xi))a_k(x,\xi).\end{gathered} \tag{14.13}$$

We now put

$$a(x,\xi):=\sum_{k=0}^\infty b_k(x,\xi). \tag{14.14}$$

The series (14.14) converges to a function $a\in C^\infty(U\times\mathbb{R}^n)$ since for any given (x,ξ) only finitely many terms are nonzero.

We claim that for each $M\in\mathbb{N}_0$ there exists a constant C_M so that

$$\begin{gathered}\left|\partial_x^\alpha\partial_\xi^\beta\left(a(x,\xi)-\sum_{k=0}^{M-1}a_k(x,\xi)\right)\right|\leq C_M\langle\xi\rangle^{m-M+1-|\beta|}\\ \text{for all } |\alpha|,|\beta|\leq M,\ (x,\xi)\in K_M\times\mathbb{R}^n.\end{gathered} \tag{14.15}$$

Note that on the surface, (14.15) appears weaker than (14.10) since we lose a power of $\langle\xi\rangle$ and restrict the α,β,K that we can take depending on M. However, in Step 3 of the proof we will show that (14.15) implies (14.10), since we access more and more of the symbol space seminorms as M grows.

To show (14.15) we write

$$a(x,\xi)-\sum_{k=0}^{M-1}a_k(x,\xi)=-\sum_{k=0}^{M-1}\chi(\varepsilon_k\xi)a_k(x,\xi)+\sum_{k=M}^{\infty}b_k(x,\xi).$$

Since χ is compactly supported, the first sum on the right-hand side lies in $S^{-\infty}(U\times\mathbb{R}^n)$ and in particular satisfies the estimate (14.13). To bound the second sum, we use (14.13) to estimate for all $|\alpha|,|\beta|\leq M$ and $(x,\xi)\in K_M\times\mathbb{R}^n$

$$\begin{aligned}\Big|\partial_x^\alpha\partial_\xi^\beta\sum_{k=M}^{\infty}b_k(x,\xi)\Big| &\leq \sum_{k=M}^{\infty}|\partial_x^\alpha\partial_\xi^\beta b_k(x,\xi)|\\ &\leq \sum_{k=M}^{\infty}2^{-k}\langle\xi\rangle^{m-M+1-|\beta|}\leq 2^{1-M}\langle\xi\rangle^{m-M+1-|\beta|}\end{aligned}$$

finishing the proof of (14.15).

3. We now show that $a\sim\sum_{k=0}^{\infty}a_k$ in the sense of Definition 14.7. Take arbitrary $N\in\mathbb{N}_0$. We need to show that $a-\sum_{k=0}^{N-1}a_k\in S^{m-N}(U\times\mathbb{R}^n)$, that is for any $K\Subset U$ and α,β there exists C such that

$$\Big|\partial_x^\alpha\partial_\xi^\beta\Big(a(x,\xi)-\sum_{k=0}^{N-1}a_k(x,\xi)\Big)\Big|\leq C\langle\xi\rangle^{m-N-|\beta|}\quad\text{for all }(x,\xi)\in K\times\mathbb{R}^n. \tag{14.16}$$

Take $M\geq N+1$ such that $K\subset K_M$ and $|\alpha|,|\beta|\leq M$. We write

$$a(x,\xi)-\sum_{k=0}^{N-1}a_k(x,\xi)=\Big(a(x,\xi)-\sum_{k=0}^{M-1}a_k(x,\xi)\Big)+\sum_{k=N}^{M-1}a_k(x,\xi).$$

The first term on the right-hand side satisfies the bound (14.16) by (14.15). The second term lies in $S^{m-N}(U\times\mathbb{R}^n)$ and thus satisfies the bound (14.16) as well.

4. Finally, if $a,b\in S^m(U\times\mathbb{R}^n)$ are such that $a,b\sim\sum_{k=0}^{\infty}a_k$, then from (14.10) we see that $a-b\in S^{m-N}(U\times\mathbb{R}^n)$ for all N, which implies that $a-b\in S^{-\infty}(U\times\mathbb{R}^n)$. ☐

14.1.3. Pseudodifferential operators and quantization. We now develop a *quantization procedure* which lets us turn symbols in $S^m(U\times\mathbb{R}^n)$ into operators on functions on U. The term 'quantization' is used because this procedure is related to the map from classical to quantum observables in quantum mechanics.

Definition 14.10. *Let $U\Subset\mathbb{R}^n$, $m\in\mathbb{R}$, and $a\in S^m(U\times\mathbb{R}^n)$. For $\varphi\in C_c^\infty(U)$, define the function $\operatorname{Op}(a)\varphi:U\to\mathbb{C}$ as follows:*

$$\operatorname{Op}(a)\varphi(x)=(2\pi)^{-n}\int_{\mathbb{R}^n}e^{ix\cdot\xi}a(x,\xi)\widehat{\varphi}(\xi)\,d\xi,\quad x\in U. \tag{14.17}$$

where $\widehat{\varphi} \in \mathscr{S}(\mathbb{R}^n)$ *is the Fourier transform of the extension of* φ *by 0 to* $C_c^\infty(\mathbb{R}^n)$*, and the integral converges absolutely since* $a(x,\xi)$ *is polynomially bounded in* ξ *and* $\widehat{\varphi}$ *is rapidly decaying.*

An example of quantization, which justifies the prefactor $(2\pi)^{-n}$, is given by

PROPOSITION 14.11. $\mathrm{Op}(1)$ *is the identity operator* $C_c^\infty(U) \to C_c^\infty(U)$*.*

PROOF. This follows immediately from the Fourier Inversion Formula, namely Theorem 11.15. □

More generally, if a is a polynomial in ξ then $\mathrm{Op}(a)$ is a differential operator – see Exercise 14.3 below. This is the reason why operators of the form $\mathrm{Op}(a)$ are called *pseudodifferential operators*. (If $a \in S^0(U \times \mathbb{R}^n)$, pseudodifferential operators are also related to *singular integral operators* studied in harmonic analysis.)

Another example is given by symbols which depend only on ξ, in which case $\mathrm{Op}(a)$ is a convolution operator:

PROPOSITION 14.12. *Assume that* $a \in S^m(\mathbb{R}^n)$ *(see Definition 12.21) and consider* $a(x,\xi) = a(\xi)$ *as a symbol in* $S^m(\mathbb{R}^n \times \mathbb{R}^n)$*. Then*

$$\mathrm{Op}(a)\varphi = (\mathcal{F}^{-1}a) * \varphi \quad \textit{for all } \varphi \in C_c^\infty(\mathbb{R}^n)$$

where $\mathcal{F}^{-1}a \in \mathscr{S}'(\mathbb{R}^n)$ *is the inverse Fourier transform of* a*, defined in §11.2.2.*

PROOF. From (14.17) and using (11.40) we see that

$$\mathrm{Op}(a)\varphi = \mathcal{F}^{-1}(a\widehat{\varphi}) = \mathcal{F}^{-1}(\widehat{\mathcal{F}^{-1}(a)}\widehat{\varphi}) = \mathcal{F}^{-1}(a) * \varphi.$$

□

We now establish the basic mapping properties of the operator $\mathrm{Op}(a)$, starting with

PROPOSITION 14.13. *Assume that* $a \in S^m(U \times \mathbb{R}^n)$*. Then* $\mathrm{Op}(a)$ *is a sequentially continuous operator* $C_c^\infty(U) \to C^\infty(U)$*.*

PROOF.**S** Let $\varphi \in C_c^\infty(U)$. Since $\widehat{\varphi}(\xi) \in \mathscr{S}(\mathbb{R}^n)$ and all the x-derivatives of $a(x,\xi)$ are polynomially bounded in ξ, we can differentiate under the integral sign in (14.17) similarly to the proof of (11.11) to get that $\mathrm{Op}(a)\varphi \in C^\infty(U)$. The sequential continuity is straightforward to verify. □

We next extend $\mathrm{Op}(a)$ to a sequentially continuous operator on distributions:

$$\mathrm{Op}(a) : \mathcal{E}'(U) \to \mathcal{D}'(U). \tag{14.18}$$

To show existence of such an extension, we use Theorem 7.15. What one needs to show is that the transpose $\mathrm{Op}(a)^t$ is a sequentially continuous operator

$$\mathrm{Op}(a)^t : C_c^\infty(U) \to C^\infty(U) \tag{14.19}$$

and we leave this as an exercise below. Note that since $\mathrm{Op}(a) : C_c^\infty(U) \to C^\infty(U)$, the transpose $\mathrm{Op}(a)^t$ acts $\mathcal{E}'(U) \to \mathcal{D}'(U)$ as well.

Let us now consider the case when a lies in the residual class $S^{-\infty}$ defined in (14.4). In this case $\mathrm{Op}(a)$ is a smoothing operator:

Proposition 14.14. *Assume that $a \in S^{-\infty}(U \times \mathbb{R}^n)$. Then $\mathrm{Op}(a)$ extends to a sequentially continuous operator $\mathcal{E}'(U) \to C^\infty(U)$.*

Proof. Since $a(x, \xi)$ is rapidly decaying in ξ, by Fubini's Theorem we see that $\mathrm{Op}(a)$ is an integral operator:

$$\mathrm{Op}(a)\varphi(x) = \int_U \mathcal{K}_a(x, y)\varphi(y)\, dy \quad \text{for all } \varphi \in C_c^\infty(U),\ x \in U$$

where the Schwartz kernel $\mathcal{K}_a$ is given by

$$\mathcal{K}_a(x, y) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{i(x-y)\cdot\xi} a(x, \xi)\, d\xi, \quad x, y \in U. \tag{14.20}$$

Since $a(x, \xi)$ and all its x-derivatives are rapidly decaying in ξ, similarly to the proof of (11.11) we can differentiate under the integral sign to see that

$$\mathcal{K}_a \in C^\infty(U \times U).$$

Now Proposition 7.10 shows that $\mathrm{Op}(a)$ extends to a sequentially continuous operator $\mathcal{E}'(U) \to C^\infty(U)$. □

For general $a \in S^m(U \times \mathbb{R}^n)$, the operator $\mathrm{Op}(a)$ is not smoothing (see e.g. Proposition 14.11). However, it is *pseudolocal* in the following sense:

Proposition 14.15. *Let $a \in S^m(U \times \mathbb{R}^n)$. Then:*

1. If $\mathcal{K}_a \in \mathcal{D}'(U \times U)$ is the Schwartz kernel of $\mathrm{Op}(a)$ (see §7.2) then its singular support is contained in the diagonal:

$$\operatorname{sing\,supp} \mathcal{K}_a \subset \{(x, x) \mid x \in U\}. \tag{14.21}$$

2. We have for all $u \in \mathcal{E}'(U)$

$$\operatorname{sing\,supp}(\mathrm{Op}(a)u) \subset \operatorname{sing\,supp} u, \tag{14.22}$$

$$\operatorname{sing\,supp}(\mathrm{Op}(a)^t u) \subset \operatorname{sing\,supp} u. \tag{14.23}$$

Remark 14.16. *In the special case when $a = a(\xi)$, the operator $\mathrm{Op}(a)$ is the convolution operator with $E := \mathcal{F}^{-1}(a)$ as shown in Proposition 14.12. By Proposition 12.25 we know that $\operatorname{sing\,supp} E \subset \{0\}$, which gives the pseudolocality property by (8.19).*

Proof. We roughly follow the proof by Proposition 12.25, showing that for any k, if $|\alpha|$ is large enough depending on k then $(y-x)^\alpha\mathcal{K}_a(x,y)$ is in $C^k(U\times U)$.

1. Assume first that $a\in S^m(U\times\mathbb{R}^n)$ and $m<-n$. From (14.3) we see that the function $\xi\mapsto a(x,\xi)$ lies in $L^1(\mathbb{R}^n)$, with a norm bound locally uniform in x. By (14.17) and Fubini's Theorem, the Schwartz kernel $\mathcal{K}_a$ has the form (14.20):

$$\mathcal{K}_a(x,y)=(2\pi)^{-n}\int_{\mathbb{R}^n}e^{i(x-y)\cdot\xi}a(x,\xi)\,d\xi,\quad x,y\in U. \tag{14.24}$$

We have $\mathcal{K}_a\in C^0(U\times U)$ similarly to Proposition 11.2.

Next, if $k\in\mathbb{N}_0$ and $m<-n-k$ then we can differentiate k times in x under the integral sign in (14.24). Each differentiation of the integrand either gives one more power of ξ or differentiates a in x, so the integral still converges by (14.3). Differentiation under the integral sign is justified similarly to the proof of (11.11). This shows that

$$a\in S^m(U\times\mathbb{R}^n),\quad m<-n-k\quad\Longrightarrow\quad\mathcal{K}_a\in C^k(U\times U). \tag{14.25}$$

2. We next show the identity

$$(y_j-x_j)\mathcal{K}_a(x,y)=\mathcal{K}_{D_{\xi_j}a}(x,y). \tag{14.26}$$

For $a\in S^{-\infty}(U\times\mathbb{R}^n)$ this can be seen from (14.20) by integrating by parts:

$$\begin{aligned}(y_j-x_j)\mathcal{K}_a(x,y)&=(2\pi)^{-n}\int_{\mathbb{R}^n}(y_j-x_j)e^{i(x-y)\cdot\xi}a(x,\xi)\,d\xi\\&=-(2\pi)^{-n}\int_{\mathbb{R}^n}\big(D_{\xi_j}e^{i(x-y)\cdot\xi}\big)a(x,\xi)\,d\xi\\&=(2\pi)^{-n}\int_{\mathbb{R}^n}e^{i(x-y)\cdot\xi}D_{\xi_j}a(x,\xi)\,d\xi=\mathcal{K}_{D_{\xi_j}a}(x,y).\end{aligned}$$

For general $a\in S^m(U\times\mathbb{R}^n)$, from the definition (7.15) of the Schwartz kernel of an operator we see that $(y_j-x_j)\mathcal{K}_a(x,y)$ is the Schwartz kernel of the commutator $[\mathrm{Op}(a),x_j]$ where $x_j:C^\infty(U)\to C^\infty(U)$ is a multiplication operator; indeed, for all $\varphi,\psi\in C^\infty_{\mathrm{c}}(U)$ we have

$$\big((y_j-x_j)\mathcal{K}_a(x,y),\psi(x)\otimes\varphi(y)\big)=(\mathrm{Op}(a)x_j\varphi,\psi)-(\mathrm{Op}(a)\varphi,x_j\psi)=([\mathrm{Op}(a),x_j]\varphi,\psi).$$

Next, for all $\varphi\in C^\infty_{\mathrm{c}}(U)$ and $x\in U$ we compute

$$\begin{aligned}[\mathrm{Op}(a),x_j]\varphi(x)&=-(2\pi)^{-n}\int_{\mathbb{R}^n}e^{ix\cdot\xi}a(x,\xi)\big(D_{\xi_j}\widehat{\varphi}(\xi)+x_j\widehat{\varphi}(\xi)\big)\,d\xi\\&=-(2\pi)^{-n}\int_{\mathbb{R}^n}a(x,\xi)D_{\xi_j}\big(e^{ix\cdot\xi}\widehat{\varphi}(\xi)\big)\,d\xi\\&=(2\pi)^{-n}\int_{\mathbb{R}^n}e^{ix\cdot\xi}(D_{\xi_j}a(x,\xi))\widehat{\varphi}(\xi)\,d\xi=\mathrm{Op}_{D_{\xi_j}a}\,\varphi(x).\end{aligned}$$

Here the first equality follows from the definition (14.17) and the formula (11.11). In the third equality we integrate by parts, which is justified similarly to (11.10) since $\widehat{\varphi}(\xi)$ is rapidly decaying and $a(x,\xi)$ is polynomially bounded in ξ. Thus the operator $[\mathrm{Op}(a), x_j]$ has Schwartz kernel $\mathcal{K}_{D_{\xi_j}a}(x,y)$, giving (14.26).

3. Iterating (14.26), we see that for any multiindex β

$$(y-x)^\beta \mathcal{K}_a(x,y) = \mathcal{K}_{D_\xi^\beta a}(x,y). \tag{14.27}$$

Assume that $a \in S^m(U \times \mathbb{R}^n)$. Then $D_\xi^\beta a \in S^{m-|\beta|}$ by (14.7), we see from (14.25) that

$$|\beta| > m+n+k \quad \Longrightarrow \quad (y-x)^\beta \mathcal{K}_a(x,y) \in C^k(U \times U).$$

In particular, if we choose $N \in \mathbb{N}_0$ such that $2N > m+n+k$ then $|y-x|^{2N}\mathcal{K}_a(x,y) \in C^k(U \times U)$, which implies that $\mathcal{K}_a \in C^k(\{(x,y) \in U \times U \mid x \neq y\})$. Since k can be taken arbitrarily large, we see that $\mathcal{K}_a$ is smooth on $U \times U$ away from the diagonal, giving (14.21).

4. We now show that (14.21) implies (14.22). (The statement (14.23) follows in a similar way, since the Schwartz kernel of $\mathrm{Op}(a)^t$ is equal to $\mathcal{K}_a(y,x)$ by (7.21).) Let $u \in \mathcal{E}'(U)$ and $x_0 \in U \setminus \operatorname{sing\,supp} u$. We need to show that

$$x_0 \notin \operatorname{sing\,supp}(\mathrm{Op}(a)u). \tag{14.28}$$

Fix cutoff functions $\chi_1, \chi_2 \in C_c^\infty(U)$ such that

$$\chi_1(x_0) \neq 0, \quad \operatorname{supp}(1-\chi_2) \cap \operatorname{sing\,supp} u = \emptyset, \quad \operatorname{supp}\chi_1 \cap \operatorname{supp}\chi_2 = \emptyset.$$

We write

$$\chi_1 \mathrm{Op}(a)u = \chi_1 \mathrm{Op}(a)\chi_2 u + \chi_1 \mathrm{Op}(a)(1-\chi_2)u.$$

The Schwartz kernel of the operator $\chi_1 \mathrm{Op}(a)\chi_2$ is given by $(\chi_1(x) \otimes \chi_2(y))\mathcal{K}_a(x,y)$ which lies in $C_c^\infty(U \times U)$ since $\operatorname{supp}(\chi_1(x) \otimes \chi_2(y)) = \operatorname{supp}\chi_1 \times \operatorname{supp}\chi_2$ does not intersect the diagonal of U, which by (14.21) contains $\operatorname{sing\,supp}\mathcal{K}_a$. Thus by Proposition 7.10 we have

$$\chi_1 \mathrm{Op}(a)\chi_2 u \in C_c^\infty(U). \tag{14.29}$$

Next, we have $(1-\chi_2)u \in C_c^\infty(U)$, so by Proposition 14.13 we get

$$\chi_1 \mathrm{Op}(a)(1-\chi_2)u \in C_c^\infty(U). \tag{14.30}$$

Adding (14.29) and (14.30) we see that $\chi_1 \mathrm{Op}(a)u \in C_c^\infty(U)$, which implies (14.28) and finishes the proof. □

14.2. Proof of Elliptic Regularity III

14.2.1. Elliptic parametrix. The proof of Theorem 14.2, given below, relies on the existence of *elliptic parametrices* which is important in its own right:

THEOREM 14.17 (Elliptic parametrix). *Let $U \Subset \mathbb{R}^n$ and assume that $P \in \operatorname{Diff}^m(U)$ is an elliptic differential operator. Then there exist sequentially continuous operators*

$$Q, \widetilde{Q} : \mathcal{E}'(U) \to \mathcal{D}'(U)$$

such that

1. *$Q, \widetilde{Q}$ are sequentially continuous $C_c^\infty(U) \to C^\infty(U)$;*
2. *$Q, \widetilde{Q}$ are pseudolocal in the sense that*

$$\operatorname{sing\,supp}(Qu), \operatorname{sing\,supp}(\widetilde{Q}u) \subset \operatorname{sing\,supp} u \quad \text{for all } u \in \mathcal{E}'(U); \tag{14.31}$$

3. *the operators $I - PQ$ and $I - \widetilde{Q}P$ are smoothing in the sense that their Schwartz kernels are in $C^\infty(U \times U)$ and thus (by Proposition 7.10) they are sequentially continuous*

$$I - PQ, I - \widetilde{Q}P : \mathcal{E}'(U) \to C^\infty(U). \tag{14.32}$$

REMARK 14.18. *We call $Q, \widetilde{Q}$ right, respectively left, parametrices of P, where the word 'parametrix' stands for an explicitly constructed operator which is a (one-sided) inverse to P modulo smoothing operators. One could actually take $Q = \widetilde{Q}$ but we do not prove this here.*

Before proceeding with the proof of Theorem 14.17, we present one more property of pseudodifferential operators, computing the composition $P \operatorname{Op}(b)$ where P is a differential operator:

LEMMA 14.19. *Let $U \Subset \mathbb{R}^n$, $P \in \operatorname{Diff}^m(U)$, and $b \in S^\ell(U \times \mathbb{R}^n)$. Denote by $p := \sigma_m(P) \in S^m(U \times \mathbb{R}^n)$ the principal symbol of P. Then*

$$P \operatorname{Op}(b) = \operatorname{Op}(P\#b) \tag{14.33}$$

for some symbol $P\#b \in S^{m+\ell}(U \times \mathbb{R}^n)$, depending linearly on b, whose leading part is just the product pb:

$$P\#b - pb \in S^{m+\ell-1}(U \times \mathbb{R}^n). \tag{14.34}$$

REMARK 14.20. *In the special case when $b(x, \xi)$ is a polynomial in ξ and thus $\operatorname{Op}(b)$ is a differential operator, Lemma 14.19 follows from the Product Rule in Proposition 13.10. See Exercise 14.5 below for an explicit example of the computation of $P\#b$.*

PROOF. Similarly to Proposition 14.13, we can differentiate under the integral sign in (14.17) to get for all $\varphi \in C_c^\infty(U)$

$$P\,\mathrm{Op}(b)\varphi(x) = (2\pi)^{-n}\int_{\mathbb{R}^n} P(e^{ix\cdot\xi}b(x,\xi))\widehat{\varphi}(\xi)\,d\xi. \tag{14.35}$$

Now, similarly to (13.54) we see that, writing $P = \sum_{|\alpha|\le m} a_\alpha(x)D_x^\alpha$,

$$\begin{aligned} P(e^{ix\cdot\xi}b(x,\xi)) &= e^{ix\cdot\xi}(P\#b)(x,\xi) \\ \text{where } (P\#b)(x,\xi) &= \sum_{|\alpha|\le m} a_\alpha(x)(D_x+\xi)^\alpha b(x,\xi). \end{aligned} \tag{14.36}$$

Here for a multiindex $\alpha = (\alpha_1,\dots,\alpha_n)$, we define $(D_x+\xi)^\alpha = (D_{x_1}+\xi_1)^{\alpha_1}\dots(D_{x_n}+\xi_n)^{\alpha_n}$. By the Leibniz Rule we get the formula (where $\alpha! = \alpha_1!\cdots\alpha_n!$)

$$P\#b(x,\xi) = \sum_{|\alpha|\le m} a_\alpha(x)\sum_{\beta+\gamma=\alpha}\frac{\alpha!}{\beta!\gamma!}\xi^\beta D_x^\gamma b(x,\xi). \tag{14.37}$$

By (14.6) and (14.7) we have $\xi^\beta D_x^\gamma b \in S^{\ell+|\beta|}(U\times\mathbb{R}^n)$. It follows that $P\#b \in S^{m+\ell}(U\times\mathbb{R}^n)$. Moreover, the term in the sum (14.37) corresponding to $\beta+\gamma=\alpha$ is in $S^{m+\ell-1}(U\times\mathbb{R}^n)$ unless $|\alpha| = m$ and $\beta = \alpha$. The terms with $|\alpha| = m$ and $\beta = \alpha$ together give

$$\sum_{|\alpha|=m} a_\alpha(x)\xi^\alpha b(x,\xi) = p(x,\xi)b(x,\xi)$$

which shows (14.34). Finally, (14.35) and (14.36) show that $P\,\mathrm{Op}(b)\varphi = \mathrm{Op}(P\#b)\varphi$ for all $\varphi \in C_c^\infty(U)$, which (since $C_c^\infty(U)$ is dense in $\mathcal{E}'(U)$) implies that $P\,\mathrm{Op}(b) = \mathrm{Op}(P\#b)$ as operators $\mathcal{E}'(U) \to \mathcal{D}'(U)$. □

We are now ready for

PROOF OF THEOREM 14.17. 1. We first construct the operator Q, taking it in the form

$$Q := \mathrm{Op}(q) \quad \text{for some } q \in S^{-m}(U\times\mathbb{R}^n).$$

Any such Q is sequentially continuous $C_c^\infty(U) \to C^\infty(U)$ (by Proposition 14.13) as well as $\mathcal{E}'(U) \to \mathcal{D}'(U)$ (by (14.18)), and it is pseudolocal (by Proposition 14.15). Thus it remains to construct q such that $I - PQ$ is a smoothing operator.

By Lemma 14.19 and since $I = \mathrm{Op}(1)$ by Proposition 14.11, we see that

$$I - PQ = \mathrm{Op}(1 - P\#q).$$

By Proposition 14.14, $I - PQ$ is a smoothing operator if we construct q such that

$$1 - P\#q \in S^{-\infty}(U\times\mathbb{R}^n). \tag{14.38}$$

Henceforth in this proof we denote the spaces $S^\bullet(U\times\mathbb{R}^n)$ by just $S^\bullet$.

2. We construct a solution q to (14.38) by an iteration procedure. We start by finding

$$q_0 \in S^{-m} \quad \text{such that } r_1 := 1 - P\#q_0 \in S^{-1}. \tag{14.39}$$

Denote by $p := \sigma_m(P) \in S^m$ the principal symbol of P. By (14.34) we see that $P\#q_0 - pq_0 \in S^{-1}$. Therefore (14.39) is equivalent to

$$1 - pq_0 \in S^{-1}. \tag{14.40}$$

It is time for us to use the fact that P is an elliptic operator. The principal symbol $p(x,\xi)$ is a homogeneous polynomial of degree m in ξ, and since P is elliptic (recalling Definition 14.1), we have $p(x,\xi) \neq 0$ for all $x \in U$ and $\xi \neq 0$. Similarly to (12.51), we see that for each $K \Subset U$ there exists $c_K > 0$ such that

$$|p(x,\xi)| \geq c_K |\xi|^m \quad \text{for all } (x,\xi) \in K \times \mathbb{R}^n. \tag{14.41}$$

Take any function

$$q_0 \in C^\infty(U \times \mathbb{R}^n), \quad q_0(x,\xi) = \frac{1}{p(x,\xi)} \quad \text{for all } x \in U,\ \xi \in \mathbb{R}^n,\ |\xi| \geq 1. \tag{14.42}$$

For example, we can put $q_0(x,\xi) := (1-\chi(\xi))/p(x,\xi)$ where $\chi \in C_c^\infty(B^\circ(0,1))$ satisfies $\chi = 1$ near 0. By Proposition 14.6 we see that (14.41) implies

$$q_0 \in S^{-m}. \tag{14.43}$$

The symbol $1 - pq_0 \in C^\infty(U \times \mathbb{R}^n)$ is supported in $\{|\xi| \leq 1\}$ and thus lies in $S^{-\infty}$. It follows that q_0 solves (14.40) and thus (14.39).

3. We next add a correction term to q_0 to improve the remainder in (14.39) from S^{-1} to S^{-2}. More precisely, we construct

$$q_1 \in S^{-m-1} \quad \text{such that } r_2 := 1 - P\#(q_0 + q_1) \in S^{-2}. \tag{14.44}$$

Let $r_1 \in S^{-1}$ be the remainder term from (14.39) and q_0 be defined in (14.42). We put

$$q_1 := q_0 r_1, \quad q_1(x,\xi) = \frac{r_1(x,\xi)}{p(x,\xi)} \quad \text{for } x \in U,\ |\xi| \geq 1. \tag{14.45}$$

From (14.43) and (14.6) we have $q_1 \in S^{-m-1}$. Then by (14.34) we have $P\#q_1 - pq_1 \in S^{-2}$. Since $pq_1 - r_1 \in S^{-\infty}$ by (14.45), we get

$$1 - P\#(q_0 + q_1) = r_1 - P\#q_1 \in S^{-2},$$

giving (14.44).

4. Iterating Step 3 of the proof, we construct symbols

$$q_k \in S^{-m-k}, \quad k \in \mathbb{N}_0,$$

such that for all $k \in \mathbb{N}_0$

$$r_{k+1} := 1 - P\# \sum_{\ell=0}^{k} q_\ell \ \in \ S^{-k-1}. \tag{14.46}$$

Indeed, q_0 and q_1 were already constructed in the previous two steps of the proof. If $k \geq 2$, then we let r_k be defined by (14.46) for $k-1$ and put

$$q_k := q_0 r_k \ \in \ S^{-m-k}. \tag{14.47}$$

Arguing as in Step 3 above, we see that (14.46) holds for k.

5. Having constructed all the symbols q_k, we use Borel's Theorem 14.9 to see that there exists

$$q \in S^{-m}, \quad q \sim \sum_{k=0}^{\infty} q_k.$$

We claim that q solves (14.38). Indeed, for any $N \in \mathbb{N}$ we have $q - \sum_{k=0}^{N-1} q_k \in S^{-m-N}$. Thus by Lemma 14.19 we have $P\#(q - \sum_{k=0}^{N-1} q_k) \in S^{-N}$. By (14.46) we have $1 - P\#\sum_{k=0}^{N-1} q_k \in S^{-N}$ as well. It follows that $1 - P\#q \in S^{-N}$. Since this is true for all N, we get $1 - P\#q \in S^{-\infty}$ as needed. This finishes the construction of the operator Q satisfying the conclusions in the theorem.

6. It remains to construct the operator $\widetilde{Q}$. Let $P^t \in \operatorname{Diff}^m(U)$ be the transpose of P. By (9.3) we get the following formula for the principal symbol of P^t:

$$\sigma_m(P^t)(x,\xi) = \sigma_m(P)(x,-\xi) = (-1)^m \sigma_m(P)(x,\xi). \tag{14.48}$$

Since P is elliptic, we see that P^t is also elliptic. Thus by Steps 1–5 above there exists $\widetilde{q} \in S^{-m}(U \times \mathbb{R}^n)$ such that $I - P^t \operatorname{Op}(\widetilde{q})$ is smoothing. Define $\widetilde{Q}$ to be the transpose of $\operatorname{Op}(\widetilde{q}) : C_c^\infty(U) \to C^\infty(U)$:

$$\widetilde{Q} := \operatorname{Op}(\widetilde{q})^t : \mathcal{E}'(U) \to \mathcal{D}'(U). \tag{14.49}$$

Then $\widetilde{Q}$ also maps $C_c^\infty(U) \to C^\infty(U)$ by (14.19), and it is pseudolocal by Proposition 14.15. We now have

$$I - \widetilde{Q}P = (I - P^t \operatorname{Op}(\widetilde{q}))^t.$$

Since $I - P^t \operatorname{Op}(\widetilde{q})$ is smoothing, so is $I - \widetilde{Q}P$ (see the beginning of Step 4 of the proof of Proposition 14.15). □

Remark 14.21.[X] *It is possible to get by without invoking Borel's Theorem. Indeed, for each $\ell \in \mathbb{N}_0$ we can take $q := \sum_{k=0}^{N} q_k$ in Theorem 14.17 for N large enough depending on ℓ to get the operator $I - P\operatorname{Op}(q)$ to have Schwartz kernel in C^ℓ. Since ℓ can be taken arbitrarily large, this is enough to show elliptic regularity. However, it is conceptually cleaner to construct a single q such that $I - P\operatorname{Op}(q)$ is smoothing.*

Remark 14.22.[X] *Theorem 14.17 features both a right parametrix Q and a left parametrix $\widetilde{Q}$. The reason is as follows: it is easier for us to construct a right parametrix because of the formula for $P\operatorname{Op}(q)$ given in Lemma 14.19. An analog of this lemma for $\operatorname{Op}(q)P$ is harder to prove, so we construct the left parametrix $\widetilde{Q}$ as the transpose of the right parametrix for the operator P^t. However, the proof of elliptic regularity in the next subsection needs a left parametrix. This is a technical point made* *necessary by our refusal to develop a proper calculus of pseudodifferential operators (which would in particular show that the transpose* $\mathrm{Op}(q)^t$ *has the form* $\mathrm{Op}(q^t)$ *for some symbol* q^t*, modulo a smoothing operator).*

14.2.2. Proof of Elliptic Regularity. We are now ready to give the proof of Theorem 14.2. We have $\operatorname{sing\,supp}(Pu) \subset \operatorname{sing\,supp} u$ for any differential operator P, so we need to show that $\operatorname{sing\,supp} u \subset \operatorname{sing\,supp}(Pu)$.

1. We first note that it is enough to consider the setting of differential operators on open subsets of $\mathbb{R}^n$. Indeed, assume that $\mathcal{M}$ is a general manifold, $P \in \mathrm{Diff}^m(\mathcal{M})$ is an elliptic differential operator, $u \in \mathcal{D}'(\mathcal{M})$, and $x_0 \in \mathcal{M}$ satisfies $x_0 \notin \operatorname{sing\,supp}(Pu)$. Take a chart $\varkappa : U_0 \to V_0$ such that $x_0 \in U_0$, and let $\varkappa_* u \in \mathcal{D}'(V_0)$ be the pushforward defined in (13.42). In terms of the operator $\varkappa_* P$ defined in (13.59), we have

$$(\varkappa_* P)(\varkappa_* u) = \varkappa_*(Pu).$$

Since $x_0 \notin \operatorname{sing\,supp}(Pu)$, we have $\varkappa(x_0) \notin \operatorname{sing\,supp} \varkappa_*(Pu)$. The operator $\varkappa_* P \in \mathrm{Diff}^m(V_0)$ is elliptic as follows from (13.60). Therefore, the version of Theorem 14.2 for $V_0 \Subset \mathbb{R}^n$ shows that $\varkappa(x_0) \notin \operatorname{sing\,supp}(\varkappa_* u)$, which implies that $x_0 \notin \operatorname{sing\,supp} u$.

