Lecture Notes

Analysis of Sublaplacians on Lie Groups

Course held by

Prof. Fulvio Ricci

Notes written by

Francesco Paolo Maiale

Scuola Normale Superiore, Pisa

Last Update: February 17, 2020

Disclaimer

I wrote these notes to summarise the content of the course Analysis of Sublaplacians on Lie Groups, held by Professor Fulvio Ricci at SNS.

I tried to include all the topics that were discussed in class and combine it with some additional information from several other courses to produce a self-contained document.

I will try to review them periodically, but I am sure that at the end there will be a large number of mistakes and oversights. To report them, feel free to send me an email at **francesco (dot) maiale (at) sns (dot) it**.

Contents

Part I

Fourier Analysis on Euclidean Spaces

Chapter 1

Fourier Transform on $\mathbb{R}^n$

The **convolution** $*$ is a fundamental operation in Mathematics and, while it is usually defined between two functions, it can be applied in a much more general setting. In fact, we can give a meaning to the expression

$$\Phi * \Psi$$

even when both Φ and Ψ are distributions, although we have to introduce mild assumptions on one of the supports. The connection between convolution and **Fourier transform** is well-known and it says that

$$\mathscr{F}(f)\mathscr{F}(g) = c_n \mathscr{F}(f * g),$$

where f and g range in an appropriate class. In particular, we will investigate properties of the Fourier transform on Schwartz spaces,

$$\mathscr{S}(\Omega) := \left\{ f \in C^\infty(\Omega) \ : \ |\partial^\alpha f(x)| \leq C_{\alpha,\, N}(1 + |x|)^{-N} \ \forall (\alpha,\, N) \in \mathbb{N}^n \times \mathbb{N} \right\}$$

such as being an invertible map (allowing us to define the so-called *inverse Fourier transform*), even when the domain is the dual $\mathscr{S}'(\Omega)$.

1.1 Introduction

In this brief section, we recollect some useful definitions and properties related to the convolution of two functions in order to set the ground for the generalisation to distributions.

Definition 1.1 (Convolution)**.** Let f and g be two functions. The *convolution* between them is defined by setting

$$f * g(x) := \int_{\mathbb{R}^n} f(x - y) g(y) \, \mathrm{d}y. \tag{1.1}$$

We did not use the term function to refer to (1.1) because it actually depends on the regularity of f and g. Indeed, it may happen that the integral

$$\int_{\mathbb{R}^n} f(x-y)g(y)\,\mathrm{d}y$$

is not well-defined as a function, although it might still belong to some larger class.

Remark 1.2. If $f,\, g \in L^1(\mathbb{R}^n)$, then (1.1) converges (absolutely) for almost every $x \in \mathbb{R}^n$ and using Fubini's theorem leads to the formula

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x-y)g(y)\,\mathrm{d}y\mathrm{d}x = \int_{\mathbb{R}^n} g(y)\int_{\mathbb{R}^n} f(x-y)\,\mathrm{d}x\mathrm{d}y.$$

A simple computation shows that

$$\|f * g\|_{L^1(\mathbb{R}^n)} \le \|f\|_{L^1(\mathbb{R}^n)}\|g\|_{L^1(\mathbb{R}^n)}. \tag{1.2}$$

Remark 1.3. The convolution operation is commutative. Indeed, given $f,\, g \in L^1(\mathbb{R}^n)$ we can use the change of variables formula to obtain the following identity

$$f * g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\,\mathrm{d}y \overset{v=x-y}{=} \int_{\mathbb{R}^n} f(v)g(x-v)\,\mathrm{d}v = g * f(x).$$

Remark 1.4. Let $f,\, g \in L^1(\mathbb{R}^n)$. Then

$$\int_{\mathbb{R}^n} f * g(x)h(x)\,\mathrm{d}x = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(u)g(v)h(u+v)\,\mathrm{d}u\mathrm{d}v \tag{1.3}$$

holds, for example, for all $h \in L^1(\mathbb{R}^n)$.

The idea is that we can use h as a test function to generalise the definition of convolution to a slightly more general setting, namely to finite Borel measures.

Definition 1.5 (Convolution)**.** Let $\mu,\, \nu \in \mathcal{M}(\mathbb{R}^n)$ be two finite Borel measures. We define their convolution, denoted by $\mu * \nu$, by setting

$$\int_{\mathbb{R}^n} h(x)\mathrm{d}(\mu * \nu)(x) := \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} h(x+y)\,\mathrm{d}\mu(x)\mathrm{d}\nu(y) \quad \text{for all } h \in C_0(\mathbb{R}^n). \tag{1.4}$$

Remark 1.6. Notice that (1.4) uniquely identifies the measure $\mu * \nu$. Indeed, it follows from Riesz (representation) theorem that $\mathcal{M}(\mathbb{R}^n)$ is the dual space of $C_0(\mathbb{R}^n)$ and

$$\|\mu\|_{C_0(\mathbb{R}^n)'} = |\mu|(\mathbb{R}^n) =: \|\mu\|_1,$$

where the right-hand side denotes the **total variation** of μ. This is defined for any complex-valued measure ν as follows:

$$|\nu|(E) = \sup_{\pi} \sum_{A\in\pi} |\nu(A)|,$$

where π ranges among all countable partition in measurable subsets of E.

Proposition 1.7. *Let $\mu,\, \nu \in \mathcal{M}(\mathbb{R}^n)$ be two finite Borel measures. Then $\mu * \nu$ is also a finite Borel measure and the analogous of* (1.2) *holds:*

$$\|\mu * \nu\|_1 \le \|\mu\|_1\|\nu\|_1. \tag{1.5}$$

Proof. Let h be a $C_0(\mathbb{R}^n)$ function. By definition $\mu * \nu$ acts on h as follows:

$$\int_{\mathbb{R}^n} h(x)\mathrm{d}(\mu * \nu)(x) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} h(x+y)\,\mathrm{d}\mu(x)\mathrm{d}\nu(y).$$

Now take the absolute value of both sides and notice that

$$\left|\int_{\mathbb{R}^n} h(x)\mathrm{d}(\mu * \nu)(x)\right| \le \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |h(x+y)|\,\mathrm{d}|\mu|(x)\mathrm{d}|\nu|(y).$$

Since $h \in C_0(\mathbb{R}^n) \subset L^\infty(\mathbb{R}^n)$, we know that $\|h\|_\infty$ is equal to a finite constant, and therefore we can write the following estimate:

$$\left|\int_{\mathbb{R}^n} h(x)\mathrm{d}(\mu * \nu)(x)\right| \le \|h\|_\infty \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \mathrm{d}|\mu|(x)\mathrm{d}|\nu|(y) = \|h\|_\infty \|\mu\|_1 \|\nu\|_1.$$

Finally take the supremum on both sides with respect to $h \in C_0(\mathbb{R}^n)$ with $\|h\|_\infty \le 1$ to achieve the desired inequality. □

Proposition 1.8. *Let $\mu, \nu \in \mathcal{M}(\mathbb{R}^n)$ be two finite Borel measures and assume that ν is absolutely continuous with respect to the Lebesgue measure $\mathcal{L}^n$ with density f. Then*

$$\mu * \nu \ll \mathcal{L}^n,$$

and its density is given by

$$\varphi(x) = \int_{\mathbb{R}^n} f(x-y)\,\mathrm{d}\mu(y) =: f * \mu(x).$$

Assume now that $f \in L^1(\mathbb{R}^n)$ and $g \in L^\infty(\mathbb{R}^n)$. Then (1.1) converges (absolutely) for *every* $x \in \mathbb{R}^n$ and using Fubini's theorem we find that

$$\|f * g\|_\infty \le \|f\|_{L^1(\mathbb{R}^n)} \|g\|_\infty. \tag{1.6}$$

In particular, the convolution $f * g$ belongs to $L^\infty(\mathbb{R}^n)$ and, with little effort, it can be proved that it is actually a bounded function (i.e., an element of $C_b(\mathbb{R}^n)$.)

Lemma 1.9. *Let $g \in L^p(\mathbb{R}^n)$, $1 < p < \infty$. Then*

$$g(x) = g_1(x) + g_2(x) \in L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n).$$

Proof. The set Ω is bounded, so we have the inclusion

$$L^p(\Omega) \subseteq L^1(\Omega)$$

for all $1 \le p \le \infty$. We can thus choose g_1 and g_2 as follows:

$$g_1(x) = \begin{cases} g(x) & \text{if } |x| \le 1, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad g_2(x) = \begin{cases} g(x) & \text{if } |x| > 1, \\ 0 & \text{otherwise.} \end{cases}$$

□

This decomposition allows us to define the convolution between a function $f \in L^1(\mathbb{R}^n)$

and another function $g \in L^p(\mathbb{R}^n)$, $1 < p < \infty$. Indeed,

$$f * g(x) = f * g_1(x) + f * g_2(x) \in L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n),$$

but it can be proved (see Theorem 1.13) that it actually belongs to $L^p(\mathbb{R}^n)$.

Definition 1.10 (Bounded Operator)**.** Let $(\mathfrak{X},\, \mu)$ and $(\mathfrak{Y},\, \nu)$ be two measures space. We say that the operator

$$T : L^p(\mathfrak{X},\, \mu) \longrightarrow L^q(\mathfrak{Y},\, \nu)$$

is *bounded* if

$$\|T(f)\|_{L^q(\mathfrak{Y},\, \nu)} \lesssim_{p,\, q} \|f\|_{L^p(\mathfrak{X},\, \mu)} \quad \text{for all } f \in L^p(\mathfrak{X},\, \mu). \tag{1.7}$$

We can introduce a slightly weaker notion of boundedness, which is useful when we deal with the so-called generalized Lebesgue spaces (or Lorentz spaces).

Definition 1.11 (Weakly Bounded Operator)**.** Let $(\mathfrak{X},\, \mu)$ and $(\mathfrak{Y},\, \nu)$ be two measures space. We say that the operator

$$T : L^p(\mathfrak{X},\, \mu) \longrightarrow L^{q,\, \infty}(\mathfrak{Y},\, \nu)$$

is *weakly bounded* (or of *weak-type* $(p,\, q)$) if

$$\nu\left(\{y \in \mathfrak{Y} \;:\; |T(f)(y)| > \lambda\}\right) \lesssim_{p,\, q} \left(\frac{\|f\|_{L^p(\mathfrak{X},\, \mu)}}{\lambda}\right)^q \tag{1.8}$$

for all $f \in L^p(\mathfrak{X},\, \mu)$ and all $\lambda > 0$.

We are now ready to state one of the most significant interpolations result in analysis. The idea is that a strongly bounded operator at the extremal of a segment, say

$$(p_0,\, q_0) \quad \text{and} \quad (p_1,\, q_1),$$

is also weakly bounded in the interior of the segment, that is,

$$(p,\, q) \in \{t(p_0,\, q_0) + (1-t)(p_1,\, q_1) \;:\; t \in [0,\, 1]\}\,.$$

Definition 1.12. Let $(\mathfrak{X},\, \mu)$ be a measure space. We say that $\mathfrak{X}$ is a σ-finite measure space if we can find a countable family $(X_n)_{n\in\mathbb{N}}$ such that

$$\mathfrak{X} = \bigcup_{n\in\mathbb{N}} X_n \quad \text{and} \quad \mu(A_n) < \infty \text{ for all } n \in \mathbb{N}.$$

Theorem 1.13 (Riesz-Thorin)**.** *Let $(\mathfrak{X},\, \mu)$ and $(\mathfrak{Y},\, \nu)$ be two σ-finite measure spaces. Consider a linear operator*

$$T : \mathcal{S}(\mathfrak{X},\, \mu) \longrightarrow \mathcal{M}(\mathfrak{Y},\, \nu),$$

where $\mathcal{S}(\mathfrak{X},\, \mu)$ here denotes the set of finitely simple functions[1]*. Let $1 \le p_0,\, p_1,\, q_0,\, q_1 \le \infty$ and assume that*

$$\|T(f)\|_{L^{q_0}(\mathfrak{Y},\, \nu)} \le M_0 \|f\|_{L^{p_0}(\mathfrak{X},\, \mu)},$$

$$\|T(f)\|_{L^{q_1}(\mathfrak{Y},\, \nu)} \le M_1 \|f\|_{L^{p_1}(\mathfrak{X},\, \mu)}$$

[1] A simple function is finitely simple if and only if it is supported in a set of finite measure. Also, recall that these are dense in $L^p(\mathfrak{X},\, \mu)$ for all $0 < p < \infty$.

for all $f \in \mathcal{S}(\mathfrak{X}, \mu)$*. Then for all* $\theta \in (0, 1)$ *we have*

$$\|T(f)\|_{L^q(\mathfrak{Y}, \nu)} \leq M_0^{1-\theta} M_1^{\theta} \|f\|_{L^p(\mathfrak{X}, \mu)} \tag{1.9}$$

for all $f \in \mathcal{S}(\mathfrak{X}, \mu)$*, where*

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1} \quad \textit{and} \quad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

Proof. The reader may consult the original paper [12], due to Thorin, which extends the result of Riesz via complex methods. □

We are now ready to apply this result to the convolution. Let $f \in L^p(\mathbb{R}^n)$ for some $p \in (1, \infty)$, and consider the operator

$$T_f(g) := f * g.$$

We know already that T is linear and satisfies the following inequalities:

$$\|T_f(g)\|_{L^p(\mathbb{R}^n)} \leq \|f\|_{L^p(\mathbb{R}^n)} \|g\|_{L^1(\mathbb{R}^n)},$$

$$\|T_f(g)\|_{L^\infty(\mathbb{R}^n)} \leq \|f\|_{L^p(\mathbb{R}^n)} \|g\|_{L^{p'}(\mathbb{R}^n)},$$

where the second is an obvious consequence of Hölder's inequality. Applying (1.9) with $M_0 = M_1 = \|f\|_{L^p(\mathbb{R}^n)}$ we find the inequality

$$\|f * g\|_{L^r(\mathbb{R}^n)} \leq \|f\|_{L^p(\mathbb{R}^n)} \|g\|_{L^q(\mathbb{R}^n)}, \tag{1.10}$$

for all triples (p, q, r) that satisfy the condition

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}.$$

This result is usually known as **Young inequality** and it arises in countless situations in harmonic analysis as well as in other areas of analysis.

1.2 Convolution of distributions

Let $\Omega \subset \mathbb{R}^n$ be an open set. We denote by $\mathcal{D}(\Omega)$ the family of C^∞-functions with compact support contained in Ω. Recall that the family of seminorms

$$p_{K, \alpha}(f) := \sup_{x \in K} |\partial^\alpha f(x)|,$$

where K ranges among all compact subsets of Ω and $\alpha \in \mathbb{N}^n$, generates a locally convex topology on $\mathcal{D}(\Omega)$. It is easy to verify that we can equivalently consider the enlarged family

$$\widetilde{\mathcal{P}} := \{\| \cdot \|_{(N)}\}_{N \in \mathbb{N}},$$

defined by setting

$$\|f\|_{(N)} := \max_{|\alpha| \leq N} \max_{x \in \Omega} |D^\alpha \varphi(x)|, \tag{1.11}$$

generates the same locally convex topology. The space of *distributions* is the dual space of $\mathcal{D}(\Omega)$, that is, the space of all linear and continuous maps

$$\Phi : \mathcal{D}(\Omega) \longrightarrow \mathbb{C}.$$

Given a distribution $\Phi \in \mathcal{D}'(\Omega)$ and a function $f \in \mathcal{D}(\Omega)$, we can introduce the following notation for the evaluation

$$\Phi(f) := \langle \Phi, f \rangle,$$

which will be useful later on to compare properties of distributions versus integrals.

Definition 1.14 (Order)**.** Let $\Phi \in \mathcal{D}'(\Omega)$. The *order* of Φ is defined as the smallest positive integer N such that for all $K \subset \Omega$ there exists $C(K) > 0$ with

$$|\Phi(f)| \leq C(K) \|f\|_{(N)} \quad \text{for all } f \in \mathcal{D}_K.$$

If such a N does not exist, then we say that Φ has order infinity.

Example 1.15 (Dirac)**.** Let $p \in \mathbb{R}^n$. The *Dirac delta* at the point p is defined by

$$\delta_p : \mathcal{D}(\mathbb{R}^n) \ni f \longmapsto f(p) \in \mathbb{C}.$$

This map defines a distribution since it is both linear and continuous. Moreover, it is easy to verify that δ_p has order zero since for each given compact set $K \subset \Omega$ we have

$$|\delta_p(f)| \leq \|f\|_{(0)}$$

for all $f \in \mathcal{D}_K$. Note that δ_p belongs to a class we know already: it is a measure.

Example 1.16. Let $p \in \mathbb{R}^n$ and set

$$\Phi_p(f) := f'(p).$$

It is easy to check that Φ_p is a distribution of order one, but the reader might find it interesting to show that Φ_p is **not** a measure in any sense so

$$\mathcal{M}(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$$

is an actual strict inclusion.

Example 1.17 (Locally Summable)**.** Given $f \in L^1_{\text{loc}}(\Omega)$, define

$$\Lambda_f(g) := \int_\Omega f(x) g(x) \, \mathrm{d}x.$$

The linearity is obvious (the integral itself is linear), while the continuity is an easy consequence of the following estimate:

$$|\Lambda_f(g)| \leq \|f\|_{L^1(\mathrm{spt}(g))} \|g\|_{(0)}.$$

Example 1.18. Given μ Borel measure (or locally finite positive measure), define

$$\Lambda_\mu(f) := \int_\Omega f(x)\,\mathrm{d}\mu(x).$$

The linearity is obvious, while the continuity is an easy consequence of the following estimate:

$$|\Lambda_\mu(g)| \le \|\mu\|_1 \|g\|_{(0)},$$

where $\|\mu\|_1$ denotes the total variation norm of μ.

We now wonder whether or not it makes sense to multiply a Borel measure μ for a function f. We expect that $f\mu$ is a measure satisfying the following identity:

$$\int_{\mathbb{R}^n} g(x)\,\mathrm{d}(f\mu)(x) = \int_{\mathbb{R}^n} g(x)f(x)\,\mathrm{d}\mu(x).$$

However, the right-hand side is well-defined if and only if $fg \in C_0(\mathbb{R}^n)$ for all $g \in C_0(\mathbb{R}^n)$. Equivalently, we might require $f \in C_b(\mathbb{R}^n)$. Therefore, the object

$$(f,\, \mu) \longmapsto f\mu$$

can be defined at least for all $f \in C_b(\mathbb{R}^n)$.