2. From now on we assume that $U \Subset \mathbb{R}^n$, $P \in \mathrm{Diff}^m(U)$ is an elliptic differential operator, $u \in \mathcal{D}'(U)$, and $x_0 \in U$ satisfies $x_0 \notin \operatorname{sing\,supp}(Pu)$. We need to show that $x_0 \notin \operatorname{sing\,supp} u$.

Fix a cutoff function

$$\chi \in C_c^\infty(U), \quad x_0 \notin \operatorname{supp}(1-\chi).$$

Let $\widetilde{Q}$ be the left elliptic parametrix constructed in Theorem 14.17. Then by (14.32)

$$I = \widetilde{Q}P + \widetilde{R} \quad \text{where } \widetilde{R} : \mathcal{E}'(U) \to C^\infty(U). \tag{14.50}$$

Applying this to $\chi u \in \mathcal{E}'(U)$, we get

$$\chi u = \widetilde{Q}P\chi u + \widetilde{R}\chi u. \tag{14.51}$$

By the pseudolocality property (14.31) for $\widetilde{Q}$, and since $\widetilde{R}\chi u \in C^\infty(U)$, we see that

$$\operatorname{sing\,supp}(\chi u) \subset \operatorname{sing\,supp}(P\chi u).$$

Similarly to the proof of Theorem 9.14, since $P\chi u = \chi P u + [P,\chi]u$ and $x_0 \notin \operatorname{sing\,supp}(Pu)$, $x_0 \notin \operatorname{supp}[P,\chi]u$, we see that $x_0 \notin \operatorname{sing\,supp}(P\chi u)$. Thus $x_0 \notin \operatorname{sing\,supp}(\chi u)$, which implies that $x_0 \notin \operatorname{sing\,supp} u$, finishing the proof.

14.2.3. The case of vector bundles. We finally give the analog of Theorem 14.2 for operators acting on vector bundles:

THEOREM 14.23 (Elliptic Regularity III′). *Assume that* $\mathcal{M}$ *is a manifold,* $\mathscr{E}, \mathscr{F}$ *are (complex) vector bundles of the same dimension over* $\mathcal{M}$*, and* $\mathbf{P} \in \mathrm{Diff}^m(\mathcal{M}; \mathscr{E} \to \mathscr{F})$ *is a differential operator (see §13.3.3). Assume that* $\mathbf{P}$ *is elliptic in the following sense:*

for each $(x,\xi) \in T^*\mathcal{M}$ *with* $\xi \neq 0$*, the principal symbol* $\sigma_m(\mathbf{P})(x,\xi) : \mathscr{E}(x) \to \mathscr{F}(x)$ *is a linear isomorphism. Then we have for all* $\mathbf{u} \in \mathcal{D}'(\mathcal{M};\mathscr{E})$

$$\operatorname{sing\,supp} \mathbf{u} = \operatorname{sing\,supp}(\mathbf{P}\mathbf{u}). \tag{14.52}$$

The proof of Theorem 14.23 is similar to that of Theorem 14.2, so we just give a brief outline here. Similarly to Step 1 in §14.2.2, we can reduce to the case when $\mathcal{M} = U \Subset \mathbb{R}^n$ and $\mathscr{E} = \mathscr{F} = U \times \mathbb{C}^\ell$ are trivial vector bundles. The operator $\mathbf{P}$ is then given by a matrix of differential operators (see (13.66)):

$$\mathbf{P} = (P_{jj'} \in \operatorname{Diff}^m(U))_{j,j'=1}^{\ell}.$$

Following the proof of Theorem 14.17, we construct pseudolocal operators

$$\mathbf{Q}, \widetilde{\mathbf{Q}} : C_c^\infty(U;\mathbb{C}^\ell) \to C^\infty(U;\mathbb{C}^\ell),\ \mathcal{E}'(U;\mathbb{C}^\ell) \to \mathcal{D}'(U;\mathbb{C}^\ell)$$

such that $I - \mathbf{P}\mathbf{Q}$, $I - \widetilde{\mathbf{Q}}\mathbf{P}$ are smoothing:

$$I - \mathbf{P}\mathbf{Q}, I - \widetilde{\mathbf{Q}}\mathbf{P} : \mathcal{E}'(U;\mathbb{C}^\ell) \to C^\infty(U;\mathbb{C}^\ell). \tag{14.53}$$

The operator $\mathbf{Q}$ is a matrix of pseudodifferential operators. More precisely, we construct a matrix of symbols

$$\mathbf{q} = (q_{jj'})_{j,j'=1}^{\ell}, \quad q_{jj'} \in S^{-m}(U \times \mathbb{R}^n),$$

and put $\mathbf{Q} = \operatorname{Op}(\mathbf{q})$ where

$$\operatorname{Op}(\mathbf{q}) := (\operatorname{Op}(q_{jj'}))_{j,j'=1}^{\ell} : \mathcal{E}'(U;\mathbb{C}^\ell) \to \mathcal{D}'(U;\mathbb{C}^\ell).$$

The matrix-valued symbol $\mathbf{q} \in S^{-m}(U \times \mathbb{R}^n; \operatorname{Hom}(\mathbb{C}^\ell \to \mathbb{C}^\ell))$ is constructed as an asymptotic sum:

$$\mathbf{q} \sim \sum_{k=0}^{\infty} \mathbf{q}_k, \quad \mathbf{q}_k \in S^{-m-k}(U \times \mathbb{R}^n; \operatorname{Hom}(\mathbb{C}^\ell \to \mathbb{C}^\ell))$$

where for each $k \in \mathbb{N}_0$ we have

$$I - \mathbf{P}\operatorname{Op}\left(\sum_{s=0}^{k} \mathbf{q}_k\right) = \operatorname{Op}(\mathbf{r}_{k+1}) \quad \text{for some } \mathbf{r}_{k+1} \in S^{-k-1}. \tag{14.54}$$

Let $\mathbf{p} := \sigma_m(\mathbf{P}) \in S^m(U \times \mathbb{R}^n; \operatorname{Hom}(\mathbb{C}^\ell \to \mathbb{C}^\ell))$ be the principal symbol of $\mathbf{P}$. From Lemma 14.19 we see that

$$\mathbf{P}\operatorname{Op}(\mathbf{a}) = \operatorname{Op}(\mathbf{P}\#\mathbf{a}) \quad \text{for any } \mathbf{a} \in S^r(U \times \mathbb{R}^n; \operatorname{Hom}(\mathbb{C}^\ell \to \mathbb{C}^\ell))$$
$$\text{where } \mathbf{P}\#\mathbf{a} \in S^{m+r}(U \times \mathbb{R}^n; \operatorname{Hom}(\mathbb{C}^\ell \to \mathbb{C}^\ell)), \quad \mathbf{P}\#\mathbf{a} - \mathbf{p}\mathbf{a} \in S^{m+r-1}$$

and $\mathbf{pa}$ is defined using multiplication of $\ell \times \ell$ matrices.

Now, in Step 2 of the proof of Theorem 14.17 in place of (14.42) we should take $\mathbf{q}_0(x,\xi) \in C^\infty(U \times \mathbb{R}^n; \operatorname{Hom}(\mathbb{C}^\ell \to \mathbb{C}^\ell))$ to be the (matrix) inverse of $\mathbf{p}$:

$$\mathbf{q}_0(x,\xi) = \mathbf{p}(x,\xi)^{-1} : \mathbb{C}^\ell \to \mathbb{C}^\ell, \quad x \in U,\ |\xi| \geq 1.$$

Since $\mathbf{P}$ is elliptic and thus $\mathbf{p}(x,\xi) : \mathbb{C}^\ell \to \mathbb{C}^\ell$ is invertible for $\xi \neq 0$, one can follow the proof of Proposition 14.6 to see that $\mathbf{q}_0 \in S^{-m}$, where (14.9) now looks more complicated and features matrix multiplication. (Alternatively one can apply Cramer's Rule and use that the scalar symbol $\det \mathbf{p}(x,\xi)$ is homogeneous of degree $m\ell$ in ξ and nonvanishing for $\xi \neq 0$.)

Next, Steps 3–4 of the proof of Theorem 14.17 adapt to the setting of matrix-valued symbols to construct the symbols $\mathbf{q}_1, \mathbf{q}_2, \dots$ such that (14.54) holds. Here in place of (14.47) we put $\mathbf{q}_k := \mathbf{q}_0 \mathbf{r}_k$, defined by matrix multiplication. Step 5 of the proof applies as well, showing that $\mathbf{Q} := \operatorname{Op}(\mathbf{q})$ satisfies (14.53), and Step 6 works as before to construct $\widetilde{\mathbf{Q}}$.

Finally, Step 2 of the proof of Theorem 14.2 in §14.2.2 applies (with $\widetilde{\mathbf{Q}}$ taking the place of $\widetilde{Q}$) to give the conclusion of Theorem 14.23.

14.3. Notes and exercises

The modern theory of pseudodifferential operators, in the form quite similar to what we present in §14.1, goes back to the work of Kohn and Nirenberg [**KN65**]. See also the slightly later paper of Hörmander [**Hör65**] which shows coordinate invariance of pseudodifferential operators and has a form of elliptic regularity [**Hör65**, Theorem 4.7] identical to Theorem 15.1 below.

The theory of pseudodifferential operators has at least two precursors: the theory of singular integral operators, which are essentially pseudodifferential operators whose symbols are homogeneous of degree 0 in ξ, and the theory of quantum/classical correspondence in quantum mechanics developed in the early XXth century (in particular, by Hermann Weyl who introduced Weyl quantization, which is an alternative to (14.17)). See the introduction to [**KN65**] and the notes to [**Hör07**, Chapter 18] for more on the history of the subject.

We treat pseudodifferential operators here as a means to an end and prove the bare minimum needed for the proof of Theorem 14.2. A proper treatment of pseudodifferential operators (including the analogues of Propositions 13.10 and 13.14 and the notion of pseudodifferential operator on a manifold) is a part of the field called *microlocal analysis* (in MIT, it is taught in 18.157). A curious reader is welcome to look at [**Hör07**, Section 18.1] or [**GS94**, Chapter 3] for a comprehensive introduction to pseudodifferential calculus.

The presentation in this chapter was partially inspired by [**Mel**, Chapter 4].

Exercise 14.1. ($2 = 1 + 0.5 + 0.5$ pts) *Let $U \Subset \mathbb{R}^n$ and $m \in \mathbb{R}$.*

(a) *Assume that $a \in C^\infty(U \times \mathbb{R}^n)$. For $t \geq 1$, define the dilated function*

$$\Lambda_t a \in C^\infty(U \times \mathbb{R}^n), \quad \Lambda_t a(x,\xi) = a(x, t\xi).$$

Show that $a \in S^m(U \times \mathbb{R}^n)$ if and only if for all $\widetilde{K} \Subset U \times (\mathbb{R}^n \setminus \{0\})$ and α, β there exists a constant $C_{\alpha\beta\widetilde{K}}$ such that

$$|\partial_x^\alpha \partial_\xi^\beta (\Lambda_t a)(x,\xi)| \leq C_{\alpha\beta\widetilde{K}} t^m \quad \text{for all } (x,\xi) \in \widetilde{K},\ t \geq 1. \tag{14.55}$$

(This can be used to show that the class S^m is invariant under changes of variables appearing in (13.56) and thus one can define invariantly the class $S^m(T^*\mathcal{M})$ where $\mathcal{M}$ is a manifold. It can also be used to give an alternative proof of Proposition 14.6.)

(b) *Assume that $a \in C^\infty(U \times \mathbb{R}^n)$ has the following homogeneity property:*

$$a(x, t\xi) = t^m a(x,\xi) \quad \text{for all } x \in U,\ |\xi| \geq 1,\ t \geq 1. \tag{14.56}$$

Show that $a \in S^m(U \times \mathbb{R}^n)$.

(c) *Let $\langle \xi \rangle$ be defined in (12.3). Show that the function $a(\xi) := \langle \xi \rangle^m$ lies in $S^m(\mathbb{R}^n)$.*

Exercise 14.2. (1 = 0.5+0.5 pt) *In this exercise you show the following version of Borel's Theorem 14.9: for any sequence $a_k \in \mathbb{C}$, $k = 0, 1, \dots,$ there exists $f \in C^\infty(\mathbb{R})$ such that $f^{(k)}(0)/k! = a_k$ for all k.*

(a) *Fix $\chi \in C_c^\infty(\mathbb{R})$ such that $\chi = 1$ near 0. Show that there exists a sequence $\varepsilon_k > 0$, $k = 0, 1, \dots,$ such that $\varepsilon_k \to 0$ and*

$$\max_{0 \leq j < k} \sup_x |\partial_x^j g_k(x)| \leq 2^{-k} \quad \text{where} \quad g_k(x) := \chi\Big(\frac{x}{\varepsilon_k}\Big) a_k x^k.$$

(b) *Show that the series*

$$f(x) := \sum_{k=0}^\infty g_k(x)$$

converges in $C_c^j(\mathbb{R})$ for every j to a function $f \in C_c^\infty(\mathbb{R})$ and $f^{(j)}(0)/j! = a_j$ for all j.

Exercise 14.3. (1 pt) *Assume that $a(x,\xi) = \sum_{|\alpha| \leq m} a_\alpha(x) \xi^\alpha$ is a polynomial of degree m in ξ with coefficients $a_\alpha(x)$ which are smooth functions on $U \mathbin{\text{⋐}} \mathbb{R}^n$. Show that $\operatorname{Op}(a)$ is a differential operator:*

$$\operatorname{Op}(a)\varphi(x) = \sum_{|\alpha| \leq m} a_\alpha(x) D_x^\alpha \varphi(x).$$

Exercise 14.4. (1 pt) *Show that if $a \in S^m(U \times \mathbb{R}^n)$, then $\operatorname{Op}(a)^t : \mathcal{E}'(U) \to \mathcal{D}'(U)$ restricts to a sequentially continuous operator $C_c^\infty(U) \to C^\infty(U)$, giving (14.19). (Hint: write $\operatorname{Op}(a)^t \varphi = \widehat{B\varphi}$ where B is a certain integral operator. Then show that if $\varphi \in C_c^\infty(U)$ then $B\varphi(\xi) = \mathcal{O}(\langle \xi \rangle^{-\infty})$, either by using Fourier transform or directly by repeated integration by parts.)*

Exercise 14.5. (1.5 = 0.5+1 pts) *This exercise carries out the elliptic parametrix construction (Theorem 14.17) for a one-dimensional Schrödinger operator on $\mathbb{R}$*

$$P := -\partial_x^2 + V(x) = D_x^2 + V(x) \quad \text{where } V \in C^\infty(\mathbb{R}). \tag{14.57}$$

(a) *Let $b \in S^m(\mathbb{R} \times \mathbb{R})$. Show that the symbol $P\#b$ from Lemma 14.19 is given by*

$$P\#b(x,\xi) = \xi^2 b(x,\xi) + 2\xi D_x b(x,\xi) + (D_x^2 + V(x))b(x,\xi).$$

(b) *Find $q \in S^{-2}(\mathbb{R} \times \mathbb{R})$ such that $1 - P\#q \in S^{-4}(\mathbb{R} \times \mathbb{R})$. Please give an explicit formula for $q(x,\xi)$ for $|\xi| \geq 1$ and do not use Borel's Theorem.*

CHAPTER 15

Elliptic operators and Sobolev spaces

In the previous chapter we established regularity for elliptic differential operators $P \in \mathrm{Diff}^m$ (Theorem 14.2): if u is a distribution and $Pu \in C^\infty$ then $u \in C^\infty$ as well. We now turn our attention to elliptic operators acting on Sobolev spaces.

We first show an analog of Theorem 12.20: if $Pu \in H^s_{\mathrm{loc}}$ then $u \in H^{s+m}_{\mathrm{loc}}$ (Theorem 15.1). We also prove the corresponding *elliptic estimate* (Proposition 15.6, Theorem 15.7). We next show the Rellich–Kondrachov Theorem on compact embeddings of Sobolev spaces (Theorems 15.8, 15.10).

We finally restrict to the case of compact manifolds and use the elliptic estimate and the Rellich–Kondrachov Theorem to show that $P : H^{s+m} \to H^s$ is a Fredholm operator (Theorem 15.13). The Fredholm mapping property means that P is invertible modulo finite dimensional spaces; under additional assumptions one can show that P is invertible, which means that the problem $Pu = f$ is well-posed in Sobolev spaces. One of the consequences of the Fredholm property is that one can study the index of P, which can be computed by the Atiyah–Singer index theorem – we mention this at the end of this chapter but do not state the theorem itself.

15.1. Elliptic regularity in Sobolev spaces

Let $\mathcal{M}$ be a manifold and $P \in \mathrm{Diff}^m(\mathcal{M})$ be a differential operator (see §13.3.2). Then P defines a sequentially continuous operator on the local/compactly supported Sobolev spaces from §13.2.3:

$$P : H^{s+m}_{\mathrm{loc}}(\mathcal{M}) \to H^s_{\mathrm{loc}}(\mathcal{M}),\ H^{s+m}_{\mathrm{c}}(\mathcal{M}) \to H^s_{\mathrm{c}}(\mathcal{M}) \quad \text{for all } s \in \mathbb{R}. \tag{15.1}$$

This follows from Definition 13.8 and the fact that for each chart $\varkappa : U \to V$ on $\mathcal{M}$, the pushforward $\varkappa_* P$ is an order m differential operator on V and thus is sequentially continuous $H^{s+m}_{\mathrm{loc}}(V) \to H^s_{\mathrm{loc}}(V)$ by Proposition 12.13.

In this section we assume that P is elliptic and establish various regularity results on P in Sobolev spaces. The simplest one to state is

Theorem 15.1. *Assume that $P \in \mathrm{Diff}^m(\mathcal{M})$ is an elliptic differential operator. Then we have for all $u \in \mathcal{D}'(\mathcal{M})$ and $s \in \mathbb{R}$*

$$Pu \in H^s_{\mathrm{loc}}(\mathcal{M}) \quad \Longrightarrow \quad u \in H^{s+m}_{\mathrm{loc}}(\mathcal{M}). \tag{15.2}$$

As an example, if g is a Riemannian metric on $\mathcal{M}$ and Δ_g is the corresponding Laplace–Beltrami operator, then $\Delta_g u \in L^2_{\mathrm{loc}}(\mathcal{M})$ implies that $u \in H^2_{\mathrm{loc}}(\mathcal{M})$.

REMARK 15.2. *Theorem 15.1, as well as Theorems 15.7 and 15.13 below, applies also to elliptic differential operators acting on sections of vector bundles, defined in §14.2.3. The proofs are exactly the same (generalizing the elliptic parametrix construction to the setting of matrices of operators as explained in §14.2.3), so to keep notation simple we state the results in the scalar setting.*

15.1.1. Pseudodifferential operators acting on Sobolev spaces. To show Theorem 15.1, we revisit the proof of Elliptic Regularity III in §14.2.2. We need to show that the operators $Q, \widetilde{Q}$ constructed in Theorem 14.17 have the right mapping properties on Sobolev spaces; this is given by

PROPOSITION 15.3 (Continuity of pseudodifferential operators on Sobolev spaces). *Assume that $U \Subset \mathbb{R}^n$, $s, m \in \mathbb{R}$, and $a \in S^m(U \times \mathbb{R}^n)$. Let $\mathrm{Op}(a), \mathrm{Op}(a)^t : \mathcal{E}'(U) \to \mathcal{D}'(U)$ be defined in §14.1.3. Then $\mathrm{Op}(a)$ and its transpose restrict to sequentially continuous operators*

$$\mathrm{Op}(a), \mathrm{Op}(a)^t : H^{s+m}_{\mathrm{c}}(U) \to H^s_{\mathrm{loc}}(U). \tag{15.3}$$

REMARK 15.4. *A special case of Proposition 15.3 is when $m \in \mathbb{N}_0$ and $a(x,\xi)$ is a polynomial in ξ, so that $\mathrm{Op}(a)$ is a differential operator – see Exercise 14.3 and (15.1). In particular, for $m = 0$ we get continuity of multiplication by smooth functions on Sobolev spaces, which was previously established as a corollary of Proposition 12.9.*

Our proof of Proposition 15.3 follows the scheme of proof of Proposition 12.9 above. We first establish the following analog of Young's convolution inequality (Lemma 12.10):

LEMMA 15.5 (Schur's bound). *Assume that $\mathcal{K} \in L^\infty(\mathbb{R}^{2n})$ and define the integral operator $A : L^1(\mathbb{R}^n) \to L^\infty(\mathbb{R}^n)$ by*

$$Af(\xi) = \int_{\mathbb{R}^n} \mathcal{K}(\xi, \eta) f(\eta)\, d\eta, \quad f \in L^1(\mathbb{R}^n), \quad \xi \in \mathbb{R}^n. \tag{15.4}$$

Assume that the following constants are finite:

$$C_1 := \sup_{\xi \in \mathbb{R}^n} \int_{\mathbb{R}^n} |\mathcal{K}(\xi, \eta)|\, d\eta, \quad C_2 := \sup_{\eta \in \mathbb{R}^n} \int_{\mathbb{R}^n} |\mathcal{K}(\xi, \eta)|\, d\xi. \tag{15.5}$$

Then we have for any $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$

$$\|Af\|_{L^2(\mathbb{R}^n)} \le \sqrt{C_1 C_2}\, \|f\|_{L^2(\mathbb{R}^n)}. \tag{15.6}$$

PROOF. For any $\xi \in \mathbb{R}^n$, we estimate

$$\begin{aligned}|Af(\xi)|^2 &= \Big|\int_{\mathbb{R}^n} \mathcal{K}(\xi,\eta) f(\eta)\,d\eta\Big|^2\\ &\le \Big(\int_{\mathbb{R}^n} |\mathcal{K}(\xi,\eta)|\,d\eta\Big)\Big(\int_{\mathbb{R}^n} |\mathcal{K}(\xi,\eta)|\cdot|f(\eta)|^2\,d\eta\Big)\\ &\le C_1\int_{\mathbb{R}^n} |\mathcal{K}(\xi,\eta)|\cdot|f(\eta)|^2\,d\eta.\end{aligned}$$

Here in the second line we write $|\mathcal{K}(\xi,\eta)f(\eta)| = \sqrt{|\mathcal{K}(\xi,\eta)|}\cdot(\sqrt{|\mathcal{K}(\xi,\eta)|}\cdot|f(\eta)|)$ and use Cauchy–Schwarz. Integrating in ξ and using Fubini's Theorem we get

$$\begin{aligned}\|Af\|^2_{L^2(\mathbb{R}^n)} &\le C_1\int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n}|\mathcal{K}(\xi,\eta)|\,d\xi\Big)|f(\eta)|^2\,d\eta\\ &\le C_1C_2\|f\|^2_{L^2(\mathbb{R}^n)}\end{aligned}$$

which gives (15.6). □

We now give

PROOF OF PROPOSITION 15.3. 1. Fix $\chi \in C^\infty_c(U)$. We will show the following bound: there exists a constant C depending on χ, a, s such that

$$\|\chi\operatorname{Op}(a)\varphi\|_{H^s(\mathbb{R}^n)} \le C\|\varphi\|_{H^{s+m}(\mathbb{R}^n)}\quad \text{for all } \varphi \in C^\infty_c(U). \tag{15.7}$$

Recalling the definition (12.5) of the norm on $H^s(\mathbb{R}^n)$, we have

$$\begin{gathered}\|\varphi\|_{H^{s+m}(\mathbb{R}^n)} = \|v\|_{L^2(\mathbb{R}^n)},\quad \|\chi\operatorname{Op}(a)\varphi\|_{H^s(\mathbb{R}^n)} = \|w\|_{L^2(\mathbb{R}^n)}\\ \text{where } v(\eta) := \langle\eta\rangle^{s+m}\widehat{\varphi}(\eta),\quad w(\xi) = \langle\xi\rangle^s\mathcal{F}(\chi\operatorname{Op}(a)\varphi)(\xi)\end{gathered} \tag{15.8}$$

By (14.17) and Fubini's Theorem we compute

$$\begin{gathered}w(\xi) = \int_{\mathbb{R}^n}\mathcal{K}(\xi,\eta)v(\eta)\,d\eta \quad \text{where}\\ \mathcal{K}(\xi,\eta) := (2\pi)^{-n}\frac{\langle\xi\rangle^s}{\langle\eta\rangle^{s+m}}\int_U e^{ix\cdot(\eta-\xi)}a(x,\eta)\chi(x)\,dx.\end{gathered} \tag{15.9}$$

2. We write

$$\mathcal{K}(\xi,\eta) = (2\pi)^{-n}\frac{\langle\xi\rangle^s}{\langle\eta\rangle^{s+m}}F(\xi-\eta,\eta) \quad \text{where } F(\zeta,\eta) := \int_U e^{-ix\cdot\zeta}a(x,\eta)\chi(x)\,dx.$$

Integrating by parts in x, we see that for each multiindex α

$$\zeta^\alpha F(\zeta,\eta) = \int_U e^{-ix\cdot\zeta}D^\alpha_x\big(a(x,\eta)\chi(x)\big)\,dx.$$

Since $a \in S^m(U \times \mathbb{R}^n)$, from (14.3) we see that this is bounded by some constant times $\langle \eta \rangle^m$. Since α can be taken arbitrary, we see that for each N there exists a constant C_N such that

$$|F(\zeta, \eta)| \leq C_N \langle \zeta \rangle^{-N} \langle \eta \rangle^m \quad \text{for all } \zeta, \eta \in \mathbb{R}^n.$$

Combining this with (12.26) we see that for each N there exists a constant C'_N such that

$$|\mathcal{K}(\xi, \eta)| \leq C'_N \langle \xi - \eta \rangle^{-N} \quad \text{for all } \xi, \eta \in \mathbb{R}^n.$$

Taking $N := n + 1$, we see that

$$\sup_{\xi \in \mathbb{R}^n} \int_{\mathbb{R}^n} |\mathcal{K}(\xi, \eta)|\, d\eta < \infty, \qquad \sup_{\eta \in \mathbb{R}^n} \int_{\mathbb{R}^n} |\mathcal{K}(\xi, \eta)|\, d\xi < \infty.$$

By Lemma 15.5, and recalling (15.8) and (15.9), we get the bound (15.7).

3. Arguing similarly to the proof of Theorem 11.29, using the bound (15.7), density of $C^\infty_c(U)$ in $H^{s+m}_c(U)$, and completeness of $H^s(\mathbb{R}^n)$, we see that for each $u \in H^{s+m}_c(U)$ and $\chi \in C^\infty_c(U)$, the distribution $\chi \operatorname{Op}(a)u$ lies in $H^s_c(U)$ and we have the norm bound

$$\|\chi \operatorname{Op}(a)u\|_{H^s(\mathbb{R}^n)} \leq C\|u\|_{H^{s+m}(\mathbb{R}^n)} \tag{15.10}$$

where the constant C depends on χ, a, s, but not on u. This shows that $\operatorname{Op}(a) : H^{s+m}_c(U) \to H^s_{\mathrm{loc}}(U)$ is sequentially continuous. Finally, sequential continuity of $\operatorname{Op}(a)^t$ follows from here by duality, using Proposition 12.14; we leave the details as an exercise below. □

15.1.2. Elliptic estimate. We now show regularity results for elliptic operators on Sobolev spaces. We start with

Proof of Theorem 15.1. We follow the proof of Theorem 14.2 in §14.2.2. As in Step 1 in that proof, we can reduce to the case of operators on $U \Subset \mathbb{R}^n$. Take arbitrary $\psi \in C^\infty_c(U)$ and fix $\chi \in C^\infty_c(U)$ such that $\operatorname{supp}(1 - \chi) \cap \operatorname{supp} \psi = \emptyset$. Let $\widetilde{Q}$ be the left elliptic parametrix for P constructed in Theorem 14.17. Multiplying both sides of (14.51) by ψ, we get

$$\psi u = \psi \widetilde{Q} \chi P u + \psi(\widetilde{Q}[P, \chi] + \widetilde{R}\chi)u. \tag{15.11}$$

By (14.49), we have $\widetilde{Q} = \operatorname{Op}(\widetilde{q})^t$ where $\widetilde{q} \in S^{-m}(U \times \mathbb{R}^n)$. Therefore, by Proposition 15.3 (with m replaced by $-m$) and since $Pu \in H^s_{\mathrm{loc}}(U)$ and thus $\chi P u \in H^s_c(U)$ we have

$$\psi \widetilde{Q} \chi P u \in H^{s+m}_c(U). \tag{15.12}$$

Next, the operator $\psi(\widetilde{Q}[P, \chi] + \widetilde{R}\chi)$ is smoothing:

$$\psi(\widetilde{Q}[P, \chi] + \widetilde{R}\chi) : \mathcal{D}'(U) \to C^\infty_c(U). \tag{15.13}$$

This uses pseudolocality of $\widetilde{Q}$ (Proposition 14.15). Indeed, recall that $\operatorname{supp}\psi\cap\operatorname{supp}(1-\chi)=\emptyset$ and the coefficients of $[P,\chi]$ are supported on $\operatorname{supp}\chi\cap\operatorname{supp}(1-\chi)\Subset U$. Take $\chi'\in C_c^\infty(U)$ such that

$$[P,\chi]=\chi'[P,\chi],\quad \operatorname{supp}\chi'\cap\operatorname{supp}\psi=\emptyset.$$

Then $\psi\widetilde{Q}[P,\chi]=\psi\widetilde{Q}\chi'[P,\chi]$, but $\psi\widetilde{Q}\chi'$ is smoothing similarly to (14.29). Thus $\psi\widetilde{Q}[P,\chi]$ is smoothing. We also know that $\widetilde{R}$ is smoothing. Together these give (15.13).

Putting together (15.12) and (15.13) and using that $C_c^\infty(U)\subset H_c^{s+m}(U)$, we see that $\psi u\in H_c^{s+m}(U)$. Since this holds for any $\psi\in C_c^\infty(U)$, we get $u\in H_{\mathrm{loc}}^{s+m}(U)$, finishing the proof. □

From the proof above we get the following quantitative version of Theorem 15.1:

Proposition 15.6 (Elliptic estimate on open subsets of $\mathbb{R}^n$). *Assume that $U\Subset\mathbb{R}^n$ and $P\in\operatorname{Diff}^m(U)$ is an elliptic differential operator. Take any $\psi,\widetilde{\psi}\in C_c^\infty(U)$ such that $\operatorname{supp}(1-\widetilde{\psi})\cap\operatorname{supp}\psi=\emptyset$. Then for each s,N there exists a constant C such that for all $u\in H_{\mathrm{loc}}^{s+m}(U)$ we have*

$$\|\psi u\|_{H^{s+m}(\mathbb{R}^n)}\leq C\|\widetilde{\psi}Pu\|_{H^s(\mathbb{R}^n)}+C\|\widetilde{\psi}u\|_{H^{-N}(\mathbb{R}^n)}. \tag{15.14}$$

Proof. Fix a cutoff

$$\chi\in C_c^\infty(U),\quad \operatorname{supp}(1-\widetilde{\psi})\cap\operatorname{supp}\chi=\operatorname{supp}(1-\chi)\cap\operatorname{supp}\psi=\emptyset.$$

We write ψu in the form (15.11); since $\widetilde{\psi}=1$ near $\operatorname{supp}\chi$, we have

$$\psi u=\mathcal{Q}\widetilde{\psi}Pu+\mathcal{R}\widetilde{\psi}u,\quad \mathcal{Q}:=\psi\widetilde{Q}\chi,\quad \mathcal{R}:=\psi(\widetilde{Q}[P,\chi]+\widetilde{R}\chi)$$

Since $\widetilde{Q}=\operatorname{Op}(\widetilde{q})^t$ for some $\widetilde{q}\in S^{-m}(U\times\mathbb{R}^n)$, by Exercise 15.1 below the operator $\mathcal{Q}$ is bounded $H^s(\mathbb{R}^n)\to H^{s+m}(\mathbb{R}^n)$.

As in (15.13), the operator $\mathcal{R}$ is smoothing, more precisely its Schwartz kernel lies in $C_c^\infty(U\times U)$. Thus this operator is bounded $H^{-N}(\mathbb{R}^n)\to H^{s+m}(\mathbb{R}^n)$. Indeed, if $\|v_k\|_{H^{-N}(\mathbb{R}^n)}\to 0$, then $v_k\to 0$ in $\mathcal{D}'(\mathbb{R}^n)$. Since $\mathcal{R}$ is smoothing and its Schwartz kernel is compactly supported, we have $\mathcal{R}v_k\to 0$ in $C_c^\infty(\mathbb{R}^n)$, which implies that $\mathcal{R}v_k\to 0$ in $H^{s+m}(\mathbb{R}^n)$. This implies the needed boundedness statement on $\mathcal{R}$.

We now have

$$\|\psi u\|_{H^{s+m}(\mathbb{R}^n)}\leq\|\mathcal{Q}\|_{H^s(\mathbb{R}^n)\to H^{s+m}(\mathbb{R}^n)}\|\widetilde{\psi}Pu\|_{H^s(\mathbb{R}^n)}+\|\mathcal{R}\|_{H^{-N}(\mathbb{R}^n)\to H^{s+m}(\mathbb{R}^n)}\|\widetilde{\psi}u\|_{H^{-N}(\mathbb{R}^n)}$$

giving the estimate (15.14). □

From Proposition 15.6 we get the following statement, which is used crucially in this and the next chapter to establish mapping and spectral properties of elliptic operators on compact manifolds. We will use this estimate as a black box, so the reader does not need to follow all the details of its proof to understand how it is applied later.

THEOREM 15.7 (Elliptic estimate on compact manifolds). *Assume that $\mathcal{M}$ is a compact manifold, $P \in \mathrm{Diff}^m(\mathcal{M})$ is an elliptic differential operator, and fix $s, N \in \mathbb{R}$. Then there exists a constant C such that for all $u \in H^{s+m}(\mathcal{M})$ we have*

$$\|u\|_{H^{s+m}(\mathcal{M})} \leq C\|Pu\|_{H^s(\mathcal{M})} + C\|u\|_{H^{-N}(\mathcal{M})}. \tag{15.15}$$

Here the norm on Sobolev spaces is defined in Exercise 13.4.

PROOF.[S] Recalling Exercise 13.4, take a finite collection of charts $\varkappa_\ell : U_\ell \to V_\ell$, $\ell = 1, \dots, N$ and a partition of unity

$$1 = \sum_{\ell=1}^{N} \psi_\ell, \quad \psi_\ell \in C_c^\infty(U_\ell).$$

Choose cutoffs

$$\widetilde{\psi}_\ell \in C_c^\infty(U_\ell), \quad \mathrm{supp}(1 - \widetilde{\psi}_\ell) \cap \mathrm{supp}\,\psi_\ell = \emptyset.$$

Applying Proposition 15.6 to the elliptic operator $\varkappa_{\ell *}P \in \mathrm{Diff}^m(V_\ell)$ and the distribution $\varkappa_{\ell *}u$ with the cutoffs $\varkappa_{\ell *}\psi_\ell$, $\varkappa_{\ell *}\widetilde{\psi}_\ell$ we get for each ℓ and some constant C (whose value will change from place to place in the argument)

$$\|\varkappa_{\ell *}(\psi_\ell u)\|_{H^{s+m}(\mathbb{R}^n)} \leq C\|\varkappa_{\ell *}(\widetilde{\psi}_\ell Pu)\|_{H^s(\mathbb{R}^n)} + C\|\varkappa_{\ell *}(\widetilde{\psi}_\ell u)\|_{H^{-N}(\mathbb{R}^n)}. \tag{15.16}$$

We have

$$\|\varkappa_{\ell *}(\widetilde{\psi}_\ell v)\|_{H^s(\mathbb{R}^n)} \leq C\|v\|_{H^s(\mathcal{M})} \quad \text{for all } v \in H^s(\mathcal{M}). \tag{15.17}$$

Indeed, if $\|v_k\|_{H^s(\mathcal{M})} \to 0$ then (by Exercise 13.4) the sequence v_k converges to 0 in $H^s_{\mathrm{loc}}(\mathcal{M})$, thus (recalling the discussion following Definition 13.8) we have $\varkappa_{\ell *}v_k \to 0$ in $H^s_{\mathrm{loc}}(V_\ell)$ and thus $\|\varkappa_{\ell *}(\widetilde{\psi}_\ell v)\|_{H^s(\mathbb{R}^n)} \to 0$. This shows (15.17). The same estimate holds for the space H^{-N}, so (15.16) implies that

$$\|\varkappa_{\ell *}(\psi_\ell u)\|_{H^{s+m}(\mathbb{R}^n)} \leq C\|Pu\|_{H^s(\mathcal{M})} + C\|u\|_{H^{-N}(\mathcal{M})}.$$

Adding these estimates for all ℓ and recalling the definition (13.74) of the norm $\|u\|_{H^{s+m}(\mathcal{M})}$, we get (15.15). □

15.2. Compact embedding in Sobolev spaces

We are almost ready to combine Theorem 15.7 with an argument from functional analysis to get the Fredholm mapping property on Sobolev spaces for elliptic differential operators on compact manifolds. The remaining ingredient is the fact that for a compact manifold $\mathcal{M}$, the space $H^s(\mathcal{M})$ is precompact inside $H^t(\mathcal{M})$ when $s > t$, and this is what we establish in this section. We start with the case of Sobolev spaces on $\mathbb{R}^n$:

THEOREM 15.8 (Rellich–Kondrachov Theorem on $\mathbb{R}^n$). *Assume that $s, t \in \mathbb{R}$, $s > t$, $u_k \in H^s_{\mathrm{c}}(\mathbb{R}^n)$ is a sequence, and there exist constants C_0, R such that for all k*

$$\operatorname{supp} u_k \subset B(0, R), \tag{15.18}$$

$$\|u_k\|_{H^s(\mathbb{R}^n)} \le C_0. \tag{15.19}$$

Then there exists a subsequence u_{k_ℓ} which converges in $H^t(\mathbb{R}^n)$.

REMARK 15.9. *We cannot completely get rid of (15.18): if $\chi \in C^\infty_{\mathrm{c}}(\mathbb{R})$ then the sequence $u_k(x) = \chi(x - k)$ satisfies (15.19) for any s but it does not have a limit in $L^2(\mathbb{R})$. We cannot get rid of (15.19) either: the sequence $u_k(x) = \sqrt{k}\chi(kx)$ satisfies (15.18) and is bounded in $L^2(\mathbb{R})$ but does not converge in this space. However, with more work one can replace conditions (15.18) and (15.19) by the weaker assumption that $\|\langle x\rangle^\delta u_k(x)\|_{H^s(\mathbb{R}^n)} \le C_0$ for some $\delta > 0$. One can summarize this statement informally as*

$$\textit{improved regularity + improved decay at infinity} \quad \Longrightarrow \quad \textit{precompactness.}$$

PROOF. 1. We carry out an Arzelà–Ascoli type argument on the side of the Fourier transform. Recall from Proposition 11.26 that, since $u_k \in \mathcal{E}'(\mathbb{R}^n)$, we have $\widehat{u}_k \in C^\infty(\mathbb{R}^n)$.

We first show that $\widehat{u}_k$ is locally bounded and Lipschitz continuous uniformly in k, namely for each $T > 0$ there exists a constant C_T such that for all k

$$|\widehat{u}_k(\xi)| \le C_T \quad \text{for all } \xi \in B(0, T), \tag{15.20}$$

$$|\widehat{u}_k(\xi) - \widehat{u}_k(\eta)| \le C_T|\xi - \eta| \quad \text{for all } \xi, \eta \in B(0, T). \tag{15.21}$$

Fix $N \in \mathbb{N}_0$ such that $s + N \ge 0$ and take a cutoff function $\chi \in C^\infty_{\mathrm{c}}(\mathbb{R}^n)$ which is equal to 1 near $B(0, R)$. For each multiindex α and $\xi \in \mathbb{R}^n$ we compute (with the constant C_α depending on α, C_0, χ but not on k)

$$\begin{aligned} |\partial^\alpha_\xi \widehat{u}_k(\xi)| &= |(u_k(x), x^\alpha \chi(x) e^{-ix\cdot\xi})| \\ &\le C_\alpha \|x^\alpha \chi(x) e^{-ix\cdot\xi}\|_{H^N(\mathbb{R}^n)} \\ &\le C_\alpha \|x^\alpha \chi(x) e^{-ix\cdot\xi}\|_{C^N(\mathbb{R}^n)} \\ &\le C_\alpha \langle\xi\rangle^N. \end{aligned} \tag{15.22}$$

Here in the first line we use Proposition 11.26 and the support condition (15.18). In the second line we use Proposition 12.7 and the bound (15.19). In the third line we use Proposition 12.1 and the fact that $\chi \in C^\infty_{\mathrm{c}}(\mathbb{R}^n)$.

Taking (15.22) with $\alpha = \emptyset$ we get (15.20). Taking (15.22) with $|\alpha| = 1$ we estimate $|d\widehat{u}_k(\xi)|$ which gives the bound (15.21).

2.[R] We now use (15.20)–(15.21) and the Arzelà–Ascoli theorem to see that there exists a subsequence u_{k_ℓ} such that for some $v \in C^0(\mathbb{R}^n)$ we have as $\ell \to \infty$

$$\widehat{u}_{k_\ell}(\xi) \to v(\xi) \quad \text{locally uniformly in } \xi \in \mathbb{R}^n. \tag{15.23}$$

Here 'locally uniformly' means 'uniformly on each compact subset of $\mathbb{R}^n$'. Since the version of the Arzelà–Ascoli theorem that we use is not the one most commonly stated in a real analysis course, we briefly review how the proof goes.

Let $\xi_m \in \mathbb{R}^n$ be a sequence of points which is dense in $\mathbb{R}^n$ (e.g. one can take all the points with rational coordinates, which form a countable set). By (15.20) we know that for each fixed m, the sequence $\widehat{u}_k(\xi_m)$ is bounded. Then there exists a sequence $k_{\ell,1} \to \infty$ such that $\widehat{u}_{k_{\ell,1}}(\xi_1) \to v_1$ for some $v_1 \in \mathbb{C}$, and we can iteratively construct sequences $k_{\ell,m} \to \infty$ for $m \geq 2$ such that $k_{\ell,m}$ is a subsequence of $k_{\ell,m-1}$ and $\widehat{u}_{k_{\ell,m}}(\xi_m) \to v_m$ for some $v_m \in \mathbb{C}$. Take the diagonal sequence $k_\ell := k_{\ell,\ell}$, then

$$\widehat{u}_{k_\ell}(\xi_m) \to v_m \quad \text{as } \ell \to \infty \quad \text{for all } m.$$

We next claim that $\widehat{u}_{k_\ell}(\xi)$ is a Cauchy sequence locally uniformly in ξ, that is for each $T > 0$ we have

$$\sup_{\ell,\ell' \geq r} \sup_{\xi \in B(0,T)} |\widehat{u}_{k_\ell}(\xi) - \widehat{u}_{k_{\ell'}}(\xi)| \to 0 \quad \text{as } r \to \infty. \tag{15.24}$$

This can be shown using the fact that for each $\varepsilon > 0$ there exists m_ε so that each point $\xi \in B(0,T)$ is ε-close to one of the points ξ_m with $m \leq m_\varepsilon$. Now $|\widehat{u}_{k_\ell}(\xi_m) - \widehat{u}_{k_{\ell'}}(\xi_m)|$ for $m \leq m_\varepsilon$ is estimated using (15.23) and $|\widehat{u}_{k_\ell}(\xi) - \widehat{u}_{k_\ell}(\xi_m)|$ (as well as the similar quantity for ℓ') is estimated by the Lipschitz bound (15.21).

Finally, (15.24) implies the local convergence statement (15.23) (since the space of continuous functions on $B(0,T)$ with the uniform norm is complete).

3. By Fatou's Lemma and the Sobolev norm bound (15.19) we have

$$\int_{\mathbb{R}^n} \langle \xi \rangle^{2s} |v(\xi)|^2 \, d\xi \leq \liminf_{\ell \to \infty} \int_{\mathbb{R}^n} \langle \xi \rangle^{2s} |\widehat{u}_{k_\ell}(\xi)|^2 \, d\xi < \infty.$$

Thus $\langle \xi \rangle^s v(\xi) \in L^2(\mathbb{R}^n)$, so there exists

$$u \in H^s(\mathbb{R}^n), \quad \widehat{u} = v.$$

It remains to show that for $t < s$ we have the convergence

$$\|u_{k_\ell} - u\|_{H^t(\mathbb{R}^n)} \to 0,$$

which is equivalent to

$$\int_{\mathbb{R}^n} \langle \xi \rangle^{2t} |\widehat{u}_{k_\ell}(\xi) - v(\xi)|^2 \, d\xi \to 0. \tag{15.25}$$

Take arbitrary $T > 1$. We have

$$\begin{aligned}\int_{\mathbb{R}^n \setminus B(0,T)} \langle\xi\rangle^{2t} |\widehat{u}_{k_\ell}(\xi) - v(\xi)|^2 \, d\xi &\leq 2T^{2(t-s)} \int_{\mathbb{R}^n} \langle\xi\rangle^{2s} \big(|\widehat{u}_{k_\ell}(\xi)|^2 + |v(\xi)|^2\big) \, d\xi \\ &\leq CT^{2(t-s)}\end{aligned} \tag{15.26}$$

where the constant C is independent of ℓ and T. Here in the second line we use the uniform Sobolev bound (15.19) and the fact that $\langle\xi\rangle^s v(\xi) \in L^2(\mathbb{R}^n)$. It follows that

$$\int_{\mathbb{R}^n} \langle\xi\rangle^{2t} |\widehat{u}_{k_\ell}(\xi) - v(\xi)|^2 \, d\xi \leq CT^{2(t-s)} + \int_{B(0,T)} \langle\xi\rangle^{2t} |\widehat{u}_{k_\ell}(\xi) - v(\xi)|^2 \, d\xi.$$

The second term on the right-hand side converges to 0 as $\ell \to \infty$ by (15.23). It follows that

$$\limsup_{\ell \to \infty} \int_{\mathbb{R}^n} \langle\xi\rangle^{2t} |\widehat{u}_{k_\ell}(\xi) - v(\xi)|^2 \, d\xi \leq CT^{2(t-s)}.$$

Since this is true for all T, and $t < s$, we get the convergence statement (15.25), finishing the proof. ☐

We now give the version on compact manifolds:

THEOREM 15.10 (Rellich–Kondrachov Theorem on compact manifolds). *Assume that $\mathcal{M}$ is a compact manifold and $s > t$. Then $H^s(\mathcal{M})$ embeds compactly into $H^t(\mathcal{M})$, in the following sense: if $u_k \in H^s(\mathcal{M})$ is a sequence such that $\|u_k\|_{H^s(\mathcal{M})}$ is bounded, then there exists a subsequence u_{k_ℓ} which converges in $H^t(\mathcal{M})$.*

PROOF.**S** Recalling Exercise 13.4, take a finite collection of charts $\varkappa_m : U_m \to V_m$, $m = 1, \dots, N$ and a partition of unity

$$1 = \sum_{m=1}^{N} \chi_m, \quad \chi_m \in C_c^\infty(U_m).$$

For each m, the sequence $\varkappa_{m*}(\chi_m u_k) \in H_c^s(V_m)$ is supported in $\varkappa_m(\operatorname{supp} \chi_m) \Subset V_m$ and is uniformly bounded in H^s norm. Applying Theorem 15.8 N times, we see that there exists a subsequence u_{k_ℓ} such that for each m, we have as $\ell \to \infty$

$$\varkappa_{m*}(\chi_m u_{k_\ell}) \to w_m \quad \text{in } H_c^t(V_m) \quad \text{for some } w_m \in H_c^t(V_m).$$

By Exercise 13.3, we see that $\chi_m u_{k_\ell} \to \varkappa_m^* w_m$ in $H_{\mathrm{loc}}^t(\mathcal{M})$, where $\varkappa_m^* w_m \in \mathcal{E}'(U_m)$ is extended by 0 to an element of $H^t(\mathcal{M})$. Summing over m and using Exercise 13.4, we see that

$$\|u_{k_\ell} - u\|_{H^t(\mathcal{M})} \to 0 \quad \text{where } u := \sum_{m=1}^{N} \varkappa_m^* w_m$$

finishing the proof. ☐

15.3. Fredholm theory

15.3.1. Fredholm property of elliptic operators on compact manifolds. We now combine the elliptic estimate (Theorem 15.7) with Rellich–Kondrachov Theorem 15.10 to show that elliptic differential operators on compact manifolds have the Fredholm property on Sobolev spaces.

We first give the definition of a Fredholm operator:

DEFINITION 15.11. *Let $\mathcal{H}_1, \mathcal{H}_2$ be Banach spaces and $P : \mathcal{H}_1 \to \mathcal{H}_2$ be a bounded linear operator. We say that P is a* Fredholm operator *if all of the following conditions are satisfied:*

(1) *the kernel* $\ker P := \{u \in \mathcal{H}_1 \mid Pu = 0\}$ *is finite dimensional;*
(2) *the range* $\operatorname{ran} P := \{Pu \mid u \in \mathcal{H}_1\}$ *is a closed subspace of $\mathcal{H}_2$;*
(3) *the range* $\operatorname{ran} P$ *has finite codimension in the sense that the quotient $\mathcal{H}_2/\operatorname{ran} P$ is a finite dimensional space.*

If P is a Fredholm operator, then we define its index *as the integer*

$$\operatorname{ind}(P) := \dim(\ker P) - \dim(\mathcal{H}_2/\operatorname{ran} P). \tag{15.27}$$

REMARK 15.12.$^{\mathbf{X}}$ *The property (2) above is actually unnecessary: one can show that $(1) + (3) \Rightarrow (2)$. However, a typical proof of the property (3), such as the one for elliptic operators given below, establishes property (2) along the way.*

Now, let $\mathcal{M}$ be a compact manifold and $P \in \operatorname{Diff}^m(\mathcal{M})$ be a differential operator. By (15.1), the operator $P : \mathcal{D}'(\mathcal{M}) \to \mathcal{D}'(\mathcal{M})$ restricts to bounded operators

$$P_s : H^{s+m}(\mathcal{M}) \to H^s(\mathcal{M}), \quad s \in \mathbb{R}. \tag{15.28}$$

The transpose P^t is a differential operator in $\operatorname{Diff}^m(\mathcal{M}; |\Omega| \to |\Omega|)$, where $|\Omega|$ is the bundle of densities over $\mathcal{M}$ (see (13.73)), and it restricts to bounded operators

$$P_s^t : H^{s+m}(\mathcal{M}; |\Omega|) \to H^s(\mathcal{M}; |\Omega|), \quad s \in \mathbb{R}. \tag{15.29}$$

We have the identity (which is immediate from the definition of P^t for $u, v \in C^\infty$ and extends to general u, v by a density argument, with the pairing defined similarly to Proposition 12.14)

$$(P_s u, v) = (u, P^t_{-s-m} v) \quad \text{for all } u \in H^{s+m}(\mathcal{M}),\ v \in H^{-s}(\mathcal{M}; |\Omega|). \tag{15.30}$$

Note that we can fix a smooth positive density ω_0 on $\mathcal{M}$ which identifies sections of $|\Omega|$ with functions, and consider P_s^t as a scalar operator. In this case (15.30) is valid for all $u \in H^{s+m}(\mathcal{M})$, $v \in H^{-s}(\mathcal{M})$ and the pairing $(\bullet, \bullet)$ extends to Sobolev spaces the integral $(f, g) := \int_{\mathcal{M}} f g \, \omega_0$.

If the differential operator P is elliptic, then by Theorem 14.2 the kernels of the operators P_s, P_s^t are independent of s:

$$\begin{aligned}\ker P_s &= \ker P := \{u \in C^\infty(\mathcal{M})\colon\ Pu = 0\},\\ \ker P_s^t &= \ker P^t := \{v \in C^\infty(\mathcal{M}; |\Omega|)\colon\ P^t v = 0\}.\end{aligned} \tag{15.31}$$

We are now ready to present the main result of this section:

THEOREM 15.13 (Fredholm property of elliptic operators). *Let $\mathcal{M}$ be a compact manifold and $P \in \mathrm{Diff}^m(\mathcal{M})$ be an elliptic differential operator. Then for each $s \in \mathbb{R}$, the operator P_s defined in (15.28) has the Fredholm property. Moreover, the range of P_s is characterized as follows:*

$$\operatorname{ran} P_s = \{w \in H^s(\mathcal{M})\colon \textit{for all } v \in \ker P^t \quad \textit{we have } (w, v) = 0\}. \tag{15.32}$$

REMARK 15.14. *A standard application of (15.32) is the following existence theorem: if g is a Riemannian metric on a compact connected manifold $\mathcal{M}$ and Δ_g is the corresponding Laplace–Beltrami operator (see (13.63)) then for each $w \in H^s(\mathcal{M})$ we have*

$$w = \Delta_g u \quad \textit{for some } u \in H^{s+2}(\mathcal{M}) \iff \int_{\mathcal{M}} w \, d\operatorname{vol}_g = 0. \tag{15.33}$$

We leave the proof of (15.33) as an exercise below.

REMARK 15.15.**X** *Even though the space* $\ker P$ *consists of smooth functions, to show that this space is finite dimensional one needs to use spaces with fixed regularity (such as Sobolev spaces) and compact embedding (such as Theorem 15.10).*

PROOF. We follow a classical argument from functional analysis, using Theorems 15.7 and 15.10 as black boxes.

1. We first show that $\ker P_s = \ker P$ is finite dimensional. We argue by contradiction. Assume that $\ker P$ is infinite dimensional. Fix $N > -s - m$. Using the Gram–Schmidt process, we construct a countable orthonormal system

$$u_k \in \ker P, \quad \langle u_k, u_{k'} \rangle_{H^{-N}(\mathcal{M})} = \begin{cases} 1, & k = k',\\ 0, & \text{otherwise}.\end{cases} \tag{15.34}$$

By Theorem 15.7 there exists a constant C such that for all k

$$\|u_k\|_{H^{s+m}(\mathcal{M})} \le C \|u_k\|_{H^{-N}(\mathcal{M})} = C.$$

Using the compact embedding $H^{s+m}(\mathcal{M}) \subset H^{-N}(\mathcal{M})$ given by Theorem 15.10, we see that u_k has a subsequence which converges in $H^{-N}(\mathcal{M})$. However, this contradicts (15.34) since $\|u_k - u_{k'}\|_{H^{-N}(\mathcal{M})} = \sqrt{2}$ and thus no subsequence of u_k can be a Cauchy sequence in $H^{-N}(\mathcal{M})$. This shows that $\ker P$ is finite dimensional.

2. We next show that $\operatorname{ran} P_s = \{Pu \mid u \in H^{s+m}(\mathcal{M})\}$ is a closed subspace of $H^s(\mathcal{M})$. Assume that w lies in the closure of $\operatorname{ran} P_s$ in $H^s(\mathcal{M})$. Then there exists a sequence

$$u_k \in H^{s+m}(\mathcal{M}), \quad \|Pu_k - w\|_{H^s(\mathcal{M})} \to 0. \tag{15.35}$$

We would like to find a subsequence of u_k which converges to some u in $H^{s+m}(\mathcal{M})$, which would show that $w = Pu$ lies in $\operatorname{ran} P_s$. This proceeds in three steps:

(a) We can add an element of $\ker P$ to any u_k without changing (15.35). Thus we replace u_k by its projection to the orthogonal complement of $\ker P$ in $H^{s+m}(\mathcal{M})$, so that

$$\langle u_k, f\rangle_{H^{s+m}(\mathcal{M})} = 0 \quad \text{for all } f \in \ker P. \tag{15.36}$$

(b) We claim that the sequence u_k is bounded in $H^{s+m}(\mathcal{M})$. We argue by contradiction. If u_k is not bounded, then we can pass to a subsequence to make $\|u_k\|_{H^{s+m}(\mathcal{M})} \to \infty$. Put

$$\widetilde{u}_k := \frac{u_k}{\|u_k\|_{H^{s+m}(\mathcal{M})}}, \quad \|\widetilde{u}_k\|_{H^{s+m}(\mathcal{M})} = 1. \tag{15.37}$$

By (15.35) we have

$$\|P\widetilde{u}_k\|_{H^s(\mathcal{M})} \to 0. \tag{15.38}$$

Using the compact embedding $H^{s+m}(\mathcal{M}) \subset H^{-N}(\mathcal{M})$ given by Theorem 15.10, we can pass to a further subsequence to make $\widetilde{u}_k$ converge in $H^{-N}(\mathcal{M})$. By Theorem 15.7 applied to $\widetilde{u}_k - \widetilde{u}_{k'}$, there exists a constant C such that for all k, k'

$$\|\widetilde{u}_k - \widetilde{u}_{k'}\|_{H^{s+m}(\mathcal{M})} \le C\|P\widetilde{u}_k - P\widetilde{u}_{k'}\|_{H^s(\mathcal{M})} + C\|\widetilde{u}_k - \widetilde{u}_{k'}\|_{H^{-N}(\mathcal{M})}. \tag{15.39}$$

Since $P\widetilde{u}_k$ is a Cauchy sequence in $H^s(\mathcal{M})$ by (15.38) and $\widetilde{u}_k$ is a Cauchy sequence in $H^{-N}(\mathcal{M})$, we see that the right-hand side of (15.39) converges to 0 as $k, k' \to \infty$. Thus $\widetilde{u}_k$ is a Cauchy sequence in $H^{s+m}(\mathcal{M})$. Since $H^{s+m}(\mathcal{M})$ is complete, we have

$$\|\widetilde{u}_k - \widetilde{u}\|_{H^{s+m}(\mathcal{M})} \to 0 \quad \text{for some } \widetilde{u} \in H^{s+m}(\mathcal{M}). \tag{15.40}$$

Passing to the limit in (15.36), (15.37), and (15.38), we see that

$$\langle \widetilde{u}, f\rangle_{H^{s+m}(\mathcal{M})} = 0 \quad \text{for all } f \in \ker P, \quad \|\widetilde{u}\|_{H^{s+m}(\mathcal{M})} = 1, \quad P\widetilde{u} = 0.$$

Thus $\widetilde{u} \in \ker P$. Taking $f := \widetilde{u}$ above, we get a contradiction.

(c) Now that the sequence u_k is bounded in $H^{s+m}(\mathcal{M})$, we use the compact embedding $H^{s+m}(\mathcal{M}) \subset H^{-N}(\mathcal{M})$ given by Theorem 15.10 to pass to a subsequence and make u_k converge in $H^{-N}(\mathcal{M})$. Since Pu_k is a Cauchy sequence in $H^s(\mathcal{M})$ by (15.35), we argue in the same way as for the proof of (15.40) to see that

$$\|u_k - u\|_{H^{s+m}(\mathcal{M})} \to 0 \quad \text{for some } u \in H^{s+m}(\mathcal{M}).$$

Passing to the limit in (15.35), we see that $w = Pu$ and thus w lies in $\operatorname{ran} P_s$. This shows that $\operatorname{ran} P_s$ is a closed subspace of $H^s(\mathcal{M})$.

3. We now show the characterization (15.32) of the range $\operatorname{ran} P_s$. First of all, if $w \in \operatorname{ran} P_s$ and $v \in \ker P^t$ then we write $w = Pu$ for some $u \in H^{s+m}(\mathcal{M})$ and compute by (15.30)

$$(w, v) = (Pu, v) = (u, P^t v) = 0.$$

It remains to show that if $w \in H^s(\mathcal{M})$ and $(w, v) = 0$ for all $v \in \ker P^t$, then $w \in \operatorname{ran} P_s$. We argue by contradiction. Assume that $w \notin \operatorname{ran} P_s$. Since $\operatorname{ran} P_s$ is a closed subspace of $H^s(\mathcal{M})$, there exists a bounded linear functional

$$F : H^s(\mathcal{M}) \to \mathbb{C}, \quad F|_{\operatorname{ran} P_s} = 0, \quad F(w) \neq 0.$$

For example, one can use the Orthogonal Complement Theorem 1.3 to find nonzero $\widetilde{w} \in H^s(\mathcal{M})$ such that $w - \widetilde{w} \in \operatorname{ran} P_s$ and $\widetilde{w}$ is orthogonal to $\operatorname{ran} P_s$ (with respect to the H^s inner product), and put $F(h) := \langle h, \widetilde{w} \rangle_{H^s(\mathcal{M})}$ for all $h \in H^s(\mathcal{M})$.

Similarly to Proposition 12.14, there exists $v \in H^{-s}(\mathcal{M}; |\Omega|)$ such that the functional F has the form

$$F(h) = (h, v) \quad \text{for all } h \in H^s(\mathcal{M}).$$

Since $F|_{\operatorname{ran} P_s} = 0$, we have for all $u \in H^{s+m}(\mathcal{M})$

$$0 = (Pu, v) = (u, P^t v)$$

where we used (15.30). In particular, this is true for all $u \in C^\infty(\mathcal{M})$, which (similarly to Theorem 1.16) shows that $P^t v = 0$. Thus $v \in \ker P^t$. But we also have $(w, v) = F(w) \neq 0$, which gives a contradiction.

(As a side remark, we could have avoided Proposition 12.14 and (15.30) by restricting F to $C^\infty(\mathcal{M}; |\Omega|)$ and constructing v as a distribution in $\mathcal{D}'(\mathcal{M})$. We did not do this to produce a more robust proof which applies to other, potentially non-elliptic, situations.)

4.**S** Denote by $(\ker P^t)'$ the space of linear functionals on $\ker P^t$. Note that $\ker P^t$ is finite dimensional by Step 1 of this proof applied to the elliptic operator P^t. The map

$$T : H^s(\mathcal{M}) \to (\ker P^t)', \quad T(w)(v) = (w, v)$$

is surjective. Indeed, if $v \in \ker P^t \subset C^\infty(\mathcal{M}; |\Omega|)$ is such that $T(w)(v) = 0$ for all $w \in H^s(\mathcal{M})$, then in particular $(w, v) = 0$ for all $w \in C^\infty(\mathcal{M})$ which (similarly to Theorem 1.16) gives $v = 0$. Since any proper subspace of $(\ker P^t)'$ annihilates some nonzero $v \in \ker P^t$, we see that the range of T is the entire $(\ker P^t)'$.

By (15.32), the kernel of T is equal to $\operatorname{ran} P_s$. Thus T induces an isomorphism

$$H^s(\mathcal{M}) / \operatorname{ran} P_s \simeq (\ker P^t)'. \tag{15.41}$$

This shows that ran P_s has finite codimension and completes the proof of the Fredholm property. □

15.3.2. General Fredholm theory$^{\mathbf{R}}$. We now give a brief review of general properties of Fredholm operators. We refer the reader to [**Hör07**, §19.1] and [**Lax02**, §21.1,24.1–24.2] for the proofs.

We first define the notion of a compact operator:

DEFINITION 15.16. *Let $\mathcal{H}_1, \mathcal{H}_2$ be Banach spaces and $A : \mathcal{H}_1 \to \mathcal{H}_2$ be a bounded linear operator. We say that A is a* compact operator *if the image of the ball $B(0, 1)$ in $\mathcal{H}_1$ under A is precompact in $\mathcal{H}_2$, that is for any bounded sequence $u_k \in \mathcal{H}_1$ the sequence $Au_k \in \mathcal{H}_2$ has a convergent subsequence.*

Some standard properties of compact operators are collected in

PROPOSITION 15.17. *The set of compact operators is a closed ideal in the space of bounded operators, namely:*

(1) *linear combinations of compact operators are compact operators;*
(2) *if $\mathcal{H}_1 \xrightarrow{A} \mathcal{H}_2 \xrightarrow{B} \mathcal{H}_3$ are bounded operators and one of the operators A, B is compact, then the composition $BA : \mathcal{H}_1 \to \mathcal{H}_3$ is a compact operator;*
(3) *if $A_k : \mathcal{H}_1 \to \mathcal{H}_2$ is a sequence of compact operators and $A : \mathcal{H}_1 \to \mathcal{H}_2$ is a bounded operator such that $\|A_k - A\|_{\mathcal{H}_1 \to \mathcal{H}_2} \to 0$, then A is a compact operator.*

We next give standard properties of Fredholm operators and their index.

PROPOSITION 15.18. *1. The set of Fredholm operators is open in the space of bounded operators and the index is a locally constant function on this set, namely for each Fredholm operator $P : \mathcal{H}_1 \to \mathcal{H}_2$ there exists $\varepsilon > 0$ such that for any bounded operator $Q : \mathcal{H}_1 \to \mathcal{H}_2$ with $\|P - Q\|_{\mathcal{H}_1 \to \mathcal{H}_2} < \varepsilon$, the operator Q has the Fredholm property and* $\operatorname{ind} P = \operatorname{ind} Q$.

2. Fredholm operators are stable under compact perturbations: if $P : \mathcal{H}_1 \to \mathcal{H}_2$ is a Fredholm operator and $W : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact operator, then $P + W$ is a Fredholm operator and $\operatorname{ind}(P + W) = \operatorname{ind} P$.

3. If $\mathcal{H}_1 \xrightarrow{P} \mathcal{H}_2 \xrightarrow{Q} \mathcal{H}_3$ are two Fredholm operators, then their composition QP is also a Fredholm operator and $\operatorname{ind}(QP) = \operatorname{ind} P + \operatorname{ind} Q$.

15.3.3. A touch of index theory. We now very briefly discuss index theory of elliptic operators. Let $\mathcal{M}$ be a compact manifold, $\mathscr{E}, \mathscr{F}$ be two vector bundles over $\mathcal{M}$ and $\mathbf{P} \in \operatorname{Diff}^m(\mathcal{M}; \mathscr{E} \to \mathscr{F})$ be an elliptic differential operator (see §14.2.3). We use operators on vector bundles since the index of a scalar differential operator is always equal to 0, see Exercise 15.5 below.

The operator $\mathbf{P}$ acts on Sobolev spaces similarly to (15.1), and the version of Theorem 15.13 for vector bundles shows that the index of $\mathbf{P}$ on each of these spaces is given by (recalling (15.41))

$$\operatorname{ind}\mathbf{P} = \dim\ker\mathbf{P} - \dim\ker\mathbf{P}^t. \tag{15.42}$$

We give here some basic properties of this index. We start with the statement that the index only depends on the principal symbol:

PROPOSITION 15.19. *Assume that* $\mathbf{P},\mathbf{Q} \in \mathrm{Diff}^m(\mathcal{M};\mathscr{E} \to \mathscr{F})$ *and* $\sigma_m(\mathbf{P}) = \sigma_m(\mathbf{Q})$. *Then* $\operatorname{ind}\mathbf{P} = \operatorname{ind}\mathbf{Q}$.