Definition 1.19 (Product)**.** Let $\Phi \in \mathcal{D}'(\Omega)$ and $f \in \mathcal{D}(\Omega)$. The product $f\Phi$ is the distribution defined by

$$\langle f\Phi,\, g\rangle := \langle \Phi,\, fg\rangle. \tag{1.12}$$

Exercise 1.1. Prove that (1.12) defines a distribution. Namely, show that for all $K \subset \Omega$ compact there exist $N = N(K) \in \mathbb{N}$ and $C(K) = C > 0$ such that

$$|f\Phi(g)| \le C\|g\|_{(N)}$$

Definition 1.20 (Derivative)**.** Let $\Phi \in \mathcal{D}'(\Omega)$ be a distribution. The *derivative* with respect to the direction x_k of Φ is given by

$$\langle \partial_{x_k}\Phi,\, g\rangle := -\langle \Phi,\, \partial_{x_k} g\rangle. \tag{1.13}$$

Note that the definition (1.13) is given in such a way that when $\langle f,\, g\rangle = \int f(x)g(x)\,\mathrm{d}x$, it is nothing but the integration by parts formula.

Exercise 1.2. Prove that (1.13) defines a distribution.

Example 1.21 (Heaviside)**.** Let $H : \mathbb{R} \longrightarrow \mathbb{R}$ be the function defined by

$$H(x) = \begin{cases} 1 & \text{if } x \ge 0, \\ 0 & \text{if } x < 0. \end{cases}$$

This function is not differentiable, but it admits a distributional derivative. The functional associated to H is given by Λ_H, and its derivative is

$$\partial_{x_k}\Lambda_H(f) = -\int_{\mathbb{R}} H(x)f'(x)\,\mathrm{d}x = f(0)$$

for all $f \in \mathcal{D}(\mathbb{R})$. It follows that the distributional derivative of the Heaviside function is the Dirac delta δ at $p = 0$.

Definition 1.22 (Support)**.** Let $\Phi \in \mathcal{D}'(\mathbb{R}^n)$. The *support* of Φ is defined by

$$\operatorname{spt}(\Phi) := \mathbb{R}^n \setminus \bigcup \{\omega \subset \mathbb{R}^n \mid \omega \text{ is an open subset such that } \Phi \equiv 0 \text{ on } \mathcal{D}'(\omega)\}. \tag{1.14}$$

Remark 1.23. The notion is well-defined since Φ is actually zero as an element of

$$\mathcal{D}'\left(\Omega \setminus \operatorname{spt}(\Phi)\right).$$

Theorem 1.24. *Let $\Phi \in \mathcal{D}'(\mathbb{R}^n)$, and let $S := \operatorname{spt}(\Phi)$. Then the following properties hold:*

- ***(a)*** *If $f \in \mathcal{D}(\mathbb{R}^n)$ has support disjoint from S, then $\langle \Phi, f\rangle = 0$.*
- ***(b)*** *If $S = \varnothing$, then $\Phi = 0$.*
- ***(c)*** *If $g \in C^\infty(\Omega)$ is a smooth function such that $g \equiv 1$ on some open subset $U \supset S$, then $g\Phi = \Phi$ as distributions.*

The goal of this section is to find a well-defined notion of convolution between a distribution $\Phi \in \mathcal{D}'$ and a function $f \in \mathcal{D}$. Set

$$\langle \Phi * f, g\rangle := \langle \Phi, \check{f} * g\rangle. \tag{1.15}$$

Here we use the symbol $\check{f}(y)$ to denote the function f evaluated at the symmetric point of y, that is, we set

$$\check{f}(y) := f(-y).$$

Introduce also the following notation:

$$\langle \tau_x \Phi, f\rangle := \langle \Phi, \tau_{-x} f\rangle \quad \text{and} \quad \langle \check{\Phi}, f\rangle := \langle \Phi, \check{f}\rangle.$$

We now claim that (1.15) is well-defined and $\Phi * f$ is a function, not a distribution as one would normally expect. In fact, if Φ is a Borel measure μ, we know that

$$f * \mu(x) = \int_{\mathbb{R}^n} f(x-y)\, \mathrm{d}\mu(y) = \langle \mu, \tau_x \check{f}\,\rangle,$$

so it makes sense to define the smooth function

$$u(x) := \langle \Phi, \tau_x \check{f}\,\rangle.$$

Then (1.15) can also be rewritten as

$$\langle \Phi * f, g\rangle = \langle u, g\rangle_{L^2(\mathbb{R}^n)},$$

which means that $\Phi * f$ applied to g is nothing but the distribution associated to the function u applied to g. In particular, we can define the convolution as

$$\Phi * f(x) := \langle \Phi, \tau_x \check{f}\,\rangle. \tag{1.16}$$

Proposition 1.25 (Associativity)**.** *Let $\Phi \in \mathcal{D}'$, and let $f, g \in \mathcal{D}$. The convolution is* *associative, that is,*

$$(\Phi * f) * g = \Phi * (f * g).$$

To conclude this section, we want to give the main ideas that are needed to properly define the convolution between two (or more) distributions - which is not always possible!

Proposition 1.26. *Let $\Phi \in \mathcal{D}'$ be a distribution. The operator defined by*

$$\Phi_* : \mathcal{D} \ni f \longmapsto \Phi * f \in \mathcal{E}$$

is linear, continuous and translation-invariant. Furthermore, if $L \in \mathcal{L}(\mathcal{D}, \mathcal{E})$ is a linear continuous operator commuting with translations, then

$$\textit{there exists a unique } \Phi \in \mathcal{D}' \textit{ such that } L = \Phi_*.$$

Definition 1.27 (Convolution)**.** Let $\Phi \in \mathcal{E}'$ be a compactly supported distribution and let $\Psi \in \mathcal{D}'$ be a generic distribution. The convolution is uniquely determined by the identity

$$\langle \Phi * \Psi, f \rangle := \langle \Phi, \check{\Psi} * f \rangle. \tag{1.17}$$

Lemma 1.28. *Let $\Phi \in \mathcal{E}'$ and $\Psi \in \mathcal{D}'$. Then the support of the convolution is contained in the sum of the supports, that is,*

$$\mathrm{spt}(\Phi * \Psi) \subseteq \mathrm{spt}(\Phi) + \mathrm{spt}(\Psi).$$

Exercise 1.3. Prove that (1.17) is well-defined using the previous result.

1.3 Fourier transform in $L^1(\mathbb{R}^n)$

Let $f \in L^1(\mathbb{R}^n)$. The *Fourier transform* of f is the function defined by

$$\hat{f}(\xi) := \int_{\mathbb{R}^n} f(x) \mathrm{e}^{-\imath x \cdot \xi} \, \mathrm{d}x, \tag{1.18}$$

where $\cdot$ denotes the standard Euclidean scalar product on $\mathbb{R}^n$.

Caution!

We will also use the symbol $\mathscr{F}(f)$ for the Fourier transform.

Notation. We denote by e_x the exponential function

$$e_x : \mathbb{C} \ni \xi \longmapsto \mathrm{e}^{\imath x \cdot \xi} \in \mathbb{C},$$

and, for any $\lambda \neq 0$, we denote by h_λ the scaling

$$h_\lambda : L^1(\mathbb{R}^n) \ni \varphi \longmapsto \varphi_\lambda \in L^1(\mathbb{R}^n),$$

where

$$\varphi_\lambda(x) := \varphi\left(\frac{x}{\lambda}\right).$$

We now recollect a few basic properties of the Fourier transform operator which are simple consequences of the definition (1.18).

Proposition 1.29. *Let $f,\, g \in L^1(\mathbb{R}^n)$. Then the following properties hold:*

(a) *For all $x \in \mathbb{R}^n$ we have*
$$\mathscr{F}(\tau_x(f)) = e_{-x}\mathscr{F}(f).$$

(b) *For all $x \in \mathbb{R}^n$ we have*
$$\mathscr{F}(e_x(f)) = \tau_x\mathscr{F}(f).$$

(c) *The Fourier transform of the convolution equals of the product of the Fourier transforms, that is,*
$$\mathscr{F}(f * g) = \mathscr{F}(f)\mathscr{F}(g). \tag{1.19}$$

(d) *Let $\lambda > 0$. The Fourier transform equals the inverse scaling of the Fourier transform, that is,*
$$\mathscr{F}\left(h_\lambda(f)\right) = \lambda^n h_{\frac{1}{\lambda}}(\mathscr{F}(f)).$$

(e) *The Fourier transform of the overturning is given by*
$$\mathscr{F}(\check{f}) = (-1)^n\check{\mathscr{F}}(\varphi).$$

A fundamental property[2] is that $\mathscr{F}$ maps $L^1(\mathbb{R}^n)$ in $C_0(\mathbb{R}^n)$, the space of continuous functions vanishing at infinity. Therefore
$$\mathscr{F} : L^1(\mathbb{R}^n) \longrightarrow C_0(\mathbb{R}^n) \subset L^\infty(\mathbb{R}^n)$$
is a bounded operator with costant one; namely,
$$\|\hat{f}\|_\infty \leq \|f\|_{L^1(\mathbb{R}^n)}. \tag{1.20}$$
The inclusion in $L^\infty(\mathbb{R}^n)$ is easy to verify since one can estimate $\mathscr{F}(f)$ as follows:
$$|\mathscr{F}(f)(\xi)| \leq \int_{\mathbb{R}^n} |f(x)| \underbrace{|\mathrm{e}^{\imath x\cdot\xi}|}_{=1} \,\mathrm{d}x = \|f\|_{L^1(\mathbb{R}^n)}.$$

1.3.1 Extension of the Fourier transform to $L^p(\mathbb{R}^n)$, $1 \leq p \leq 2$

We now want to show that the Fourier transform $\mathscr{F}$ can be extended to $L^2(\mathbb{R}^n)$-functions, in such a way that $\mathscr{F}(f)$ coincides with (1.18) whenever $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$.

Remark 1.30. If $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then its Fourier transform $\mathscr{F}(f)$ belongs to $L^2(\mathbb{R}^n)$ as a consequence of the Plancherel's identity
$$\|\mathscr{F}(f)\|^2_{L^2(\mathbb{R}^n)} = 2\pi\|f\|^2_{L^2(\mathbb{R}^n)}. \tag{1.21}$$

Theorem 1.31 (Plancherel)**.** *The Fourier transform $\mathscr{F} : L^1(\mathbb{R}^n) \to L^1(\mathbb{R}^n)$ can be uniquely extended to an operator $\tilde{\mathscr{F}} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ such that*
$$\tilde{\mathscr{F}}(f) = \mathscr{F}(f) \quad \textit{for all } f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$$

[2] See the Lebesgue-Riemann lemma.

and (1.21) holds for all $f \in L^2(\mathbb{R}^n)$ with $\tilde{\mathscr{F}}(f)$ in place of $\mathscr{F}(f)$.

Proof. We first notice that for all $f,\, g \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ we have

$$2\pi \langle f,\, g\rangle_2 = \langle \mathscr{F}(f),\, \mathscr{F}(f)\rangle_2$$

as a consequence of the polarisation formula applied to (1.21), where

$$\langle f,\, g\rangle_2 = \int_{\mathbb{R}^n} f(x)\overline{g(x)}\, \mathrm{d}x.$$

The space $L^1 \cap L^2$ is dense in L^2 with respect to $\| \cdot \|_{L^2(\mathbb{R}^n)}$, and hence

$$\mathscr{F} : L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n) \longrightarrow L^2(\mathbb{R}^n)$$

can be continuously extended in a unique way to an operator $\tilde{\mathscr{F}} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ that satisfies (1.21). It remains to verify that $\tilde{\mathscr{F}}$ is well-defined. Let $f \in L^2(\mathbb{R}^n)$ and set

$$f_n(x) := f(x)\chi_{B(0,\, n)}.$$

Then $f_n \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and f_n converges to f with respect to the L^2-topology. Therefore, using the Plancherel's identity and the linearity of $\mathscr{F}$ we obtain

$$\|\mathscr{F}(f_n) - \mathscr{F}(f_m)\|^2_{L^2(\mathbb{R}^n)} = 2\pi \|f_n - f_m\|^2_{L^2(\mathbb{R}^n)},$$

which means that $\{\mathscr{F}(f_n)\}_{n\in\mathbb{N}}$ is a Cauchy sequence (because $\{f_n\}_{n\in\mathbb{N}}$ is). By completeness, we can find $g \in L^2(\mathbb{R}^n)$ such that

$$\mathscr{F}(f_n) \xrightarrow{L^2(\mathbb{R}^n)} g \implies \tilde{\mathscr{F}}(f) := g$$

is well-defined. □

Corollary 1.32. *The Fourier transform $\mathscr{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is a continuous bounded operator with constant 2π.*

As a consequence of this fact, we can apply Riesz-Thorin interpolation inequality (1.9) to infer that for all $p \in [1,\, 2]$ the Fourier transform

$$\mathscr{F} : L^p(\mathbb{R}^n) \longrightarrow L^{p'}(\mathbb{R}^n)$$

is a continuous bounded operator. Moreover, the **Hausdorff-Young** inequality

$$\|\mathscr{F}(f)\|_{L^{p'}(\mathbb{R}^n)} \leq (2\pi)^{\frac{1}{p'}} \|f\|_{L^p(\mathbb{R}^n)} \tag{1.22}$$

holds, where $(p,\, p')$ is a Hölder-conjugate couple, which means that

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

1.4 Schwartz spaces $\mathscr{S}(\mathbb{R}^n)$

The goal of this section is to introduce a class of smooth functions on which the Fourier transform operator has particularly unusual properties such as invertibility.

Definition 1.33 (Schwartz)**.** Let $f \in C^\infty(\mathbb{R}^n)$ be a smooth function. We say that f belongs to the *Schwartz space* $\mathscr{S}(\mathbb{R}^n)$ if, for all $\alpha \in \mathbb{N}^n$ and $N \in \mathbb{N}$, it turns out that

$$|\partial^\alpha f(x)| \leq c(\alpha,\, N)(1+|x|)^{-N}. \tag{1.23}$$

In the literature, functions in $\mathscr{S}(\mathbb{R}^n)$ are also referred to as *rapidly decreasing* as a consequence of the fact that any derivative goes to zero faster than any polynomial.

We will now briefly discuss the topological properties of $\mathscr{S}(\mathbb{R}^n)$ and show that it is a Fréchet space endowed with a special family of seminorms.

Definition 1.34 (F-Space)**.** We say that a topological vector space $(X,\, \tau)$ is a *F-space* if the following properties are satisfied:

(a) The topology τ is induced by a translation-invariant metric d.

(b) The metric space $(X,\, d)$ is complete.

Definition 1.35 (Fréchet Space)**.** We say that a topological vector space $(X,\, \tau)$ is a *Fréchet space* if $(X,\, \tau)$ is a locally convex F-space.

A locally convex topology is usually characterised[3] via a family of seminorms satisfying certain properties. In the case of $\mathscr{S}(\mathbb{R}^n)$, we consider

$$p_{\alpha,\,N}(f) := \sup_{x\in\mathbb{R}^n} \left|(1+|x|)^N \partial^\alpha f(x)\right|,$$

where $\alpha \in \mathbb{N}^n$ and $N \in \mathbb{N}$. The collection of seminorms $\mathcal{P} := \{p_{\alpha,\,N}\}_{\alpha\in\mathbb{N}^n,\, N\in\mathbb{N}}$ defines the Schwartz space as expected:

$$\mathscr{S}(\mathbb{R}^n) := \{f \in C^\infty(\mathbb{R}^n) \mid p_{\alpha,\,N}(f) < \infty \text{ for all } \alpha \in \mathbb{N}^n \text{ and all } N \in \mathbb{N}\}.$$

We will not prove it (see *Theorem 2.33* here) but $\mathcal{P}$ induces on $\mathscr{S}(\mathbb{R}^n)$ a structure of topological vector space $(\mathscr{S}(\mathbb{R}^n),\, \tau)$, where τ is a locally convex topology.

Lemma 1.36. *The inclusions*

$$(\mathcal{D}(\mathbb{R}^n),\, \tau_{\mathcal{D}}) \hookrightarrow (\mathscr{S}(\mathbb{R}^n),\, \tau) \hookrightarrow \left(L^1(\mathbb{R}^n),\, \|\cdot\|_1\right)$$

are continuous.

Proof. We divide the proof into two steps.

[3] The interested reader can find the theory of topological vector spaces in these lecture notes.

Step 1. Let f be a smooth compactly supported function. Then

$$p_{\alpha,\,N}(f) = \sup_{x \in \operatorname{spt}(f)} \left|(1+|x|)^N \partial^\alpha f(x)\right| \lesssim \|f\|_{(|\alpha|)}$$

for all $\alpha \in \mathbb{N}^n$ and $N \in \mathbb{N}$, which means that the inclusion $\mathcal{D}(\mathbb{R}^n) \hookrightarrow \mathscr{S}(\mathbb{R}^n)$ is continuous.

Step 2. Let $f \in \mathscr{S}(\mathbb{R}^n)$ be a rapidly decreasing function. Recall that

$$H(x) := \left(\frac{1}{1+|x|^2}\right)^n$$

belongs to $L^1(\mathbb{R}^n)$. To obtain a bound for the L^1-norm of f, we multiply and divide it by $H(x)$ as follows:

$$\|f\|_{L^1(\mathbb{R}^n)} = \int_{\mathbb{R}^n} |f(x)| \, \mathrm{d}x = \int_{\mathbb{R}^n} \frac{(1+|x|^2)^n \, |f(x)|}{(1+|x|^2)^n} \, \mathrm{d}x.$$

The numerator is uniformly bounded and the denominator is integrable. Thus, if we expand the binomial n-th power, then we find that

$$\|f\|_{L^1(\mathbb{R}^n)} \lesssim \left\| \frac{1}{(1+|x|^2)^n} \right\|_{L^1(\mathbb{R}^n)} \sum_{|\alpha| \leq 2n} p_{\alpha,\,0}(f) < \infty,$$

which gives us the desired result. $\square$

Remark 1.37. The inclusion $\mathcal{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$ is proper. Indeed, it is easy to see that the Gaussian function $\mathrm{e}^{-|x|^2}$ is smooth and rapidly decreasing while its support is not compact.

Theorem 1.38. *The Schwartz space $\mathscr{S}(\mathbb{R}^n)$ is a Fréchet space.*

Proof. See **Theorem 5.3** in the lecture notes of the course *Advanced Analysis.* $\square$

Notice that the collection of seminorms $\mathcal{P}$ introduced above can be replaced by a simpler one that depends on a unique parameter.