PROOF. Take any $s \in \mathbb{R}$. The operator $\mathbf{P} - \mathbf{Q}$ lies in $\mathrm{Diff}^{m-1}(\mathcal{M};\mathscr{E} \to \mathscr{F})$ and thus is bounded $H^{s+m}(\mathcal{M};\mathscr{E}) \to H^{s+1}(\mathcal{M};\mathscr{F})$. Now, the inclusion $H^{s+1}(\mathcal{M};\mathscr{F}) \to H^s(\mathcal{M};\mathscr{F})$ is a compact operator by Theorem 15.10, thus by Proposition 15.17(2) we see that $\mathbf{P}_s - \mathbf{Q}_s : H^{s+m}(\mathcal{M};\mathscr{E}) \to H^s(\mathcal{M};\mathscr{F})$ is a compact operator. By part 2 of Proposition 15.18, we have $\operatorname{ind}\mathbf{P}_s = \operatorname{ind}\mathbf{Q}_s$. $\square$

Given Proposition 15.19, we can define the index associated to a symbol: for a degree m homogeneous polynomial $\mathbf{p} \in C^\infty(T^*\mathcal{M};\pi^*\operatorname{Hom}(\mathscr{E} \to \mathscr{F}))$ (see (13.72)) which is elliptic in the sense of Theorem 14.23, define

$$\operatorname{ind}\mathbf{p} \in \mathbb{Z}$$

to be the index of any $\mathbf{P} \in C^\infty(T^*\mathcal{M};\mathscr{E} \to \mathscr{F})$ such that $\sigma_m(\mathbf{P}) = \mathbf{p}$.

We remark that $\operatorname{ind}\mathbf{p}$ is homotopy invariant: if $\mathbf{p}_r$, $0 \le r \le 1$ is a continuous family of elliptic symbols, then $\operatorname{ind}\mathbf{p}_r$ is independent of r. Indeed, we can choose the family of corresponding differential operators $\mathbf{P}_{(r)}$, $\sigma_m(\mathbf{P}_{(r)}) = \mathbf{p}_r$, depending continuously on r. Then the function $r \mapsto \operatorname{ind}\mathbf{P}_{(r)} = \operatorname{ind}\mathbf{p}_r$ is locally constant, and thus constant. Indeed, fix $r \in [0,1]$ and take some $s \in \mathbb{R}$. We have

$$\|\mathbf{P}_{(r')} - \mathbf{P}_{(r)}\|_{H^{s+m}(\mathcal{M};\mathscr{E}) \to H^s(\mathcal{M};\mathscr{F})} \to 0 \quad \text{as } r' \to r.$$

Therefore, by part 1 of Proposition 15.18 there exists $\varepsilon > 0$ such that for $|r' - r| < \varepsilon$ we have $\operatorname{ind}\mathbf{P}_{(r')} = \operatorname{ind}\mathbf{P}_{(r)}$.

There is a general formula for the index of any elliptic differential operator, known as the *Atiyah–Singer index theorem*. It states that the index $\operatorname{ind}\mathbf{p}$ is the integral of the product of the Chern character of $\mathbf{p}$ (which is a cohomology class with compact support on T^*M) with a certain cohomology class on M pulled back to T^*M. It takes considerable effort to define the two objects above, so we refrain from stating the index theorem here. We instead refer the reader to [**Hör07**, Chapter 19] and [**Tay11b**, Theorem 10.5.1] for details. See also [**Mel93**] whose focus is the case of manifolds with boundary.

One application of the Atiyah–Singer index theorem is to the Gauss–Bonnet theorem for Riemannian surfaces, where the Euler characteristic of the surface appears as the index of the even-to-odd Dirac operator $(d + d^*)_{\text{even}} : C^\infty(\mathcal{M}; \Omega^0 \oplus \Omega^2) \to C^\infty(\mathcal{M}; \Omega^1)$, see Proposition 17.20 below for a bit more information and [**Tay11b**, §10.7] for a detailed presentation (which includes the case of general even dimensions, known as the Chern–Gauss–Bonnet theorem). Another application is the Riemann–Roch theorem on Riemann surfaces, see [**Tay11b**, §10.9].

15.4. Notes and exercises

Our presentation partially follows [**Hör07**, §19.1–19.2] and [**Mel**, Chapter 6].

The Fredholm property of Theorem 15.13 holds for elliptic differential operators on compact manifolds with boundary, if we incorporate the boundary conditions into the operator and assume that they satisfy what is known as the Lopatinski–Shapiro condition. See [**Hör07**, §20.1] for a general treatment. An important special case is that of the Dirichlet boundary value problem for the Laplace–Beltrami operator: if $(\mathcal{M}, g)$ is a compact connected Riemannian manifold with nonempty boundary $\partial\mathcal{M}$, then the operator

$$u \in \overline{H}^2(\mathcal{M}) \ \mapsto\ (\Delta_g u, u|_{\partial\mathcal{M}}) \in L^2(\mathcal{M}) \oplus H^{\frac{3}{2}}(\mathcal{M})$$

is invertible. Here $\overline{H}^2(\mathcal{M})$ is a Sobolev space on $\mathcal{M}$ as a manifold with boundary, defined by requiring that $\partial^\alpha u \in L^2(\mathcal{M})$ for $|\alpha| \leq 2$. The boundary restriction operator $u \mapsto u|_{\partial M}$ can be defined following Exercise 12.9. One does not need the machinery of general elliptic differential operators to solve the Dirichlet problem, there is a much simpler approach using the Dirichlet principle – see for example [**Tay11b**, §5.1], [**Eva10**, Chapter 6], or [**Hör07**, pp.28–29].

Exercise 15.1. (0.5 pt) *Assume that $U \Subset \mathbb{R}^n$, $m, s \in \mathbb{R}$, $a \in S^m(U \times \mathbb{R}^n)$, and $\chi, \psi \in C_c^\infty(U)$. Show that $\psi \operatorname{Op}(a)^t \chi$ is a bounded operator $H^{s+m}(\mathbb{R}^n) \to H^s(\mathbb{R}^n)$, and thus $\operatorname{Op}(a)^t$ is a sequentially continuous operator $H_c^{s+m}(U) \to H_{\text{loc}}^s(U)$. (Hint: use the mapping property of $\operatorname{Op}(a)$ proved in Proposition 15.3 and the duality statement, Proposition 12.7.)*

Exercise 15.2. (1 pt) *Show that the following elliptic estimate for the Laplacian Δ on $\mathbb{R}^2$,*

$$\|\psi u\|_{H^2(\mathbb{R}^2)} \leq C\|\chi \Delta u\|_{L^2(\mathbb{R}^2)} + C\|\chi u\|_{L^2(\mathbb{R}^2)}$$

does not hold when $\psi = \chi$. (You may choose $\chi \in C_c^\infty(\mathbb{R}^2)$ as you want. Hint: try to construct a sequence of solutions to $\Delta u = 0$ of the form $f(x_1)g(x_2)$.)

Exercise 15.3. (1.5 pts) *Let $a \in \mathbb{R}$. Fix $\chi \in C_c^\infty(\mathbb{R})$ such that $\int_{\mathbb{R}} \chi(x)\,dx = 1$ and put $u_k(x) = k^a \chi(kx)$ for $k \in \mathbb{N}$. For which $s \in \mathbb{R}$ is the sequence u_k bounded in $H^s(\mathbb{R})$? For which s does it have a limit in $H^s(\mathbb{R})$, and what is this limit?*

EXERCISE 15.4. (1 pt) *Show* (15.33). *(Hint: use* (13.63) *to understand the kernel of* Δ_g*, and note that* Δ_g *is its own transpose with respect to the density* $d\,\mathrm{vol}_g$*.)*

EXERCISE 15.5. (1 pt) *Let* $\mathcal{M}$ *be a compact manifold and* $P \in \mathrm{Diff}^m(\mathcal{M})$ *be an elliptic differential operator. Show that* $\operatorname{ind} P = 0$*. (Hint: fixing a smooth positive density, we can think of the transpose* P^t *as an operator in* $\mathrm{Diff}^m(\mathcal{M})$*. What is its principal symbol?)*

EXERCISE 15.6. (1 pt) *This exercise gives a basic example of a* 0^{th} *order pseudo-differential operator on the circle* $\mathbb{S}^1 = \mathbb{R}/2\pi\mathbb{Z}$ *which has nonzero index. Consider the operators* $\Pi^\pm$ *on* $L^2(\mathbb{S}^1)$ *defined using Fourier series as follows:*

$$\Pi^\pm\Big(\sum_{k\in\mathbb{Z}} c_k e^{ikx}\Big) = \sum_{\substack{k\in\mathbb{Z}\\ \pm k>0}} c_k e^{ikx}$$

for any sequence $(c_k) \in \ell^2(\mathbb{Z})$*. Let* $\ell \in \mathbb{Z}$ *and define the operator* P *on* $L^2(\mathbb{S}^1)$ *by*

$$Pf(x) = e^{i\ell x}\Pi^+ f(x) + \Pi^- f(x), \quad f \in L^2(\mathbb{S}^1).$$

Show that P *is a Fredholm operator of index* $-\ell$*. (With more knowledge of microlocal analysis, one could actually show that this is true with* $e^{i\ell x}$ *replaced by any nonvanishing function* $a \in C^\infty(\mathbb{S}^1)$*, and* $\operatorname{ind} P = -\frac{1}{2\pi}\int_{\mathbb{S}^1} \frac{a'(x)}{a(x)}\,dx$ *is minus the winding number of the curve* $a : \mathbb{S}^1 \to \mathbb{C}$ *about the origin – this is a 'baby index theorem'.)*

CHAPTER 16

Spectral theory

In this chapter we study the spectral theory of elliptic self-adjoint operators on compact manifolds. The main result is Theorem 16.1 which shows that the spectrum is given by a sequence of eigenvalues going to infinity and there is a Hilbert basis of corresponding eigenfunctions. The proof uses the Fredholm property (Theorem 15.13), the Rellich–Kondrachov Theorem 15.10, and the Hilbert–Schmidt Theorem 16.3 below.

16.1. Spectral theorem for self-adjoint elliptic operators

16.1.1. Self-adjoint operators and statement of the spectral theorem. We first introduce adjoints of differential operators. Let $\mathcal{M}$ be a compact manifold. We fix a positive density (see §13.1.7)

$$\omega_0 \in C^\infty(\mathcal{M}; |\Omega|), \quad \omega_0 > 0.$$

For two functions $f, g \in L^2(\mathcal{M})$, define their inner product by

$$\langle f, g\rangle_{L^2(\mathcal{M};\omega_0)} := \int_{\mathcal{M}} f\overline{g}\,\omega_0. \tag{16.1}$$

The resulting space $L^2(\mathcal{M};\omega_0)$ is a separable Hilbert space.

If $P \in \operatorname{Diff}^m(\mathcal{M})$ is a differential operator, then the adjoint $P^* \in \operatorname{Diff}^m(\mathcal{M})$ with respect to the density ω_0 is defined by the identity

$$\langle Pf, g\rangle_{L^2(\mathcal{M};\omega_0)} = \langle f, P^* g\rangle_{L^2(\mathcal{M};\omega_0)} \quad \text{for all } f, g \in C^\infty(\mathcal{M}). \tag{16.2}$$

Using the density ω_0 to identify densities on $\mathcal{M}$ with functions, we view the transpose P^t (see (13.73)) as an operator in $\operatorname{Diff}^m(\mathcal{M})$. Then P^* and P^t are related by the formula

$$P^*\overline{u} = \overline{P^t u} \quad \text{for all } u \in \mathcal{D}'(\mathcal{M}). \tag{16.3}$$

For $s \in \mathbb{R}$, let $P_s : H^{s+m}(\mathcal{M}) \to H^s(\mathcal{M})$ be the action of the operator P on Sobolev spaces, see (15.28). The *spectrum* of P_s is defined as follows:

$$\operatorname{Spec}(P_s) := \{\lambda \in \mathbb{C} \mid P_s - \lambda \text{ is not invertible}\}. \tag{16.4}$$

Here by Banach's bounded inverse theorem, if $P_s - \lambda : H^{s+m}(\mathcal{M}) \to H^s(\mathcal{M})$ is invertible, then the inverse is a bounded operator $H^s(\mathcal{M}) \to H^{s+m}(\mathcal{M})$.

We are now ready to state the main result of this chapter. We use the term 'formally self-adjoint' to keep in line with general spectral theory, see §16.3 below.

THEOREM 16.1 (Spectral Theorem). *Assume that $\mathcal{M}$ is a compact manifold with a given positive density ω_0, $m \geq 1$, and $P \in \mathrm{Diff}^m(\mathcal{M})$ is an elliptic differential operator. Assume that P is formally self-adjoint on $L^2(\mathcal{M};\omega_0)$ in the sense that $P^* = P$ where P^* is defined in (16.2). Then there exist sequences indexed by $k \in \mathbb{N}$*

$$u_k \in C^\infty(\mathcal{M}), \quad \lambda_k \in \mathbb{R}, \quad |\lambda_k| \to \infty,$$

such that u_k is an eigenfunction of P with eigenvalue λ_k:

$$Pu_k = \lambda_k u_k,$$

and $\{u_k\}$ is a Hilbert basis of $L^2(\mathcal{M};\omega_0)$, namely it is an orthonormal system in $L^2(\mathcal{M};\omega_0)$ and the span of $\{u_k\}$ is dense in $L^2(\mathcal{M};\omega_0)$. Moreover, the spectrum of P_s for any s is given by

$$\mathrm{Spec}(P_s) = \mathrm{Spec}(P) = \{\lambda_k \mid k \in \mathbb{N}\}. \tag{16.5}$$

REMARK 16.2.$^{\mathbf{X}}$ *Theorem 16.1 does not hold for $m = 0$. In this case P is a multiplication operator: $Pu = au$ for some $a \in C^\infty(\mathcal{M};\mathbb{R})$. The spectrum of P is the range of a, which is typically an interval in $\mathbb{R}$, and is not a discrete set unlike (16.5). Theorem 16.1 also does not apply to noncompact manifolds: for example, the spectrum of the Laplacian Δ on $\mathbb{R}^n$ is the half-line $(-\infty, 0]$, since the Fourier transform conjugates Δ to the multiplication operator by $-|\xi|^2$. See Exercise 16.1 below for another concrete example.*

The standard example of an operator to which Theorem 16.1 applies is the Laplace–Beltrami operator $\Delta_g \in \mathrm{Diff}^2(\mathcal{M})$ associated to a Riemannian metric g on $\mathcal{M}$. Here we put $\omega_0 := d\,\mathrm{vol}_g$ and formal self-adjointness of Δ_g follows from (13.63). We discuss more advanced results on the eigenvalues and eigenfunctions of Δ_g in §16.2 below.

As a consequence of Theorem 16.1, we can write any $f \in L^2(\mathcal{M})$ as the sum of a generalized Fourier series

$$f = \sum_{k=1}^{\infty} f_k u_k \quad \text{where } f_k \in \mathbb{C},\ \sum_{k=1}^{\infty} |f_k|^2 < \infty. \tag{16.6}$$

This makes it possible to write down solutions for the heat and the wave equation on $\mathbb{R}_t \times \mathcal{M}_x$. For example, if $\{u_k\}$ is a Hilbert basis of eigenfunctions of $-\Delta_g$ with eigenvalues $\lambda_k \geq 0$ (see §16.2 below), then the solution to the initial value problem for the heat equation

$$(\partial_t - \Delta_g)u(t,x) = 0, \quad t \geq 0,\ x \in \mathcal{M},$$
$$u(0,x) = f(x)$$

is given by the Fourier series

$$u(t,x)=\sum_{k=1}^{\infty}e^{-t\lambda_k}f_k u_k(x)\quad\text{where }f(x)=\sum_{k=1}^{\infty}f_k u_k(x).$$

Using this, one can show for example that, assuming that $\mathcal{M}$ is connected, we have exponential convergence to equilibrium

$$u(t,\bullet)=\frac{1}{\operatorname{vol}_g(\mathcal{M})}\int_{\mathcal{M}}f\,d\operatorname{vol}_g+\mathcal{O}(e^{-\delta t})_{L^2(\mathcal{M})}$$

where $\delta>0$, sometimes called the *spectral gap* of $\mathcal{M}$, is the smallest positive eigenvalue of $-\Delta_g$.

The proof of Theorem 16.1 below generalizes to the case of an elliptic formally self-adjoint operator $\mathbf{P}\in\operatorname{Diff}^m(\mathcal{M};\mathscr{E}\to\mathscr{E})$ acting on sections of a complex vector bundle $\mathscr{E}$ (see Remark 15.2). Here to make sense of the adjoint $\mathbf{P}^*$, in addition to ω_0 we fix an Hermitian inner product $\langle\bullet,\bullet\rangle_{\mathscr{E}(x)}$ on each fiber $\mathscr{E}(x)$ which depends smoothly on $x\in\mathcal{M}$. Then we put

$$\langle f,g\rangle_{L^2(\mathcal{M};\mathscr{E})}:=\int_{\mathcal{M}}\langle f(x),g(x)\rangle_{\mathscr{E}(x)}\,\omega_0\quad\text{for all }f,g\in L^2(\mathcal{M};\mathscr{E})\tag{16.7}$$

and define $\mathbf{P}^*$ similarly to (16.2). The standard examples of formally self-adjoint operators on vector bundles are the Hodge Laplacian and the Dirac operator on differential forms, see §17.3.3 below.

16.1.2. Compact self-adjoint operators$^{\mathbf{R}}$. The proof of Theorem 16.1 uses the following general statement from functional analysis on Hilbert spaces:

THEOREM 16.3 (Hilbert–Schmidt Theorem). *Assume that $\mathcal{H}$ is a Hilbert space and $A:\mathcal{H}\to\mathcal{H}$, $A\neq 0$, is a compact self-adjoint operator (see Definition 15.16). Then A has a nonzero real eigenvalue, that is there exist $\mu\in\mathbb{R}\setminus\{0\}$ and $v\in\mathcal{H}$, $v\neq 0$, such that $Av=\mu v$.*

REMARK 16.4.$^{\mathbf{X}}$ *The full Hilbert–Schmidt theorem states that there exists an orthonormal basis of eigenvectors of A. We do not give this part of the statement because we do not use it in the proof of Theorem 16.1 below.*

PROOF. 1. We first use self-adjointness of A to show the following identity:

$$\|A\|_{\mathcal{H}\to\mathcal{H}}=r\quad\text{where }r:=\sup_{u\in\mathcal{H},\,u\neq 0}\frac{|\langle Au,u\rangle|}{\|u\|_{\mathcal{H}}^2}.\tag{16.8}$$

The $\geq$ inequality in (16.8) follows from Cauchy–Schwarz. To show the $\leq$ inequality, we estimate for all $u, v \in \mathcal{H}$

$$\begin{aligned} 4\,\mathrm{Re}\langle Au, v\rangle_{\mathcal{H}} &= \langle A(u+v), u+v\rangle_{\mathcal{H}} - \langle A(u-v), u-v\rangle_{\mathcal{H}} \\ &\leq r\big(\|u+v\|_{\mathcal{H}}^2 + \|u-v\|_{\mathcal{H}}^2\big) \\ &= 2r\big(\|u\|_{\mathcal{H}}^2 + \|v\|_{\mathcal{H}}^2\big). \end{aligned}$$

Here in the first line we used that A is self-adjoint and thus $\langle Av, u\rangle_{\mathcal{H}} = \langle v, Au\rangle_{\mathcal{H}}$.

Putting $v := tAu$, we see that for all $t \geq 0$

$$4t\|Au\|_{\mathcal{H}}^2 \leq 2r\big(\|u\|_{\mathcal{H}}^2 + t^2\|Au\|_{\mathcal{H}}^2\big).$$

Assuming that $Au \neq 0$ and putting $t := \|u\|_{\mathcal{H}}/\|Au\|_{\mathcal{H}}$, we get from here that $\|Au\|_{\mathcal{H}} \leq r\|u\|_{\mathcal{H}}$, finishing the proof of (16.8).

2. Since $A \neq 0$, we know from (16.8) that $r > 0$. Take a sequence

$$u_k \in \mathcal{H}, \quad \|u_k\|_{\mathcal{H}} = 1, \quad |\langle Au_k, u_k\rangle_{\mathcal{H}}| \to r.$$

Note that $\langle Au, u\rangle_{\mathcal{H}}$ is always real. Thus we may assume that $\langle Au_k, u_k\rangle_{\mathcal{H}}$ converges to either r or $-r$. Without loss of generality (replacing A with $-A$ if necessary) we then assume that

$$\langle Au_k, u_k\rangle_{\mathcal{H}} \to r. \tag{16.9}$$

Since A is a compact operator and $\|u_k\|_{\mathcal{H}}$ is bounded, we can pass to a subsequence to make

$$Au_k \to v \quad \text{in } \mathcal{H} \quad \text{for some } v \in \mathcal{H}.$$

We claim that v is an eigenvector of A with eigenvalue r. To show this, we bound

$$\begin{aligned} \|Au_k - ru_k\|_{\mathcal{H}}^2 &= \|Au_k\|_{\mathcal{H}}^2 - 2r\langle Au_k, u_k\rangle_{\mathcal{H}} + r^2\|u_k\|_{\mathcal{H}}^2 \\ &\leq 2r^2 - 2r\langle Au_k, u_k\rangle_{\mathcal{H}} \to 0. \end{aligned} \tag{16.10}$$

Here in the inequality we used (16.8) and in the limiting statement we used (16.9).

Since $Au_k \to v$ in $\mathcal{H}$, (16.10) implies that

$$ru_k \to v \quad \text{in } \mathcal{H}.$$

Therefore $\|v\|_{\mathcal{H}} = r > 0$. Moreover, $rAu_k = Aru_k$ converges to both rv and Av in $\mathcal{H}$, thus $Av = rv$. This shows that v is an eigenvector of A with eigenvalue r. □

16.1.3. Proof of the spectral theorem. We now prove Theorem 16.1. The proof proceeds in several steps. To simplify notation, we denote $L^2(\mathcal{M}) := L^2(\mathcal{M}; \omega_0)$.

1. We first show that the spectrum of P_s is real, independent of s, and consists only of eigenvalues:

$$\mathrm{Spec}(P_s) = \{\lambda \in \mathbb{R} \mid \exists u \in C^\infty(\mathcal{M}),\ u \neq 0,\ Pu = \lambda u\}. \tag{16.11}$$

To see this, let $\lambda \in \mathbb{C}$. Since $P \in \mathrm{Diff}^m(\mathcal{M})$ and $m \geq 1$, the operator $P-\lambda \in \mathrm{Diff}^m(\mathcal{M})$ has the same principal symbol as P. In particular, since P is an elliptic differential operator, so is $P-\lambda$. By Theorem 15.13, $P_s - \lambda : H^{s+m}(\mathcal{M}) \to H^s(\mathcal{M})$ is a Fredholm operator, and we have

$$\lambda \notin \mathrm{Spec}(P_s) \iff \ker(P-\lambda) = \ker(P^t - \lambda) = 0 \tag{16.12}$$

where $\ker(P-\lambda), \ker(P^t-\lambda) \subset C^\infty(\mathcal{M})$ are defined in (15.31) and we use the density ω_0 to identify densities on $\mathcal{M}$ with functions. Since P is formally self-adjoint, by (16.3) we see that $\ker(P^t - \lambda)$ is the complex conjugate of the vector space $\ker(P - \bar\lambda)$, so in particular

$$\ker(P^t - \lambda) = 0 \iff \ker(P - \bar\lambda) = 0. \tag{16.13}$$

Assume that $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Then for each $u \in \ker(P-\lambda)$ we compute

$$(\mathrm{Im}\,\lambda)\|u\|^2_{L^2(\mathcal{M})} = \mathrm{Im}\langle Pu, u\rangle_{L^2(\mathcal{M})} = 0$$

where in the last equality we used that P is formally self-adjoint, so by (16.2) with $f = g = u$ we have $\langle Pu, u\rangle_{L^2(\mathcal{M})} = \langle u, Pu\rangle_{L^2(\mathcal{M})} = \overline{\langle Pu, u\rangle_{L^2(\mathcal{M})}}$. It follows that $\ker(P-\lambda) = 0$. Similarly we have $\ker(P-\bar\lambda) = 0$, so by (16.12) and (16.13) we have $\lambda \notin \mathrm{Spec}(P_s)$. We have thus shown that the spectrum $\mathrm{Spec}(P_s)$ is contained inside $\mathbb{R}$.

Assume now that $\lambda \in \mathbb{R}$. Then by (16.12) and (16.13) we see that $\lambda \notin \mathrm{Spec}(P_s)$ if and only if $\ker(P-\lambda) = 0$. This finishes the proof of (16.11).

Henceforth we denote for each $\lambda \in \mathbb{R}$ the eigenspace

$$E_\lambda := \ker(P-\lambda) \ \subset \ C^\infty(\mathcal{M}),$$

which by (16.11) is nontrivial if and only if $\lambda \in \mathrm{Spec}(P) = \mathrm{Spec}(P_s)$. Since $P_s - \lambda$ is a Fredholm operator, each space E_λ is finite dimensional.

2. We next claim that the spectrum $\mathrm{Spec}(P)$ is a discrete subset of $\mathbb{R}$. Fix $\lambda_0 \in \mathrm{Spec}(P)$; we need to show that there exists $\varepsilon > 0$ such that $\lambda \notin \mathrm{Spec}(P)$ for all $\lambda \in \mathbb{R}$ such that $0 < |\lambda - \lambda_0| < \varepsilon$.

Define the L^2-orthogonal complements

$$\begin{aligned} L^2_\perp &:= \{u \in L^2(\mathcal{M}) \mid \text{for all } v \in E_{\lambda_0} \quad \text{we have } \langle u, v\rangle_{L^2(\mathcal{M})} = 0\}, \\ H^m_\perp &:= L^2_\perp \cap H^m(\mathcal{M}). \end{aligned}$$

We have

$$L^2(\mathcal{M}) = L^2_\perp \oplus E_{\lambda_0}, \quad H^m(\mathcal{M}) = H^m_\perp \oplus E_{\lambda_0} \tag{16.14}$$

where the latter statement follows from the fact that the orthogonal projector $L^2(\mathcal{M}) \to L^2_\perp$ maps $H^m(\mathcal{M}) \to H^m_\perp$ since $E_{\lambda_0} \subset C^\infty(\mathcal{M})$.

The operator $P_0 - \lambda_0 : H^m(\mathcal{M}) \to L^2(\mathcal{M})$ has range equal to $L^2_\perp$ by (15.32) and the discussion preceding (16.13). Its restriction to $H^m_\perp$,

$$P_\perp - \lambda_0 : H^m_\perp \to L^2_\perp$$

is an invertible operator. Indeed, the kernel of $P_\perp - \lambda_0$ is given by $H^m_\perp \cap E_{\lambda_0} = \{0\}$, and the range of $P_\perp - \lambda_0$ is equal to $\operatorname{ran}(P_0 - \lambda_0) = L^2_\perp$. By Banach's Bounded Inverse Theorem, $P_\perp - \lambda_0$ has bounded inverse.

For any $\lambda \in \mathbb{R}$, we can write the operator $P_0 - \lambda : H^m(\mathcal{M}) \to L^2(\mathcal{M})$ in block-diagonal form with respect to the decompositions (16.14):

$$P_0 - \lambda = \begin{pmatrix} P_\perp - \lambda & 0 \\ 0 & \lambda_0 - \lambda \end{pmatrix}.$$

Since $P_\perp - \lambda_0$ has bounded inverse, there exists $\varepsilon > 0$ such that if $|\lambda - \lambda_0| < \varepsilon$ then $P_\perp - \lambda$ is invertible. If additionally $\lambda \neq \lambda_0$, then we see that $P_0 - \lambda$ is invertible, that is $\lambda \notin \operatorname{Spec}(P)$. This finishes the proof of discreteness of $\operatorname{Spec}(P)$.

3. The spaces E_λ are orthogonal to each other for different λ. Indeed, if $\lambda \neq \lambda'$ and $u \in E_\lambda$, $u' \in E_{\lambda'}$ then we compute

$$\lambda \langle u, u' \rangle_{L^2(\mathcal{M})} = \langle Pu, u' \rangle_{L^2(\mathcal{M})} = \langle u, Pu' \rangle_{L^2(\mathcal{M})} = \lambda' \langle u, u' \rangle_{L^2(\mathcal{M})},$$

implying that $\langle u, u' \rangle_{L^2(\mathcal{M})} = 0$.

Putting together orthonormal bases of all the spaces E_λ, $\lambda \in \operatorname{Spec}(P)$, we arrive to an L^2-orthonormal system

$$u_k \in C^\infty(\mathcal{M}), \quad k \in \mathscr{K}, \tag{16.15}$$

and a collection of numbers $\lambda_k \in \mathbb{R}$ such that $Pu_k = \lambda_k u_k$. Here the index set $\mathscr{K}$ is at most countable, in fact since each E_λ is finite dimensional and the set $\operatorname{Spec}(P)$ is discrete we have

$$\#\{k \in \mathscr{K} : |\lambda_k| \leq R\} < \infty \quad \text{for all } R \in \mathbb{R}. \tag{16.16}$$

However, at this stage in the argument we have not excluded the possibility that $\mathscr{K}$ is finite or even empty.

4. We now show that the system u_k constructed in (16.15) is a Hilbert basis, that is the span of this system is dense in $L^2(\mathcal{M})$. Since $L^2(\mathcal{M})$ is infinite dimensional[1] this implies that the index set $\mathscr{K}$ is infinite, so we can just take $\mathscr{K} = \mathbb{N}$. The sequence λ_k satisfies $|\lambda_k| \to \infty$ by (16.16), so the Hilbert basis property finishes the proof of Theorem 16.1.

To show the Hilbert basis property, it suffices to prove that the orthogonal complement of the span of $\{u_k\}$, given by

$$\mathcal{H} := \{u \in L^2(\mathcal{M}) \mid \text{for all } \lambda \in \operatorname{Spec}(P),\ v \in E_\lambda \quad \text{we have } \langle u, v \rangle_{L^2(\mathcal{M})} = 0\},$$

is equal to $\{0\}$. We argue by contradiction. Assume that $\mathcal{H} \neq \{0\}$. Since $\mathcal{H}$ is a closed subspace in $L^2(\mathcal{M})$, it is a Hilbert space.

[1] For a particularly pedantic reader, we should have assumed in the statement of Theorem 16.1 that the manifold $\mathcal{M}$ has positive dimension and is nonempty.

Since $\mathrm{Spec}(P)$ is discrete, we can fix some $\lambda_\emptyset \in \mathbb{R} \setminus \mathrm{Spec}(P)$. Consider the inverse

$$Q := (P_0 - \lambda_\emptyset)^{-1} : L^2(\mathcal{M}) \to H^m(\mathcal{M}).$$

Then Q is a compact operator $L^2(\mathcal{M}) \to L^2(\mathcal{M})$, since $H^m(\mathcal{M})$ embeds compactly into $L^2(\mathcal{M})$ by Theorem 15.10. Moreover, Q is a self-adjoint operator on $L^2(\mathcal{M})$ since for each $u, v \in L^2(\mathcal{M})$ we have

$$\langle Qu, v\rangle_{L^2(\mathcal{M})} = \langle Qu, (P - \lambda_\emptyset)Qv\rangle_{L^2(\mathcal{M})} = \langle (P - \lambda_\emptyset)Qu, Qv\rangle_{L^2(\mathcal{M})} = \langle u, Qv\rangle_{L^2(\mathcal{M})}$$

where the second equality above follows from (16.2), the formal self-adjointness of P, and the fact that $C^\infty(\mathcal{M})$ is dense in $H^m(\mathcal{M})$.

The operator Q maps the space $\mathcal{H}$ into itself. Indeed, assume that $u \in \mathcal{H}$. Then we have for all $\lambda \in \mathrm{Spec}(P)$ and $v \in E_\lambda$

$$\langle Qu, v\rangle_{L^2(\mathcal{M})} = \langle u, Qv\rangle_{L^2(\mathcal{M})} = 0$$

since $Qv = (\lambda - \lambda_\emptyset)^{-1}v$. Thus $Qu \in \mathcal{H}$.

The discussion above shows that the restriction $A := Q|_{\mathcal{H}}$ is a compact self-adjoint operator on the Hilbert space $\mathcal{H}$. We also have $A \neq 0$ since $\mathcal{H} \neq \{0\}$ and for any $v \in \mathcal{H}$ we have $(P - \lambda_\emptyset)Av = v$. Now Theorem 16.3 applies and shows that A has an eigenvalue, more precisely there exist

$$v \in \mathcal{H} \setminus \{0\}, \quad \mu \in \mathbb{R} \setminus \{0\}, \quad Av = \mu v.$$

We have $(P - \lambda_\emptyset)\mu v = v$, which implies that $v \in E_\lambda$ with $\lambda := \lambda_\emptyset + \mu^{-1}$. This gives a contradiction with the fact that $v \in \mathcal{H}$ and finishes the proof that $\{u_k\}$ is a Hilbert basis of $L^2(\mathcal{M})$.

16.2. Advanced results on Laplacian eigenvalues and eigenfunctions

We now discuss various classical and recent results on the spectrum of the Laplace–Beltrami operator Δ_g.

16.2.1. Basics and examples. Let $(\mathcal{M}, g)$ be a compact Riemannian manifold. We assume that $\mathcal{M}$ is connected. It will be convenient for us to use the operator $-\Delta_g$ instead of Δ_g. One advantage is that, as follows from (13.63) (taking both f, g there to be equal to an eigenfunction) the spectrum of $-\Delta_g$ is contained in $[0, \infty)$, and the eigenspace at $\lambda = 0$ consists of constant functions. Ordering the numbers λ_k from Theorem 16.1 in increasing order, we get sequences

$$\begin{gathered} 0 = \lambda_1 < \lambda_2 \leq \lambda_3 \leq \dots, \quad \lambda_k \to \infty, \\ u_k \in C^\infty(\mathcal{M}), \quad -\Delta_g u_k = \lambda_k u_k, \quad \|u_k\|_{L^2(\mathcal{M})} = 1. \end{gathered} \tag{16.17}$$

It is generally impossible to give a formula for the eigenfunctions and eigenvalues of the Laplacian on a given Riemannian manifold. However, it is possible to explicitly describe the Laplacian spectrum for the torus and for the sphere. We give this description in dimension 2 to simplify the formulas, but it generalizes to any dimension.

The case of torus corresponds to Fourier series:

PROPOSITION 16.5. *Assume that $\mathcal{M} = \mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ is the two-dimensional torus, and g is the metric induced by the Euclidean metric on $\mathbb{R}^2$. Then a Hilbert basis of eigenfunctions of $-\Delta_g$ is given by*

$$\lambda_{\mathbf{k}} = 4\pi^2(k_1^2 + k_2^2), \quad \mathbf{k} = (k_1, k_2) \in \mathbb{Z}^2, \quad u_{\mathbf{k}}(x_1, x_2) = e^{2\pi i \mathbf{k}\cdot x}. \tag{16.18}$$

The spectrum of the Laplacian on the 2-sphere was studied in Exercise 13.8 above. Examining the solution to this exercise we can also compute the multiplicities of eigenvalues, yielding

PROPOSITION 16.6. *Assume that $\mathcal{M} = \mathbb{S}^2$ and g is the round metric. Then the spectrum of $-\Delta_g$ is given by*

$$\lambda_k = k(k+1), \quad k \in \mathbb{N}_0, \quad \textit{with multiplicity } 2k+1. \tag{16.19}$$

The corresponding eigenfunctions are the restrictions to $\mathbb{S}^2 \subset \mathbb{R}^3$ of polynomials U on $\mathbb{R}^3$ which are homogeneous of degree k and harmonic (that is, $\Delta_0 U = 0$ where Δ_0 is the Euclidean Laplacian).

From this point on, no proofs are provided in this section, instead we give references to various articles on the topics covered.

16.2.2. Weyl Law$^{\mathbf{X}}$. We first discuss asymptotic behavior of eigenvalues: once you have an infinite discrete set of numbers, it is hard to resist counting them. Let λ_k be given by (16.17). For $R > 0$, define the counting function

$$\mathcal{N}(R) := \{k \in \mathbb{N} \mid \lambda_k \le R^2\}.$$

Let $n := \dim \mathcal{M}$; denote by ω_n the volume of the Euclidean unit ball in $\mathbb{R}^n$ and by $\operatorname{vol}_g(\mathcal{M}) = \int_{\mathcal{M}} d\operatorname{vol}_g$ the Riemannian volume of $\mathcal{M}$. Then the asymptotic behavior of $\mathcal{N}(R)$ is given by

THEOREM 16.7 (Weyl Law). *We have as $R \to \infty$*

$$\mathcal{N}(R) = (2\pi)^{-n}\omega_n \operatorname{vol}_g(\mathcal{M}) R^n + \mathcal{O}(R^{n-1}). \tag{16.20}$$

The reader is encouraged to check that Theorem 16.7 holds in the special cases in Propositions 16.5 and 16.6.

Theorem 16.7 with an $o(R^n)$ remainder was first proved by Weyl [**Wey12**] in the related case of Dirichlet eigenfunctions for domains in $\mathbb{R}^n$. The $\mathcal{O}(R^{n-1})$ remainder

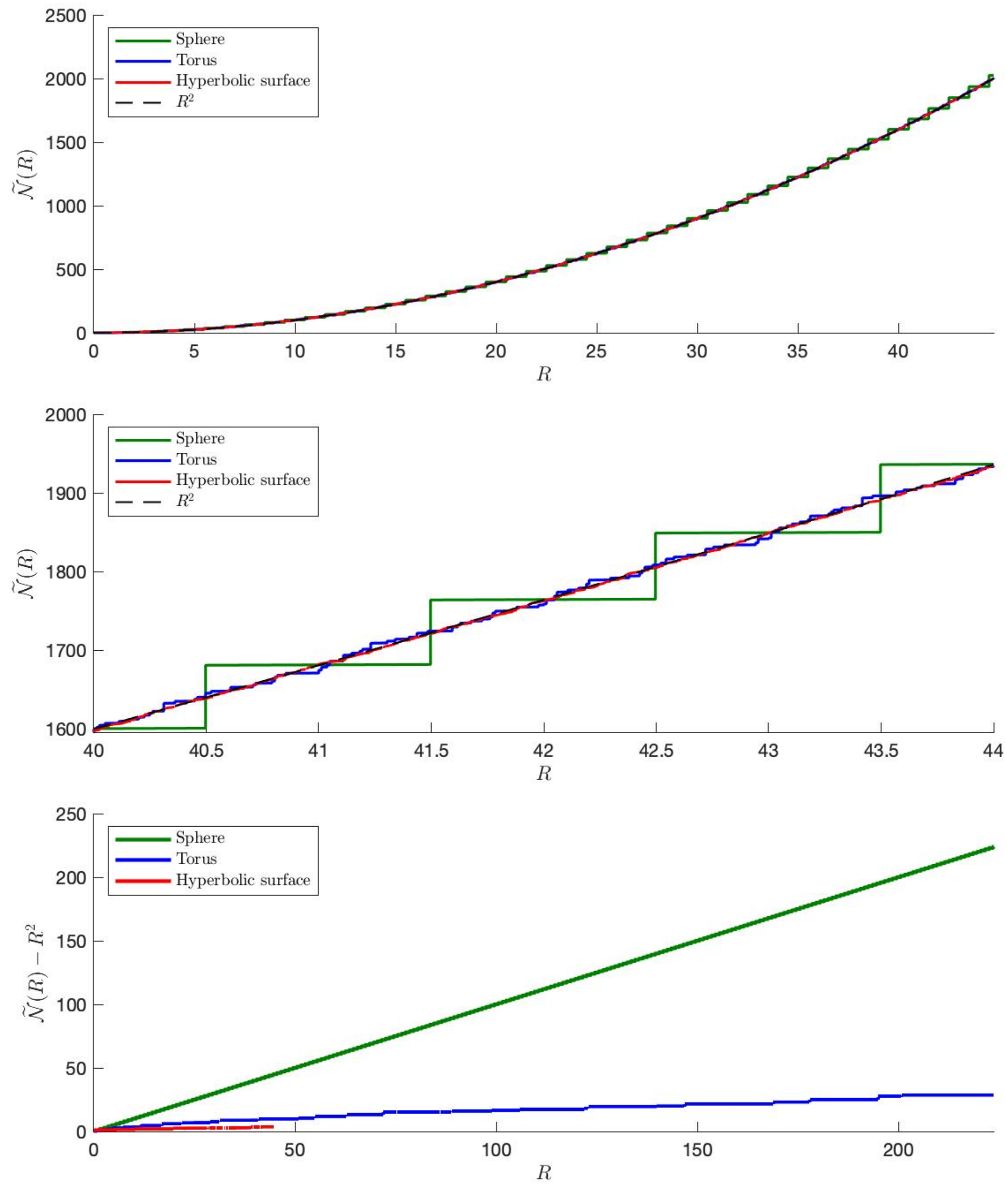

FIGURE 16.1. Eigenvalue counting function for the sphere $\mathbb{S}^2$, the torus $\mathbb{T}^2$, and a hyperbolic surface. The latter uses data provided by Alexander Strohmaier and computed using the method in Strohmaier–Uski [**SU13**]. Top: a plot of the function $\widetilde{\mathcal{N}}(R) := \mathcal{N}(R)/c$ where $c := (2\pi)^{-n}\omega_n \operatorname{vol}_g(\mathcal{M})$. Middle: A zoomed in version of the top plot. Bottom: a plot of the maximum of the error $|\widetilde{\mathcal{N}}(R) - R^2|$ in the Weyl law over the interval $[0, R]$, as a function of R.

is due to Levitan [**Lev53**] and Avakumovič [**Ava56**]. Other proofs can be found for example in [**Hör09**, §29.1] and [**Zwo12**, Theorem 14.11].

A natural question is whether the $\mathcal{O}(R^{n-1})$ remainder in (16.20) can be improved:

- In general this is not possible since on the sphere $\mathbb{S}^2$, the high multiplicity of the spectrum (see Proposition 16.6) means that remainder is not $o(R^{n-1})$.
- However, for most manifolds the remainder can be improved. More precisely, Duistermaat–Guillemin [**DG75**] showed that if the volume of the set of periodic geodesics (considered as a subset of the sphere bundle $S\mathcal{M}$) is equal to 0, then (16.20) holds with remainder $o(R^{n-1})$.
- There are quantitative forms of the remainder under various assumptions on the geodesic flow, see for example [**CG22**] for a review of the literature. In particular, Bérard [**Bér77**] showed that on manifolds without conjugate points (which includes manifolds of nonpositive sectional curvature) the remainder is $\mathcal{O}(R^{n-1}/\log R)$.
- It is a folk conjecture that on negatively curved manifolds (a prime example of which is given by hyperbolic surfaces, which have Gauss curvature -1), the remainder in (16.20) should be $O(R^{n-1-\delta})$ for some $\delta > 0$. This is widely open.
- For the 2-torus, by (16.18) the function $\mathcal{N}(R)$, up to rescaling, is just the number of integer points in the disk of radius R. The question about the remainder is known as the *Gauss circle problem*. It is conjectured that the remainder is $\mathcal{O}(R^{\frac{1}{2}+\varepsilon})$ for any $\varepsilon > 0$, and the best known upper bound to date is $\mathcal{O}(R^{\frac{131}{208}+\varepsilon})$, due to Huxley [**Hux03**].

See Figure 16.1 for a numerical illustration of the Weyl law.

16.2.3. Nodal sets. For this topic we will restrict to *real* eigenfunctions u_k, which is not an issue since $-\Delta_g$ has real coefficients. (So for example, the formula (16.18) would have to be replaced with one featuring sines and cosines.)

The *nodal set* of an eigenfunction u_k is defined to be its zero set:

$$u_k^{-1}(0) = \{x \in \mathcal{M} \mid u_k(x) = 0\}. \tag{16.21}$$

We will present here some results on the area of nodal sets. Generically we expect the nodal set to be a smooth submanifold of $\mathcal{M}$ of codimension 1. In this case, define

$$\operatorname{Area}(u_k^{-1}(0)) \geq 0 \tag{16.22}$$

to be the Riemannian volume of $u_k^{-1}(0)$ (with respect to the restriction of the metric g to $u_k^{-1}(0)$). In general, $\operatorname{Area}(u_k^{-1}(0))$ is defined as the $(n-1)$-dimensional Hausdorff measure of $u_k^{-1}(0)$.

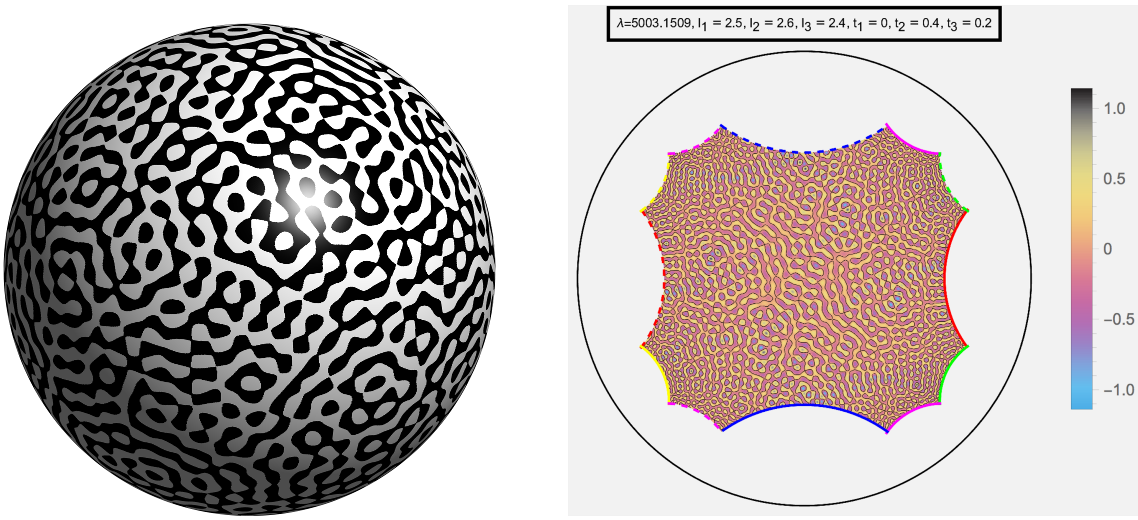

FIGURE 16.2. Left: an eigenfunction on $\mathbb{S}^2$ with the sets $\{u_k > 0\}$ and $\{u_k < 0\}$ separated by the nodal set. Source: https://www.pngwing.com/en/free-png-sooby. Right: a high energy ($\lambda \approx 5000$) eigenfunction on a hyperbolic surface. Here the surface is a quotient of the Poincaré disk model of the hyperbolic space by a group of isometries, and the eigenfunction is drawn on a fundamental domain of this group action. The thin lines are the nodal sets. Picture courtesy of Alexander Strohmaier, computed using the method in Strohmaier–Uski [**SU13**].

In 1982 Shing-Tung Yau made the following conjecture: for each compact Riemannian manifold there exist constants $0 < c \le C$ such that we have for all $k \ge 2$

$$c\sqrt{\lambda_k} \le \operatorname{Area}(u_k^{-1}(0)) \le C\sqrt{\lambda_k}. \tag{16.23}$$

While the upper bound in the conjecture is still open, there have been many significant results on it. We list here only two such results, referring the reader to [**LM20**] for a comprehensive review of the history of the conjecture:

- Donnelly–Fefferman [**DF88**]: the conjecture (16.23) holds if the metric g is *real analytic*;
- Logunov [**Log18a, Log18b**]: for any C^∞ metric we have the bounds

$$c\sqrt{\lambda_k} \le \operatorname{Area}(u_k^{-1}(0)) \le C\lambda_k^N \tag{16.24}$$

  where N is a large constant depending only on the dimension of the manifold $\mathcal{M}$. This in particular settles the lower bound in Yau's conjecture.

See Figure 16.2 for numerical illustrations of nodal sets.

16.2.4. Macroscopic behavior of eigenfunctions. We finally discuss Quantum Ergodicity and related results. To state these we use the notion of weak limits from probability:

DEFINITION 16.8. *Assume that u_{k_ℓ} is a sequence of Laplacian eigenfunctions. We say that it converges weakly to a probability measure μ on $\mathcal{M}$, if*

$$\int_{\mathcal{M}} a(x)|u_{k_\ell}(x)|^2\,d\operatorname{vol}_g(x) \to \int_{\mathcal{M}} a\,d\mu \quad \text{as } \ell\to\infty \quad \text{for all } a\in C^0(\mathcal{M}). \tag{16.25}$$

Note that here we first fix the observable a and then let $\ell \to \infty$, which is why this is a macroscopic limit.

We now assume that $(\mathcal{M}, g)$ has *negative sectional curvature*. This has the consequence that the geodesic flow on $\mathcal{M}$ has strongly chaotic behavior, see for example [**Kli95**, Theorem 3.9.1]. A fundamental example is given by hyperbolic surfaces, see the right half of Figure 16.2. (The theorems below actually apply under various weaker conditions.) We refer the reader to the author's reviews [**Dya21a, Dya21b**] for more information.

- The *Quantum Ergodicity* theorem of Shnirelman [**Shn74a, Shn74b**], Zelditch [**Zel87**], and Colin de Verdière [**CdV85**] states that there exists a *density 1 sequence* of eigenfunctions u_{k_ℓ} which converges weakly to the volume measure $d\operatorname{vol}_g / \operatorname{vol}_g(\mathcal{M})$. This means that in the high energy macroscopic limit, eigenfunctions *equidistribute*.
- The *Quantum Unique Ergodicity* conjecture of Rudnick–Sarnak [**RS94**] states that the *entire sequence* of eigenfunctions converges weakly to the volume measure. It is widely open except for Hecke eigenfunctions in arithmetic cases, see Lindenstrauss [**Lin06**].
- The *entropy bounds* of Anantharaman and Nonnenmacher [**Ana08, AN07**] show that eigenfunctions cannot concentrate too much. For example, no sequence of eigenfunctions can converge weakly to the delta measure on a geodesic (settling a conjecture of Colin de Verdière [**CdV85**]).
- The *lower bounds on mass* of Dyatlov–Jin–Nonnenmacher [**DJN21**] show that in dimension 2, for any nonempty open set $U \Subset \mathcal{M}$ there exists a constant $c_U > 0$ such that for all k

$$\|\mathbf{1}_U u_k\|_{L^2(\mathcal{M};d\operatorname{vol}_g)} \geq c_U. \tag{16.26}$$

  This also shows that no sequence of eigenfunctions can converge weakly to the delta measure on a closed geodesic; in fact, it shows that each weak limit of a sequence of eigenfunctions is a measure of full support. Proving the bound (16.26) in dimensions higher than 2 is currently an open problem.

16.3. Notes and exercises

Our presentation in §16.1 partially follows [**Mel**, §6.5]. We only invoke the theory of *bounded* compact operators on Hilbert spaces at the end of the proof, in the form of Theorem 16.3. Other proofs reduce to this theory earlier on, often by taking $(P+1)^{-1}$ which is possible if P is nonnegative, e.g. $P = -\Delta_g$, but would not work for the Dirac operator studied in §17.3.3 below; see for example [**Eva10**, §6.5].

A differential operator P of positive order does not map $L^2(\mathcal{M})$ to itself. The right abstract theory to study these operators is the spectral theory of unbounded self-adjoint operators. In this theory we have a Hilbert space $\mathcal{H}$, a dense subspace $\mathcal{D} \subset \mathcal{H}$, and a linear operator $P : \mathcal{D} \to \mathcal{H}$ whose graph is a closed subspace of $\mathcal{H} \times \mathcal{H}$. The operator P is called *formally self-adjoint* if

$$\langle Pu, v\rangle_{\mathcal{H}} = \langle u, Pv\rangle_{\mathcal{H}} \quad \text{for all } u, v \in \mathcal{D}.$$

There is a stronger property called *self-adjointness*, which is equivalent to additionally requiring that both operators $P \pm i : \mathcal{D} \to \mathcal{H}$ be invertible. (If P is a bounded operator, i.e. $\mathcal{D} = \mathcal{H}$, then self-adjointness follows from formal self-adjointness.) For self-adjoint operators there is a spectral theorem which shows that they are unitarily conjugated to a multiplication operator on the space L^2 with respect to some measure; the latter is related to the *spectral measure* of the operator P.

In our setting, $P \in \mathrm{Diff}^m(\mathcal{M})$ is a differential operator on a compact manifold without boundary, $\mathcal{H} = L^2(\mathcal{M}; \omega_0)$, and $\mathcal{D} = H^m(\mathcal{M})$. We assumed formal self-adjointness of P and we established actual self-adjointness in Step 1 of the proof of Theorem 16.1. The operator P has compact resolvent, which is why the spectrum is purely discrete. In more general settings, the situation might be different. For example, if $\mathcal{M}$ has a boundary, then one has many choices of the domain of P on which it will be self-adjoint (the two common choices correspond to Dirichlet and Neumann boundary conditions), and if $\mathcal{M}$ is not compact, then the spectrum may no longer be discrete (in fact, typically one would expect that P has no eigenvalues). We refer to the book of Davies [**Dav95**] for details.

EXERCISE 16.1. (1.5 pts) *This exercise gives an example of a self-adjoint operator with non-discrete spectrum. Consider the following multiplication operator on $L^2(\mathbb{S}^1)$ where $\mathbb{S}^1 = \mathbb{R}/2\pi\mathbb{Z}$:*

$$Pf(x) = (\cos x) f(x), \quad f \in L^2(\mathbb{S}^1).$$

Compute the spectrum $\mathrm{Spec}(P)$ *(see (16.4)). For $\lambda \in \mathrm{Spec}(P)$, does the operator $P - \lambda : L^2(\mathbb{S}^1) \to L^2(\mathbb{S}^1)$ have a nontrivial kernel? Is the range of this operator closed?*

(We could actually take any other generic real function in place of $\cos x$*. However, one nice thing about* $\cos x$ *is that using Fourier series, one can see that the operator P is unitarily conjugated to a shifted discrete Laplacian.)*

CHAPTER 17

Differential forms and Hodge theory

In this chapter we discuss differential forms and develop Hodge theory. Most of the material here belongs to differential geometry (differential forms, the operators of Hodge theory) and algebraic topology (de Rham cohomology, Hodge's Theorem). However, the key ingredient bringing de Rham cohomology and Hodge theory together relies on the Fredholm property for elliptic operators (Theorem 15.13). This is why this ingredient, while indispensable to Hodge theory, is often missing from textbooks on it.[1] We also give an application of pullbacks of distributions to degree theory.

17.1. Differential forms$^{\mathbf{R}}$

We first briefly review the standard theory of differential forms and operations on them such as wedge product, differentiation, and integration. We give few details, referring the reader to [**Lee13**, Chapters 14–16] for a comprehensive treatment.

17.1.1. Exterior powers of vector spaces. We start by defining the ℓ-th exterior power of a vector space $\mathcal{V}$, which is the space of all antisymmetric maps $\mathcal{V}^\ell \to \mathbb{R}$:

Definition 17.1. *Let $\mathcal{V}$ be a finite dimensional real vector space and $\ell \in \mathbb{N}_0$. We say that a multilinear map $\mathbf{u} : \mathcal{V}^\ell \to \mathbb{R}$ is* antisymmetric *if for each $v_1, \dots, v_\ell \in \mathcal{V}$ and a permutation γ on $\{1, \dots, \ell\}$ of sign $\operatorname{sgn}\gamma \in \{1, -1\}$ we have*

$$\mathbf{u}(v_{\gamma(1)}, \dots, v_{\gamma(\ell)}) = (\operatorname{sgn}\gamma)\mathbf{u}(v_1, \dots, v_\ell). \tag{17.1}$$

Denote by $\wedge^\ell \mathcal{V}^*$ the space of all antisymmetric maps $\mathcal{V}^\ell \to \mathbb{R}$.

Note that we put $\wedge^0 \mathcal{V}^* = \mathbb{R}$ by definition, and $\wedge^1 \mathcal{V}^* = \mathcal{V}^*$ is the space of linear functionals $\mathcal{V} \to \mathbb{R}$.

All the spaces $\wedge^\ell \mathcal{V}^*$ are finite dimensional. More precisely, if $\dim \mathcal{V} = n$ then

[1] Chern [**Che56**, p.128] writes '*This is the main theorem and we shall not give a proof as the details would take too much time.*'. Huybrechts [**Huy05**, p.285] writes '*The following however requires some hard, but by now standard, analysis.*'. Voisin [**Voi07**] refers to Demailly [**BDIP02**], who in turn reviews the material of our §§12.1,14,15 on 5 pages, replacing most of the details by the comment '*We will need the following fundamental facts, that the reader will be able to find in many of the specialized works devoted to the theory of partial differential equations*' not followed by any references. Other books such as Warner [**War71**] do give an earnest proof of the theorem.

- for $\ell \le n$, the dimension of $\wedge^\ell \mathcal{V}^*$ is equal to the binomial coefficient $\binom{n}{\ell}$;
- for $\ell > n$, the space $\wedge^\ell \mathcal{V}^*$ is trivial.

This can be seen by fixing a basis on $\mathcal{V}$ to identify it with $\mathbb{R}^n$. Then for $\ell \le n$, we have the following basis of $\wedge^\ell (\mathbb{R}^n)^*$, indexed by size ℓ subsets of $\{1, \dots, n\}$:

$$\begin{aligned} &dx_I \in \wedge^\ell (\mathbb{R}^n)^* \quad \text{where } I = \{i_1, \dots, i_\ell\} \subset \{1, \dots, n\}, \quad i_1 < i_2 < \dots < i_\ell, \\ &dx_I(v_1, \dots, v_\ell) = \det\big((dx_{i_j}(v_r))_{j,r=1}^\ell\big) \quad \text{for all } v_1, \dots, v_\ell \in \mathbb{R}^n \end{aligned} \tag{17.2}$$

where $dx_i(v)$ denotes the i-th coordinate of $v \in \mathbb{R}^n$. We often denote $dx_\emptyset = 1$.

The space $\wedge^n (\mathbb{R}^n)^*$ is one-dimensional, spanned by the element $dx_{1\dots n}$ defined by

$$dx_{1\dots n}(v_1, \dots, v_n) = \det([v_1 \dots v_n]) \tag{17.3}$$

where $[v_1 \dots v_n]$ is the matrix with columns $v_1, \dots, v_n$. It follows that for any n-dimensional space $\mathcal{V}$ and any linear map $A : \mathcal{V} \to \mathcal{V}$ we have

$$\mathbf{u}(Av_1, \dots, Av_n) = (\det A)\mathbf{u}(v_1, \dots, v_n) \quad \text{for all } \mathbf{u} \in \wedge^n \mathcal{V}^*, \quad v_1, \dots, v_n \in \mathcal{V}. \tag{17.4}$$

As an example, we describe $\wedge^\ell (\mathbb{R}^3)^*$ for $\ell = 1, 2, 3$:

- a basis of $\wedge^1 (\mathbb{R}^3)^* = (\mathbb{R}^3)^*$ is given by dx_1, dx_2, dx_3 where $dx_j(u) = u_j$ is the j-th coordinate of $u \in \mathbb{R}^3$;
- a basis of $\wedge^2 (\mathbb{R}^3)^*$ is given by $dx_{12}, dx_{13}, dx_{23}$ where for $u, v \in \mathbb{R}^3$

$$\begin{aligned} dx_{12}(u, v) &= \det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}, \\ dx_{13}(u, v) &= \det \begin{pmatrix} u_1 & v_1 \\ u_3 & v_3 \end{pmatrix}, \\ dx_{23}(u, v) &= \det \begin{pmatrix} u_2 & v_2 \\ u_3 & v_3 \end{pmatrix}; \end{aligned} \tag{17.5}$$

- a basis of $\wedge^3 (\mathbb{R}^3)^*$ is given by dx_{123} where for $u, v, w \in \mathbb{R}^3$

$$dx_{123}(u, v, w) = \det \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}.$$

One can define the *wedge product*

$$\mathbf{u} \in \wedge^\ell \mathcal{V}^*,\ \mathbf{v} \in \wedge^m \mathcal{V}^* \quad \mapsto \quad \mathbf{u} \wedge \mathbf{v} \in \wedge^{\ell+m} \mathcal{V}^* \tag{17.6}$$

with the following properties:

- $\mathbf{u} \wedge \mathbf{v}$ is bilinear in $\mathbf{u}, \mathbf{v}$;
- $(\mathbf{u} \wedge \mathbf{v}) \wedge \mathbf{w} = \mathbf{u} \wedge (\mathbf{v} \wedge \mathbf{w})$;
- $\mathbf{v} \wedge \mathbf{u} = (-1)^{\ell m} \mathbf{u} \wedge \mathbf{v}$ where $\mathbf{u} \in \wedge^\ell \mathcal{V}^*$ and $\mathbf{v} \in \wedge^m \mathcal{V}^*$;

- on $\mathbb{R}^n$, we have

$$dx_I = dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_\ell} \quad \text{where } I = \{i_1, \dots, i_\ell\}, \quad i_1 < \ldots < i_\ell.$$

So for example, on $\mathbb{R}^3$ we have

$$dx_1 \wedge dx_2 = -dx_2 \wedge dx_1 = dx_{12}, \quad dx_1 \wedge dx_1 = 0, \quad dx_{13} \wedge dx_2 = dx_2 \wedge dx_{13} = -dx_{123}.$$

It will often be convenient for us to consider the total exterior algebra (where $n = \dim \mathcal{V}$)

$$\wedge^\bullet \mathcal{V}^* := \bigoplus_{\ell=0}^{n} \wedge^\ell \mathcal{V}^* \tag{17.7}$$

which is a 2^n-dimensional space. A basis of $\wedge^\bullet \mathbb{R}^n$ is given by dx_I where I ranges over all subsets of $\{1, \dots, n\}$. One can think of the wedge product as a bilinear map $\wedge^\bullet \mathcal{V}^* \times \wedge^\bullet \mathcal{V}^* \to \wedge^\bullet \mathcal{V}^*$.

17.1.2. Differential forms and exterior derivative. Let $\mathcal{M}$ be an n-dimensional manifold. We define the vector bundle Ω^ℓ of *ℓ-forms* over $\mathcal{M}$:

$$\Omega^\ell(x) := \wedge^\ell T_x^* \mathcal{M}, \quad x \in \mathcal{M}. \tag{17.8}$$

That is, an element of $\Omega^\ell(x)$ is an antisymmetric multilinear map from $(T_x\mathcal{M})^\ell$ to $\mathbb{R}$. Note that $\Omega^\ell(x) = \{0\}$ when $\ell > n$. We use real valued forms here, but one can easily consider instead complex valued forms, which we do without further discussion later.

Similarly to (17.7) we define the total form bundle $\Omega^\bullet$ over $\mathcal{M}$ by

$$\Omega^\bullet(x) := \bigoplus_{\ell=0}^{n} \Omega^\ell(x), \quad x \in \mathcal{M}. \tag{17.9}$$

Smooth sections of Ω^ℓ, that is elements of $C^\infty(\mathcal{M}; \Omega^\ell)$, are called *$\ell$-forms*, while distributional sections in $\mathcal{D}'(\mathcal{M}; \Omega^\ell)$ are called *ℓ-currents*. Note that 0-forms are just functions and 1-forms are sections of the cotangent bundle $T^*\mathcal{M}$ described in §13.1.5.

The wedge product (17.6) is defined on forms as follows: for $\mathbf{u} \in C^\infty(\mathcal{M}; \Omega^\ell)$ and $\mathbf{v} \in C^\infty(\mathcal{M}; \Omega^m)$ we have

$$\mathbf{u} \wedge \mathbf{v} \in C^\infty(\mathcal{M}; \Omega^{\ell+m}), \quad (\mathbf{u} \wedge \mathbf{v})(x) = \mathbf{u}(x) \wedge \mathbf{v}(x), \quad x \in \mathcal{M}. \tag{17.10}$$

Any smooth map $\Phi : \mathcal{M} \to \mathcal{N}$ between two manifolds induces the *pullback operator*

$$\begin{gathered} \Phi^* : C^\infty(\mathcal{N}; \Omega^\ell) \to C^\infty(\mathcal{M}; \Omega^\ell), \\ \Phi^*\mathbf{v}(x)(v_1, \dots, v_\ell) = \mathbf{v}(\Phi(x))(d\Phi(x)v_1, \dots, d\Phi(x)v_\ell) \\ \text{for all } x \in \mathcal{M},\ v_1, \dots, v_\ell \in T_x\mathcal{M}. \end{gathered} \tag{17.11}$$

Note that wedge product is equivariant under pullback:

$$\Phi^*(\mathbf{v} \wedge \mathbf{w}) = (\Phi^*\mathbf{v}) \wedge (\Phi^*\mathbf{w}) \quad \text{for all } \mathbf{v}, \mathbf{w} \in C^\infty(\mathcal{N}; \Omega^\bullet).$$

If $\varkappa : U \to V$ is a chart on $\mathcal{M}$, then to each $\mathbf{u} \in C^\infty(\mathcal{M}; \Omega^\ell)$ corresponds its *pushforward*

$$\varkappa_* \mathbf{u} := (\varkappa^{-1})^* (\mathbf{u}|_U) \ \in \ C^\infty(V; \Omega^\ell). \tag{17.12}$$

We can write $\varkappa_* \mathbf{u}$ in the basis (17.2):

$$\varkappa_* \mathbf{u} = \sum_{\substack{I \subset \{1,\dots,n\} \\ \#(I) = \ell}} u_I(x)\, dx_I, \quad u_I \in C^\infty(V; \mathbb{R}) \tag{17.13}$$

which gives coordinate representations for differential forms.

We next define the *exterior derivative*. This is a family of differential operators (see §13.3.2)

$$d_\ell \in \operatorname{Diff}^1(\mathcal{M}; \Omega^\ell \to \Omega^{\ell+1}), \quad d_\ell : C^\infty(\mathcal{M}; \Omega^\ell) \to C^\infty(\mathcal{M}; \Omega^{\ell+1}), \tag{17.14}$$

defined using coordinate representations (17.13) as follows: for any chart $\varkappa : U \to V$ and $\mathbf{u} \in C^\infty(\mathcal{M}; \Omega^\ell)$ we have

$$\begin{gathered} \varkappa_* d_\ell \mathbf{u} = \sum_I du_I \wedge dx_I \\ \text{where } \varkappa_* \mathbf{u} = \sum_I u_I(x)\, dx_I, \quad du_I = \sum_{j=1}^n \partial_{x_j} u_I(x)\, dx_j. \end{gathered} \tag{17.15}$$

Note that for $\ell = 0$ and $f \in C^\infty(\mathcal{M}; \Omega^0) = C^\infty(\mathcal{M})$, $d_0 f \in C^\infty(\mathcal{M}; \Omega^1) = C^\infty(\mathcal{M}; T^*\mathcal{M})$ is just the differential of the function f defined in (13.17). On the other hand, we have $d_n = 0$ where $n = \dim \mathcal{M}$.