Exercise 1.4. Show that the collection of seminorms

$$\|f\|_{(N)} := \sup_{x \in \mathbb{R}^n} \sup_{|\alpha| \leq N} \left|(1+|x|)^N \partial^\alpha f(x)\right|$$

indexed by $N \in \mathbb{N}$ generates the same topology τ given by $\mathcal{P}$.

This characterization of the topology is useful because it allows us to give a notion of continuity that is easy to check each time. Indeed, a linear functional

$$\lambda : \mathscr{S}(\mathbb{R}^n) \longrightarrow \mathbb{C}$$

is continuous if and only if there are $N \in \mathbb{N}$ and $C > 0$ such that

$$|\lambda|(f) \leq C \|f\|_{(N)} \quad \text{for all } f \in \mathscr{S}(\mathbb{R}^n).$$

Similarly, a linear operator $T : \mathscr{S}(\mathbb{R}^n) \to B$, where $(B, \|\cdot\|_B)$ is a normed space, is continuous if and only if there are $N \in \mathbb{N}$ and $C > 0$ such that

$$\|Tf\|_B \le C\|f\|_{(N)} \quad \text{for all } f \in \mathscr{S}(\mathbb{R}^n).$$

This result is particularly important when B is $\mathscr{S}(\mathbb{R}^n)$ itself since it gives an easy criterion to check whether or not a linear operator is continuous.

Proposition 1.39. *A linear operator $T : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is continuous if and only if for all $M \in \mathbb{N}$ we can find $N(M) \in \mathbb{N}$ and $C(M) > 0$ such that*

$$\|Tf\|_{(M)} \le C(M)\|f\|_{(N)} \quad \textit{for all } f \in \mathscr{S}(\mathbb{R}^n).$$

To conclude this section, we collect a couple of properties of the Fourier transform on Schwartz spaces. Notice that $\mathscr{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$, so $\mathscr{F}\big|_{\mathscr{S}(\mathbb{R}^n)}$ is well-defined.

Proposition 1.40. *Let $f \in \mathscr{S}(\mathbb{R}^n)$ and $\alpha \in \mathbb{N}^n$. Then*

$$\mathscr{F}\left(\partial^\alpha f\right)(\xi) = \imath^{|\alpha|}\xi^\alpha \hat{f}(\xi) \quad \textit{and} \quad \mathscr{F}\left(x^\alpha f\right)(\xi) = (-\imath)^{|\alpha|}\partial^\alpha \hat{f}(\xi).$$

Theorem 1.41. *The inclusion $\mathscr{F}(\mathscr{S}(\mathbb{R}^n)) \subseteq \mathscr{S}(\mathbb{R}^n)$ is continuous.*

Proof. Let $f \in \mathscr{S}(\mathbb{R}^n)$ and fix $M \in \mathbb{N}$. Then

$$\begin{aligned}
\|\mathscr{F}(f)\|_{(M)} &= \sup_{\xi\in\mathbb{R}^n} \sup_{|\alpha|\le M} \left|(1+|\xi|)^M \partial^\alpha \mathscr{F}(f)(\xi)\right| = \\
&= \sup_{\xi\in\mathbb{R}^n} \sup_{|\alpha|\le M} \left|(1+|\xi|)^M \mathscr{F}(x^\alpha f)(\xi)\right| = \\
&= \sum_{|\beta|\le M} c_\beta \sup_{\xi\in\mathbb{R}^n} \sup_{|\alpha|\le M} \left|\mathscr{F}(\partial^\beta x^\alpha f)(\xi)\right| \le \\
&\le \sum_{|\beta|\le M} c_\beta \sup_{x\in\mathbb{R}^n} \left|\partial^\beta (1+|x|)^M f(x)\right| \le \\
&\le c(M)\|f\|_{(N)},
\end{aligned}$$

and this is enough to conclude that the inclusion is continuous. □

1.5 Inverse Fourier transform

The numerous properties of the Fourier operator may be exploited, for example, to find solutions to certain PDEs. However, even if we were able to find u such that

$$\mathscr{F}(P(D)u) = \mathscr{F}(f),$$

where $P(D)$ is a linear differential operator, at this point nothing guarantees us that we can recover u from $\mathscr{F}(u)$. The goal of this section is to show that

$$\mathscr{F}^{-1} : \mathscr{S}(\mathbb{R}^n) \longrightarrow \mathscr{S}(\mathbb{R}^n)$$

is well-defined and can be represented by an integral formula. We will also show that this phenomenon also happens in a more general framework, that of *tempered distributions.*

1.5.1 Inverse operator on Schwartz spaces

Recall that $\mathscr{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is well-defined. We know already that the operator is an isometry (up to the 2π constant) so it suffices to show that $\mathscr{F}$ is onto to conclude that

$$\mathscr{F}^{-1} : L^2(\mathbb{R}^n) \longrightarrow L^2(\mathbb{R}^n)$$

is well-defined and continuous. For this, first recall that

$$2\pi \langle f, g \rangle_2 = \langle \mathscr{F}(f), \mathscr{F}(g) \rangle_2$$

as a consequence of the polarisation formula applied to Plancherel identity (1.21). Taking the adjoint operator on the right-hand side leads to

$$2\pi \langle f, g \rangle_2 = \langle \mathscr{F}^* \mathscr{F} f, g \rangle_2 ,$$

which is equivalent to the operator identity

$$\frac{1}{2\pi} \mathscr{F}^* \mathscr{F} = \mathrm{Id}_{L^2(\mathbb{R}^n)}.$$

This shows that the Fourier transform on $L^2(\mathbb{R}^n)$ is invertible and there results

$$\mathscr{F}^{-1} = \frac{1}{2\pi} \mathscr{F}^*.$$

Since $\mathscr{S}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$ we know that the Fourier transform is invertible on the Schwartz space, but we do not know yet whether or not

$$\mathscr{F}^{-1}(\mathscr{S}(\mathbb{R}^n)) \subseteq \mathscr{S}(\mathbb{R}^n).$$

Remark 1.42. Suppose that we have an integral operator T with kernel K, namely

$$Tf(\xi) = \int_{\mathbb{R}^n} f(x) K(x, \xi) \,\mathrm{d}x.$$

Then a standard result in harmonic analysis asserts that the adjoint of T is given by

$$T^* f(x) = \int_{\mathbb{R}^n} f(\xi) \overline{K(\xi, x)} \,\mathrm{d}\xi.$$

We can apply this general result to our case ($T = \mathscr{F}$), with kernel which is given by the complex exponential

$$K(x, \xi) = \mathrm{e}^{-\imath x \cdot \xi}.$$

This quantity is invariant under the transformation $(x, \xi) \mapsto (\xi, x)$, and hence

$$\mathscr{F}^{-1} f(x) = \frac{1}{2\pi} \int_{\mathbb{R}^n} f(\xi) \mathrm{e}^{\imath x \cdot \xi} \, \mathrm{d}\xi \tag{1.24}$$

is the *inverse Fourier transform.* A straightforward consequence is the *inversion formula* which asserts that

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}^n} \hat{f}(\xi) \mathrm{e}^{\imath x \cdot \xi} \, \mathrm{d}\xi. \tag{1.25}$$

It is easy to verify that

$$f \in \mathscr{S}(\mathbb{R}^n) \implies \mathscr{F}^{-1} f \in \mathscr{S}(\mathbb{R}^n)$$

using (1.24); therefore

$$\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \longrightarrow \mathscr{S}(\mathbb{R}^n)$$

is an isomorphism of Fréchet spaces. We summarize the results obtained so far for the inverse operator in the following theorem:

Theorem 1.43 (Inversion Theorem)**.** *Let $f \in \mathscr{S}(\mathbb{R}^n)$ be a Schwartz function. Then*

$$\mathscr{F} \circ \mathscr{F}(f)(x) = \check{f}(x),$$

where $\check{f}(x) := f(-x)$.

Remark 1.44. The inclusion $\mathcal{D}(\mathbb{R}^n) \hookrightarrow \mathscr{S}(\mathbb{R}^n)$ implies that $\mathscr{F}(\mathcal{D}(\mathbb{R}^n)) \subset \mathscr{S}(\mathbb{R}^n)$. However, it is **not** true that

$$\mathscr{F}(\mathcal{D}(\mathbb{R}^n)) \subseteq \mathcal{D}(\mathbb{R}^n).$$

Indeed, the Fourier transform of a compactly supported smooth function has compact support if and only if it is *identically zero.*

1.6 Tempered Distributions

In this section, our goal is to introduce a suitable notion of *Fourier transform* which makes sense for a special class of distributions and is also compatible with the one introduced already for functions. The naive idea would be to set

$$\langle \mathscr{F}\Phi, f \rangle := \langle \Phi, \mathscr{F} f \rangle, \tag{1.26}$$

but it is easy to verify that for $\Phi \in \mathcal{D}'(\mathbb{R}^n)$ it might not make any sense. To avoid this issue we restrict ourselves to a special class of distributions, called *tempered distributions.*

Definition 1.45. A *tempered distribution* Φ is a linear continuous functional defined on the Schwartz space, that is,

$$\Phi : \mathscr{S}(\mathbb{R}^n) \longrightarrow \mathbb{C},$$

in which case we write $\Phi \in \mathscr{S}'(\mathbb{R}^n)$.

Remark 1.46. Notice that[4], if Φ is a tempered distribution, then the composition

$$\mathcal{D}(\mathbb{R}^n) \hookrightarrow \mathscr{S}(\mathbb{R}^n) \xrightarrow{\Phi} \mathbb{C}$$

[4]The inclusion $\mathcal{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$ is dense. We only proved that it is continuous, but the reader is encouraged to do it using a cutoff function and rescalings.

identifies Φ with an element of $\mathcal{D}'(\mathbb{R}^n)$. This shows that $\mathscr{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$.

Unless otherwise stated, we shall always endow $\mathscr{S}'(\mathbb{R}^n)$ with the weak topology in such a way that a linear map $T : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ is continuous if and only if

$$\mathscr{S}(\mathbb{R}^n) \ni f \longmapsto \langle Tf,\, g\rangle \in \mathbb{C}$$

is continuous for all $g \in \mathscr{S}(\mathbb{R}^n)$. Consequently, the bilinear mapping

$$B(f,\, g) := \langle Tf,\, g\rangle$$

is separately continuous on both variables, and thus it is jointly continuous. In other words, there are $N,\, M \in \mathbb{N}$ and $C > 0$ such that

$$|B(f,\, g)| \leq C\|f\|_{(N)}\|g\|_{(M)} \quad \text{for all } f,\, g \in \mathscr{S}(\mathbb{R}^n).$$

Theorem 1.47 (Bourbaki). *Let X be a complete metric space, and let Y and Z be topological vector spaces. If a bilinear mapping*

$$B : X \times Y \longrightarrow Z$$

is separately sequentially continuous, then B is also jointly sequentially continuous.

Example 1.48 (of tempered distributions)**.**

(1) Let μ be a measure defined on $\mathbb{R}^n$. If there exists $N \in \mathbb{N}$ such that

$$\int_{\mathbb{R}^n} \frac{1}{(1+|x|^2)^N}\, \mathrm{d}\mu(x) \leq C < \infty, \tag{1.27}$$

then $\Lambda_\mu \in \mathscr{S}'(\mathbb{R}^n)$ is a tempered distribution.

(2) Let $f \in L^p(\mathbb{R}^n)$ for some $1 \leq p \leq \infty$. Then $\Lambda_f \in \mathscr{S}'(\mathbb{R}^n)$.

(3) Let $\varphi(x) = x^\alpha$ be a monomial. Then $\Lambda_\varphi \in \mathscr{S}'(\mathbb{R}^n)$ for all $\alpha \in \mathbb{N}^n$.

We now want to generalise the multiplication rule (1.12) to hold for a generic tempered distribution $\Phi \in \mathscr{S}'(\mathbb{R}^n)$. The naive approach does not work since the implication

$$\Phi \in \mathscr{S}'(\mathbb{R}^n),\, f \in C^\infty(\mathbb{R}^n) \implies f\Phi \in \mathscr{S}'(\mathbb{R}^n)$$

is false, unless f satisfies some extra assumptions.

Definition 1.49. Let $f \in C^\infty(\mathbb{R}^n)$. We say that f is of *moderate growth* if for all $\alpha \in \mathbb{N}^n$ there exists $m_\alpha \in \mathbb{N}$ such that

$$|\partial^\alpha f(x)| \lesssim_\alpha (1+|x|)^{m_\alpha}. \tag{1.28}$$

Proposition 1.50. *Let $\Phi \in \mathscr{S}'(\mathbb{R}^n)$ and let f be a function of moderate growth and define the product with a distribution as*

$$\langle f\Phi,\, g\rangle := \langle \Phi,\, fg\rangle.$$

Then

$$f\Phi \in \mathscr{S}'(\mathbb{R}^n) \quad \textit{and} \quad \partial^\alpha \Phi \in \mathscr{S}'(\mathbb{R}^n)$$

for all $\alpha \in \mathbb{N}^n$.

Definition 1.51 (Convolution)**.** Let $\Phi \in \mathscr{S}'(\mathbb{R}^n)$ be a tempered distribution and let $f \in \mathscr{S}(\mathbb{R}^n)$. The convolution is the tempered distribution defined by setting

$$\langle f * \Phi,\, g\rangle := \langle \Phi,\, \check{f} * g\rangle. \tag{1.29}$$

This definition is coherent with the one we have already given in the case $\Phi \in \mathcal{D}'(\mathbb{R}^n)$, but we still need to check that

$$f,\, g \in \mathscr{S}(\mathbb{R}^n) \implies \check{f} * g \in \mathscr{S}(\mathbb{R}^n).$$

Lemma 1.52. *The convolution of two Schwartz functions is also a Schwartz function.*

Proof. See [11]. □

To conclude this section, notice that the expression in (1.29) does not really define a tempered distribution but rather a function. Namely,

$$f * \Phi(x) = \langle \Phi,\, \tau_x f\rangle \tag{1.30}$$

and, as the reader might check by herself, it has moderate growth. This observation comes into play when we define the convolution between tempered distributions. Indeed,

$$\langle \Phi * \Psi,\, f\rangle := \langle \Phi,\, \check{\Psi} * f\rangle$$

is well-defined and belongs to $\mathscr{S}'(\mathbb{R}^n)$ if, for example, one of them has compact support.

1.6.1 Fourier transform on tempered distributions

Let Φ be the distribution associated to a summable function $h \in L^1(\mathbb{R}^n)$. Namely,

$$\langle \Phi,\, f\rangle = \Lambda_h(f) := \langle h,\, f\rangle_2.$$

If we take the Fourier transform of h we obtain

$$\langle \mathscr{F}\Phi,\, f\rangle = \langle \mathscr{F}h\, f\rangle_2 = \langle h,\, \mathscr{F}f\rangle_2 = \langle \Phi,\, \mathscr{F}f\rangle,$$

so it makes sense to generalise the Fourier transform to tempered distributions $\Phi \in \mathscr{S}'(\mathbb{R}^n)$ by setting

$$\langle \mathscr{F}\Phi,\, f\rangle := \langle \Phi,\, \mathscr{F}f\rangle. \tag{1.31}$$

The functional $\mathscr{F}\Phi$ is obviously linear and continuous since it can be written as the composition of two continuous maps. More precisely, we have

$$\mathscr{F}\Phi = \Phi \circ \mathscr{F}.$$

Furthermore, it is possible to find an extension of the property (1.19) to (1.31). The idea is that using the bracket notation we obtain the identity

$$\langle \mathscr{F}(f * \Phi),\, g\rangle = \left\langle \Phi,\, \check{f} * \mathscr{F}g \right\rangle.$$

The function on the right-hand side can be rewritten as the Fourier transform of a product using Fubini-Tonelli's theorem:

$$\begin{aligned}\check{f} * \mathscr{F}g(x) &= \int_{\mathbb{R}^n} f(y-x) \int_{\mathbb{R}^n} g(\xi) \mathrm{e}^{-\imath y\cdot\xi}\, \mathrm{d}\xi \mathrm{d}y = \\ &= \int_{\mathbb{R}^n} g(\xi) \mathrm{e}^{-\imath x\cdot\xi} \int_{\mathbb{R}^n} f(y-x) \mathrm{e}^{-\imath (y-x)\cdot\xi}\, \mathrm{d}y = \\ &= \mathscr{F}[g\mathscr{F}(f)].\end{aligned}$$

It follows that

$$\mathscr{F}(f * \Phi) = \mathscr{F}f\mathscr{F}\Phi,$$

which is the equivalent of (1.19) for $\mathscr{S}'(\mathbb{R}^n)$ and $\mathscr{S}(\mathbb{R}^n)$.

Remark 1.53. The inclusion $\mathcal{D}'(\mathbb{R}^n) \supset \mathscr{S}'(\mathbb{R}^n)$ is proper. For example, consider

$$\Phi_0 := \sum_{n=0}^{+\infty} \delta_n^{(n)}$$

and notice that Φ_0 is a distribution (that is, Φ belongs to $\mathcal{D}'(\mathbb{R}^n)$). On the other hand, it is easy to see that

$$\Phi_0(f) = \sum_{n=0}^{+\infty} (-1)^n f^{(n)}(n)$$

makes sense when f has compact support, but it might not even when $f \in \mathscr{S}'(\mathbb{R}^n)$.

The reason behind all this is that all $\Psi \in \mathscr{S}'(\mathbb{R}^n)$ have **finite order** while Φ_0 has infinite order (as many other distributions in $\mathcal{D}'(\mathbb{R}^n)$).

1.7 Approximate identity

A natural question arising from the theory developed above is whether or not it is possible to find a function - or a distribution - u on $\mathbb{R}^n$ that is the identity element for the convolution, that is,

$$u * f = f \quad \text{for all } f \in C_c^\infty(\mathbb{R}^n).$$

Recall that the space of all distributions $\mathcal{D}'(\mathbb{R}^n)$ is closed under convolution. Thus, taking the Dirac delta at the origin as u, we find that that

$$\langle \Phi * \delta_0,\, f\rangle = \langle \Phi,\, \delta_0 * f\rangle = \langle \Phi,\, f\rangle \implies \Phi * \delta_0 = \Phi$$

holds for all $\Phi \in \mathcal{D}'(\mathbb{R}^n)$. In particular,

$$(\mathcal{D}'(\mathbb{R}^n),\, *)$$

is a unitary algebra. On the other hand, it is not possible to find $u \in L^1(\mathbb{R}^n)$ such that

$$u * f = f \quad \text{for all } f \in L^1(\mathbb{R}^n).$$

Therefore, we can formulate a slightly weaker version of the question above. Namely, if there exist a sequence $\{\varphi_n\}_{n\in\mathbb{N}}$ (or a collection $\{\varphi_t\}_{t>0}$) of summable functions such that

$$\varphi_n * f \xrightarrow{n\to+\infty} f$$

with respect to an adequate notion of convergence.

Definition 1.54 (Mollifier)**.** Let $\varphi \in L^1(\mathbb{R}^n)$ be a nonnegative function with

$$\|\varphi\|_1 = \int_{\mathbb{R}^n} \varphi(x)\,\mathrm{d}x = 1.$$

For $t > 0$ define the rescaling

$$\varphi_t(x) := t^{-n}\varphi\left(\frac{x}{t}\right).$$

Then $\{\varphi_t\}_{t>0}$ is called a family of *mollifiers.*

Remark 1.55. If $\{\varphi_t\}_{t>0}$ is a collection of mollifiers, it is not hard to see that

$$\|\varphi_t\|_1 = 1 \quad \text{for all } t > 0.$$

Lemma 1.56. *Let $\Phi \in \mathcal{D}'(\mathbb{R}^n)$ and let $f \in \mathcal{D}(\mathbb{R}^n)$. The support of the convolution is contained in the sum of the supports, that is,*

$$\operatorname{spt}(\Phi * f) \subseteq \operatorname{spt}(\Phi) + \operatorname{spt}(f).$$

Theorem 1.57. *Let $\{\varphi_t\}_{t>0}$ be a family of mollifiers. Then for all $f \in L^1(\mathbb{R}^n)$ we have*

$$\|\varphi_t * f - f\|_{L^1(\mathbb{R}^n)} \xrightarrow{t\to 0^+} 0.$$

The proof of this result is moderately straightforward, but we first need to state and prove a technical lemma which asserts that the map

$$\mathbb{R}^n \ni h \longmapsto \tau_h f \in L^1(\mathbb{R}^n),$$

is continuous for all $f \in L^1(\mathbb{R}^n)$ fixed, where $\tau_h f(x) = f(x-h)$.

Lemma 1.58. *Let $f \in L^1(\mathbb{R}^n)$. Then*

$$\|f - \tau_h f\|_{L^1(\mathbb{R}^n)} \xrightarrow{h\to 0} 0.$$

Proof. First, assume that $f \in C_c^\infty(\mathbb{R}^n)$. Then f is uniformly continuous, which means that for every $\epsilon > 0$ we can find a uniform $\delta > 0$ such that

$$|x - x'| < \delta \implies |f(x) - f(x')| < \epsilon.$$

Let K be the support of f. Then $K+h$ is the support of $\tau_h f$, and thus

$$\|f-\tau_h f\|_{L^1(\mathbb{R}^n)} = \int_{K\cup(K+h)} |f(x-h)-f(x)|\,\mathrm{d}x.$$

By uniform continuity of f, provided that $|h|$ is small enough, we readily obtain the estimate

$$\|f-\tau_h f\|_{L^1(\mathbb{R}^n)} \le 2\epsilon|K| \lesssim \epsilon.$$

Now suppose that $f\in L^1(\mathbb{R}^n)$. By density we can find $g_\epsilon \in C_c^\infty(\mathbb{R}^n)$ such that

$$\|g_\epsilon - f\|_{L^1(\mathbb{R}^n)} \le \frac{\epsilon}{3}.$$

Finally, the triangular inequality implies that

$$\|f-\tau_h f\|_{L^1(\mathbb{R}^n)} \le \|f-g_\epsilon\|_{L^1(\mathbb{R}^n)} + \|g_\epsilon - \tau_h g_\epsilon\|_{L^1(\mathbb{R}^n)} + \|\tau_h(g_\epsilon - f)\|_{L^1(\mathbb{R}^n)} \le \epsilon,$$

and this concludes the proof. □

Proof of Theorem 1.57. First, we evaluate the absolute value of the difference:

$$\begin{aligned}
|\varphi_t * f(x) - f(x)| &= \left|\int_{\mathbb{R}^n} \varphi_t(y) f(x-y)\,\mathrm{d}y - f(x)\right| = \\
&= \left|\int_{\mathbb{R}^n} \varphi_t(y) f(x-y)\,\mathrm{d}y - \int_{\mathbb{R}^n} f(x)\varphi_t(y)\,\mathrm{d}y\right| \le \\
&\le \int_{\mathbb{R}^n} |\tau_y f(x) - f(x)||\varphi_t(y)|\,\mathrm{d}y.
\end{aligned}$$

Next, we take the L^1-norm of the left-hand side. It turns out that

$$\begin{aligned}
\|\varphi_t * f(x) - f(x)\|_{L^1(\mathbb{R}^n)} &\le \int_{\mathbb{R}^n} \mathrm{d}x \int_{\mathbb{R}^n} \mathrm{d}y\, |\tau_y f(x) - f(x)|\, |\varphi_t(y)| = \\
&= \int_{\mathbb{R}^n} \mathrm{d}x \int_{\mathbb{R}^n} \mathrm{d}y'\, |\tau_{ty'} f(x) - f(x)|\, |\varphi(y')| = \\
&= \int_{\mathbb{R}^n} \mathrm{d}y' |\varphi(y')| \int_{\mathbb{R}^n} \mathrm{d}x\, |\tau_{ty'} f(x) - f(x)| = \\
&= \int_{\mathbb{R}^n} \|\tau_{ty'} f - f\|_{L^1(\mathbb{R}^n)} |\varphi(y')|\,\mathrm{d}y'.
\end{aligned}$$

For $t\to 0$ we can apply the continuity lemma (proved above) and the Lebesgue's dominated convergence theorem to infer that the right-hand side goes to zero. □

Remark 1.59. The red equality follows from the change of variables $y \mapsto ty'$, while the orange one from the *Fubini-Tonelli's theorem.*

This approximation via mollifiers is not a peculiarity of $L^1(\mathbb{R}^n)$ but, as we can see from the proof, it is true for all Banach spaces $\mathfrak{X}$ satisfying specific properties.

Theorem 1.60. *Let $\{\varphi_t\}_{t>0}$ be a family of mollifiers and let $\mathfrak{X}$ be a Banach space of measurable real-valued functions such that the following properties hold:*

- *($\star$) If $f \in \mathfrak{X}$, then $\tau_h f \in \mathfrak{X}$ and $\|f\|_{\mathfrak{X}} = \|\tau_h f\|_{\mathfrak{X}}$ for all $h \in \mathbb{R}^n$.*
- *($\star$) The mapping $h \longmapsto \tau_h f$ is continuous for all $f \in \mathfrak{X}$.*

Then

$$\|\varphi_t * f - f\|_{\mathfrak{X}} \xrightarrow{t\to 0^+} 0.$$

Remark 1.61. We can apply this theorem with $\mathfrak{X} = L^p(\mathbb{R}^n)$ for all $p < \infty$. However, in the case with $p = \infty$ it is easy to verify that the mapping

$$h \longmapsto \tau_h f$$

is not continuous for all $f \in L^\infty(\mathbb{R}^n)$. It works, though, if we replace L^∞ with the space of vanishing functions $C_0(\mathbb{R}^n)$, which is equal to the closure of $C_c^\infty(\mathbb{R}^n)$ with respect to $\|\cdot\|_\infty$.

Nevertheless, we can recover the result in $L^\infty(\mathbb{R}^n)$ if we require $\varphi_t * f$ to converge to f with respect to the weak-$*$ topology induced by

$$L^\infty(\mathbb{R}^n) = \left[L^1(\mathbb{R}^n)\right]'.$$

Theorem 1.62. *Let $\{\varphi_t\}_{t>0}$ be a family of mollifiers. Then for all $f \in L^\infty(\mathbb{R}^n)$ we have*

$$\varphi_t * f \xrightarrow{\text{weak-}*\text{ topology}} f.$$

Theorem 1.63. *Let $\{\varphi_t\}_{t>0}$ be a family of mollifiers. Then for all $\mu \in \mathcal{M}(\mathbb{R}^n)$ we have*

$$\varphi_t * \mu \xrightarrow{\text{weak-}*\text{ topology}} \mu.$$

Here $\mathcal{M}(\mathbb{R}^n)$ denotes the set of all $\mathbb{R}^n$-valued finite Borel measures.

The next step would be to extend this approximation result to $\mathscr{S}(\mathbb{R}^n)$. Unfortunately, it is rather easy to verify that

$$f \in \mathscr{S}(\mathbb{R}^n) \quad \text{and} \quad \varphi \in C_c^\infty(\mathbb{R}^n) \implies \varphi * f \in C^\infty(\mathbb{R}^n),$$

but, a priori, there is no guarantee that it also belongs to $\mathscr{S}(\mathbb{R}^n)$ and, in general, it will not be. Nonetheless, we can prove the following result with little effort:

Theorem 1.64. *Let $\varphi \in \mathscr{S}(\mathbb{R}^n)$ and let $\{\varphi_t\}_{t>0}$ be the associated mollifiers. Then for all $f \in \mathscr{S}(\mathbb{R}^n)$ it turns out that*

$$\varphi_t * f \xrightarrow{\mathscr{S}(\mathbb{R}^n)} f.$$

In other words, for all $N \in \mathbb{N}$ it turns out that

$$\|\varphi_t * f - f\|_{(N)} \xrightarrow{t\to 0^+} 0.$$

Corollary 1.65. *Let $\Phi \in \mathscr{S}'(\mathbb{R}^n)$ and let $\{\varphi_t\}_{t>0}$ be a Schwartz mollifiers collection. Then*

$$\varphi_t * \Phi \longrightarrow \Phi$$

with respect to the weak topology of $\mathscr{S}'(\mathbb{R}^n)$.

1.8 Paley-Wiener theorem

The primary goal of this section is to show that the Fourier transform is the restriction on the real line $\mathbb{R}^n$ of a complex-valued operator which, under mild assumptions, sends compactly supported distributions to entire functions on the whole $\mathbb{C}^n$.

1.8.1 Introduction to complex analysis

We first recollect some basic definitions in complex analysis. Since we are mainly interested in the notion of **entire** function, we will focus on being holomorphic and on the identity principle.

Definition 1.66. Let $\Omega \subseteq \mathbb{C}^n$ be an open subset of the complex n-plane.

(1) A function $f_i : \Omega_i \subset \mathbb{C} \to \mathbb{C}$ is *holomorphic* on Ω_i if it is *complex differentiable* at all $z_0 \in \Omega_i$. In other words,
$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$
exists at all $z_0 \in \Omega_i$.

(2) A function $f : \Omega \to \mathbb{C}$ is *holomorphic* if it is continuous at every point of Ω, and the coordinate functions
$$\Omega_i \ni z_i \longmapsto f(z_1, \ldots, z_i, \ldots, z_n)$$
are holomorphic in the sense of (**1**) for all $i = 1, \ldots, n$.

(3) A function $f : \mathbb{C}^n \to \mathbb{C}$ is *entire* if it is holomorphic at every point $z \in \mathbb{C}^n$.

Lemma 1.67 (Identity Principle)**.** *Let $f : \mathbb{C}^n \to \mathbb{C}$ be an entire function which is identically zero on the real line, that is,*

$$f(z) = 0 \quad \text{for all } z = x + \imath y \text{ with } y = 0.$$

Then f is the function identically zero on $\mathbb{C}^n$.

Proof. Consider the kth predicate

$$P_k : f(z) = 0 \text{ for every } z = (z_1, \ldots, z_n) \in \mathbb{C}^n \text{ such that } z_1, \ldots, z_k \in \mathbb{R}.$$

By assumption P_n is true. Therefore, it is enough to prove that

$$P_k \implies P_{k-1}$$

to conclude since n is finite. But this is trivially true because

$$\mathbb{C} \ni z_k \longmapsto f(z_1, \ldots, z_k, \ldots, z_n) \in \mathbb{C}$$

is a complex function of a single variable, and we can hence apply the identity theorem[5]. $\square$

Theorem 1.68 (Moreira). *Let $f : \Omega \subseteq \mathbb{C} \to \mathbb{C}$ be a continuous function defined on a open subset of the complex plane satisfying*

$$\oint_\gamma f(z)\,\mathrm{d}z = 0$$

for every closed piecewise C^1-curve γ in Ω. Then f is holomorphic on Ω.

Theorem 1.69 (Cauchy). *Let Ω be an open subset of $\mathbb{C}$ which is simply connected. Let $f : \Omega \to \mathbb{C}$ be a holomorphic function, and let γ be a rectifiable closed path in Ω. Then*

$$\oint_\gamma f(z)\,\mathrm{d}z = 0.$$

1.8.2 Paley-Wiener theorem

We are now ready to state and prove the main result of this section, starting from the case in which f is a compactly supported function.

Theorem 1.70 (Paley-Wiener). *Let $f \in \mathcal{D}(\mathbb{R}^n)$ be a function with $\mathrm{spt}(f) \subseteq \bar{B}_r$. Then*

$$F(\zeta) := \int_{\mathbb{R}^n} f(\xi)\mathrm{e}^{-\imath\xi\cdot\zeta}\,\mathrm{d}\xi \tag{1.32}$$

is an entire function which satisfies the estimate

$$|\partial^\alpha F(\zeta)| \lesssim_{\alpha,\,N} (1+|\zeta|)^{-N}\mathrm{e}^{r|\mathfrak{Im}(\zeta)|} \tag{1.33}$$

for all $\zeta \in \mathbb{C}^n$, $\alpha \in \mathbb{N}^n$ and $N \in \mathbb{N}$.

Vice versa, if F is an entire function on $\mathbb{C}^n$ satisfying the estimate (1.33), then we can find a smooth function $f \in \mathcal{D}(\mathbb{R}^n)$ such that

$$\mathrm{spt}(f) \subseteq \bar{B}_r.$$

Furthermore, F is exactly given by the formula (1.32).

Proof. The argument is moderately involved. Hence we break down the proof into several small steps, to ease the notation for the reader.

Step 1. The function given by (1.32) is well-defined and continuous, as a consequence of the dominated convergence theorem. Moreover, the coordinate function

$$F_k : \mathbb{C} \ni \zeta_k \longmapsto F(\zeta_1, \ldots, \zeta_k, \ldots, \zeta_n) \in \mathbb{C}$$

[5]**Identity Theorem.** Let $g,\, h : D \to \mathbb{C}$ be complex functions defined on a connected open set $D \subseteq \mathbb{C}$. If $f(x) = g(x)$ for every $x \in S$, where S is a nonempty open subset of D, then $f(x) = g(x)$ for every $x \in D$.

is holomorphic for any $k = 1, \dots, n$ as a simple application of Moreira's theorem. Indeed, given a piecewise differentiable closed curve γ, it is easy to verify that

$$\oint_\gamma \left[\zeta_k \mapsto \int_{\mathbb{R}^n} f(\xi)\mathrm{e}^{-\imath\xi\cdot\zeta}\,\mathrm{d}\xi\right]\mathrm{d}\zeta_k = \int_{\mathbb{R}^n}\left[\oint_\gamma \left(\zeta_k \mapsto \mathrm{e}^{-\imath\xi\cdot\zeta}\right)\mathrm{d}\zeta_k\right]\mathrm{d}\xi,$$

where the latter is equal to 0, using Fubini-Tonelli's theorem together with the fact that the complex exponential is a holomorphic function.

Step 2. We now want to show that the estimate (1.33) holds. Fix two indexes $\alpha, \beta \in \mathbb{N}^n$ and notice that

$$\begin{aligned}\left|\zeta^\beta \partial_\zeta^\alpha F(\zeta)\right| &= \left|\zeta^\beta \int_{\mathbb{R}^n} x^\alpha f(x)\mathrm{e}^{-\imath x\cdot\zeta}\,\mathrm{d}x\right| = \\ &= \left|\int_{\mathbb{R}^n} x^\alpha f(x)\partial_x^\beta(\mathrm{e}^{-\imath x\cdot\zeta})\,\mathrm{d}x\right| = \\ &= \left|\int_{\mathbb{R}^n} \partial_x^\beta(x^\alpha f(x))\mathrm{e}^{-\imath x\cdot\zeta}\,\mathrm{d}x\right|.\end{aligned}$$

Identify $\mathbb{C}^n$ with $\mathbb{R}^n \times \mathbb{R}^n$, write $\zeta = \xi + \imath\eta$ and notice that

$$\left|\mathrm{e}^{-\imath x\cdot\zeta}\right| = \mathrm{e}^{x\cdot\eta}.$$

It follows that

$$\begin{aligned}\left|\zeta^\beta \partial_\zeta^\alpha F(\zeta)\right| &\le \int_{\mathbb{R}^n} |\partial_x^\beta(x^\alpha f(x))||\mathrm{e}^{-\imath x\cdot\zeta}|\,\mathrm{d}x \le \\ &\le \|\partial_x^\beta(x^\alpha f(x))\|_{L^\infty(\mathbb{R}^n)}\mathrm{e}^{r|\mathfrak{Im}(\zeta)|}.\end{aligned}$$

It remains to estimate the L^∞-norm of $\partial_x^\beta(x^\alpha f(x))$, but this is an immediate consequence of the following (up to constant) identity

$$(1+|\zeta|)^N|\partial_\zeta^\alpha F(\zeta)| \simeq \sum_{|\beta|\le N}\left||\zeta|^\beta \partial^\alpha F(\zeta)\right|,$$

which allows us to obtain (1.32) with a precise value for the constant:

$$\left|(1+|\zeta|)^N \partial_\zeta^\alpha F(\zeta)\right| \lesssim \sum_{|\beta|\le N}\left[\|\partial_x^\beta(x^\alpha f(x))\|_\infty\right]\mathrm{e}^{r|\mathfrak{Im}(\zeta)|}.$$

Step 3. Vice versa, let F be the entire function defined on $\mathbb{C}^n$ satisfying (1.33) and denote by $\hat{f}$ the restriction of F to the real line, that is,

$$\hat{f} : \mathbb{R}^n \ni x \longmapsto F(x + \imath 0) \in \mathbb{C}.$$

Suppose for the sake of simplicity that $n = 1$. The estimate (1.33), applied to the function $\hat{f}$, shows immediately that $\hat{f}$ belongs to $\mathscr{S}(\mathbb{R}^n)$. It follows that

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(\xi)\mathrm{e}^{\imath\xi\cdot x}\,\mathrm{d}\xi$$

is the unique possible choice for the function f since the Fourier transform is an invertible operator restricted to the Schwartz space.