Put together, the operators d_ℓ give the total exterior derivative operator

$$d \in \operatorname{Diff}^1(\mathcal{M}; \Omega^\bullet \to \Omega^\bullet). \tag{17.16}$$

As an example, we compute the exterior derivative on forms in $\mathbb{R}^3$ (using the more commonly used basis $dx_2 \wedge dx_3 = dx_{23}$, $dx_3 \wedge dx_1 = -dx_{13}$, $dx_1 \wedge dx_2 = dx_{12}$ rather than (17.5)):

$$\begin{aligned} df(x) &= \partial_{x_1} f(x)\, dx_1 + \partial_{x_2} f(x)\, dx_2 + \partial_{x_3} f(x)\, dx_3, \\ d(u_1(x)\, dx_1) &= \partial_{x_3} u_1(x)\, dx_3 \wedge dx_1 - \partial_{x_2} u_1(x)\, dx_1 \wedge dx_2, \\ d(u_2(x)\, dx_2) &= \partial_{x_1} u_2(x)\, dx_1 \wedge dx_2 - \partial_{x_3} u_2(x)\, dx_2 \wedge dx_3, \\ d(u_3(x)\, dx_3) &= \partial_{x_2} u_3(x)\, dx_2 \wedge dx_3 - \partial_{x_1} u_3(x)\, dx_3 \wedge dx_1, \\ d(u_{23}(x)\, dx_2 \wedge dx_3) &= \partial_{x_1} u_{23}(x)\, dx_1 \wedge dx_2 \wedge dx_3, \\ d(u_{31}(x)\, dx_3 \wedge dx_1) &= \partial_{x_2} u_{31}(x)\, dx_1 \wedge dx_2 \wedge dx_3, \\ d(u_{12}(x)\, dx_1 \wedge dx_2) &= \partial_{x_3} u_{12}(x)\, dx_1 \wedge dx_2 \wedge dx_3. \end{aligned} \tag{17.17}$$

Note that $d_\ell : C^\infty(\mathbb{R}^3; \Omega^\ell) \to C^\infty(\mathbb{R}^3; \Omega^{\ell+1})$ correspond to the gradient ($\ell = 0$), curl ($\ell = 1$), and divergence ($\ell = 2$) operators in multivariable calculus. (However, this requires us to identify forms of various degrees using the Euclidean inner product and so a more correct interpretation would feature the Hodge star operator, see (17.41) below).

We list below the standard properties of the operator d:

- $d^2 = 0$, that is
$$d_{\ell+1} d_\ell \mathbf{u} = 0 \quad \text{for all } \mathbf{u} \in C^\infty(\mathcal{M}; \Omega^\ell); \tag{17.18}$$
- if $\mathbf{u} \in C^\infty(\mathcal{M}; \Omega^\ell)$ and $\mathbf{v} \in C^\infty(\mathcal{M}; \Omega^m)$ then we have the Leibniz rule
$$d(\mathbf{u} \wedge \mathbf{v}) = (d\mathbf{u}) \wedge \mathbf{v} + (-1)^\ell \mathbf{u} \wedge (d\mathbf{v}); \tag{17.19}$$
- if $\Phi : \mathcal{M} \to \mathcal{N}$ is a smooth map, then for each $\mathbf{v} \in C^\infty(\mathcal{N}; \Omega^\bullet)$ we have
$$d(\Phi^* \mathbf{v}) = \Phi^*(d\mathbf{v}). \tag{17.20}$$

Note that in the example of $\mathbb{R}^3$ considered above, (17.18) corresponds to the curl of the gradient, and the divergence of the curl, being zero. For the case $\ell = 0$ and $m = 1, 2$, the formula (17.19) corresponds to the formulas for the curl and the divergence of the product of a function and a vector field.

17.1.3. Orientation, integration, and Stokes's Theorem. We now discuss integration of differential forms, which is perhaps the main reason why they are so useful. To integrate a differential form in a coordinate invariant way one has to fix only one additional piece of data: a choice of *orientation* on a manifold.

This is not a purely technical issue: certain quantities in multivariable calculus/physics such as work and flux are naturally dependent on the choice of direction of travel. Mathematically this choice is expressed by fixing an orientation, and work and flux are best thought of as integrals of differential forms. On the other hand, quantities such as length, area, mass, or charge do not depend on the choice of orientation, and are best thought of as integrals of densities (see §13.1.7).

We first define the concept of orientation on a vector space:

DEFINITION 17.2. *Let $\mathcal{V}$ be an n-dimensional real vector space. Denote by $\mathscr{B}(\mathcal{V}) \subset \mathcal{V}^n$ the set of bases on $\mathcal{V}$. An* orientation *on $\mathcal{V}$ is a map*
$$o : \mathscr{B}(\mathcal{V}) \to \{-1, 1\}$$
such that for any linear isomorphism $A : \mathcal{V} \to \mathcal{V}$ and any $(v_1, \dots, v_n) \in \mathscr{B}(\mathcal{V})$ we have
$$o(Av_1, \dots, Av_n) = \operatorname{sgn}(\det A) o(v_1, \dots, v_n). \tag{17.21}$$
We say that $(v_1, \dots, v_n) \in \mathscr{B}(\mathcal{V})$ is positively oriented *with respect to o, if $o(v_1, \dots, v_n) = 1$, and* negatively oriented *otherwise.*

Each finite dimensional vector space has exactly two orientations. For $\mathbb{R}^n$, these are the *standard orientation*

$$o(v_1, \dots, v_n) = \operatorname{sgn}\det[v_1 \dots v_n],$$

where $[v_1 \dots v_n]$ denotes the matrix with columns $v_1, \dots, v_n$, and the opposite orientation $-o$.

Next, let $\mathcal{M}$ be a manifold. A basis of $T_x\mathcal{M}$ for some x is called a *frame*. An *orientation* on $\mathcal{M}$ is defined to be a choice of orientation on each tangent space,

$$o(x) : \mathscr{B}(T_x\mathcal{M}) \to \{-1, 1\}, \quad x \in \mathcal{M},$$

which is continuous in x (where continuity can be defined, for instance, using charts on $\mathcal{M}$). We say $\mathcal{M}$ is *oriented* if it is endowed with an orientation, and *orientable* if there exists an orientation on $\mathcal{M}$. There exist nonorientable manifolds, of which the simplest example is the Möbius strip (or if one prefers a compact manifold without boundary, the Klein bottle). If $\mathcal{M}$ is a connected orientable manifold, then it has exactly two possible orientations.

For a diffeomorphism $\Phi : \mathcal{M} \to \mathcal{N}$ of oriented manifolds $\mathcal{M}, \mathcal{N}$, we say that it is *orientation preserving* if its differential maps any positively oriented frame on $\mathcal{M}$ to a positively oriented frame on $\mathcal{N}$, and *orientation reversing* if positively oriented frames are mapped to negatively oriented ones.

We are now ready to define the integral of a differential form. Let $\mathcal{M}$ be an oriented n-dimensional manifold with orientation o and $\mathbf{u} \in C^\infty_{\mathrm{c}}(\mathcal{M}; \Omega^n)$ be an n-form on $\mathcal{M}$. For any point $x \in \mathcal{M}$, basis $(v_1, \dots, v_n) \in \mathscr{B}(T_x\mathcal{M})$, and linear isomorphism $A : T_x\mathcal{M} \to T_x\mathcal{M}$ we have

$$\begin{aligned} \mathbf{u}(x)(Av_1, \dots, Av_n) &= (\det A)\, \mathbf{u}(x)(v_1, \dots, v_n), \\ o(x)(Av_1, \dots, Av_n) &= \operatorname{sgn}(\det A)\, o(x)(v_1, \dots, v_n) \end{aligned}$$

where the first equality follows from (17.4). This implies that the map $o\mathbf{u}(x) : \mathscr{B}(T_x\mathcal{M}) \to \mathbb{R}$ satisfies

$$(o\mathbf{u})(x)(Av_1, \dots, Av_n) = |\det A|\, (o\mathbf{u})(x)(v_1, \dots, v_n),$$

which means that $o\mathbf{u}$ is a density on $\mathcal{M}$ (see §13.1.7):

$$o\mathbf{u} \in C^\infty_{\mathrm{c}}(\mathcal{M}; |\Omega|).$$

Now we can define the integral of the differential form $\mathbf{u}$ by

$$\int_{\mathcal{M}} \mathbf{u} := \int_{\mathcal{M}} o\mathbf{u} \tag{17.22}$$

where the integral on the right is an integral of a density, defined in §13.1.7.

Alternatively one can repeat the procedure in §13.1.7, breaking $\mathbf{u}$ into pieces supported in charts. If $\varkappa : U \to V$ is a chart on $\mathcal{M}$ and $\operatorname{supp} \mathbf{u} \Subset U$, then we have

$$\int_{\mathcal{M}} \mathbf{u} = \pm \int_V u(x)\,dx \quad \text{where } u \in C^\infty_c(V), \quad \varkappa_* \mathbf{u} = u(x)\,dx_1 \wedge \cdots \wedge dx_n,$$

the integral on the right-hand side is the usual integral with respect to Lebesgue measure, and the $\pm$ sign is $+$ when $\varkappa$ is orientation preserving and $-$ if $\varkappa$ is orientation reversing.

The notion of integration extends naturally to $\mathbf{u} \in L^1_c(\mathcal{M})$ and in fact to compactly supported distributions, yielding the linear map

$$\mathbf{u} \in \mathcal{E}'(\mathcal{M}; \Omega^n) \;\mapsto\; \int_{\mathcal{M}} \mathbf{u} \in \mathbb{R}. \tag{17.23}$$

This map is equivariant under diffeomorphisms: if $\Phi : \mathcal{M} \to \mathcal{N}$ is an orientation preserving diffeomorphism then

$$\int_{\mathcal{M}} \Phi^* \mathbf{v} = \int_{\mathcal{N}} \mathbf{v} \quad \text{for all } \mathbf{v} \in \mathcal{E}'(\mathcal{N}; \Omega^n). \tag{17.24}$$

The key result on integration of differential forms is the following

THEOREM 17.3 (Stokes's Theorem for differential forms). *Assume that $\mathcal{M}$ is a compact n-dimensional oriented manifold with boundary $\partial\mathcal{M}$ and $\mathbf{u} \in C^\infty(\mathcal{M}; \Omega^{n-1})$. Then*

$$\int_{\mathcal{M}} d\mathbf{u} = \int_{\partial\mathcal{M}} \mathbf{u}. \tag{17.25}$$

Here $d\mathbf{u} = d_{n-1}\mathbf{u} \in C^\infty(\mathcal{M}; \Omega^n)$ is defined in (17.14).

Theorem 17.3 implies as special cases several theorems from multivariable calculus:

- the usual Fundamental Theorem of Calculus ($\mathcal{M} \subset \mathbb{R}$ is an interval, $n = 1$);
- Fundamental Theorem of Calculus for line integrals ($\mathcal{M} \subset \mathbb{R}^N$ is a curve, $n = 1$);
- Green's Theorem ($\mathcal{M} \subset \mathbb{R}^2$ is a domain, $n = 2$);
- Stokes's Theorem ($\mathcal{M} \subset \mathbb{R}^3$ is a surface, $n = 2$);
- Divergence Theorem in 3D ($\mathcal{M} \subset \mathbb{R}^3$ is a domain, $n = 3$).

To make the statement of 17.3 precise we would have to define the notion of a manifold with boundary and explain how an orientation on $\mathcal{M}$ naturally induces orientation on $\partial\mathcal{M}$ – see for instance [**Lee13**, Theorem 16.11] for details. We do not develop this here, since all we need is the following version:

Theorem 17.4 (Stokes's Theorem without boundary). *Assume that $\mathcal{M}$ is an n-dimensional oriented manifold without boundary and $\mathbf{u} \in C^\infty_c(\mathcal{M}; \Omega^{n-1})$ is an $n-1$-form on $\mathcal{M}$. Then*

$$\int_{\mathcal{M}} d\mathbf{u} = 0. \tag{17.26}$$

In fact, the main case of Theorem 17.4 that we use below is when $\mathcal{M}$ is a compact manifold without boundary, when every $\mathbf{u} \in C^\infty(\mathcal{M}; \Omega^{n-1})$ is automatically compactly supported.

17.2. De Rham cohomology

We now review the de Rham cohomology theory. We again keep the presentation brief and omit most proofs, sending the reader to [Lee13, Chapters 17–18] for details. This theory belongs to *algebraic topology*, which is a big field associating various topological invariants to manifolds and more general topological spaces.

Let $\mathcal{M}$ be a manifold. Recall the exterior derivative $d_\ell : C^\infty(\mathcal{M}; \Omega^\ell) \to C^\infty(\mathcal{M}; \Omega^{\ell+1})$ defined in (17.14). The starting point of the de Rham theory is the identity (17.18), namely $d_\ell d_{\ell-1} = 0$. It implies that the space of *exact* ℓ-forms, defined as

$$d_{\ell-1}(C^\infty(\mathcal{M}; \Omega^{\ell-1})) := \{d_{\ell-1}\mathbf{v} \mid \mathbf{v} \in C^\infty(\mathcal{M}; \Omega^{\ell-1})\} \tag{17.27}$$

is contained inside the space of *closed* ℓ-forms, defined as

$$\ker_{C^\infty} d_\ell := \{\mathbf{u} \in C^\infty(\mathcal{M}; \Omega^\ell) \mid d_\ell \mathbf{u} = 0\}. \tag{17.28}$$

The order ℓ *de Rham cohomology* space is defined as the quotient of the space of closed forms by the space of exact forms:

$$\mathbf{H}^\ell(\mathcal{M}; \mathbb{R}) := \frac{\ker_{C^\infty} d_\ell}{d_{\ell-1}(C^\infty(\mathcal{M}; \Omega^{\ell-1}))}, \tag{17.29}$$

which is a real vector space. The ℓ-th *Betti number* is defined as

$$b_\ell(\mathcal{M}) := \dim \mathbf{H}^\ell(\mathcal{M}; \mathbb{R}) \ \in \ \mathbb{N}_0 \cup \{\infty\}. \tag{17.30}$$

To each closed ℓ-form $\mathbf{u} \in \ker_{C^\infty} d_\ell$ corresponds its *cohomology class*

$$[\mathbf{u}] \in \mathbf{H}^\ell(\mathcal{M}; \mathbb{R}).$$

Here are some basic properties of de Rham cohomology:

- We have $\mathbf{H}^\ell(\mathcal{M}; \mathbb{R}) = \{0\}$ when $\ell > \dim \mathcal{M}$, simply because $\Omega^\ell = 0$.
- If $\mathcal{M}$ is connected, then $b_0(\mathcal{M}) = 1$. Indeed, we have $d_{-1}(C^\infty(\mathcal{M}; \Omega^{-1})) = \{0\}$, since $\Omega^{-1} = 0$ by convention. On the other hand, $\ker_{C^\infty} d_0$ consists of functions $f \in C^\infty(\mathcal{M})$ which satisfy $df = 0$; since $\mathcal{M}$ is connected, such functions are constant.

- If $\mathcal{M} \subset \mathbb{R}^n$ is an open subset which is convex (or more generally, star-shaped), then $\mathbf{H}^\ell(\mathcal{M};\mathbb{R}) = \{0\}$ for all $\ell \geq 1$. This is commonly known as the *Poincaré Lemma*. In particular, this applies to $\mathcal{M} = \mathbb{R}^n$.
- If $\mathcal{M}$ is connected and $\mathbf{u} \in \ker_{C^\infty} d_1$, then $[\mathbf{u}] = 0$ if and only if the integral of $\mathbf{u}$ on every closed curve in $\mathcal{M}$ is equal to 0. (In fact, the space $\mathbf{H}^1(\mathcal{M};\mathbb{R})$ is related to the fundamental group of $\mathcal{M}$, as a corollary of Hurewicz's theorem in algebraic topology.)

Additional information about the cohomology of $\mathcal{M}$ is given by the *product structure*. It follows from (17.19) that the wedge product of two closed forms is closed, and the wedge product of a closed and an exact form is exact. Thus the wedge product descends to cohomology:

$$[\mathbf{u}] \in \mathbf{H}^\ell(\mathcal{M};\mathbb{R}),\ [\mathbf{v}] \in \mathbf{H}^m(\mathcal{M};\mathbb{R}) \ \mapsto\ [\mathbf{u} \wedge \mathbf{v}] \in \mathbf{H}^{\ell+m}(\mathcal{M};\mathbb{R}).$$

17.2.1. Some examples. A basic example of a manifold with nontrivial cohomology $\mathbf{H}^1$ is given by

PROPOSITION 17.5. *Let $\mathcal{M} := \mathbb{R}^2 \backslash \{0\}$. Then $b_1(\mathcal{M}) = 1$ and $\mathbf{H}^1(\mathcal{M};\mathbb{R})$ is spanned by the cohomology class $[\mathbf{u}]$ of the closed 1-form*

$$\mathbf{u} := \frac{x_1\,dx_2 - x_2\,dx_1}{x_1^2 + x_2^2}. \tag{17.31}$$

The part of Proposition 17.5 that is easier to prove is the fact that $[\mathbf{u}] \neq 0$, that is $\mathbf{u}$ is not exact. Indeed, if $\mathbf{u} = df$ for some $f \in C^\infty(\mathbb{R}^2 \setminus \{0\})$, then by Stokes's Theorem 17.4 we see that $\int_\gamma \mathbf{u} = 0$ for any closed curve γ. However, for the curve $\gamma = \{(\cos t, \sin t) \mid 0 \leq t \leq 2\pi\}$, oriented in the direction of increasing t, one can compute $\int_\gamma \mathbf{u} = 2\pi$. In fact, for a general immersed closed curve $\gamma \subset \mathbb{R}^2 \setminus \{0\}$, the integral $\frac{1}{2\pi}\int_\gamma \mathbf{u}$ is the (signed) *winding number* of γ about the origin.

A fundamental example of cohomology computation is that of the sphere $\mathbb{S}^n$: we have

$$b_\ell(\mathbb{S}^n) = \begin{cases} 1, & \ell = 0 \text{ or } \ell = n, \\ 0, & \text{otherwise.} \end{cases} \tag{17.32}$$

Another example is given by compact connected oriented surfaces. Such surfaces are characterized by their *genus* $\mathbf{g} \in \mathbb{N}_0$ which can be interpreted as the 'number of holes': in particular the sphere $\mathbb{S}^2$ has genus 0 and the torus $\mathbb{T}^2$ has genus 1. If $\mathcal{M}$ is a compact connected oriented surface of genus $\mathbf{g}$, then

$$b_0(\mathcal{M}) = b_2(\mathcal{M}) = 1, \quad b_1(\mathcal{M}) = 2\mathbf{g}.$$

As an application of Hodge theory, we will show below (see §17.3.5) that on a *compact* manifold $\mathcal{M}$ (without boundary) of dimension n the spaces $\mathbf{H}^\ell(\mathcal{M};\mathbb{R})$ are finite dimensional and if $\mathcal{M}$ is additionally orientable, then we have *Poincaré duality* $b_{n-\ell}(\mathcal{M}) = b_\ell(\mathcal{M})$; in particular if $\mathcal{M}$ is connected then $b_n(\mathcal{M}) = 1$.

17.3. Hodge theory

Hodge theory is the theory of differential forms on oriented Riemannian manifolds. The choice of metric and orientation induces several operations on differential forms:

- an inner product on the fibers of $\Omega^\bullet$,
- the *Hodge star operators* $\star_\ell : C^\infty(\mathcal{M};\Omega^\ell) \to C^\infty(\mathcal{M};\Omega^{n-\ell})$ where $n = \dim \mathcal{M}$,
- the *codifferentials* $d_\ell^* : C^\infty(\mathcal{M};\Omega^{\ell+1}) \to C^\infty(\mathcal{M};\Omega^\ell)$,
- the *Dirac operator* $d + d^* : C^\infty(\mathcal{M};\Omega^\bullet) \to C^\infty(\mathcal{M};\Omega^\bullet)$,
- and the *Hodge Laplacian* $\Delta_g = (d + d^*)^2$.

We define these in §§17.3.1–17.3.3.

We next specialize to compact manifolds and prove *Hodge's Theorem* which gives a bijection between de Rham cohomology classes and *harmonic forms*, which are elements of the kernel of the Dirac operator (or equivalently, the Hodge Laplacian). The key ingredient of the proof is the Fredholm property of $d+d^*$ which follows from Theorem 15.13. As an application of Hodge's Theorem, we prove Poincaré duality. We also discuss degree theory.

17.3.1. Inner product on exterior powers and the Hodge star. We first define an inner product on exterior powers of vector spaces studied in §17.1.1 above.

Let $\mathcal{V}$ be a finite dimensional real vector space with a fixed inner product $\langle \bullet, \bullet \rangle_{\mathcal{V}}$. Then we have a natural inner product on each exterior power $\wedge^\ell \mathcal{V}^*$, defined as follows:

LEMMA 17.6. *There exists unique inner product $\langle \bullet, \bullet \rangle_{\wedge^\ell \mathcal{V}^*}$ on $\wedge^\ell \mathcal{V}^*$ such that for all $\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{v}_1, \dots, \mathbf{v}_\ell \in \mathcal{V}^*$ we have*

$$\langle \mathbf{u}_1 \wedge \cdots \wedge \mathbf{u}_\ell, \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_\ell \rangle_{\wedge^\ell \mathcal{V}^*} = \det\big((\langle \mathbf{u}_j, \mathbf{v}_k \rangle_{\mathcal{V}^*})_{j,k=1}^\ell \big) \tag{17.33}$$

where $\langle \bullet, \bullet \rangle_{\mathcal{V}^*}$ is the inner product on the dual space $\mathcal{V}^*$ induced by $\langle \bullet, \bullet \rangle_{\mathcal{V}}$.

PROOF. Fixing an orthonormal basis of $\mathcal{V}$, we identify it with the space $\mathbb{R}^n$ with the standard Euclidean inner product. Both sides of (17.33) are linear in each of the vectors $\mathbf{u}_j$, $\mathbf{v}_k$, thus for (17.33) to hold in general it is enough for it to hold when each of the vectors $\mathbf{u}_1, \dots, \mathbf{u}_\ell, \mathbf{v}_1, \dots, \mathbf{v}_\ell$ is equal to one of the canonical basis vectors $dx_1, \dots, dx_n \in (\mathbb{R}^n)^*$. In this case each $\mathbf{u}_1 \wedge \cdots \wedge \mathbf{u}_\ell$ is either zero or equal to $\pm dx_I$ where dx_I is an element of the basis (17.2), and same is true for $\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_\ell$. It follows that (17.33) holds if and only if

$$\{dx_I \mid I \subset \{1,\dots,n\},\ \#(I)=\ell\} \quad \text{is orthonormal with respect to } \langle \bullet, \bullet \rangle_{\wedge^\ell \mathcal{V}^*} \tag{17.34}$$

which gives the desired existence and uniqueness statement. □

One can also define the inner product on the total exterior algebra $\wedge^\bullet \mathcal{V}^*$, by making the spaces $\wedge^\ell \mathcal{V}^*$ orthogonal to each other for different ℓ. If $\mathcal{V} = \mathbb{R}^n$, this is the inner product in which the basis $\{dx_I \mid I \subset \{1,\dots,n\}\}$ is orthonormal.

Assume next that in addition to an inner product, we fix an orientation o on $\mathcal{V}$ (see Definition 17.2). Define the *volume form* $d\operatorname{vol}_{\mathcal{V}} \in \wedge^n \mathcal{V}^*$, where $n = \dim \mathcal{V}$, as follows: for any basis $v_1,\dots,v_n \in \mathcal{V}$

$$d\operatorname{vol}_{\mathcal{V}}(v_1,\dots,v_n) = o(v_1,\dots,v_n)\sqrt{\det\big((\langle v_j, v_k\rangle_{\mathcal{V}})_{j,k=1}^n\big)}. \tag{17.35}$$

To see that this is indeed an element of $\wedge^n \mathcal{V}^*$, note that if $\mathcal{V}$ is equal to $\mathbb{R}^n$ with the Euclidean inner product, then $d\operatorname{vol}_{\mathcal{V}} = dx_1 \wedge \cdots \wedge dx_n$, that is $d\operatorname{vol}_{\mathcal{V}}(v_1,\dots,v_n) = \det[v_1 \dots v_n]$ where $[v_1 \dots v_n]$ is the matrix with columns $v_1,\dots,v_n \in \mathbb{R}^n$ (see (17.3)).

We now define the Hodge star operator:

Lemma 17.7. *Assume that $\mathcal{V}$ is an n-dimensional vector space with given inner product and orientation. Let $\mathbf{v} \in \wedge^\ell \mathcal{V}^*$. Then there exists unique $\star_\ell \mathbf{v} \in \wedge^{n-\ell}\mathcal{V}^*$ such that*

$$\mathbf{u} \wedge (\star_\ell \mathbf{v}) = \langle \mathbf{u}, \mathbf{v}\rangle_{\wedge^\ell \mathcal{V}^*}\, d\operatorname{vol}_{\mathcal{V}} \quad \text{for all } \mathbf{u} \in \wedge^\ell \mathcal{V}^*. \tag{17.36}$$

We call $\star_\ell \mathbf{v}$ the Hodge star *of $\mathbf{v}$.*

Proof. For each $\mathbf{u} \in \wedge^\ell \mathcal{V}^*$ and $\mathbf{w} \in \wedge^{n-\ell}\mathcal{V}^*$, the wedge product $\mathbf{u} \wedge \mathbf{w}$ lies in $\wedge^n \mathcal{V}^*$ which is a one-dimensional vector space spanned by $d\operatorname{vol}_{\mathcal{V}}$. The map

$$\mathbf{u} \in \wedge^\ell \mathcal{V}^* \ \mapsto\ \frac{\mathbf{u} \wedge \mathbf{w}}{d\operatorname{vol}_{\mathcal{V}}} \in \mathbb{R}$$

is a linear functional, therefore it is equal to $\mathbf{u} \mapsto \langle \mathbf{u}, T\mathbf{w}\rangle_{\wedge^\ell \mathcal{V}^*}$ for some $T\mathbf{w} \in \wedge^\ell \mathcal{V}^*$. In other words,

$$\mathbf{u} \wedge \mathbf{w} = \langle \mathbf{u}, T\mathbf{w}\rangle_{\wedge^\ell \mathcal{V}^*}\, d\operatorname{vol}_{\mathcal{V}} \quad \text{for all } \mathbf{u} \in \wedge^\ell \mathcal{V}^*.$$

This defines a linear map $T : \wedge^{n-\ell}\mathcal{V}^* \to \wedge^\ell \mathcal{V}^*$. This map is invertible: indeed, the spaces $\wedge^{n-\ell}\mathcal{V}^*$ and $\wedge^\ell \mathcal{V}^*$ have the same dimension $\binom{n}{\ell}$, and T is injective since a computation using the basis (17.2) shows that if $\mathbf{u} \wedge \mathbf{w} = 0$ for all $\mathbf{u} \in \wedge^\ell \mathcal{V}^*$, then $\mathbf{u} = 0$.

Now the Hodge star of $\mathbf{v}$ is equal to $\star_\ell \mathbf{v} = T^{-1}\mathbf{v}$. □

Lemma 17.7 defines the linear operators

$$\star_\ell : \wedge^\ell \mathcal{V}^* \to \wedge^{n-\ell}\mathcal{V}^*. \tag{17.37}$$

Putting these operators together, we get the Hodge star operator on the total exterior algebra

$$\star : \wedge^\bullet \mathcal{V}^* \to \wedge^\bullet \mathcal{V}^*. \tag{17.38}$$

The next proposition lists the standard properties of Hodge star. We leave the proof as an exercise below.

PROPOSITION 17.8. *1. The operator $\star_\ell$ is equal to its own inverse up to sign, more precisely we have*

$$\star_\ell \star_{n-\ell} = (-1)^{\ell(n-\ell)}. \tag{17.39}$$

2. The operator $\star_\ell$ is an isometry with respect to the inner products $\langle \bullet, \bullet \rangle_{\wedge^\ell \mathcal{V}^*}$ and $\langle \bullet, \bullet \rangle_{\wedge^{n-\ell} \mathcal{V}^*}$.

We now give explicit formulas for the Hodge star operator on $\mathbb{R}^2$ and $\mathbb{R}^3$ with the Euclidean inner product and the standard orientation. We leave the verification as an exercise below.

- On $\mathbb{R}^2$, we have

$$\begin{aligned} \star(1) &= dx_1 \wedge dx_2, \quad & \star(dx_1 \wedge dx_2) &= 1, \\ \star(dx_1) &= dx_2, \quad & \star(dx_2) &= -dx_1. \end{aligned} \tag{17.40}$$

  In particular, the operator $\star_1 : (\mathbb{R}^2)^* \to (\mathbb{R}^2)^*$ is the counterclockwise rotation by angle $\frac{\pi}{2}$.
- On $\mathbb{R}^3$, we have

$$\begin{aligned} \star(1) &= dx_1 \wedge dx_2 \wedge dx_3, \quad & \star(dx_1 \wedge dx_2 \wedge dx_3) &= 1, \\ \star(dx_1) &= dx_2 \wedge dx_3, \quad & \star(dx_2 \wedge dx_3) &= dx_1, \\ \star(dx_2) &= dx_3 \wedge dx_1, \quad & \star(dx_3 \wedge dx_1) &= dx_2, \\ \star(dx_3) &= dx_1 \wedge dx_2, \quad & \star(dx_1 \wedge dx_2) &= dx_3. \end{aligned} \tag{17.41}$$

  In particular, if $\mathbf{u}, \mathbf{v} \in (\mathbb{R}^3)^*$, then $\star_2(\mathbf{u} \wedge \mathbf{v})$ is equal to the cross product of $\mathbf{u}$ and $\mathbf{v}$.

17.3.2. The d^* operator. Let $(\mathcal{M}, g)$ be an n-dimensional oriented Riemannian manifold. For each $x \in \mathcal{M}$, the tangent space $T_x\mathcal{M}$ has the inner product given by $g(x)$ and the orientation coming from the orientation on $\mathcal{M}$. Thus Lemma 17.7 defines the Hodge star operator

$$\star_\ell(x) : \Omega^\ell(x) \to \Omega^{n-\ell}(x), \quad x \in \mathcal{M}.$$

This operator depends smoothly on x and thus defines a bundle homomorphism $\star_\ell \in C^\infty(\mathcal{M}; \mathrm{Hom}(\Omega^\ell \to \Omega^{n-\ell}))$, see §13.1.8. This bundle homomorphism acts on differential forms as a 0-th order differential operator:

$$\star_\ell : C^\infty(\mathcal{M}; \Omega^\ell) \to C^\infty(\mathcal{M}; \Omega^{n-\ell}), \quad \star : C^\infty(\mathcal{M}; \Omega^\bullet) \to C^\infty(\mathcal{M}; \Omega^\bullet). \tag{17.42}$$

We now study the differential operator

$$d_\ell^* \in \mathrm{Diff}^1(\mathcal{M}; \Omega^{\ell+1} \to \Omega^\ell)$$

which is the adjoint of the exterior derivative d_ℓ defined in (17.14). (Note that the is the same as the transpose of d_ℓ, since all the operators studied act on real-valued forms.)

To make sense of d_ℓ^*, we first have to fix an inner product on $L^2_{\mathrm{c}}(\mathcal{M}; \Omega^\ell)$ for all ℓ. Using (16.7), we see that one needs to fix a positive density on $\mathcal{M}$ and an inner product on each fiber $\Omega^\ell(x)$. We take the density $d\,\mathrm{vol}_g$ induced by the metric g (see (13.29)) and the inner product $\langle \bullet, \bullet \rangle_{\Omega^\ell(x)}$ induced by $g(x)$, defined in Lemma 17.6. This gives the following inner product on $L^2_{\mathrm{c}}(\mathcal{M}; \Omega^\ell)$:

$$\langle \mathbf{u}, \mathbf{v} \rangle_{L^2(\mathcal{M};\Omega^\ell)} := \int_{\mathcal{M}} \langle \mathbf{u}(x), \mathbf{v}(x) \rangle_{\Omega^\ell(x)} \, d\,\mathrm{vol}_g(x) \quad \text{for all } \mathbf{u}, \mathbf{v} \in L^2_{\mathrm{c}}(\mathcal{M}; \Omega^\ell). \tag{17.43}$$

The integral in (17.43) can be interpreted as an integral of a differential form, if we let $d\,\mathrm{vol}_g \in C^\infty(\mathcal{M}; \Omega^n)$ be the volume form defined in (17.35), which is just the product of the Riemannian volume density given by (13.29) and the orientation on $\mathcal{M}$. Thus Lemma 17.7 gives the following formula for the inner product (17.43):

$$\langle \mathbf{u}, \mathbf{v} \rangle_{L^2(\mathcal{M};\Omega^\ell)} = \int_{\mathcal{M}} \mathbf{u} \wedge (\star_\ell \mathbf{v}) \quad \text{for all } \mathbf{u}, \mathbf{v} \in L^2_{\mathrm{c}}(\mathcal{M}; \Omega^\ell). \tag{17.44}$$

Here the right-hand side is an integral of a differential n-form.