Step 4. We now claim that we can equivalently evaluate f at x by integrating F on any line parallel to the x-axis. In other words, we have

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} F(\xi + \imath\eta) e^{\imath x \dot{(}\xi + \imath\eta)} \, d\xi \tag{1.34}$$

for all $\eta \in \mathbb{R}$. This will allow us to extend the function f to the whole $\mathbb{C}$ since it only depends on the values attained on the real line.

Proof. Let γ_R be the curve defined as in Figure 1.1. The integrand is holomorphic and γ_R is rectifiable and closed so, using Cauchy's theorem, we find that

$$0 = \int_{-R}^{R} F(\xi) e^{\imath x \cdot \xi} \, d\xi - \int_{-R}^{R} F(\xi + \imath\eta) e^{\imath x \cdot (\xi + \imath\eta)} \, d\xi + \dots$$

$$\dots + \int_{0}^{\eta} f(R + \imath\sigma) e^{\imath x \cdot (R + \imath\sigma)} \, d\sigma - \int_{0}^{\eta} F(-R + \imath\sigma) e^{\imath x \cdot (-R + \imath\sigma)} \, d\sigma.$$

The vertical terms go to zero as $R \to +\infty$ as a consequence of the estimate (1.33). Therefore

$$\int_{-\infty}^{\infty} F(\xi) e^{\imath x \cdot \xi} \, d\xi - \int_{-\infty}^{\infty} F(\xi + \imath\eta) e^{\imath x \cdot (\xi + \imath\eta)} \, d\xi = 0,$$

and this proves the claim. □

Step 5. To conclude, we need to show that the support of f is contained in $\bar{B}_r$. Let x be a real number such that $|x| > r$, and notice that

$$|f(x)| \leq \frac{1}{2\pi} \int_{\mathbb{R}} |F(\xi + \imath\eta)| \left| e^{\imath x \cdot (\xi + \imath\eta)} \right| \, d\xi \lesssim_n$$

$$\lesssim_n e^{-x \cdot \eta} \int_{\mathbb{R}} |F(\xi + \imath\eta)| \, d\xi \lesssim_{\alpha, n}$$

$$\lesssim_{\alpha, n} e^{r|\eta| - x \cdot \eta}.$$

If we choose $\eta := \lambda \frac{x}{|x|}$, then

$$\lambda(r - |x|) < 0 \implies \lim_{\lambda \to +\infty} e^{r|\eta| - x \cdot \eta} = -\infty,$$

which means that $f(x) = 0$ for all $|x| > r$.

Step 6. The final assertion, that the estimate (1.32) holds true for any $z \in \mathbb{C}$, follows trivially from Lemma 1.67 and the fact that it holds already at all points $x \in \mathbb{R}$. □

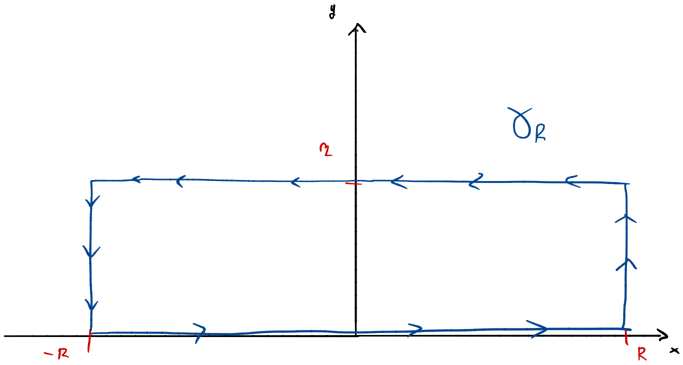

Figure 1.1: A picture of the integral path γ_R.

Theorem 1.71 (Paley-Wiener, II)**.** *Let $\Phi \in \mathcal{D}'(\mathbb{R}^n)$ be a distribution with compact support contained in the ball $\bar{B}_r$. Then the function*

$$F(\zeta) := \Phi\left(\mathrm{e}_{-\imath\zeta}\right), \tag{1.35}$$

is entire on $\mathbb{C}^n$ and satisfies the estimate

$$|F(\zeta)| \lesssim (1+|\zeta|)^N \, \mathrm{e}^{r|\mathfrak{Im}(\zeta)|}, \tag{1.36}$$

for all $\zeta \in \mathbb{C}^n$ and $\alpha \in \mathbb{N}^n$, where N is the order of the distribution Φ.

Vice versa, if F is an entire function that satisfies the estimate (1.36), then there exists a compactly supported distribution $\Phi \in \mathcal{D}'(\mathbb{R}^n)$ with

$$\operatorname{spt}(\Phi) \subset \bar{B}_r.$$

Moreover, the function F is given by the formula (1.35).

Remark 1.72. The estimate (1.33), for the vice versa, can be weakened to only require

$$|F(\zeta)| \lesssim_N (1+|\zeta|)^{-N} \mathrm{e}^{r|\mathfrak{Im}(\zeta)|} \tag{1.37}$$

for all $\zeta \in \mathbb{C}^n$. The reason is that the α-derivative of F at any point is given by Cauchy's formula as follows:

$$\partial^\alpha F(0) = \frac{\alpha!}{(2\pi\imath)^n} \oint_{\gamma_\rho^1} \cdots \oint_{\gamma_\rho^n} \frac{F(\zeta_1, \ldots, \zeta_n)}{\zeta_1^{\alpha_1+1} \ldots \zeta_n^{\alpha_n+1}} \, \mathrm{d}\zeta,$$

where γ_ρ^j is a curve parameterising the circumference of radius ρ in the jth $\mathbb{C}$. We can hence estimate the absolute value of $\partial^\alpha F(0)$:

$$|\partial^\alpha F(0)| \leq c_\alpha \rho^n \max_{|\zeta_1| = \cdots = |\zeta_n| = \rho} |F(\zeta_1, \ldots, \zeta_n)| \rho^{-|\alpha|-n} \leq$$

$$\leq c_\alpha \rho^{-|\alpha|} (1+\rho)^N.$$

Since ρ is arbitrary, we may send it to $+\infty$ and find that $\partial^\alpha F(0) = 0$ for all multi-indexes α of length strictly greater than N. This means that

$$F(\zeta) = \sum_{|\alpha|\leq N} \frac{1}{\alpha!}\partial^\alpha F(0)\zeta^\alpha,$$

and thus F is given by a finite sum of derivatives. In particular, the estimate (1.37) is equivalent to the apparently stronger one (1.33).

Corollary 1.73. *Let $\Phi \in \mathscr{S}'(\mathbb{R}^n)$. Then*

$$\operatorname{spt}(\Phi) \subseteq \{0\} \iff \mathscr{F}\Phi \textit{ is a polynomial.}$$

Proof. First, notice that if Φ is the Dirac delta δ_0 centred at the origin, then

$$\langle \mathscr{F}\delta_0,\, f\rangle = \mathscr{F}f(0) = \int_{\mathbb{R}^n} f(x)\,\mathrm{d}x = \langle 1,\, f\rangle_2.$$

It follows that $\mathscr{F}\delta_0$ is equal, in a distributional sense, to the constant polynomial 1. Moreover, the α-derivative of the Dirac delta satisfies the identity

$$\langle \partial^\alpha\delta_0,\, f\rangle = (-1)^{|\alpha|}\langle \delta_0,\, \partial^\alpha f\rangle = (-1)^{|\alpha|}\partial f(0),$$

and thus any linear combination of derivatives of Dirac deltas $\partial^\alpha\delta_0$ has support contained in $\{0\}$. On the other hand, we have

$$\langle \mathscr{F}(\partial^\alpha\delta_0),\, f\rangle = (-1)^{|\alpha|}\partial_\xi^\alpha\mathscr{F}(f)(0),$$

so, if $\mathscr{F}\Phi$ is a polynomial, then Φ is a combination of $\partial^\alpha\delta_0$ since

$$\mathscr{F}(\partial^\alpha\delta_0)(\xi) = (\imath\xi)^\alpha = \imath^{|\alpha|}\xi_1^{\alpha_1}\dots\xi_n^{\alpha_n}.$$

Vice versa, if $\Phi \in \mathscr{S}'(\mathbb{R}^n)$ has support contained in the singlet $\{0\}$, then Paley-Wiener theorem shows that $\mathscr{F}\Phi$ must be a polynomial, and therefore

$$\Phi = \sum_{|\alpha|\leq N} c_\alpha\partial^\alpha\delta_0.$$

□

Chapter 2

Linear Differential Operators

In the first half of this chapter, we will focus on linear operators defined on $\mathscr{S}(\mathbb{R}^n)$ and taking values in the space of tempered distributions. More precisely, given

$$T : \mathscr{S}(\mathbb{R}^n) \longrightarrow \mathscr{S}'(\mathbb{R}^n),$$

we will show that the following characterisation holds:

Theorem A. Suppose that T is a continuous and translation-invariant. Then there exists a unique $\Psi \in \mathscr{S}'(\mathbb{R}^n)$ such that

$$Tf = \Psi * f \quad \text{for all } f \in \mathscr{S}(\mathbb{R}^n).$$

Next, we consider translation-invariant operators between L^p spaces, show that they identify a particular class $C_{p,\,q}$, which can be connected with known spaces for specific values of p and q. For example, we shall prove the following result:

Theorem B. The space $C_{2,\,2}$ is isometrically isomorphic to $\mathscr{F}(L^\infty)$.

In the last part of the chapter, we discuss a completely different topic: linear differential operators and fundamental solutions. Recall that

$$Lf = \sum_{|\alpha| \le m} c_\alpha \partial^\alpha f,$$

where c_α are functions, eventually constant, is a linear differential operator. Our goal is to prove a result of the utmost importance, which is due to Malgrange and Ehrenpreis.

Theorem C. Let L be a linear differential operator with constant coefficients. Then there exists $\psi \in \mathcal{D}'(\mathbb{R}^n)$ such that

$$L\psi = \delta_0.$$

2.1 Distributional kernels

Recall that an operator $T : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ is continuous if the duality coupling

$$f \longmapsto \langle Tf, g\rangle \tag{2.1}$$

is a continuous linear functional for all $g \in \mathscr{S}(\mathbb{R}^n)$. We already mentioned that (2.1) can be identified to the following bilinear form

$$(f, g) \longmapsto \langle Tf, g\rangle,$$

which is separately continuous, provided that T is continuous. By **Banach-Alaoglu**, it is also jointly continuous and thus we can find $N \in \mathbb{N}$ and $C > 0$ such that

$$|\langle Tf, g\rangle| \le C\|f\|_{(N)}\|g\|_{(N)}.$$

Example 2.1. Let us consider an integral operator with kernel k, namely

$$Tf(x) := \int_{\mathbb{R}^n} k(x, y)f(y)\,\mathrm{d}y, \tag{2.2}$$

and assume that k is continuous and supported on a compact set K. The associated linear functional (2.1) is given by

$$\langle Tf, g\rangle = \int_{\mathbb{R}^n} g(x) \int_{\mathbb{R}^n} k(x, y)f(y)\,\mathrm{d}y\mathrm{d}x.$$

Now introduce the symbol $g\otimes f(x, y)$ to indicate the product $f(y)g(x)$. Then we can rewrite the linear functional as follows:

$$\langle Tf,\, g\rangle = \iint_{\mathbb{R}^n\times\mathbb{R}^n} k(x,\, y)(g\otimes f)(x,\, y)\,\mathrm{d}y\mathrm{d}x.$$

If f and g are Schwartz functions we easily infer that T is a continuous operator since

$$|\langle Tf,\, g\rangle| \le \|k\|_{L^1(\Omega)}\|f\|_{(0)}\|g\|_{(0)},$$

and $\|k\|_{L^1(\Omega)}$ is finite because we assumed k to be (more than) summable.

Now let $\Phi \in \mathcal{S}'(\mathbb{R}^m_x \times \mathbb{R}^n_y)$ be a tempered distribution, $f \in \mathcal{S}(\mathbb{R}^m)$ and $g \in \mathcal{S}(\mathbb{R}^n)$ in such a way that $g \otimes f \in \mathcal{S}(\mathbb{R}^m_x \times \mathbb{R}^n_y)$. We can define the equivalent of (2.2), a *distributional kernel*, by setting

$$\langle T_\Phi f,\, g\rangle := \langle \Phi,\, g\otimes f\rangle.$$

Since Φ is continuous, we can find $N \in \mathbb{N}$ and $C > 0$ such that

$$|\langle T_\Phi f,\, g\rangle| \lesssim \|g\otimes f\|_{(N)} \lesssim \|f\|_{(N)}\|g\|_{(N)}.$$

This shows that the operator T_Φ is continuous from $\mathcal{S}(\mathbb{R}^m)$ to $\mathcal{S}'(\mathbb{R}^n)$.

Theorem 2.2. *Let $T : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^m)$ be a linear continuous operator. Then there exists a unique $\Phi \in \mathscr{S}'(\mathbb{R}^m \times \mathbb{R}^n)$ such that*

$$T = T_\Phi.$$

We will not give full proof of this result, but instead, we wish to discuss the main ideas behind it. The first step is to find a tensorial decomposition for Schwartz functions.

Lemma 2.3. *For every $n \in \mathbb{N}$ there exists $m \in \mathbb{N}$ such that for every function $h \in \mathscr{S}(\mathbb{R}^m \times \mathbb{R}^n)$ we can decompose h as follows:*

$$h = \sum_{j=0}^{\infty} g_j \otimes f_j,$$

where f_j and g_j are functions satisfying the estimate

$$\sum_{j=0}^{\infty} \|g_j\|_{(N)} \|f_j\|_{(N)} \lesssim \|h\|_{(M)}.$$

Now suppose that there exists Φ such that $T = T_\Phi$. By definition, given an arbitrary function h, we can write

$$\langle \Phi, h \rangle = \sum_{j=0}^{\infty} \langle \Phi, g_j \otimes f_j \rangle = \sum_{j=0}^{\infty} \langle T f_j, g_j \rangle$$

as a consequence of the previous technical lemma. The idea is to start from this identity and construct a Φ that satisfies it for every function $h \in \mathcal{S}(\mathbb{R}^m_x \times \mathbb{R}^n_y)$.

Example 2.4. Consider the non-linear differential operator

$$Lf(x) := \sum_{|\alpha| \le d} c_\alpha(x) \partial^\alpha f(x),$$

where c_α are smooth and bounded functions. It is easy to verify that

$$L(\mathscr{S}(\mathbb{R}^n)) \subseteq \mathscr{S}(\mathbb{R}^n) \subset \mathscr{S}'(\mathbb{R}^n),$$

so we can try to write explicitly the associated distributional kernel. If Φ exists, it must satisfy

$$\langle \Phi, g \otimes f \rangle = \langle Lf, g \rangle = \int_{\mathbb{R}^n} \left(\sum_{|\alpha| \le d} c_\alpha(x) \partial^\alpha f(x) \right) g(x) \, \mathrm{d}x.$$

The reader might prove, as an exercise, that

$$\langle \Phi, h \rangle = \int_{\mathbb{R}^n} c_\alpha(x) (\partial_2^\alpha h)(x, x) \, \mathrm{d}x,$$

where ∂_2^α indicates the derivative with respect to the second variable. Finally, notice that this formula defines a distribution since it is linear and there results:

$$|\langle \Phi, h \rangle| \le \sum_{|\alpha| \le d} \|c_\alpha\|_\infty \int_{\mathbb{R}^n} \frac{c\|h\|_{(N+d)}}{(1 + |x|)^N} \, \mathrm{d}x,$$

which for N big enough (so that the integral converges) leads to

$$|\langle \Phi, h \rangle| \lesssim_\alpha \|h\|_{(N+d)}.$$

2.2 Translation-invariant operators

The goal of this section is to characterise linear operators $T : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ that are continuous and *translation-invariant*.

Definition 2.5. We say that the operator $T : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ is *translation-invariant* if

$$\langle \tau_h(Tf),\, g\rangle = \langle T(\tau_h f),\, g\rangle$$

holds for all $h \in \mathbb{R}^n$ and all $g \in \mathscr{S}(\mathbb{R}^n)$. The translation of a distribution is given by

$$\langle \tau_h \Phi,\, f\rangle := \langle \Phi,\, \tau_{-h} f\rangle.$$

Let Φ be the distributional kernel given by Theorem 2.2. The operator $T = T_\Phi$ is translation-invariant if and only if

$$\langle \Phi,\, g \otimes (\tau_h f)\rangle = \langle \Phi,\, (\tau_{-h} g) \otimes f\rangle,$$

and, if we replace g with $\tau_h g$, it is also equivalent to

$$\langle \Phi,\, (\tau_h g) \otimes (\tau_h f)\rangle = \langle \Phi,\, g \otimes f\rangle. \tag{2.3}$$

Note that, if $\varphi \in \mathcal{S}(\mathbb{R}^m \times \mathbb{R}^n)$ is not given by the elementary product $g \otimes f$, we can exploit the decomposition given by Lemma 2.3. Then (2.3) implies that

$$\begin{aligned}
\langle \Phi,\, \varphi\rangle &= \sum_{j=0}^{\infty} \langle \Phi,\, g_j \otimes f_j\rangle = \\
&= \sum_{j=0}^{\infty} \langle \Phi,\, (\tau_h g_j) \otimes (\tau_h f_j)\rangle = \\
&= \langle \Phi,\, \tau_h \varphi\rangle,
\end{aligned}$$

which means that (2.3) completely characterise the property of being translation-invariant for the operator T_Φ.

Example 2.6. Let $\Psi \in \mathscr{S}'(\mathbb{R}^n)$. The *convolution operator* associated to Ψ is defined by setting

$$\mathscr{S}(\mathbb{R}^n) \ni f \longmapsto Tf := \Psi * f.$$

A straightforward computation, which follows from the definitions, shows that

$$\begin{aligned}
\langle \Psi * (\tau_h f),\, g\rangle &= \langle \Psi,\, \widetilde{\tau_h f} * g\rangle = \\
&= \langle \Psi,\, \check{f} * (\tau_{-h} g)\rangle = \\
&= \langle \Psi * f,\, \tau_{-h} g\rangle = \\
&= \langle \tau_h(\Psi * f),\, g\rangle.
\end{aligned}$$

From this chain of equalities we easily deduce that

$$T(\tau_h f) = \Psi * (\tau_h f) = \tau_h(\Psi * f) = \tau_h(Tf),$$

which means that T is translation-invariant as claimed.

The following result asserts that all linear continuous translation-invariant operators are necessarily equal to a *convolution operator*.

Theorem 2.7. *Let $T : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ be a linear continuous translation-invariant operator. Then there exists a unique $\Psi \in \mathscr{S}'(\mathbb{R}^n)$ such that*

$$Tf = \Psi * f.$$

We can now investigate the relation between Ψ and the kernel Φ associated to T. More precisely, we will show that in a special case the following formula holds:

$$\Psi(x-y) := \Phi(x-y,\, 0),$$

and then we will try to generalise it via an approximation argument.

Example 2.8. Let T be a linear continuous operator and let Φ be its kernel, that is,

$$\langle Tf,\, g\rangle = \langle \Phi,\, g \otimes f\rangle.$$

Suppose that Φ is a continuous **bounded function**. We can rewrite the above expression via integrals as follows:

$$\langle Tf,\, g\rangle = \iint \Phi(x,\, y) g(x) f(y) \,\mathrm{d}x\mathrm{d}y.$$

If T is also translation-invariant, then

$$\iint \Phi(x,\, y) g(x) f(y-h) \,\mathrm{d}x\mathrm{d}y = \iint \Phi(x,\, y) g(x+h) f(y) \,\mathrm{d}x\mathrm{d}y,$$

which is equivalent, via two change of variables, to the condition

$$\iint \Phi(x,\, y+h) g(x) f(y) \,\mathrm{d}x\mathrm{d}y = \iint \Phi(x-h,\, y) g(x) f(y) \,\mathrm{d}x\mathrm{d}y.$$

Since Φ is a continuous function, it is easy to see that it must satisfy

$$\Phi(x+h,\, y) = \Phi(x,\, y-h),$$

for all $x,\, y,\, h \in \mathbb{R}^n$. We can evaluate it at the point $y = y' + h$ and find that

$$\Phi(x+h,\, y+h) = \Phi(x,\, y).$$

In particular, the function $\Phi(x,\, y) = \Phi(x-y,\, 0)$ depends on a single variable, the difference between x and y, and therefore the following function is well-defined:

$$\Psi(x-y) := \Phi(x-y,\, 0).$$

It is now easy to check that

$$Tf = \Psi * f,$$

and this concludes the proof in the special case.

In the general case, the computation above is not valid. Nevertheless, we can approximate Φ via a family of mollifiers $\{\eta_\epsilon \otimes \eta_\epsilon\}_{\epsilon>0}$ and use the fact that

$$(\eta_\epsilon \otimes \eta_\epsilon) * \Phi$$

is a smooth function that converges to Φ when ϵ tends to zero.

Corollary 2.9. *If T is a linear continuous translation-invariant operator, then*

$$T(f * g) = (Tf) * g \quad \textit{for all } f,\, g \in \mathscr{S}(\mathbb{R}^n).$$

Proof. By Theorem 2.7, we know that there exists a unique tempered distribution Ψ such that $Tf = \Psi * f$, so it suffices to show that convolution is associative, that is,

$$\Psi * (f * g) \overset{?}{=} (\Psi * f) * g.$$

Let $h \in \mathscr{S}(\mathbb{R}^n)$ be a test function. Then

$$\begin{aligned} \langle \Phi * (f * g),\, h \rangle &= \langle \Phi,\, (f * g) * h \rangle = \\ &= \langle \Phi,\, \check{f} * \check{g} * h \rangle = \\ &= \langle \Phi * f,\, \check{g} * h \rangle = \langle (\Phi * f) * g,\, h \rangle, \end{aligned}$$

and this concludes the proof since h is arbitrary. $\square$

2.3 Translation-invariant operators in L^p

Let $1 \leq p < q \leq \infty$. If we can show that a tempered distribution $\Phi \in \mathscr{S}'(\mathbb{R}^n)$ satisfies an inequality of the type

$$\|\Phi * f\|_{L^q(\mathbb{R}^n)} \lesssim_{p,\,q} \|f\|_{L^p(\mathbb{R}^n)} \tag{2.4}$$

for all $f \in \mathscr{S}(\mathbb{R}^n)$, then the associated linear operator

$$T_\Phi : \mathscr{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n) \longrightarrow L^q(\mathbb{R}^n) \subset \mathscr{S}'(\mathbb{R}^n)$$

is continuous. In particular, it extends uniquely to the closure of the domain,

$$\widetilde{T_\Phi} : \overline{\mathscr{S}(\mathbb{R}^n)}^{\|\cdot\|_p} \longrightarrow L^q(\mathbb{R}^n).$$

Remark 2.10. The closure of $\mathscr{S}(\mathbb{R}^n)$ with respect to $\|\cdot\|_{L^p(\mathbb{R}^n)}$ coincides with $L^p(\mathbb{R}^n)$ if $p \geq 1$ is finite and is given by

$$C_0(\mathbb{R}^n) := \{f \in C(\mathbb{R}^n) \,:\, \lim_{|x|\to\infty} f(x) = 0\}$$

if $p = \infty$.

Consequently, it makes sense to wonder whether or not a linear continuous translation-invariant operator T from $L^p(\mathbb{R}^n)$ to $L^q(\mathbb{R}^n)$, $p < \infty$, is necessarily a convolution operator,

$$T(f) = \Phi * f,$$

where Φ belongs to an appropriate subspace of $\mathcal{D}'(\mathbb{R}^n)$. Notice that

$$T\big|_{\mathscr{S}(\mathbb{R}^n)} : \mathscr{S}(\mathbb{R}^n) \longrightarrow \mathscr{S}'(\mathbb{R}^n) \supset L^q(\mathbb{R}^n),$$

so by Theorem 2.7 we can always find a tempered distribution Φ such that

$$T(f) = \Phi * f \quad \text{for all } f \in \mathscr{S}(\mathbb{R}^n).$$

Set $p = 1$ and let $\{\varphi_n\}_{n\in\mathbb{N}} \subset \mathscr{S}(\mathbb{R}^n)$ be an approximation identity (see section 1.7) such that whenever $f \in \mathscr{S}(\mathbb{R}^n)$

$$\|\varphi_n * f - f\|_{L^1(\mathbb{R}^n)} \xrightarrow{n\to+\infty} 0.$$

The operator T is continuous, which means that Tf is the limit (w.r.t the L^q strong topology) of $T(\varphi_n * f)$; in other words, we have

$$\|Tf - (T\varphi_n) * f\|_{L^q(\mathbb{R}^n)} \xrightarrow{n\to+\ infty} 0.$$

Moreover, it is not hard to see that the family $\{T\varphi_n\}_{n\in\mathbb{N}}$ is bounded in $L^q(\mathbb{R}^n)$. In fact, there exists a uniform constant $C(q) := C > 0$ such that

$$\|T\varphi_n\|_{L^q(\mathbb{R}^n)} \le C < \infty \quad \text{for all } n \in \mathbb{N}.$$

By Banach-Alaoglu, we can find a subsequence $(n_k)_{k\in\mathbb{N}}$ for which $T\varphi_{n_k}$ converges to $\Phi \in L^q(\mathbb{R}^n)$ with respect to the **weak topology**. It follows that, given $g \in \mathscr{S}(\mathbb{R}^n)$, we have

$$\begin{aligned}\langle Tf,\, g\rangle &= \langle T,\, \check{f} * g\rangle = \\ &= \langle T,\, (f * \varphi_n) * g\rangle = \\ &= \langle T\varphi_n,\, \check{f} * g\rangle = \\ &= \langle T\varphi_n * f,\, g\rangle,\end{aligned}$$

which means that, up to subsequences, we have

$$\langle Tf,\, g\rangle = \lim_{n\to+\infty} \langle T\varphi_n * f,\, g\rangle = \langle \Phi * f,\, g\rangle,$$

and this concludes the proof that T is a convolution operator in the case $p = 1$.

Remark 2.11. The computation above requires the additional assumption $q \neq 1, \infty$ since otherwise we end up with $L^\infty(\mathbb{R}^n)$ and $L^1(\mathbb{R}^n)$, which are not reflexive and hence Banach-Alaoglu theorem does not apply.

We now summarise what we have obtained above in the next proposition, which asserts that translation-invariant operators from L^1 to L^q are convolution operators.

Proposition 2.12. *Let $T : L^1(\mathbb{R}^n) \to L^q(\mathbb{R}^n)$, $1 < q < \infty$, be a linear continuous translation-invariant operator. Then there exists $\Psi \in L^q(\mathbb{R}^n)$ such that*

$$Tf = \Psi * f \quad \textit{for all } f \in L^1(\mathbb{R}^n).$$

Moreover, for all translations $h \in \mathbb{R}^n$ there results

$$\Psi * (\tau_h f) = \tau_h(\Psi * f).$$

Definition 2.13 ($C_{p,\,q}$-spaces)**.** The $(p,\,q)$-convolution space is defined as

$$C_{p,\,q} := \{T : L^p(\mathbb{R}^n) \to L^q(\mathbb{R}^n) \,:\, T \text{ linear, continuous and translation-invariant}\}$$

if p is finite, and

$$C_{\infty,\,q} := \{T : C_0(\mathbb{R}^n) \to L^q(\mathbb{R}^n) \,:\, T \text{ linear, continuous and translation-invariant}\}$$

otherwise.

Remark 2.14. We proved earlier that we can associate to each element T of $C_{p,\,q}$ a unique tempered distribution Ψ_T in such a way that

$$Tf = \Psi_T * f.$$

This induces an isomorphism between $C_{p,\,q}$ and a subset of $\mathscr{S}'(\mathbb{R}^n)$; more precisely,

$$C_{p,\,q} = \left\{\Phi \in \mathscr{S}'(\mathbb{R}^n) \,:\, \|\Phi * f\|_{L^q(\mathbb{R}^n)} \lesssim_{p,\,q} \|f\|_{L^p(\mathbb{R}^n)} \text{ for all } f \in \mathscr{S}(\mathbb{R}^n)\right\}.$$

Notice that Proposition 2.12 can be reformulated by saying that the convolution space $C_{1,\,q}$ is isomorphic to $L^q(\mathbb{R}^n)$ for all $1 < q < \infty$. Indeed, by Young's inequality

$$\|\Psi * f\|_{L^q(\mathbb{R}^n)} \le \|f\|_{L^1(\mathbb{R}^n)}\|\Psi\|_{L^q(\mathbb{R}^n)},$$

which, in turn, implies that the operator $\Psi * f$ is continuous from $L^1(\mathbb{R}^n)$ to $L^q(\mathbb{R}^n)$ as long as Ψ belongs to $L^q(\mathbb{R}^n)$.

Unfortunately, **Banach-Alaoglu's theorem** does not apply when $q = 1$, but it seems plausible to expect that measures will take the place of $L^q(\mathbb{R}^n)$.

Proposition 2.15. *Let $\mathcal{M}(\mathbb{R}^n)$ be the set of finite complex-valued Borel measures. Then*

$$C_{1,\,1} \cong \mathcal{M}(\mathbb{R}^n).$$

Exercise 2.1. Prove Proposition 2.15.

Hint. First, show that $L^1(\mathbb{R}^n)$ is isometrically contained in $\mathcal{M}(\mathbb{R}^n)$, which is the dual space of $C_0(\mathbb{R}^n)$. Next, use it to find a subsequence converging to some $\mu \in \mathcal{M}(\mathbb{R}^n)$ such that

$$\|\mu\|_1 \le \|T\|.$$

Finally, the same argument used above shows that $C_{1,\,1} \cong \mathcal{M}(\mathbb{R}^n)$. □

To investigate further properties of the $C_{p,q}$-spaces we need a technical lemma, which asserts that L^p-functions behave in a "nice" way with respect to infinite translations.

Lemma 2.16. *Let $f \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$. Then*

$$\lim_{|h| \to +\infty} \|\tau_h f - f\|_{L^p(\mathbb{R}^n)} = 2^{\frac{1}{p}} \|f\|_{L^p(\mathbb{R}^n)}. \tag{2.5}$$

Similarly, if $f \in C_0(\mathbb{R}^n)$, then

$$\lim_{|h| \to +\infty} \|\tau_h f - f\|_\infty = \|f\|_\infty. \tag{2.6}$$

Proof. First, assume that $f \in C_c(\mathbb{R}^n)$. If $|h|$ is large enough, the support of f and $\tau_h f$ are necessarily disjoint, and thus

$$\|\tau_h f - f\|^p_{L^p(\mathbb{R}^n)} = \int_{\mathrm{spt}(f)} |f(x)|^p \, \mathrm{d}x + \int_{\mathrm{spt}(f)+h} |f(x-h)|^p \, \mathrm{d}x = 2\|f\|^p_{L^p(\mathbb{R}^n)}.$$

Taking the pth square root, we obtain (2.5). For a generic $f \in L^p(\mathbb{R}^n)$, we consider the restrictions $f_r := f\chi_{B_r}$ and, given $\epsilon > 0$, we choose r in such a way that

$$\|f_r - f\|_{L^p(\mathbb{R}^n)} \le \epsilon.$$

Finally, if $|h|$ is big enough (e.g., $|h| > 2r$ is enough), we infer that

$$\begin{aligned}
\left| \|\tau_h f - f\|_{L^p(\mathbb{R}^n)} - 2^{\frac{1}{p}} \|f\|_{L^p(\mathbb{R}^n)} \right| &\le \left| \|\tau_h f - f\|_{L^p(\mathbb{R}^n)} - \|\tau_h f_r - f_r\|_{L^p(\mathbb{R}^n)} \right| + \dots \\
&\dots + 2^{\frac{1}{p}} \left| \|f_r\|_{L^p(\mathbb{R}^n)} - \|f\|_{L^p(\mathbb{R}^n)} \right| \le \\
&\le \|\tau_h (f - f_r) - (f - f_r)\|_{L^p(\mathbb{R}^n)} + 2^{\frac{1}{p}} \|f - f_r\|_{L^p(\mathbb{R}^n)} \le \\
&\le (2 + 2^{\frac{1}{p}}) \|f_r\|_{L^p(\mathbb{R}^n)} \le 4\epsilon.
\end{aligned}$$

□

Theorem 2.17. *For all $q < p$ there results $C_{p,q} = \{0\}$.*

Proof. Suppose that there exists $T \in C_{p,q}$ such that $T \neq 0$. Then

$$\|Tf\|_{L^q(\mathbb{R}^n)} \le \|T\| \|f\|_{L^p(\mathbb{R}^n)}$$

for all $f \in L^p(\mathbb{R}^n)$. Take $g := \tau_h f - f$ and exploit the linearity of T to obtain the estimate

$$\|T(\tau_h f) - Tf\|_{L^q(\mathbb{R}^n)} \le \|T\| \|\tau_h f - f\|_{L^p(\mathbb{R}^n)}.$$

Since T is translation-invariant, the left-hand side can be rewritten as

$$\|\tau_h (Tf) - Tf\|_{L^q(\mathbb{R}^n)}.$$

Now let $|h|$ go to ∞ and apply (2.5). It turns out that

$$2^{\frac{1}{q}}\|Tf\|_{L^q(\mathbb{R}^n)} \le 2^{\frac{1}{p}}\|T\|\|f\|_{L^p(\mathbb{R}^n)}. \tag{2.7}$$

The operator norm $\|T\|$ is defined as the supremum of $\|Tf\|_{L^q(\mathbb{R}^n)}$ as f ranges in the closed ball of radius one in $L^p(\mathbb{R}^n)$. Thus taking the supremum in (2.7) leads to

$$\|T\| \le 2^{\frac{1}{p}-\frac{1}{q}}\|T\|,$$

and this is absurd because $2^{\frac{1}{p}-\frac{1}{q}}$ is strictly less than one. $\square$

Theorem 2.18. *The space $C_{p,\,q}$ is isometrically isomorphic to $C_{q',\,p'}$, where*

$$\frac{1}{p} + \frac{1}{p'} = \frac{1}{q} + \frac{1}{q'} = 1.$$

Proof. Let $T \in C_{p,\,q}$ and denote its dual operator by

$$T' : L^{q'}(\mathbb{R}^n) \longrightarrow L^{p'}(\mathbb{R}^n).$$

At this point, it is not clear whether or not $T\big|_{\mathscr{S}(\mathbb{R}^n)}$ coincides with $T'\big|_{\mathscr{S}'(\mathbb{R}^n)}$. Consider the bilinear form $B_p : L^p(\mathbb{R}^n) \times L^{p'}(\mathbb{R}^n) \to \mathbb{C}$ given by

$$B_p(f,\,g) := \int_{\mathbb{R}^n} f(x)g(-x)\,\mathrm{d}x.$$

This characterise, for all $1 \le p \le \infty$, the L^p-norm of f via duality as follows:

$$\|f\|_{L^p(\mathbb{R}^n)} = \sup_{\|g\|_{p'} \le 1} B_p(f,\,g).$$

In particular, the dual operator T' can be defined via this bilinear form in such a way that $T'g$ is the unique element satisfying

$$B_p(f,\,T'g) = B_q(Tf,\,g) \quad \text{for all } f \in L^p(\mathbb{R}^n).$$

Thanks to Theorem 2.17, we can assume $p \le q$, $q > 1$ and $p < \infty$. Then

$$\begin{aligned}
\|T'\| &= \sup_{\substack{g \in \mathscr{S}(\mathbb{R}^n) \\ \|g\|_{L^{q'}(\mathbb{R}^n)} \le 1}} \|T'g\|_{L^{p'}(\mathbb{R}^n)} = \\
&= \sup_{\substack{f,\,g \in \mathscr{S}(\mathbb{R}^n) \\ \|f\|_{L^p(\mathbb{R}^n)},\|g\|_{L^{q'}(\mathbb{R}^n)} \le 1}} |B_p(f,\,T'g)| = \\
&= \sup_{\substack{f,\,g \in \mathscr{S}(\mathbb{R}^n) \\ \|f\|_{L^p(\mathbb{R}^n)},\|g\|_{L^{q'}(\mathbb{R}^n)} \le 1}} |B_q(Tf,\,g)| = \|T\|.
\end{aligned}$$

Now, if $Tf = K * f$, for $f,\, g \in \mathscr{S}(\mathbb{R}^n)$ we have

$$\begin{aligned} B_p(f,\, T'g) = B_q(Tf,\, g) = & \\ &= \int_{\mathbb{R}^n} (K * f)(x) g(-x)\,\mathrm{d}x = \\ &= \langle K * f,\, \check{g} \rangle = \\ &= \langle f,\, \check{K} * \check{g} \rangle = B_p(f,\, Tg). \end{aligned}$$

This shows that $T = T'$ and concludes the proof. The case $p = q = \infty$ and $p = q = 1$ are left to the reader as an exercise. □

Theorem 2.19. *The space $C_{2,\,2}$ is isometrically isomorphic to $\mathscr{F}(L^\infty(\mathbb{R}^n))$.*

Proof. Let $T \in C_{2,\,2}$. Then there exists $\Phi \in \mathscr{S}'(\mathbb{R}^n)$ such that T is the convolution operator associated to Φ. In particular,

$$Tf = \Phi * f,$$

which, taking the Fourier transform, leads to

$$\mathscr{F}(Tf) = \mathscr{F}(\Phi)\mathscr{F}(f).$$

Denote $\mathscr{F}(\Phi)$ by M (which stands for *multiplicative operator*) and apply Theorem 1.31 to infer that M is bounded on $L^2(\mathbb{R}^n)$ with

$$\|M\| = \|T\|.$$

Therefore, given $f \in \mathscr{S}(\mathbb{R}^n)$ it is easy to verify that $Mf \in L^2(\mathbb{R}^n)$ and

$$\|Mf\|_{L^2(\Omega)} \leq \|T\| \|f\|_{L^2(\Omega)}.$$

Let ψ be a cutoff function with compact support, $\chi_{B_1} \leq \psi \leq \chi_{B_{\frac{3}{2}}}$, and call $\psi_j(\xi)$ the rescaling $\psi(2^{-j}\xi)$. The functions $m_j := M\psi_j$ belong to $L^2(\mathbb{R}^n)$ and

$$m_j m_{j+1} = m_j$$

holds for all $j \in \mathbb{N}$. We can easily construct $m \in L^2_{\mathrm{loc}}(\mathbb{R}^n)$ which coincides with each m_j on B_{2^j} so $Mf = mf$ for all $f \in C_c^\infty(\mathbb{R}^n)$. We now claim that

$$m \in L^\infty(\mathbb{R}^n) \quad \text{and} \quad \|m\|_\infty \leq \|M\|.$$

We argue by contradiction. Suppose that there is $\delta > 0$ such that $|m| > \|M\| + \delta$ on a set E of positive measure (which we may assume to be bounded). If we take

$$f_j \in C_c^\infty \;:\; \|f_j - \chi_E\|_{L^2(\mathbb{R}^n)} \xrightarrow{j \to \infty} 0,$$

then

$$M\chi_E = \lim_{j \to +\infty} Mf_j = m\chi_E.$$

It follows that $\|m\chi_E\|_{L^2(\mathbb{R}^n)} > (\|M\| + \delta)\|\chi_E\|_{L^2(\mathbb{R}^n)}$, which is absurd. Conversely, given

$m \in L^\infty(\mathbb{R}^n)$, the associated operator is given by

$$Tf = (\mathscr{F}^{-1}m) * f,$$

and it is bounded on $L^2(\mathbb{R}^n)$ by Plancherel's identity, with $\|T\| \leq \|m\|_\infty$. □

Corollary 2.20.