We are now ready to compute the adjoint (which again, is the same as the transpose) of d_ℓ:

Lemma 17.9. *Define the differential operator*

$$d_\ell^* := (-1)^{n\ell+1} \star_{n-\ell} d_{n-\ell-1} \star_{\ell+1} \;\in\; \mathrm{Diff}^1(\mathcal{M}; \Omega^{\ell+1} \to \Omega^\ell). \tag{17.45}$$

Then d_ℓ^* is the adjoint of d_ℓ in the following sense:

$$\begin{gathered} \langle d_\ell \mathbf{u}, \mathbf{v} \rangle_{L^2(\mathcal{M};\Omega^{\ell+1})} = \langle \mathbf{u}, d_\ell^* \mathbf{v} \rangle_{L^2(\mathcal{M};\Omega^\ell)} \\ \textit{for all } \mathbf{u} \in C^\infty(\mathcal{M}; \Omega^\ell),\ \mathbf{v} \in C^\infty_{\mathrm{c}}(\mathcal{M}; \Omega^{\ell+1}). \end{gathered} \tag{17.46}$$

Proof. Let $\mathbf{u}, \mathbf{v}$ be as in (17.46). We compute

$$\begin{aligned} \langle d_\ell \mathbf{u}, \mathbf{v} \rangle_{L^2(\mathcal{M};\Omega^{\ell+1})} - \langle \mathbf{u}, d_\ell^* \mathbf{v} \rangle_{L^2(\mathcal{M};\Omega^\ell)} &= \int_{\mathcal{M}} (d_\ell \mathbf{u}) \wedge (\star_{\ell+1} \mathbf{v}) - \mathbf{u} \wedge (\star_\ell\, d_\ell^* \mathbf{v}) \\ &= \int_{\mathcal{M}} d_{n-1}\big(\mathbf{u} \wedge (\star_{\ell+1} \mathbf{v})\big) = 0, \end{aligned}$$

giving (17.46). Here in the first equality we used (17.44), in the second equality we used (17.19), (17.39), and (17.45), and the last equality follows from Stokes's Theorem 17.4. □

Putting the operators d_ℓ^* together, we get the total *codifferential operator*

$$d^* \in \operatorname{Diff}^1(\mathcal{M};\Omega^\bullet \to \Omega^\bullet). \tag{17.47}$$

We remark that by definition, $d_0^* = 0$. Moreover, since $d^2 = 0$ by (17.18), we have $(d^*)^2 = 0$ as well:

$$d_{\ell-1}^* d_\ell^* = 0. \tag{17.48}$$

As an example, we use (17.17) (and its analog on $\mathbb{R}^2$) together with (17.40) and (17.41) to compute the codifferential in $\mathbb{R}^2$ and $\mathbb{R}^3$ with the Euclidean metric:

- On $\mathbb{R}^2$, we have

$$\begin{aligned} d_0^*(u_1(x)\,dx_1 + u_2(x)\,dx_2) &= -(\partial_{x_1}u_1(x) + \partial_{x_2}u_2(x)),\\ d_1^*(f(x)\,dx_1\wedge dx_2) &= \partial_{x_2}f(x)\,dx_1 - \partial_{x_1}f(x)\,dx_2 \end{aligned} \tag{17.49}$$

  That is, d_0^* corresponds to divergence and d_1^* corresponds to gradient rotated by angle $\frac{\pi}{2}$.
- On $\mathbb{R}^3$, we have

$$\begin{aligned} d_0^*(u_1(x)\,dx_1 + u_2(x)\,dx_2 + u_3(x)\,dx_3) &= -(\partial_{x_1}u_1(x) + \partial_{x_2}u_2(x) + \partial_{x_3}u_3(x)),\\ d_1^*(u_{23}(x)\,dx_2\wedge dx_3) &= \partial_{x_3}u_{23}(x)\,dx_2 - \partial_{x_2}u_{23}(x)\,dx_3,\\ d_1^*(u_{31}(x)\,dx_3\wedge dx_1) &= \partial_{x_1}u_{31}(x)\,dx_3 - \partial_{x_3}u_{31}(x)\,dx_1,\\ d_1^*(u_{12}(x)\,dx_1\wedge dx_2) &= \partial_{x_2}u_{12}(x)\,dx_1 - \partial_{x_1}u_{12}(x)\,dx_2, \end{aligned} \tag{17.50}$$

  and

$$\begin{aligned} &d_2^*(f(x)\,dx_1\wedge dx_2\wedge dx_3)\\ &= -(\partial_{x_1}f(x)\,dx_2\wedge dx_3 + \partial_{x_2}f(x)\,dx_3\wedge dx_1 + \partial_{x_3}f(x)\,dx_1\wedge dx_2). \end{aligned} \tag{17.51}$$

  We see that d_0^* corresponds to divergence, d_1^* to curl, and d_2^* to gradient.

REMARK 17.10.$^{\mathbf{X}}$ *An attentive reader might have noticed that we actually do not need to fix an orientation on $\mathcal{M}$ to define the inner product (17.43) and the codifferential operator d^*. However, having an orientation lets us access the Hodge star operator, which makes the formulas nicer, and it will also be used in the proof of Poincaré duality (Theorem 17.18) below.*

REMARK 17.11.$^{\mathbf{X}}$ *For readers familiar with Riemannian geometry, one can check that on any oriented Riemannian manifold $(\mathcal{M}, g)$, if $X \in C^\infty(\mathcal{M}; T\mathcal{M})$ is a vector field and $X^\flat$ is the 1-form corresponding to X by the metric, then the function $d_0^* X^\flat$ is equal to minus the divergence of X with respect to the Levi–Civita connection.*

17.3.3. The Dirac operator and the Hodge Laplacian. As in §17.3.2 above, let $(\mathcal{M}, g)$ be an n-dimensional oriented Riemannian manifold. The operators $d, d^* \in \mathrm{Diff}^1(\mathcal{M}; \Omega^\bullet \to \Omega^\bullet)$ cannot be elliptic if $n \geq 2$: for example, d maps functions (sections of the 1-dimensional trivial bundle) to 1-forms (sections of the n-dimensional cotangent bundle). However, if we add these together then we obtain the *Dirac operator*

$$d + d^* \;\in\; \mathrm{Diff}^1(\mathcal{M}; \Omega^\bullet \to \Omega^\bullet) \tag{17.52}$$

which, as we show below, is elliptic. Note that $d + d^*$ does not respect the degree of a differential form: for example, if $\mathbf{u}$ is a 1-form then $(d + d^*)\mathbf{u}$ is the sum of a 0-form and a 2-form. Nevertheless, it switches the parity of the degree. More precisely, if we decompose the bundle $\Omega^\bullet$ defined in (17.9) as

$$\Omega^\bullet = \Omega^{\mathrm{even}} \oplus \Omega^{\mathrm{odd}} \quad \text{where } \Omega^{\mathrm{even}} := \bigoplus_{\ell \text{ even}} \Omega^\ell,\ \Omega^{\mathrm{odd}} := \bigoplus_{\ell \text{ odd}} \Omega^\ell \tag{17.53}$$

then the operator $d + d^*$ is the direct sum of its restrictions

$$\begin{aligned} (d + d^*)_{\mathrm{even}} &\in \mathrm{Diff}^1(\mathcal{M}; \Omega^{\mathrm{even}} \to \Omega^{\mathrm{odd}}), \\ (d + d^*)_{\mathrm{odd}} &\in \mathrm{Diff}^1(\mathcal{M}; \Omega^{\mathrm{odd}} \to \Omega^{\mathrm{even}}). \end{aligned} \tag{17.54}$$

Before proceeding with the general properties of the operator $d + d^*$, let us compute it for $\mathbb{R}^2$ and $\mathbb{R}^3$ with the Euclidean metric, using (17.17) and (17.49)–(17.51). Since $\Omega^\bullet$ in these cases is the trivial bundle, we fix a frame on $\Omega^\bullet$ and think of the operator $d + d^*$ as a matrix of first order differential operators.

- On $\mathbb{R}^2$ in the frame $1, dx_1 \wedge dx_2; dx_1, dx_2$, we have

$$(d + d^*)_{\mathrm{even}} = \begin{pmatrix} \partial_{x_1} & \partial_{x_2} \\ \partial_{x_2} & -\partial_{x_1} \end{pmatrix}, \quad (d + d^*)_{\mathrm{odd}} = \begin{pmatrix} -\partial_{x_1} & -\partial_{x_2} \\ -\partial_{x_2} & \partial_{x_1} \end{pmatrix}. \tag{17.55}$$

In other words,

$$d + d^* = \begin{pmatrix} 0 & 0 & -\partial_{x_1} & -\partial_{x_2} \\ 0 & 0 & -\partial_{x_2} & \partial_{x_1} \\ \partial_{x_1} & \partial_{x_2} & 0 & 0 \\ \partial_{x_2} & -\partial_{x_1} & 0 & 0 \end{pmatrix}. \tag{17.56}$$

- On $\mathbb{R}^3$ in the frame $1, dx_2\wedge dx_3, dx_3\wedge dx_1, dx_1\wedge dx_2; dx_1\wedge dx_2\wedge dx_3, dx_1, dx_2, dx_3$, we have

$$
\begin{aligned}
(d+d^*)_{\mathrm{even}} &= \begin{pmatrix} 0 & \partial_{x_1} & \partial_{x_2} & \partial_{x_3} \\ \partial_{x_1} & 0 & -\partial_{x_3} & \partial_{x_2} \\ \partial_{x_2} & \partial_{x_3} & 0 & -\partial_{x_1} \\ \partial_{x_3} & -\partial_{x_2} & \partial_{x_1} & 0 \end{pmatrix}, \\
(d+d^*)_{\mathrm{odd}} &= \begin{pmatrix} 0 & -\partial_{x_1} & -\partial_{x_2} & -\partial_{x_3} \\ -\partial_{x_1} & 0 & -\partial_{x_3} & \partial_{x_2} \\ -\partial_{x_2} & \partial_{x_3} & 0 & -\partial_{x_1} \\ -\partial_{x_3} & -\partial_{x_2} & \partial_{x_1} & 0 \end{pmatrix}.
\end{aligned}
\tag{17.57}
$$

The following lemma is used in the proof of Hodge's Theorem in §17.3.4 below.

LEMMA 17.12. *The Dirac operator $d + d^* \in \mathrm{Diff}^1(\mathcal{M}; \Omega^\bullet \to \Omega^\bullet)$ is elliptic in the sense of Theorem 14.23.*

PROOF. We need to show that for each $(x^0, \xi^0) \in T^*\mathcal{M}$ with $\xi^0 \neq 0$, the principal symbol $\sigma_1(d+d^*)(x^0, \xi^0)$ is an invertible linear map on (the complexification of) $\Omega^\bullet(x^0)$. Since $\sigma_1(d + d^*)(x, \xi)$ is homogeneous of degree 1 in ξ, without loss of generality we may assume that $\langle \xi^0, \xi^0\rangle_{g(x^0)} = 1$.

To simplify the computation, fix a chart $\varkappa : U \to V$ on $\mathcal{M}$ such that $x^0 \in U$ and

$$\varkappa(x^0) = 0, \quad d\varkappa(x^0)^{-T}\xi^0 = dx_1, \quad \varkappa_* g(0) = \sum_{j=1}^n dx_j^2.$$

Then we need to compute the principal symbol of $\varkappa_*(d + d^*)$ at $(0, dx_1) \in T^*V$ and show that it is an invertible linear map on (the complexification of) $\wedge^\bullet(\mathbb{R}^n)^*$.

We start with the principal symbol of $\varkappa_* d$, which is equal to the operator d on V by (17.20). Using the formula (17.15) we compute

$$\sigma_1(\varkappa_* d)(0, dx_1)\mathbf{v} = i\, dx_1 \wedge \mathbf{v} \quad \text{for all } \mathbf{v} \in \wedge^\bullet(\mathbb{R}^n)^*.$$

In the standard basis of $\wedge^\bullet(\mathbb{R}^n)^*$ given by (17.2) we have for all $I \subset \{1, \dots, n\}$

$$\sigma_1(\varkappa_* d)(0, dx_1)dx_I = \begin{cases} 0, & 1 \in I, \\ i\, dx_{I\cup\{1\}}, & 1 \notin I. \end{cases} \tag{17.58}$$

To compute the symbol $\sigma_1(\varkappa_* d^*)(0, dx_1)dx_I$, one could use the fact that $\varkappa_* d^*$ is the d^* operator for the metric $\varkappa_* g$ and the formula (17.45). Here the Hodge star is a 0-th order differential operator, and its symbol at $(0, dx_1)$ depends only on $\varkappa_* g(0)$ which is the Euclidean metric.

However, we instead use that $\varkappa_* d^*$ is the adjoint of $\varkappa_* d$ with respect to the inner product on $L^2(V; \Omega^\bullet)$ induced by the metric $\varkappa_* g$. Then the vector bundle version of the Adjoint Rule (13.52) implies that $\sigma_1(\varkappa_* d^*)(0, dx_1)$ is the adjoint of $\sigma_1(\varkappa_* d)(0, dx_1)$ with respect to the Hermitian inner product on (the complexification of) $\wedge^\bullet(\mathbb{R}^n)^*$ induced by the Euclidean inner product $\varkappa_* g(0)$. Recalling (17.34) we see that this inner product has an orthonormal basis $\{dx_I\}_{I\subset\{1,\dots,n\}}$. Using (17.58) we then compute

$$\sigma_1(\varkappa_* d^*)(0, dx_1)dx_I = \begin{cases} -i\, dx_{I\setminus\{1\}}, & 1 \in I, \\ 0, & 1 \notin I. \end{cases} \tag{17.59}$$

Adding (17.58) and (17.59) we see that

$$\sigma_1(\varkappa_*(d + d^*))(0, dx_1)dx_I = \begin{cases} -i\, dx_{I\setminus\{1\}}, & 1 \in I, \\ i\, dx_{I\cup\{1\}}, & 1 \notin I, \end{cases} \tag{17.60}$$

which implies that $\sigma_1(\varkappa_*(d + d^*))(0, dx_1)$ is an invertible linear operator on the complexification of $\wedge^\bullet(\mathbb{R}^n)^*$, finishing the proof.

(The above computation is somewhat abstract and relies on the version of Adjoint Rule (13.52) for vector bundles, which has not been properly developed even though it follows by a direct computation from the usual Adjoint Rule. The reader is strongly encouraged to look at the coefficients of ∂_{x_1} in (17.56) and (17.57) to see that the formula (17.60) does hold in these special cases.) □

We next define the *Hodge Laplacian* as the square of the Dirac operator:

$$\mathbf{\Delta}_g := (d + d^*)^2 \ \in \ \mathrm{Diff}^2(\mathcal{M}; \Omega^\bullet \to \Omega^\bullet). \tag{17.61}$$

Since $d^2 = (d^*)^2 = 0$ by (17.18) and (17.48), we have the following alternative formula for the Hodge Laplacian:

$$\mathbf{\Delta}_g = dd^* + d^*d. \tag{17.62}$$

Since d maps ℓ-forms to $\ell+1$-forms and d^* maps $\ell+1$-forms to ℓ-forms, we see that $\mathbf{\Delta}_g$ maps ℓ-forms to ℓ-forms, for any ℓ. That is, the Hodge Laplacian is the direct sum of its restrictions

$$\mathbf{\Delta}_{g,\ell} \in \mathrm{Diff}^2(\mathcal{M}; \Omega^\ell \to \Omega^\ell). \tag{17.63}$$

Note that, as d^* is the adjoint of d (see Lemma 17.9), the Dirac operator and the Hodge Laplacian are both self-adjoint (or equivalently, equal to their own transposes since we are working on real bundles):

$$(d + d^*)^* = d + d^*, \qquad \mathbf{\Delta}_g^* = \mathbf{\Delta}_g. \tag{17.64}$$

A direct calculation using (17.55) shows that on $\mathbb{R}^2$ with the Euclidean metric, the Hodge Laplacian is just a diagonal matrix with entries $-\Delta_0$ where $\Delta_0 = \partial_{x_1}^2 + \partial_{x_2}^2$ is the usual scalar Laplacian:

$$\mathbf{\Delta}_g = \begin{pmatrix} -\Delta_0 & 0 & 0 & 0 \\ 0 & -\Delta_0 & 0 & 0 \\ 0 & 0 & -\Delta_0 & 0 \\ 0 & 0 & 0 & -\Delta_0 \end{pmatrix}. \tag{17.65}$$

A similar calculation using (17.57) shows that the same is true for $\mathbb{R}^3$. We note that for $\mathbb{R}^3$, the formula (17.62) on 1-forms gives the curl of curl identity in multivariable calculus. The same form of the Hodge Laplacian is valid on any $\mathbb{R}^n$ with the Euclidean metric, which follows by a direct computation which we do not give here.

For general $(\mathcal{M}, g)$, the operator $\mathbf{\Delta}_{g,0} = d_0^* d_0 \in \mathrm{Diff}^2(\mathcal{M})$ is equal to $-\Delta_g$ where Δ_g is the Laplace–Beltrami operator, as follows from (13.63) and (17.46). Moreover, as one can show using the formula (17.60) for the principal symbol of the Dirac operator in the proof of Lemma 17.12 above and the vector bundle version of the Product Rule (13.49), the principal symbol of the Hodge Laplacian is given by

$$\sigma_2(\mathbf{\Delta}_g)(x,\xi) = \langle \xi, \xi \rangle_{g(x)} I \quad \text{for all } (x,\xi) \in T^*\mathcal{M} \tag{17.66}$$

where I is the identity homomorphism on $\Omega^\bullet(x)$.

REMARK 17.13.$^{\mathbf{X}}$ *This remark is not directly relevant to the analytical results of this chapter, since the metric used is not positive definite, so the 'Laplacian' obtained here is really a wave operator, and in particular not elliptic. However, it is a connection of Hodge theory to physics which I could not resist including in these notes.*

Consider the space $\mathbb{R}^4 = \mathbb{R}_t \times \mathbb{R}^3_x$ with the Minkowski 'metric'

$$g = -dt^2 + dx_1^2 + dx_2^2 + dx_3^2.$$

Note that g is a nondegenerate quadratic form on $\mathbb{R}^4$ but it is not positive definite; it gives what is known a Lorentzian metric *instead of a Riemannian metric. The formula (17.33) produces the quadratic form on $\Omega^\bullet = \wedge^\bullet(\mathbb{R}^4)^*$ in which the basis elements dx_I from (17.2) are orthogonal to each other and we have in particular*

$$\langle \mathbf{u}, \mathbf{u} \rangle_{\wedge^2(\mathbb{R}^4)^*} = \begin{cases} 1, & \mathbf{u} \in \{dx_2 \wedge dx_3,\ dx_3 \wedge dx_1,\ dx_1 \wedge dx_2\}, \\ -1, & \mathbf{u} \in \{dt \wedge dx_1,\ dt \wedge dx_2,\ dt \wedge dx_3\}. \end{cases}$$

The volume form is $d\,\mathrm{vol} = dt \wedge dx_1 \wedge dx_2 \wedge dx_3$. The Hodge star operator is defined in the same way as in Lemma 17.7 and we compute

$$\begin{aligned} \star(dx_2 \wedge dx_3) &= dt \wedge dx_1, \quad \star(dx_3 \wedge dx_1) = dt \wedge dx_2, \quad \star(dx_1 \wedge dx_2) = dt \wedge dx_3, \\ \star(dt \wedge dx_1) &= -dx_2 \wedge dx_3, \quad \star(dt \wedge dx_2) = -dx_3 \wedge dx_1, \quad \star(dt \wedge dx_3) = -dx_1 \wedge dx_2. \end{aligned}$$

Now, let $\mathbf{u} \in C^\infty(\mathbb{R}^4; \Omega^2)$ be a 2-form. We write the analogue of the harmonic form equations $d\mathbf{u} = d^*\mathbf{u} = 0$ from (17.73) below; by Lemma 17.9 these are equivalent to

$$d\mathbf{u} = 0, \quad d(\star \mathbf{u}) = 0. \tag{17.67}$$

Let us write $\mathbf{u}$ in the form

$$\mathbf{u} = dt \wedge (E_1\, dx_1 + E_2\, dx_2 + E_3\, dx_3) + (B_1\, dx_2 \wedge dx_3 + B_2\, dx_3 \wedge dx_1 + B_3\, dx_1 \wedge dx_2)$$

where $\mathbf{E} = (E_1, E_2, E_3)$ *and* $\mathbf{B} = (B_1, B_2, B_3)$ *are smooth maps from* $\mathbb{R}^4$ *to* $\mathbb{R}^3$*. Then the equations* (17.67) *can be rewritten using the vector calculus operators* curl *and* div *(acting in the* x *variables only) as follows:*

$$\begin{aligned} \partial_t \mathbf{B} &= -\operatorname{curl} \mathbf{E}, \\ \operatorname{div} \mathbf{B} &= 0, \\ \partial_t \mathbf{E} &= \operatorname{curl} \mathbf{B}, \\ \operatorname{div} \mathbf{E} &= 0. \end{aligned} \tag{17.68}$$

These are (in the right choice of units) the vacuum Maxwell's equations*, describing electrodynamics: here* $\mathbf{E}$ *is the electric field and* $\mathbf{B}$ *is the magnetic field. Note that* (17.67) *implies the 'Hodge d'Alembertian equation'* $(dd^* + d^*d)\mathbf{u} = 0$*, which is the same as each component of* $\mathbf{E}, \mathbf{B}$ *solving the wave equation.*

17.3.4. Harmonic forms and Hodge's Theorem. As before, let $(\mathcal{M}, g)$ be an n-dimensional oriented Riemannian manifold. We now additionally assume that $\mathcal{M}$ is *compact*. The Dirac operator $d + d^* \in \operatorname{Diff}^1(\mathcal{M}; \Omega^\bullet \to \Omega^\bullet)$ is elliptic by Lemma 17.12, so by Theorem 14.23 we see that any solution $\mathbf{u} \in \mathcal{D}'(\mathcal{M}; \Omega^\bullet)$ to the equation $(d + d^*)\mathbf{u} = 0$ is smooth. These solutions are called *harmonic forms* and they are characterized by

Lemma 17.14. *Let* $(\mathcal{M}, g)$ *be a compact oriented Riemannian manifold and* $\mathbf{u} \in C^\infty(\mathcal{M}; \Omega^\bullet)$ *be a differential form on* $\mathcal{M}$*. Then the following are equivalent:*

$$d\mathbf{u} = d^*\mathbf{u} = 0; \tag{17.69}$$

$$(d + d^*)\mathbf{u} = 0; \tag{17.70}$$

$$\boldsymbol{\Delta}_g \mathbf{u} = 0. \tag{17.71}$$

Proof. Since $\boldsymbol{\Delta}_g = (d + d^*)^2$, it is easy to see that (17.69) ⇒ (17.70) ⇒ (17.71). Thus it remains to assume that $\Delta_g \mathbf{u} = 0$ and show that $d\mathbf{u} = d^*\mathbf{u} = 0$. By (17.62) and since d^* is the adjoint of d by (17.46), we compute

$$\begin{aligned} 0 = \langle \Delta_g \mathbf{u}, \mathbf{u} \rangle_{L^2(\mathcal{M};\Omega^\bullet)} &= \langle dd^*\mathbf{u}, \mathbf{u} \rangle_{L^2(\mathcal{M};\Omega^\bullet)} + \langle d^*d\mathbf{u}, \mathbf{u} \rangle_{L^2(\mathcal{M};\Omega^\bullet)} \\ &= \langle d^*\mathbf{u}, d^*\mathbf{u} \rangle_{L^2(\mathcal{M};\Omega^\bullet)} + \langle d\mathbf{u}, d\mathbf{u} \rangle_{L^2(\mathcal{M};\Omega^\bullet)} \end{aligned} \tag{17.72}$$

implying that $d\mathbf{u} = d^*\mathbf{u} = 0$ as needed. □

Denote by $\mathscr{H}^\bullet(\mathcal{M})$ the space of harmonic forms of all degrees:

$$\mathscr{H}^\bullet(\mathcal{M}) := \{\mathbf{u} \in C^\infty(\mathcal{M}; \Omega^\bullet) \mid d\mathbf{u} = d^*\mathbf{u} = 0\}. \tag{17.73}$$

This space is finite dimensional by Theorem 15.13 applied to the operator $d + d^*$, which is elliptic by Lemma 17.12:

$$\dim \mathscr{H}^\bullet(\mathcal{M}) < \infty. \tag{17.74}$$

As follows from its definition, the space $\mathscr{H}^\bullet(\mathcal{M})$ is the direct sum of the spaces of harmonic forms of specific degrees:

$$\begin{gathered}\mathscr{H}^\bullet(\mathcal{M}) = \bigoplus_{\ell=0}^{n} \mathscr{H}^\ell(\mathcal{M})\\ \text{where } \mathscr{H}^\ell(\mathcal{M}) := \{\mathbf{u} \in C^\infty(\mathcal{M}; \Omega^\ell) \mid d_\ell \mathbf{u} = 0,\ d_{\ell-1}^* \mathbf{u} = 0\}.\end{gathered} \tag{17.75}$$

The spaces $\mathscr{H}^\ell(\mathcal{M})$ depend on the choice of the metric g; however, as we will see in Theorem 17.17 below, their dimension is independent of g.

We are now ready to prove the main results of this chapter.

THEOREM 17.15 (Hodge's Theorem). *Assume that $(\mathcal{M}, g)$ is a compact oriented Riemannian manifold. For any ℓ, we have the* Hodge decomposition

$$C^\infty(\mathcal{M}; \Omega^\ell) = \mathscr{H}^\ell(\mathcal{M}) \oplus d_{\ell-1}\big(C^\infty(\mathcal{M}; \Omega^{\ell-1})\big) \oplus d_\ell^*\big(C^\infty(\mathcal{M}; \Omega^{\ell+1})\big). \tag{17.76}$$

That is, any smooth differential form can be written in a unique way as the sum of a harmonic form, an exact form (an element of the range of d), and a coexact form (an element of the range of d^*).

REMARK 17.16.**X** *As in Remark 17.10, orientability is not actually necessary for Theorem 17.15 and Theorem 17.17 below to hold.*

PROOF. 1. First of all, arguing similarly to (17.72) (taking the L^2-inner product with $\mathbf{v}$) we see that for each $\mathbf{v} \in C^\infty(\mathcal{M}; \Omega^\bullet)$

$$d^* d\mathbf{v} = 0 \quad \Longrightarrow \quad d\mathbf{v} = 0, \tag{17.77}$$

$$dd^*\mathbf{v} = 0 \quad \Longrightarrow \quad d^*\mathbf{v} = 0. \tag{17.78}$$

We now show that the sum in (17.76) is direct, that is if

$$0 = \mathbf{h} + d\mathbf{v} + d^*\mathbf{w}, \quad \mathbf{h} \in \mathscr{H}^\bullet, \quad \mathbf{v}, \mathbf{w} \in C^\infty(\mathcal{M}; \Omega^\bullet), \tag{17.79}$$

then $\mathbf{h} = d\mathbf{v} = d^*\mathbf{w} = 0$. To see this, we apply d^* and d to (17.79) and use that $d^2 = (d^*)^2 = 0$ by (17.18) and (17.48) to get

$$d^* d\mathbf{v} = 0, \quad dd^*\mathbf{w} = 0,$$

which by (17.77)–(17.78) gives $d\mathbf{v} = d^*\mathbf{w} = 0$ and thus $\mathbf{h} = 0$ as well.

2. We next show that for each $\mathbf{u} \in C^\infty(\mathcal{M}; \Omega^\bullet)$ we can write

$$\mathbf{u} = \mathbf{h} + d\mathbf{v} + d^*\mathbf{v} \quad \text{for some } \mathbf{h} \in \mathscr{H}^\bullet(\mathcal{M}), \quad \mathbf{v} \in C^\infty(\mathcal{M}; \Omega^\bullet). \tag{17.80}$$

This is the step where we use the theory of elliptic operators developed in the earlier chapters. The Dirac operator $d + d^* \in \mathrm{Diff}^1(\mathcal{M}; \Omega^\bullet \to \Omega^\bullet)$ is elliptic by Lemma 17.12 and the manifold $\mathcal{M}$ is compact, so by the formula (15.32) in the statement of the Fredholm property (Theorem 15.13) we see that for any $s \in \mathbb{R}$ and $\widetilde{\mathbf{u}} \in H^s(\mathcal{M}; \Omega^\bullet)$

we have (noting that there is no difference between transpose and adjoint for operators on real-valued forms and the operator $d + d^*$ is self-adjoint by (17.64); see also Remark 15.2)

$$\begin{gathered}\widetilde{\mathbf{u}} = (d+d^*)\mathbf{v} \quad \text{for some } \mathbf{v} \in H^{s+1}(\mathcal{M};\Omega^\bullet) \\ \Updownarrow \\ \langle \widetilde{\mathbf{u}}, \widetilde{\mathbf{h}}\rangle_{L^2(\mathcal{M};\Omega^\bullet)} = 0 \quad \text{for all } \widetilde{\mathbf{h}} \in \ker(d+d^*).\end{gathered} \tag{17.81}$$

By Lemma 17.14, the space $\ker(d+d^*)$ is the same as the space $\mathscr{H}^\bullet$ of harmonic forms. Take arbitrary $\mathbf{u} \in C^\infty(\mathcal{M};\Omega^\bullet)$ and decompose it as

$$\mathbf{u} = \mathbf{h} + \widetilde{\mathbf{u}} \quad \text{where } \mathbf{h} \in \mathscr{H}^\bullet(\mathcal{M})$$

and $\widetilde{\mathbf{u}} \in C^\infty(\mathcal{M};\Omega^\bullet)$ is orthogonal to $\mathscr{H}^\bullet(\mathcal{M})$ with respect to the inner product on $L^2(\mathcal{M};\Omega^\bullet)$. Fix $s \in \mathbb{R}$. Then (17.81) implies that $\widetilde{\mathbf{u}} = (d+d^*)\mathbf{v}$ for some $\mathbf{v} \in H^{s+1}(\mathcal{M};\Omega^\bullet)$, and we have $\mathbf{v} \in C^\infty(\mathcal{M};\Omega^\bullet)$ by Theorem 14.23. This gives existence of the decomposition (17.80).

3.$^{\mathbf{S}}$ We now show that $C^\infty(\mathcal{M};\Omega^\ell)$ is contained in the right-hand side of (17.76). Let $\mathbf{u} \in C^\infty(\mathcal{M};\Omega^\ell)$ and write it in the form (17.80) for some $\mathbf{h} \in \mathscr{H}^\bullet(\mathcal{M})$ and $\mathbf{v} \in C^\infty(\mathcal{M};\Omega^\bullet)$. We decompose

$$\mathbf{h} = \sum_{r=0}^{n} \mathbf{h}_r, \quad \mathbf{v} = \sum_{r=0}^{n} \mathbf{v}_r$$

where $\mathbf{h}_r \in \mathscr{H}^r(\mathcal{M})$ and $\mathbf{v}_r \in C^\infty(\mathcal{M};\Omega^r)$. Taking the part of (17.80) corresponding to Ω^ℓ we see that

$$\mathbf{u} = \mathbf{h}_\ell + d\mathbf{v}_{\ell-1} + d^*\mathbf{v}_{\ell+1} \tag{17.82}$$

which shows that $\mathbf{u}$ lies in the right-hand side of (17.76). □

As an application of Theorem 17.15 we show that harmonic ℓ-forms are in one to one correspondence with de Rham cohomology classes reviewed in §17.2, in fact each cohomology class contains a unique harmonic form:

THEOREM 17.17. *Assume that $(\mathcal{M}, g)$ is a compact oriented Riemannian manifold. For each ℓ we have the decomposition featuring the spaces of closed and exact forms introduced in (17.28) and (17.27)*

$$\ker_{C^\infty} d_\ell = \mathscr{H}^\ell(\mathcal{M}) \oplus d_{\ell-1}(C^\infty(\mathcal{M};\Omega^{\ell-1})). \tag{17.83}$$

Moreover, the space $\mathscr{H}^\ell(\mathcal{M})$ of harmonic forms is isomorphic to the de Rham cohomology space $\mathbf{H}^\ell(\mathcal{M};\mathbb{R})$ defined in (17.29).

PROOF. The sum on the right-hand side of (17.83) is direct by (17.76) and it is contained in the left-hand side of (17.83). Thus it remains to show that each $\mathbf{u} \in$

$C^\infty(\mathcal{M};\Omega^\ell)$ such that $d\mathbf{u} = 0$ can be written as the sum of a harmonic form and a closed form. Let us write the decomposition (17.76) for $\mathbf{u}$:

$$\mathbf{u} = \mathbf{h} + d\mathbf{v} + d^*\mathbf{w} \quad \text{where } \mathbf{h} \in \mathscr{H}^\ell(\mathcal{M}),\ \mathbf{v} \in C^\infty(\mathcal{M};\Omega^{\ell-1}),\ \mathbf{w} \in C^\infty(\mathcal{M};\Omega^{\ell+1}).$$

Applying d to both sides and using that $\mathbf{d}u = d\mathbf{h} = 0$ and $d^2 = 0$ by (17.18), we see that $dd^*\mathbf{w} = 0$. By (17.78) this shows that $d^*\mathbf{w} = 0$. Thus $\mathbf{u} = \mathbf{h} + d\mathbf{v}$, which finishes the proof of (17.83).

Finally, consider the map from harmonic forms to their cohomology classes:

$$\mathbf{h} \in \mathscr{H}^\ell(\mathcal{M}) \ \mapsto\ [\mathbf{h}] \in \mathbf{H}^\ell(\mathcal{M};\mathbb{R}). \tag{17.84}$$

By (17.83) this map is an isomorphism. □

17.3.5. Applications of Hodge theory. Theorem 17.17 and (17.74) imply that when $\mathcal{M}$ is a compact manifold, its de Rham cohomology groups are finite dimensional:

$$b_\ell(\mathcal{M}) = \dim \mathbf{H}^\ell(\mathcal{M};\mathbb{R}) = \dim \mathscr{H}^\ell(\mathcal{M}) < \infty. \tag{17.85}$$

Another consequence of this theorem is the de Rham version of *Poincaré duality* for the Betti numbers (17.30):

THEOREM 17.18. *Assume that $\mathcal{M}$ is a compact orientable manifold. Then we have for all $\ell = 0,\dots,n$*

$$b_\ell(\mathcal{M}) = b_{n-\ell}(\mathcal{M}). \tag{17.86}$$

REMARK 17.19.$^{\mathbf{X}}$ *Unlike Theorems 17.15 and 17.17, Theorem 17.18 uses orientability of $\mathcal{M}$ in an essential way, via the existence of the Hodge star operator. In fact, if $\mathcal{M}$ is connected and not orientable, then $H^n(\mathcal{M};\mathbb{R}) = \{0\}$ in contrast with (17.89) below, see for example* [**Lee13**, Theorem 17.34].