(a) *If $1 < p < q < 2$, then $C_{1,1} \subset C_{p,p} \subset C_{q,q} \subset C_{2,2}$.*

(b) *If $\frac{1}{p} - \frac{1}{q} = \frac{1}{r'}$, then $L^{(}\Omega) \subset C_{p,q}$.*

2.4 Differential operators and fundamental solutions

In this section, we focus on a special class of operators, which are usually known as *differential operators*. The main goal is to study existence of solutions to the equation

$$Lu = f$$

passing via *fundamental solutions*, which is a natural notion arising from the fact that we can study the same problem in the phases space via Fourier transform,

$$\mathscr{F}(Lu) = \mathscr{F}(f).$$

Thus, if we can find a solution of the transformed equation, we can obtain a solution of the initial problem via the inverse Fourier transform (when it is invertible).

Definition 2.21 (Differential operator)**.** Let $\alpha \in \mathbb{N}^n$ be a multi-index. The α-derivative of a function f is defined by setting

$$\partial^\alpha f(x) = \partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n} f(x).$$

A *linear differential operator* L is a linear combination of derivatives, that is,

$$(Lf)(x) = \sum_{|\alpha| \leq m} c_\alpha(x) \partial^\alpha f(x),$$

where $c_\alpha : \mathbb{R}^n \to \mathbb{R}$ are real-valued functions. If all c_α are constant, we can associate a *polynomial* to L which is given by

$$P(\xi) := \sum_{|\alpha| \leq m} c_\alpha (\imath\xi)^\alpha,$$

and the corresponding principal part

$$P_m(\xi) := \sum_{|\alpha| = m} c_\alpha (\imath\xi)^\alpha.$$

Notice that L is translation-invariant and, if the c_α are all constant, it sends $\mathscr{S}(\mathbb{R}^n)$ into

$\mathscr{S}(\mathbb{R}^n)$. Therefore, there exists a tempered distribution Φ such that

$$Lf = f * \Phi. \tag{2.8}$$

A simple computation shows that Φ must be a linear combination of derivatives of the Dirac delta centred at the origin; more precisely, we have

$$\Phi = \sum_{|\alpha| \le m} c_\alpha \partial^\alpha \delta_0.$$

Definition 2.22 (Fundamental solution)**.** Let L be a linear differential operator. We say that a distribution $\Psi \in \mathcal{D}'(\mathbb{R}^n)$ is a *fundamental solution* for L if

$$L\psi = \delta_0. \tag{2.9}$$

The existence of a fundamental solution is one of the critical features in the theory of differential operators. In the next proposition, we explain why this is the case.

Proposition 2.23. *Let L be a linear differential operator and let Ψ be a fundamental solution. Then for any given $f \in C_c^\infty(\mathbb{R}^n)$ it turns out that*

$$L(\Psi * f) = f \implies u = \Psi * f \text{ is a solution.}$$

Proof. It is clearly sufficient to show that

$$L(\Psi * f) = (L\Psi) * f. \tag{2.10}$$

For a test function $g \in \mathcal{D}(\mathbb{R}^n)$ it turns out that

$$\begin{aligned} \langle L(\Psi * f),\, g \rangle &= \langle \Psi * f,\, {}^{\mathrm{t}}Lg \rangle = \\ &= \langle \Psi,\, {}^{\mathrm{t}}L(\check{f} * g) \rangle = \\ &= \langle L\Psi,\, \check{f} * g \rangle = \langle (L\Psi) * f,\, g \rangle, \end{aligned}$$

where ${}^{\mathrm{t}}L$ denotes the transpose operator. This shows that (2.10) holds and concludes the proof. □

However, finding a fundamental solution for an arbitrary operator L is no easy task. Let us consider a linear differential operator of the form

$$(Lf)(x) = \sum_{|\alpha| \le m} c_\alpha \partial^\alpha f(x),$$

and suppose that we have a fundamental solution $\Psi \in \mathscr{S}'(\mathbb{R}^n)$. Since its Fourier transform is well-defined, it makes sense to compute

$$\langle \mathscr{F}(L\Psi),\, f \rangle = \langle L\Psi\, \mathscr{F} f \rangle = \langle \Psi,\, {}^{\mathrm{t}}L \mathscr{F} f \rangle.$$

One can check that the transpose operator ${}^{\mathrm{t}}L$ is given by

$$ {}^{\mathrm{t}}Lf(x) = \sum_{|\alpha|\leq m} (-1)^{|\alpha|} c_\alpha \partial^\alpha f(x). $$

We can apply ${}^{\mathrm{t}}L$ to the Fourier transform of f and find the following chain of equalities:

$$ \begin{aligned} {}^{\mathrm{t}}L\mathscr{F}f(x) &= \sum_{|\alpha|\leq m} (-1)^{|\alpha|} c_\alpha \partial_x^\alpha \int_{\mathbb{R}^n} f(\xi) \mathrm{e}^{-\imath \xi\cdot x} \, \mathrm{d}\xi = \\ &= \sum_{|\alpha|\leq m} (-1)^{|\alpha|} c_\alpha \int_{\mathbb{R}^n} f(\xi)(-\imath\xi)^\alpha \mathrm{e}^{-\imath \xi\cdot x} \, \mathrm{d}\xi = \\ &= \int_{\mathbb{R}^n} f(\xi) P(\xi) \mathrm{e}^{\imath \xi\cdot x} \, \mathrm{d}\xi = \mathscr{F}(fP)(x), \end{aligned} $$

where P is the polynomial associated to the operator L, that is,

$$ P(\xi) := P_L(\xi) = \sum_{|\alpha|\leq m} c_\alpha (\imath\xi)^\alpha. $$

It follows that

$$ \langle \mathscr{F}(L\Psi),\, f\rangle = \langle \Psi,\, \mathscr{F}(fP)\rangle = \langle P\mathscr{F}\Psi,\, f\rangle, $$

which (using the fact that Ψ is a fundamental solution) immediately leads to

$$ P\mathscr{F}(\Psi) = 1. \tag{2.11} $$

The naïve approach would suggest to simply take $\mathscr{F}\Psi$ as $\frac{1}{P}$ but, as we will see in the next examples, this is not always possible (because P might fail to be entire - see Theorem 1.70).

Example 2.24 (Laplace operator)**.** Let $L := \Delta$ be the Laplace operator on $\mathbb{R}^n$, i.e.,

$$ Lf(x) = \left(\partial_{x_1}^2 + \cdots + \partial_{x_n}^2\right) f(x). $$

The associated polynomial is $P(\xi) = -|\xi|^2$ and, clearly, we cannot define $\mathscr{F}\Psi$ as above because

$$ \left\langle -\frac{1}{|\cdot|^2},\, f \right\rangle = -\int_{\mathbb{R}^n} \frac{f(x)}{|x|^2} \, \mathrm{d}x $$

is not always a well-defined distribution. Indeed, if $n \geq 3$ then everything works fine, but for $n = 2$ the functional is not continuous.

Example 2.25 (Heat operator)**.** Let $L := \partial_t - \Delta_{\mathbb{R}^n}$ be the *heat operator* in $\mathbb{R}_t \times \mathbb{R}_x^n$. Then the associated polynomial is

$$ P(\xi) = \imath\xi_0^2 + |\xi|^2 = \imath\xi_0^2 + \sum_{j=1}^n \xi_j^2. $$

In this case, the reciprocal of P is well-defined and it allows us to define $\mathscr{F}\Psi$ as above, obtaining a distribution that is a fundamental solution for L.

Example 2.26 (Wave operator)**.** Let $L := \partial_t^2 - \Delta_{\mathbb{R}^2}$ be the *wave operator* in $\mathbb{R}_t \times \mathbb{R}_x^2$. Then the associated polynomial is

$$P(\xi) = -\xi_0^2 + \xi_1^2 + \xi_2^2.$$

This polynomial vanishes on the whole cone

$$\{\xi \in \mathbb{R}^3 \,:\, \xi_0^2 = \xi_1^2 + \xi_2^2\},$$

and hence we cannot expect, in general, that the reciprocal defines the Fourier transform of a distribution.

That said, it is possible to show the existence of a fundamental solution for a special class of linear operators; more precisely, the ones with constant coefficients.

Theorem 2.27 (Malgrange–Ehrenpreis)**.** *Let L be a linear differential operator with constant coefficients. Then there exists $\Phi \in \mathcal{D}'(\mathbb{R}^n)$ satisfying (2.9).*

Sketch of the proof. Let P be the polynomial associated to the linear differential operator L. Then we can rewrite it as

$$P(\xi) = \sum_{|\alpha| \leq m} a_\alpha \xi^\alpha,$$

where $a_\alpha = \imath^\alpha c_\alpha$. In a similar fashion, we can define

$$P_m(\xi) = \sum_{|\alpha| = m} a_\alpha \xi^\alpha,$$

to be the principal part of P.

Step 1. In particular, there must be a point $\bar{\xi} \in \mathbb{R}^n$ such that $P_m(\bar{\xi}) \neq 0$ and, up to a change of variables, we can always assume that

$$\bar{\xi} = (1, \mathbf{0}) \text{ and } P_m(\bar{\xi}) = 1.$$

It follows that

$$P_m(\xi) = \xi_1^m + \sum_{j=1}^{m-1} \xi_1^j q_{m-j}(\xi'),$$

where $\xi' = (\xi_2, \ldots, \xi_n) \in \mathbb{R}^{n-1}$ and q_{m-j} is a $(m-j)$-homogeneous multinomial.

Step 2. Now fix ξ' and denote by $z_1(\xi'), \ldots, z_p(\xi')$ the distinct roots, with multiplicities $m_j(\xi')$, of the equation

$$\xi_1^m + \sum_{j=1}^{m-1} \xi_1^j q_{m-j}(\xi') = 0.$$

A further change of variables, namely $\Gamma(\xi) := (\xi_1, \xi') + (\imath\varphi(\xi'), 0)$, is necessary to avoid points where P_m vanishes. The naïve idea would be to choose φ in such a way that

$$|\varphi(\xi') - z_j(\xi')| > 1 \quad \text{for all } j \in \{1, \ldots, n\}.$$

Let $B_j(\xi')$ be the ball centred at $z_j(\xi')$ with radius $\epsilon_j > 0$ in such a way that

$$B_j(\xi') \cap B_\ell(\xi') = \varnothing \quad \text{for all } j \neq \ell.$$

Applying **Rouché's theorem**, we conclude that for a given ξ'', close enough to ξ', the roots of $P_m(\cdot,\, \xi'')$ remain confined inside the collection of balls $\{B_j\}$. Now define

$$\varphi \text{ constant function satisfying } \varphi(\xi') := a \in [1,\, m+1],$$

so that the estimate

$$|a - \mathfrak{Im}(z_j(\xi''))| > 1$$

holds for all $\xi'' \in V$, where V is the small neighbourhood of ξ' mentioned before. Now Γ allows us to avoid all the zeros of the polynomial, and thus we can write

$$\langle \Phi,\, g\rangle = \int_{\mathbb{R}^n} \frac{\mathscr{F}g(-\xi_1 - \imath\varphi(\xi'),\, -\xi')}{P(\xi_1 + \imath\varphi(\xi'),\, \xi')}\, \mathrm{d}\xi \tag{2.12}$$

for any compactly supported function $g \in \mathcal{D}(\mathbb{R}^n)$. But now

$$|P(\xi_1 + \imath\varphi(\xi'),\, \xi')| = \prod_{j=1}^{m} |\xi_1 + \imath\varphi(\xi') - z_j(\xi')| \geq \prod_{j=1}^{m} |\varphi(\xi') - \mathfrak{Im}(z_j(\xi'))| > 1,$$

so the denominator on the right-hand side of (2.12) is smaller than the constant 1. In a similar fashion, we can estimate the numerator as follows:

$$\begin{aligned} |\mathscr{F}g(-\xi_1 - \imath\varphi(\xi'),\, -\xi')| &= \left| \int_{\mathbb{R}^n} g(x_1,\, x')\mathrm{e}^{\imath(x_1(\xi_1 + \imath\varphi(\xi')) + x'\xi'}\, \mathrm{d}x \right| \leq \\ &\leq \int_{\mathbb{R}^n} |g(x_1,\, x')|\mathrm{e}^{x_1\varphi(\xi')}\, \mathrm{d}x_1\mathrm{d}x' \leq \\ &\leq \|g\|_\infty r^n \mathrm{e}^{r|\varphi(\xi')|}. \end{aligned}$$

Notice that the last inequality follows from Theorem 1.70 because g belongs to $\mathcal{D}(\mathbb{R}^n)$ and thus we can find $r > 0$ such that $\mathrm{spt}(g) \subset B_r$. More in general, there results

$$|z^\alpha \mathscr{F}g(z)| \lesssim_\alpha \|\partial^\alpha g\|_\infty r^n \mathrm{e}^{r|\mathfrak{Im}(z)|}.$$

Combining this with the estimate above leads to

$$|\mathscr{F}g(-\xi_1 - \imath\varphi(\xi'),\, -\xi')| \leq \|g\|_{(N)}(1 + |\xi|)^{-N} r^n \mathrm{e}^{r|\varphi(\xi')|}$$

and, if we choose $N \in \mathbb{N}$ big enough, we readily find that the right-hand side of (2.12) is an absolutely convergent integral. More precisely, we have

$$|\langle \Phi,\, g\rangle| \lesssim_r \|g\|_{(N)},$$

which means that Φ is continuous (hence belongs to $\mathcal{D}'(\mathbb{R}^n)$), and this concludes the proof. □

To conclude this section, we state a stronger result concerning the existence of fundamental solutions which is due to Łojasiewicz-Hörmander.

Theorem 2.28 (Łojasiewicz-Hörmander)**.** *Let L be a linear differential operator with constant coefficients. Then there exists $\Phi \in \mathscr{S}'(\mathbb{R}^n)$ satisfying (2.9).*

The proof of this result is rather involved. The reader may refer to [5], where a simple proof is presented, and the reference therein of the original works by Łojasiewicz and Hörmander.

2.4.1 Applications of Malgrange–Ehrenpreis theorem

Let L be a linear differential operator with constant coefficients and $\Omega \subset \mathbb{R}^n$ an open set. We are interested in finding a solution of the problem

$$Lu = f \tag{2.13}$$

when f belongs to $\mathcal{D}'(\Omega)$. However, all we can do is to find a solution in any bounded subsets Ω' which is relatively compact in Ω. To do it, consider a cutoff function

$$\varphi \in \mathcal{D}(\Omega) \;:\; \varphi\big|_{\Omega'} \equiv 1$$

and let $f_0 := \varphi f$. Clearly, f_0 is still a distribution that coincides with f on Ω', but it has a big advantage; namely, its support is compact. Thus

$$L(\Phi * f_0) = f_0 \implies u = \Phi * f_0 \text{ is a solution of (2.13) in } \Omega'.$$

Corollary 2.29. *Linear differential operators with constant coefficients are locally solvable*[1].

We will now show that if we drop the assumption on the coefficients, even first-order operators with complex coefficients may not be locally solvable.

Theorem 2.30 (Lewy)**.** *Let $f \in C^1(\mathbb{R}, \mathbb{R})$ and consider the problem*

$$Au(z, t) = f'(t), \tag{2.14}$$

where A is the first-order linear differential operator defined by

$$A = 2\partial_{\bar{z}} - 2\imath z \partial_t.$$

Suppose that (2.14) has a solution of class C^1 defined in a neighbourhood of the form

$$D(0, \delta) \times (-\epsilon, \epsilon).$$

Then f must be analytic in the interval $(-\epsilon, \epsilon)$.

Remark 2.31. There are no solutions even in the distributional sense for some initial data ψ. In particular, the operator A is not locally solvable in a neighbourhood of the origin.

[1] In this set of notes, locally solvable means that we can find a solution in any relative compact subset. However, this notation is not standard and the reader may find a completely different notion of locally solvable in most books.