PROOF. Fix an orientation and a Riemannian metric on $\mathcal{M}$. By Theorem 17.17, we have $b_\ell(\mathcal{M}) = \dim \mathscr{H}^\ell(\mathcal{M})$. Recalling the formula (17.45) for the operator d^* and the identity (17.39), we see that for each $\mathbf{u} \in C^\infty(\mathcal{M};\Omega^\ell)$ we have

$$d\mathbf{u} = 0 \iff d^*(\star\mathbf{u}) = 0,$$
$$d^*\mathbf{u} = 0 \iff d(\star\mathbf{u}) = 0.$$

Thus the Hodge star operator $\star_\ell$ restricts to a linear isomorphism

$$\star_\ell : \mathscr{H}^\ell(\mathcal{M}) \to \mathscr{H}^{n-\ell}(\mathcal{M}) \tag{17.87}$$

which implies (17.86). □

As a corollary of (17.87), if $\mathcal{M}$ is a compact connected oriented Riemannian manifold then we can compute the highest and lowest degree harmonic forms on $\mathcal{M}$:

$$\mathscr{H}^0(\mathcal{M}) = \mathbb{R}1, \quad \mathscr{H}^n(\mathcal{M}) = \mathbb{R}\, d\,\mathrm{vol}_g \tag{17.88}$$

where $d\,\mathrm{vol}_g \in C^\infty(\mathcal{M};\Omega^n)$ is the volume form induced by the metric g and the choice of orientation, see the paragraph following (17.43). In particular, we have

$$b_0(\mathcal{M}) = b_n(\mathcal{M}) = 1. \tag{17.89}$$

We finish this subsection with a simple formula for the index of the even and odd parts of the Dirac operator $d + d^*$, see §15.3.3 and (17.54):

Proposition 17.20. *Let $\mathcal{M}$ be a compact oriented Riemannian manifold. Then*

$$\operatorname{ind}(d+d^*)_{\mathrm{even}} = -\operatorname{ind}(d+d^*)_{\mathrm{odd}} = \sum_{\ell=0}^{n}(-1)^\ell b_\ell(\mathcal{M}). \tag{17.90}$$

The expression on the right-hand side of (17.90) is called the Euler characteristic *of $\mathcal{M}$.*

Proof. Since $d + d^*$ is self-adjoint (see (17.64)) and the spaces $C^\infty(\mathcal{M};\Omega^{\mathrm{even}})$ and $C^\infty(\mathcal{M};\Omega^{\mathrm{odd}})$ are orthogonal to each other with respect to the $L^2(\mathcal{M};\Omega^\bullet)$-inner product, we have

$$\big((d+d^*)_{\mathrm{even}}\big)^* = (d+d^*)_{\mathrm{odd}}.$$

By (15.42) (where there is no difference between transpose and adjoint since we are working with real bundles), we have

$$\operatorname{ind}(d+d^*)_{\mathrm{even}} = -\operatorname{ind}(d+d^*)_{\mathrm{odd}} = \dim\ker(d+d^*)_{\mathrm{even}} - \dim\ker(d+d^*)_{\mathrm{odd}}.$$

Next, Lemma 17.14 shows that

$$\ker(d+d^*)_{\mathrm{even}} = \bigoplus_{\ell\ \mathrm{even}} \mathscr{H}^\ell(\mathcal{M}), \quad \ker(d+d^*)_{\mathrm{odd}} = \bigoplus_{\ell\ \mathrm{odd}} \mathscr{H}^\ell(\mathcal{M}).$$

Now (17.90) follows from Theorem 17.17. □

17.3.6. Degree theory. Assume that $\mathcal{M},\mathcal{N}$ are two compact connected oriented manifolds of the same dimension. Let

$$\Phi : \mathcal{M} \to \mathcal{N}$$

be a smooth map (not necessarily a diffeomorphism). In this section, we define a topological invariant corresponding to Φ, called the *degree* of Φ. We next use this invariant to give a proof of the Hairy Ball Theorem.

We first make a few preliminary definitions:

Definition 17.21. *1. Let $x \in \mathcal{M}$. We say that x is a* regular point *for Φ, if the linear map $d\Phi(x) : T_x\mathcal{M} \to T_{\Phi(x)}\mathcal{N}$ is invertible. In this case, we define*

$$\operatorname{sgn}\det d\Phi(x) \in \{1,-1\}$$

to be equal to 1 if $d\Phi(x)$ is orientation preserving (with respect to the orientations fixed on $\mathcal{M}$, $\mathcal{N}$) and -1 otherwise.

2. Let $y \in \mathcal{N}$. We say that y is a regular value *for Φ if each $x \in \Phi^{-1}(y)$ is a regular point for Φ. (This includes the case when $\Phi^{-1}(y) = \emptyset$.)*

Any smooth map Φ has a regular value. In fact, Sard's Theorem [**Lee13**, Theorem 6.10] shows that the set of non-regular values of Φ has Lebesgue measure 0 in $\mathcal{N}$. Note that if y is a regular value, then $\Phi^{-1}(y)$ has to consist of isolated points and thus is a finite set.

We can now give the definition of the degree. We in fact give two definitions and the statement below shows that they coincide.

THEOREM 17.22. *Let $\mathcal{M}, \mathcal{N}$ be compact connected oriented manifolds of the same dimension n and $\Phi : \mathcal{M} \to \mathcal{N}$ be a smooth map. Then there exists an integer*

$$\deg \Phi \in \mathbb{Z}, \tag{17.91}$$

called the degree of Φ, with the following properties:

(1) *for any n-form $\mathbf{v} \in C^\infty(\mathcal{N}; \Omega^n)$, we have*

$$\int_{\mathcal{M}} \Phi^* \mathbf{v} = (\deg \Phi) \int_{\mathcal{N}} \mathbf{v} \tag{17.92}$$

where $\Phi^* \mathbf{v} \in C^\infty(\mathcal{M}; \Omega^n)$ is the pullback of $\mathbf{v}$ by Φ, defined in (17.11);

(2) *for any $y \in \mathcal{N}$ which is a regular value of Φ, we have*

$$\deg \Phi = \sum_{x \in \Phi^{-1}(y)} \operatorname{sgn} \det d\Phi(x). \tag{17.93}$$

Before proceeding to the proof, let us give a few remarks:

- If Φ is an orientation preserving diffeomorphism, then $\deg \Phi = 1$ as follows from either (17.24) or (17.93). Similarly, if Φ is an orientation reversing diffeomorphism, then $\deg \Phi = -1$.
- If Φ is not onto then, taking $y \in \mathcal{N} \setminus \Phi(\mathcal{M})$ in (17.93), we see that $\deg \Phi = 0$.
- From (17.92) we see that the degree of the composition of two maps is the product of their degrees.
- As an example, if $\mathcal{M} = \mathcal{N} = \mathbb{S}^1 := \mathbb{R}/\mathbb{Z}$ and $\Phi(x) = mx \bmod \mathbb{Z}$ for some $m \in \mathbb{Z}$, then $\deg \Phi = m$, as can be seen from either (17.92) (taking $\mathbf{v} = dx$, in which case $\Phi^* \mathbf{v} = m\, dx$) or (17.93). If we think of $\mathbb{S}^1$ as the unit circle in $\mathbb{C}$, then Φ corresponds to the map $z \mapsto z^m$, explaining the use of the word 'degree'.
- **X** The degree of a map is a special case of the induced maps on de Rham cohomology spaces defined in (17.29). More precisely, if $\mathbf{u} \in C^\infty(\mathcal{N}; \Omega^\ell)$ is a closed form, then its pullback $\Phi^* \mathbf{u}$ is closed, and if $\mathbf{u}$ is an exact form then $\Phi^* \mathbf{u}$ is exact (both follow from (17.20)). Thus Φ^* descends to a linear map $\Phi_\ell^* : \mathbf{H}^\ell(\mathcal{N}; \mathbb{R}) \to \mathbf{H}^\ell(\mathcal{M}; \mathbb{R})$. By (17.89), we have $\mathbf{H}^n(\mathcal{M}; \mathbb{R}) \simeq \mathbf{H}^n(\mathcal{N}; \mathbb{R}) \simeq$

$\mathbb{R}$, with the identification being canonical using the integral of an n-form over the entire manifold. Then the map Φ_n^* is the multiplication by the degree of Φ. The fact that the degree is an integer is related to the existence of cohomology groups with integer coefficients, which we do not study here.

Our proof of Theorem 17.22 relies on the following characterization of exact n-forms:

LEMMA 17.23. *Assume that $\mathcal{M}$ is an n-dimensional compact connected oriented manifold and $\mathbf{v} \in C^\infty(\mathcal{M}; \Omega^n)$. Then*

$$\mathbf{v} = d\mathbf{w} \quad \text{for some } \mathbf{w} \in C^\infty(\mathcal{M}; \Omega^{n-1}) \iff \int_{\mathcal{M}} \mathbf{v} = 0.$$

PROOF. The $\Rightarrow$ direction follows immediately from Stokes's Theorem 17.4, so it remains to show the $\Leftarrow$ direction. Assume that $\int_{\mathcal{M}} \mathbf{v} = 0$. Fix a metric g on $\mathcal{M}$. By (17.82) we write

$$\mathbf{v} = \mathbf{h} + d\mathbf{w} \quad \text{for some } \mathbf{h} \in \mathscr{H}^n(\mathcal{M}),\ \mathbf{w} \in C^\infty(\mathcal{M}; \Omega^{n-1}).$$

Taking the integrals of both sides over $\mathcal{M}$ and using Theorem 17.4 to see that the integral of $d\mathbf{w}$ is equal to 0, we see that $\int_{\mathcal{M}} \mathbf{h} = 0$. But $\mathscr{H}^n(\mathcal{M})$ is spanned by $d\,\mathrm{vol}_g$ by (17.88), and $\int_{\mathcal{M}} d\,\mathrm{vol}_g = \mathrm{vol}_g(\mathcal{M}) > 0$. It follows that $\mathbf{h} = 0$ and thus $\mathbf{v} = d\mathbf{w}$ as needed. □

We are now ready to give

PROOF OF THEOREM 17.22. 1. Fix any

$$\mathbf{v}_0 \in C^\infty(\mathcal{N}; \Omega^n), \quad \int_{\mathcal{N}} \mathbf{v}_0 = 1.$$

For example, one can fix a Riemannian metric on $\mathcal{N}$ and let $\mathbf{v}_0$ be a multiple of the corresponding volume form. Now, put

$$\deg \Phi := \int_{\mathcal{M}} \Phi^* \mathbf{v}_0 \ \in\ \mathbb{R}.$$

We first show that (17.92) holds. Take arbitrary $\mathbf{v} \in C^\infty(\mathcal{N}; \Omega^n)$ and put $c := \int_{\mathcal{N}} \mathbf{v}$. Then $\int_{\mathcal{N}} \mathbf{v} - c\mathbf{v}_0 = 0$ and thus by Lemma 17.23 we have

$$\mathbf{v} - c\mathbf{v}_0 = d\mathbf{w} \quad \text{for some } \mathbf{w} \in C^\infty(\mathcal{N}; \Omega^{n-1}).$$

Then by (17.20) we get

$$\Phi^*\mathbf{v} - c\Phi^*\mathbf{v}_0 = d\Phi^*\mathbf{w}$$

and thus by Theorem 17.4 we have $\int_{\mathcal{M}} \Phi^*\mathbf{v} - c\Phi^*\mathbf{v}_0 = 0$ which gives (17.92).

2. It remains to show that (17.93) holds, which (since Φ has a regular value) also implies that $\deg \Phi \in \mathbb{Z}$. Let $y \in \mathcal{N}$ be a regular value for Φ. Denote by $\delta_y \in \mathcal{D}'(\mathcal{N}; \Omega^n)$ the delta-current at y. It is similar to the delta-density defined in (13.46) but incorporates

the orientation fixed on $\mathcal{M}$. More precisely, $\operatorname{supp}\delta_y = \{y\}$ and if $\varkappa : U \to V$ is a chart on $\mathcal{N}$ such that $y \in U$, then we have

$$\varkappa_* \delta_y = (\operatorname{sgn}\det d\varkappa(y))\delta_{\varkappa(y)}\, dx_1 \wedge \cdots \wedge dx_n. \tag{17.94}$$

The currents defined on the right-hand side of (17.94) satisfy compatibility conditions analogous to (13.43), as follows from (10.10), and thus give rise to a current δ_y on $\mathcal{N}$.

Now (17.93) follows by applying (17.92) with $\mathbf{v} := \delta_y$. Since δ_y is not a smooth form, some explanations are in order. Since y is a regular value for Φ, we have

$$\Phi^{-1}(y) = \{x_1, \dots, x_N\} \quad \text{for some } x_1, \dots, x_N \in \mathcal{M}$$

and $d\Phi(x_k)$ is invertible for each k. By the Inverse Mapping Theorem, we can fix a neighborhood $\mathcal{V} \Subset \mathcal{N}$ of y such that

$$\Phi^{-1}(\mathcal{V}) = \bigsqcup_{k=1}^{N} \mathcal{U}_k, \quad x_k \in \mathcal{U}_k,$$

and the restriction of Φ to each $\mathcal{U}_k$ is a diffeomorphism $\mathcal{U}_k \to \mathcal{V}$. Using the pullback operators on distributions by these diffeomorphisms, we can define the operators of pullback on differential forms $(\Phi|_{\mathcal{U}_k})^* : \mathcal{E}'(\mathcal{V}; \Omega^n) \to \mathcal{E}'(\mathcal{U}_k; \Omega^n)$. Adding these together and extending by zero to $\mathcal{M}$, we get the pullback operator

$$\Phi^* : \mathcal{E}'(\mathcal{V}; \Omega^n) \to \mathcal{D}'(\mathcal{M}; \Omega^n)$$

which is sequentially continuous and agrees with the usual pullback of differential forms on $C_c^\infty(\mathcal{V}; \Omega^n)$. Since $C_c^\infty(\mathcal{V}; \Omega^n)$ is dense in $\mathcal{E}'(\mathcal{V}; \Omega^n)$, from (17.92) we see that

$$\int_{\mathcal{M}} \Phi^* \mathbf{v} = (\deg \Phi) \int_{\mathcal{N}} \mathbf{v} \quad \text{for all } \mathbf{v} \in \mathcal{E}'(\mathcal{V}; \Omega^n).$$

Now apply this with $\mathbf{v} := \delta_y$. The right-hand side is equal to $\deg \Phi$. Using (17.94) we see that the left-hand side is equal to

$$\int_{\mathcal{M}} \sum_{k=1}^{N} (\operatorname{sgn}\det d\Phi(x_k))\delta_{x_k} = \sum_{k=1}^{N} \operatorname{sgn}\det d\Phi(x_k).$$

This gives (17.93). □

A classical application of Theorem 17.22 is the following

THEOREM 17.24 (Hairy Ball Theorem).[X] *Assume that n is even and $X \in C^\infty(\mathbb{S}^n; T\mathbb{S}^n)$ is a vector field on the sphere $\mathbb{S}^n$. Then there exists $x \in \mathbb{S}^n$ such that $X(x) = 0$.*

PROOF. Fix the standard metric on $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ coming from the Euclidean metric and take the standard orientation o on $\mathbb{S}^n$ corresponding to the outward normal: that is, for each $x \in \mathbb{S}^n$ and a basis $v_1, \dots, v_n \in T_x\mathbb{S}^n \subset \mathbb{R}^{n+1}$, we have

$o(v_1, \ldots, v_n) = \operatorname{sgn}\det[xv_1 \ldots v_n]$ where $[xv_1 \ldots v_n]$ is the $(n+1) \times (n+1)$ matrix with columns $x, v_1, \ldots, v_n$.

We argue by contradiction. Assume that the vector field X is nonvanishing. Dividing it by its length, we may assume that $|X| = 1$ everywhere. Then for each $x \in \mathbb{S}^n$, $X(x) \in \mathbb{R}^{n+1}$ is a unit vector orthogonal to x. Now define the family of smooth maps

$$\Phi_t : \mathbb{S}^n \to \mathbb{S}^n,\ t \in [0, \pi], \quad \Phi_t(x) = (\cos t)x + (\sin t)X(x).$$

That is, $\Phi_t(x)$ is the result of following for time t the geodesic on $\mathbb{S}^n$ with initial position x and initial velocity $X(x)$. Consider the degree

$$\deg \Phi_t \in \mathbb{Z}.$$

It is a continuous function of t as can be seen from (17.92) and takes integer values. Thus $\deg \Phi_t$ is constant. This gives a contradiction since we can compute

$$\deg \Phi_0 = 1, \quad \deg \Phi_\pi = -1.$$

Indeed, Φ_0 is the identity map (which is an orientation preserving diffeomorphism) and $\Phi_\pi(x) = -x$ is the antipodal map, which (as n is even and thus $n+1$ is odd) is an orientation reversing diffeomorphism. □

17.4. Notes and exercises

For an introduction to the de Rham cohomology theory and some of its applications (including degree theory and the Brouwer fixed point theorem) see Lee [**Lee13**, Chapters 17–18]. For a more comprehensive treatment of differential topology see Bott and Tu [**BT82**].

For Hodge theory and its applications, the reader is referred to the books [**Che56, BDIP02, Huy05, Voi07, War71**]. Many of these consider the case of complex or Kähler manifolds, in which the manifold has a complex structure which gives additional structure for the de Rham cohomology spaces and for the operators d, d^*.

EXERCISE 17.1. (1 pt) *Prove Proposition 17.8.*

EXERCISE 17.2. (0.5 pt) *Verify* (17.40) *and* (17.41)*.*

List of notation

$*$	convolution, see §§1.3.1,6.1,8.1,8.2
$\star$	Hodge star, see §17.3.1
$\otimes$	tensor product of distributions, see §7.1
$\sim$	asymptotic sum, see §14.1.2
$\wedge$	wedge product, see §17.1.1
$\wedge^\ell \mathcal{V}$	exterior power, see §17.1.1
$\wedge^\bullet \mathcal{V}$	total exterior algebra, see §17.1.1
a.e.	Lebesgue almost everywhere, see §1.2.1
$b_\ell(\mathcal{M})$	Betti number, see §17.2
$B^\circ(x,r)$	the open ball in some metric space centered at a point x and of radius r
$B(x,r)$	the closed ball in some metric space centered at a point x and of radius r
$C^k(U)$	the space of k-times differentiable functions on U, see §1.2.3
$C^k_\mathrm{c}(U)$	the space of compactly supported functions in $C^k(U)$, see §1.2.3
$C^\infty(U)$	the space of infinitely differentiable functions on U, see §1.2.4
$C^\infty(\mathcal{M};\mathscr{E})$	the space of smooth sections of a vector bundle, see §13.1.8
$C^\infty_\mathrm{c}(U)$	the space of compactly supported functions in $C^\infty(U)$, see §1.2.4
$\lVert \bullet \rVert_{C^N(U,K)}$	see (4.1)
$\cdot$	inner product on $\mathbb{R}^n$, see §11.1
$\mathcal{D}'(U)$	the space of distributions on U, see §2.1
$\mathcal{D}'(\mathcal{M})$	the space of distributions on a manifold $\mathcal{M}$, see §13.2
d	exterior derivative, see §17.1.2
dx_I	see §17.1.1
d^*	codifferential, see §17.3.2
∂_{x_j}	the partial derivative operator in j-th variable
∂_x^α	the higher order partial derivative w.r.t. multiindex α, see §1.2.3
D_{x_j}, D_x^α	the operators $-i\partial_{x_j}$, $(-i)^{\lvert\alpha\rvert}\partial_x^\alpha$, see §11.1.3
δ_y	Dirac delta function, see §2.1
Δ	the Laplace operator $\partial_{x_1}^2 + \cdots + \partial_{x_n}^2$ on $\mathbb{R}^n$
Δ_g	Laplace–Beltrami operator, see §13.3.2
$\boldsymbol{\Delta}_g$	Hodge Laplacian, see §17.3.3
$\mathrm{Den}(\mathcal{V})$	the space of densities of a vector space $\mathcal{V}$, see §13.1.7
Diff^m	the space of differential operators, see §§9.1.1,13.3

$d\,\mathrm{vol}_g$	Riemannian volume density or volume form, see §§13.1.7,17.3.2
$\mathcal{E}'(U)$	the space of distributions with compact support on U, see §4.2
$\widehat{f}, \mathcal{F}(f)$	the Fourier transform of f, see §§11.1–11.2
H	the Heaviside function, see (3.3)
H^s	Sobolev space, see §§12.1.2,13.2.3
$H^s_{\mathrm{loc}}, H^s_{\mathrm{c}}$	see §§12.1.5,13.2.3
$\mathbf{H}^\ell(\mathcal{M};\mathbb{R})$	de Rham cohomology space, see §17.2
$\mathscr{H}^\ell(\mathcal{M})$	the space of harmonic forms, see §17.3.4
$\mathrm{Hom}(\mathscr{E}\to\mathscr{F})$	the bundle of linear homomorphisms, see §13.1.8
$\mathrm{ind}\,P$	the index of an operator P, see §15.3.1
$\mathbf{1}_A$	the indicator function of a subset, see (1.17)
$\mathcal{J}_\Phi$	the Jacobian of a diffeomorphism Φ, see §10.1.3
$\varkappa_*$	pushforward by a chart, see §§13.1.3–13.1.7,13.2.2,13.3.2
$\ker P$	the kernel of an operator P, see §15.3.1
Λ_t	dilation operator, see §5.1.2
L^p_{c}	compactly supported L^p functions, see §1.2.1
L^p_{loc}	locally L^p functions, see §1.2.1
$(\bullet,\bullet)$	distributional pairing, see §2.1
$\langle\bullet,\bullet\rangle_{L^2}$	L^2 Hermitian inner product, see (1.19)
$\langle\bullet\rangle$	see §12.1.2
$\lvert\Omega\rvert$	the bundle of densities on a manifold, see §13.1.7
Ω^ℓ	the bundle of ℓ-forms on a manifold, see §17.1.2
$\Omega^\bullet$	the total bundle of differential forms on a manifold, see §17.1.2
$\Subset$ (open)	open subset, see Definition 1.1
Op	quantization procedure, see §14.1.3
P^t	the transpose of an operator, see §§7.3,13.3.3
P^*	the adjoint of an operator, see §§7.3,16.1.1
Φ^*	pullback of a function or a distribution by a map Φ, see §10.1
$\mathrm{p.v.}\frac{1}{x}$	principal value integral, see §5.2.3
$\mathrm{ran}\,P$	the range of an operator P, see §15.3.1
$\mathscr{S}(\mathbb{R}^n)$	the space of Schwartz functions, see §11.1.2
$\mathscr{S}'(\mathbb{R}^n)$	the space of tempered distributions, see §11.2.1
S^m	the space of Kohn–Nirenberg symbols, see §§12.2.3,14.1.1
$S^{-\infty}$	rapidly decaying symbols, see §14.1.1
$\mathbb{S}^n$	the n-dimensional sphere, see §13.1.2
σ_m	the principal symbol of a differential operator, see §13.3
$\mathrm{sing\,supp}$	singular support of a distribution, see §8.3
$\mathrm{Spec}(P)$	spectrum of an operator, see §16.1.1
$\Subset$	compactly contained, see Definition 1.1
supp	support of a function (see §1.2.3) or a distribution (see §4.1)

$\mathbb{T}^n$	the n-dimensional torus, see §13.1.2
$T\mathcal{M}$	tangent bundle, see §13.1.4
$T^*\mathcal{M}$	cotangent bundle, see §13.1.5
x_+^a	see §5.2
x^α	see §1.2.3

Bibliography

[AN07] Nalini Anantharaman and Stéphane Nonnenmacher. Half-delocalization of eigenfunctions for the Laplacian on an Anosov manifold. *Ann. Inst. Fourier (Grenoble)*, 57(7):2465–2523, 2007. Festival Yves Colin de Verdière.

[Ana08] Nalini Anantharaman. Entropy and the localization of eigenfunctions. *Ann. of Math. (2)*, 168(2):435–475, 2008.

[Ava56] Vojislav G. Avakumović. Über die Eigenfunktionen auf geschlossenen Riemannschen Mannigfaltigkeiten. *Math. Z.*, 65:327–344, 1956.

[BDIP02] José Bertin, Jean-Pierre Demailly, Luc Illusie, and Chris Peters. *Introduction to Hodge theory*, volume 8 of *SMF/AMS Texts and Monographs*. American Mathematical Society, Providence, RI; Société Mathématique de France, Paris, 2002. Translated from the 1996 French original by James Lewis and Peters.

[Bea02] H.S. Bear. *A Primer of Lebesgue Integration*. Elsevier Science, 2002.

[Bér77] Pierre H. Bérard. On the wave equation on a compact Riemannian manifold without conjugate points. *Math. Z.*, 155(3):249–276, 1977.

[BT82] Raoul Bott and Loring W. Tu. *Differential forms in algebraic topology*, volume 82 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982.

[CdV85] Yves Colin de Verdière. Ergodicité et fonctions propres du laplacien. In *Bony-Sjöstrand-Meyer seminar, 1984–1985*, pages Exp. No. 13, 8. École Polytech., Palaiseau, 1985.

[CG22] Yaiza Canzani and Jeffrey Galkowski. Logarithmic improvements in the weyl law and exponential bounds on the number of closed geodesics are predominant, 2022.

[Che56] S.S. Chern. *Complex Manifolds: Lectures*. Lu-Chiang, 1956.

[Dav95] E. B. Davies. *Spectral theory and differential operators*, volume 42 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995.

[DF88] Harold Donnelly and Charles Fefferman. Nodal sets of eigenfunctions on Riemannian manifolds. *Invent. Math.*, 93(1):161–183, 1988.

[DG75] J. J. Duistermaat and V. W. Guillemin. The spectrum of positive elliptic operators and periodic bicharacteristics. *Invent. Math.*, 29(1):39–79, 1975.

[DJN21] Semyon Dyatlov, Long Jin, and Stéphane Nonnenmacher. Control of eigenfunctions on surfaces of variable curvature. *J. Amer. Math. Soc.*, 2021.

[DS88] N. Dunford and J.T. Schwartz. *Linear Operators, Part 1: General Theory*. Wiley Classics Library. Wiley, 1988.

[Dya21a] Semyon Dyatlov. Around quantum ergodicity. *Annales mathématiques du Québec*, 2021.

[Dya21b] Semyon Dyatlov. Macroscopic limits of chaotic eigenfunctions, 2021. https://arxiv.org/abs/2109.09053.

[Eva10] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[FJ98] F. Gerard Friedlander and Mark Joshi. *Introduction to the Theory of Distributions*. Cambridge University Press, 1998.

[GS94] Alain Grigis and Johannes Sjöstrand. *Microlocal analysis for differential operators*, volume 196 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1994. An introduction.

[Hör65] Lars Hörmander. Pseudo-differential operators. *Comm. Pure Appl. Math.*, 18:501–517, 1965.

[Hör03] Lars Hörmander. *The analysis of linear partial differential operators. I*. Classics in Mathematics. Springer-Verlag, Berlin, 2003. Distribution theory and Fourier analysis, Reprint of the second (1990) edition [Springer, Berlin; MR1065993 (91m:35001a)].

[Hör05] Lars Hörmander. *The analysis of linear partial differential operators. II*. Classics in Mathematics. Springer-Verlag, Berlin, 2005. Differential operators with constant coefficients, Reprint of the 1983 original.

[Hör07] Lars Hörmander. *The analysis of linear partial differential operators. III*. Classics in Mathematics. Springer, Berlin, 2007. Pseudo-differential operators, Reprint of the 1994 edition.

[Hör09] Lars Hörmander. *The analysis of linear partial differential operators. IV*. Classics in Mathematics. Springer-Verlag, Berlin, 2009. Fourier integral operators, Reprint of the 1994 edition.

[Hux03] M. N. Huxley. Exponential sums and lattice points. III. *Proc. London Math. Soc. (3)*, 87(3):591–609, 2003.

[Huy05] Daniel Huybrechts. *Complex geometry*. Universitext. Springer-Verlag, Berlin, 2005. An introduction.

[Jon01] F. Jones. *Lebesgue Integration on Euclidean Space*. Jones and Bartlett books in mathematics. Jones and Bartlett, 2001.

[Kli95] Wilhelm P. A. Klingenberg. *Riemannian geometry*, volume 1 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, second edition, 1995.

[KN65] J. J. Kohn and L. Nirenberg. An algebra of pseudo-differential operators. *Comm. Pure Appl. Math.*, 18:269–305, 1965.

[Lax02] P.D. Lax. *Functional Analysis*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2002.

[Lee13] John M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2013.

[Lev53] B. M. Levitan. On expansion in characteristic functions of the Laplace operator. *Doklady Akad. Nauk SSSR (N.S.)*, 90:133–135, 1953.

[Lin06] Elon Lindenstrauss. Invariant measures and arithmetic quantum unique ergodicity. *Ann. of Math. (2)*, 163(1):165–219, 2006.

[LM20] Alexander Logunov and Eugenia Malinnikova. Review of Yau's conjecture on zero sets of Laplace eigenfunctions. In *Current developments in mathematics 2018*, pages 179–212. Int. Press, Somerville, MA, [2020] ©2020.

[Log18a] Alexander Logunov. Nodal sets of Laplace eigenfunctions: polynomial upper estimates of the Hausdorff measure. *Ann. of Math. (2)*, 187(1):221–239, 2018.

[Log18b] Alexander Logunov. Nodal sets of Laplace eigenfunctions: proof of Nadirashvili's conjecture and of the lower bound in Yau's conjecture. *Ann. of Math. (2)*, 187(1):241–262, 2018.

[Mel] Richard Melrose. Differential analysis, lecture notes for 18.155 and 156. https://math.mit.edu/~rbm/18.155-F16/GradAnal.pdf.

[Mel93] Richard B. Melrose. *The Atiyah-Patodi-Singer index theorem*, volume 4 of *Research Notes in Mathematics*. A K Peters, Ltd., Wellesley, MA, 1993.

[RS81] M. Reed and B. Simon. *I: Functional Analysis*. Methods of Modern Mathematical Physics. Elsevier Science, 1981.

[RS94] Zeév Rudnick and Peter Sarnak. The behaviour of eigenstates of arithmetic hyperbolic manifolds. *Comm. Math. Phys.*, 161(1):195–213, 1994.

[Rud64] W. Rudin. *Principles of Mathematical Analysis*. International series in pure and applied mathematics. McGraw-Hill, 1964.

[Rud87] W. Rudin. *Real and Complex Analysis*. Higher Mathematics Series. McGraw-Hill Education, 1987.

[Rud91] W. Rudin. *Functional Analysis*. International series in pure and applied mathematics. McGraw-Hill, 1991.

[Sch50] L. Schwartz. *Théorie des distributions*. Number v. 1 in Actualités scientifiques et industrielles : Publications de l'Institut de Mathématique de l'Université de Strasbourg. Jouve, 1950.

[Sch57] L. Schwartz. *Théorie des distributions*. Number v. 2 in Actualités scientifiques et industrielles. Hermann, 1957.

[Shn74a] Alexander Shnirelman. Ergodic properties of eigenfunctions. *Uspehi Mat. Nauk*, 29(6(180)):181–182, 1974.

[Shn74b] Alexander Shnirelman. Statistical properties of eigenfunctions. In *Proceedings of the All-USSR School in Differential Equations with Infinite Number of Independent Variables and in Dynamical Systems with Infinitely Many Degrees of Freedom, Dilijan, Armenia, May 21–June 3, 1973*. Armenian Academy of Sciences, Erevan, 1974.

[Sob36] Sergeĭ Sobolev. Méthode nouvelle à résoudre le problème de cauchy pour les équations linéaires hyperboliques normales. *Matematiceskij sbornik*, 43(1):39–72, 1936.

[Sob91] Sergeĭ Sobolev. *Some applications of functional analysis in mathematical physics*, volume 90 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1991. Translated from the third Russian edition by Harold H. McFaden, With comments by V. P. Palamodov.

[Spi65] Michael Spivak. *Calculus on manifolds. A modern approach to classical theorems of advanced calculus*. W. A. Benjamin, Inc., New York-Amsterdam, 1965.

[Str11] D.W. Stroock. *Essentials of Integration Theory for Analysis*. Graduate Texts in Mathematics. Springer New York, 2011.

[SU13] Alexander Strohmaier and Ville Uski. An algorithm for the computation of eigenvalues, spectral zeta functions and zeta-determinants on hyperbolic surfaces. *Comm. Math. Phys.*, 317(3):827–869, 2013.

[Tay11a] Michael E. Taylor. *Partial differential equations I. Basic theory*, volume 115 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2011.

[Tay11b] Michael E. Taylor. *Partial differential equations II. Qualitative studies of linear equations*, volume 116 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2011.

[Voi07] Claire Voisin. *Hodge theory and complex algebraic geometry. I*, volume 76 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, english edition, 2007. Translated from the French by Leila Schneps.

[War71] Frank W. Warner. *Foundations of differentiable manifolds and Lie groups*. Scott, Foresman & Co., Glenview, Ill.-London, 1971.

[Wey12] Hermann Weyl. Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung). *Math. Ann.*, 71(4):441–479, 1912.

[Zel87] Steven Zelditch. Uniform distribution of eigenfunctions on compact hyperbolic surfaces. *Duke Math. J.*, 55(4):919–941, 1987.

[Zwo12] Maciej Zworski. *Semiclassical analysis*, volume 138 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.