Proof. We will try to follow the original argument given in [6], although the complex variable z will usually be replaced by (x, y). Using polar coordinates, we have

$$x + \imath y = r^{\frac{1}{2}} \mathrm{e}^{\imath\theta},$$

so that we can rewrite f as

$$f(x, y) = f(r^{\frac{1}{2}} \cos\theta, r^{\frac{1}{2}} \sin\theta) =: g_f(r, \theta).$$

A simple computation shows that the partial derivatives in this new coordinate system are given by

$$\partial_r g_f(r, \theta) = \frac{1}{2} r^{-\frac{1}{2}} \left(\cos\theta \partial_x f + \sin\theta \partial_y f\right),$$

$$\partial_\theta g_f(r, \theta) = r^{\frac{1}{2}} \left(-\sin\theta \partial_x f + \cos\theta \partial_y f\right),$$

and this readily leads to the Jacobian matrix of the change of variables:

$$r^{\frac{1}{2}} \partial_x f = (2r\cos\theta \partial_r - \sin\theta \partial_\theta) g_f,$$

$$r^{\frac{1}{2}} \partial_y f = (2r\sin\theta \partial_r + \cos\theta \partial_\theta) g_f.$$

The complex derivative $\partial_{\bar{z}}$ is thus given by

$$2\partial_{\bar{z}} = \partial_x + \imath\partial_y = 2r^{\frac{1}{2}} \mathrm{e}^{\imath\theta} \partial_r + \imath r^{-\frac{1}{2}} \mathrm{e}^{\imath\theta} \partial_\theta,$$

except at the point $(0, 0)$ where this is not well-defined since $r^{-\frac{1}{2}}$ is singular. We now exploit the polar coordinates to infer that

$$\begin{aligned}
\int_0^{2\pi} \left[(\partial_x + \imath\partial_y) f\right] (r^{\frac{1}{2}} \mathrm{e}^{\imath\theta}) \,\mathrm{d}\theta = \int_0^{2\pi} \left[r^{\frac{1}{2}} \mathrm{e}^{\imath\theta} \partial_r + \imath r^{-\frac{1}{2}} \mathrm{e}^{\imath\theta} \partial_\theta\right] g(r, \theta) \,\mathrm{d}\theta &= \\
= \int_0^{2\pi} [r^{\frac{1}{2}} \mathrm{e}^{\imath\theta} \partial_r g](r, \theta) \,\mathrm{d}\theta + r^{-\frac{1}{2}} \int_0^{2\pi} g(r, \theta) \mathrm{e}^{\imath\theta} \,\mathrm{d}\theta &= \\
= \int_0^{2\pi} \mathrm{e}^{\imath\theta} \left[2r^{\frac{1}{2}} \partial_r + r^{-\frac{1}{2}}\right] g(r, \theta) \,\mathrm{d}\theta &= \\
= 2 \int_0^{2\pi} \mathrm{e}^{\imath\theta} \partial_r (r^{\frac{1}{2}} g)(r, \theta) \,\mathrm{d}\theta &= \\
2\partial_r \int_0^{2\pi} r^{\frac{1}{2}} g(r, \theta) \mathrm{e}^{\imath\theta} \,\mathrm{d}\theta &
\end{aligned}$$

where the orange identity follows from integration by parts of $\mathrm{e}^{\imath\theta}\partial_\theta g$. Now set

$$U(t, r) := \int_0^{2\pi} r^{\frac{1}{2}} \mathrm{e}^{\imath\theta} u(r^{\frac{1}{2}} \mathrm{e}^{\imath\theta}, t) \,\mathrm{d}\theta,$$

and notice that

$$\partial_t U + \imath \partial_r U = \int_0^{2\pi} r^{\frac{1}{2}} \mathrm{e}^{\imath\theta} \partial_t u(r^{\frac{1}{2}} \mathrm{e}^{\imath\theta},\, t)\, \mathrm{d}\theta + \frac{1}{2}\imath \int_0^{2\pi} \left[(\partial_x + \imath \partial_y) f\right] (r^{\frac{1}{2}} \mathrm{e}^{\imath\theta})\, \mathrm{d}\theta =$$

$$= \imath \int_0^{2\pi} (\partial_{\bar{z}} - \imath z \partial_t)\, u(r^{\frac{1}{2}} \mathrm{e}^{\imath\theta},\, t)\, \mathrm{d}\theta.$$

Let F be a primitive of f. Then the function defined by setting

$$\tilde{U}(t + \imath r) := U(t,\, r) - 2\pi\imath F(t)$$

is holomorphic (since $\partial_{\bar{z}} \tilde{U} = 0$) in $(-\epsilon,\, \epsilon) \times (0,\, \delta^2)$, which means that it extends holomorphically to $z = 0$ by setting

$$\tilde{U}(t) := -2\pi\imath F(t).$$

It follows that F is a real analytic function, and by definition so is its derivative f. This concludes the proof of the theorem. ☐

Remark 2.32. A linear differential operator with constant real-valued coefficient of order one is necessarily locally solvable so we must go to order at least 2 to find a counterexample.

Chapter 3

Hypoelliptic Operators and Hörmander Theorem

In this chapter, we introduce a notion that generalises the one of elliptic operators. Namely, we say that a linear differential operator L with smooth coefficients in Ω is **hypoelliptic** if

$$Lu \in C^\infty(\Omega') \implies u \in C^\infty(\Omega')$$

for all $\Omega' \subset \Omega$. The main goal of this chapter is to state and prove the well-known Hörmander theorem which gives a sufficient condition for an operator of the form

$$L = \sum_{j=1}^{k} X_j^2$$

to be hypoelliptic, where the X_j are vector fields.

Theorem A. Assume that, at each $x \in \Omega$, the vector fields

$$\{X_j\}_{j=1,\dots,k},\ \{[X_j,\, X_i]\}_{1\le i<j\le k},\ \{[X_j,\, [X_i,\, X_\ell]]\}_{i,j,\ell},\ \dots$$

span $\mathbb{R}^n$. Then the operator $L = \sum_{j=1}^{k} X_j^2$ is hypoelliptic.

3.1 Local solvability for constant coefficients operators

We can characterise hypoelliptic operators via the *singular support*, which is well-defined for all distributions in $\mathcal{D}'(\mathbb{R}^n)$.

Definition 3.1. The *singular support* of a distribution $\phi \in \mathcal{D}'(\Omega)$ is defined as the complement of

$$\{\Omega' \subset \Omega \,:\, \Omega \text{ open and } \phi \in C^\infty(\Omega')\}\,,$$

and it will be denoted by $\mathrm{spt}_{\mathrm{sing}}(\phi)$.

Proposition 3.2. *A linear differential operator L is hypoelliptic if and only if for all $u \in$*

$\mathcal{D}'(\Omega)$ *the following inclusion holds:*

$$\mathrm{spt}_{\mathrm{sing}}(u) \subseteq \mathrm{spt}_{\mathrm{sing}}(Lu).$$

The next result describes the local behaviour of hypoelliptic operators and allows one to estimate the C^k-norm of compactly supported functions. Recall that

$$\|f\|_k := \sum_{|\alpha|\leq k} \|\partial^\alpha f\|_\infty.$$

Theorem 3.3. *Let L be a hypoelliptic operator and fix $x \in \Omega$ and $k \in \mathbb{N}$. There exist a compact neighbourhood U_k of x and $k' = k'(k) \in \mathbb{N}$ such that*

$$\|f\|_k \lesssim_k \|Lf\|_{k'} \quad \textit{for all } f \in \mathcal{D}_{U_k}. \tag{3.1}$$

We first need to introduce some notations and state a couple of technical results. For a compact subset K of Ω and $\nu \in \mathbb{N}$ define

$$V^\nu(K) := \{f \in C^\nu(K) \,:\, Lf \in \mathcal{D}_K\},$$

equipped with the family of norms

$$\|f\|_{V^\nu,\,k} := \|f\|_\nu + \|Lf\|_k.$$

Lemma 3.4. *The couple $(V^\nu(K), \|\cdot\|_{V^\nu,\,k})$ is a Fréchet space. If in addition L is hypoelliptic, then it coincides with $\mathcal{D}_K$.*

Proof. The first assertion is left to the reader as an exercise. If L is hypoelliptic, then

$$Lf \in \mathcal{D}_K \implies f \in \mathcal{D}_U \quad \text{for all } U \subset K \text{ open}.$$

By compactness, we can find $U_1, \ldots, U_n \subset K$ such that $K = \cup_{i=1}^n U_i$ and, applying the implication above, yields

$$Lf \in \mathcal{D}_K \implies f \in \mathcal{D}_K \implies V^\nu(K) = \mathcal{D}_K.$$

□

Lemma 3.5. *Let $f \in C^\infty(\Omega)$ with* $\mathrm{spt}(f) \subset B(r)$*. Then*

$$\|f\|_k \leq 2r\|f\|_{k+1} \quad \textit{for all } k \in \mathbb{N}. \tag{3.2}$$

Proof. To simplify the notations, we can use the equivalent norm

$$\|f\|_k := \|f\|_\infty + \max_{|\alpha|=k} \|\partial^\alpha f\|_\infty.$$

We can assume that the support of f is contained in the cube $Q := [-r,\, r]^n$. For any multi-index α of length less than or equal to k, let $\alpha' = \alpha + (1, 0, \ldots, 0)$. If $x \in Q$,

$$|\partial^\alpha f(x)| = \left| \int_{-r}^{x_1} \partial^{\alpha'} f(t, x_2, \ldots, x_n)\,\mathrm{d}t \right| \leq$$

$$\leq 2r\|\partial^{\alpha'} f\|_\infty,$$

so that passing to the supremum yields (3.2). □

Proof of Theorem 3.3. Let K be a compact neighbourhood of x and fix $k \in \mathbb{N}$. The identity map

$$\iota : \mathcal{D}_K \longrightarrow V^{k-1}(K)$$

is continuous and, as a consequence of what we have proved above, also surjective. By the **open mapping theorem** ι is a homeomorphism. Therefore, the inclusion

$$j : V^{k-1}(K) \hookrightarrow C^k(K)$$

is continuous. Consequently, we can find $k' \in \mathbb{N}$ and $C_k > 0$ such that

$$\|f\|_k \lesssim_k \|f\|_{k-1} + \|Lf\|_{k'} \quad \text{for every } f \in \mathcal{D}_K.$$

Finally, let U_k be a ball centered at x of radius $\frac{1}{4C_k}$; it follows from the estimate (3.2) that

$$\|f\|_k \lesssim_k \|Lf\|_{k'},$$

and this concludes the proof. □

Corollary 3.6. *If L is hypoelliptic, then L is injective on $\mathcal{D}(U_k)$.*

Definition 3.7 (Locally solvable)**.** Let L be a linear differential operator with smooth coefficients on Ω. We say that L is *locally solvable* at $x \in \Omega$ if

$$\forall k \in \mathbb{N},\ \exists V_k \in \mathcal{N}(x)\ :\ \forall \psi \in \mathcal{D}'_k(\Omega),\ \exists u \in \mathcal{D}'(V_k)\ :\ Lu = \psi \text{ on } V_k.$$

In other words, for each $k \in \mathbb{N}$ we can find a neighbourhood of x, V_k, such that for each compactly supported distribution ψ of order k the equation

$$Lu = \psi$$

admits a solution $u \in \mathcal{D}'(V_k)$.

Theorem 3.8. *Let L be a hypoelliptic operator. Then ${}^{\mathrm{t}}L$ is locally solvable.*

Proof. Given $x \in \Omega$, let U_0 be a compact neighbourhood of x. Since $\psi \in \mathcal{D}'(\Omega)$ is continuous, we know that there are $k \in \mathbb{N}$ and $C > 0$ such that

$$|\langle \psi, f\rangle| \leq C\|f\|_{(k)} \quad \text{for all } f \in \mathcal{D}(U_0).$$

Let $U := U_k$ be given as in Theorem 3.3 and assume that $U \subset U_0$. Let

$$\mathfrak{X} := \{Lg\ :\ g \in \mathcal{D}_U\} \subseteq \mathcal{D}_U,$$

and define the linear functional $\lambda : \mathfrak{X} \to \mathbb{C}$ by setting

$$\lambda(Lg) := \langle \psi,\, g \rangle.$$

This is well-defined because L is injective and, using Theorem 3.3, we can find $k' \in \mathbb{N}$ which provides a bound to the functional λ:

$$|\lambda(Lg)| \le C\|g\|_{(k)} \lesssim_k \|Lg\|_{(k')}.$$

Applying Hahn-Banach we can extend λ to a continuous linear functional on $\mathcal{D}_{k',U}$. For $g \in \mathcal{D}_U$ there is a distribution u on V of order k' such that

$$\langle u,\, Lg \rangle = \lambda(Lg) = \langle \psi,\, g \rangle,$$

which means that ${}^{t}Lu = \psi$ on $V := U$. □

Let L be a constant coefficient differential operator with symbol σ. The following characterisation of hypoelliptic operators on $\mathbb{R}^n$ is due to Hörmander in [4].

Theorem 3.9. *A linear differential operator L with smooth coefficients is hypoelliptic on $\mathbb{R}^n$ if and only if, for some $\delta > 0$, the polynomial p satisfies the inequality*

$$\left| \frac{\partial^\alpha p(\xi)}{p(\xi)} \right| \le C|\xi|^{-\delta|\alpha|} \quad \text{for } |\xi| \text{ sufficiently big,}$$

for all $\alpha \in \mathbb{N}^n$ with length less than or equal to the degree of p as a polynomial.

This condition is satisfied by two fundamental classes of operators which also contains the Laplace operator (first one) and the heat operator (second one):

(i) Elliptic operators whose principal polynomial p_0 only vanishes at the origin, e.g., the Laplace operator with
$$p_0(\xi) = \xi_1^2 + \cdots + \xi_n^2.$$

(ii) Operators with polynomial only vanishes at the origin and is homogeneous with respect to some non-isotropic dilations, that is,
$$p(\xi) = \sum_\alpha c_\alpha \xi^\alpha,$$
where the sum is restricted to multi-indices α satisfying the affine relation $b \cdot \alpha = m$.

3.2 Hörmander theorem

The primary goal of this section is to characterise hypoelliptic operators of the form $X_j X_j$, where $X_1, \ldots, X_k$ are smooth real vector fields defined on Ω.

Caution!

The reader who is not familiar with the notions of vector field, commutator, flow, etc. is encouraged to read section A.3 before going any further.

Commutators. Let X denote the vector field on $\mathbb{R}^n$ defined by

$$X = \sum_{j=1}^{n} a_j(x)\partial_{x_j},$$

where $a_j \in C^\infty(\Omega)$ for all j. Let Y be another vector field, given by

$$Y = \sum_{j=1}^{n} b_j(x)\partial_{x_j},$$

with $b_j \in C^\infty(\Omega)$. We can now compute XY and YX explicitly:

$$XYf = \sum_{j,\,k=1}^{n} a_j(x)b_k(x)\partial_{x_j}\partial_{x_k}f(x) + \sum_{j,\,k=1}^{n} a_j(x)\partial_{x_j}b_k(x)\partial_{x_k}f(x),$$

$$YXf = \sum_{j,\,k=1}^{n} a_j(x)b_k(x)\partial_{x_j}\partial_{x_k}f(x) + \sum_{j,\,k=1}^{n} b_j(x)\partial_{x_j}a_k(x)\partial_{x_k}f(x).$$

This shows that $XY \neq YX$ or, in other words, X and Y do not commute. On the other hand, the commutator between X and Y is given by

$$[X,\,Y]f(x) = \sum_{k=1}^{n}\left(\sum_{j=1}^{n}(a_j(x)\partial_{x_j}b_k(x) - b_j(x)\partial_{x_j}a_k(x))\right)\partial_{x_k}f(x),$$

and it is easy to see that it only depends on the value of a_j, b_j and f at the point x. If $L = \sum_{j=1}^{k} X_j^2$, then L is an *elliptic* operator on Ω if and only if

$$\{X_j(x)\}_{j=1,\,\dots,\,k}$$

spans all of $\mathbb{R}^n$ at all points $x \in \Omega$. In particular, k must be greater than or equal to n.

Theorem 3.10 (Hörmander)**.** *Assume that, at each $x \in \Omega$, the vector fields*

$$\{X_j\}_{j=1,\,\dots,\,k},\ \{[X_j,\,X_i]\}_{1\le i<j\le k},\ \{[X_j,\,[X_i,\,X_\ell]]\}_{i,\,j,\,\ell},\ \dots$$

span $\mathbb{R}^n$. Then the operator $L = \sum_{j=1}^{k} X_j^2$ is hypoelliptic.

Remark 3.11. Under the same assumptions, the operator $\sum_{j=1}^{k-1} X_j^2 + X_k$ is also hypoelliptic.

Example 3.12 (Grushin plane)**.** Let $X = \partial_x$ and $Y = x\partial_y$ defined on $\mathbb{R}^2$. The operator

$$L = \partial_x^2 + x^2\partial_y^2$$

is *elliptic* away from $\{(x,\,y) \in \mathbb{R}^2 \,:\, x = 0\}$. However, the commutator is given by

$$[X,\,Y]\,\big|_{x=0} = \partial_y,$$

and hence L is hypoelliptic on the whole plane $\mathbb{R}^2$, as a consequence of Theorem 3.10.

Example 3.13 (Heisenberg space)**.** The operator defined in $\mathbb{R}^3$ by

$$L = \underbrace{(\partial_x - 2y\partial_z)^2}_{:=X^2} + \underbrace{(\partial_y + 2x\partial_z)^2}_{:=Y^2}$$

is hypoelliptic and it is usually referred to as sublaplacian. Notice that X and Y are always linear independent and

$$[X, Y] = 4\partial_z \implies \{X(x), Y(x), [X, Y](x)\} \text{ basis of } \mathbb{R}^3.$$

3.3 Besov potential spaces

In this section, we introduce a few technical tools that are needed to prove the Hörmander hypoellipticity theorem. To be more precise, given X vector field on Ω, we would like to exploit the flow to define a normed space with

$$\|f\|_{X,\,\alpha,\,\delta},$$

which supposedly generalises the usual Lipschitz spaces, and investigate them.

3.3.1 Besov spaces of functions

Definition 3.14 (Lipschitz)**.** Let f be a function defined on $\mathbb{R}^n$. We say that f is α-Lipschitz[1], where $\alpha \in (0, 1]$, if

$$|f(x+h) - f(x)| \lesssim |h|^\alpha \quad \text{for all } x,\, h \in \mathbb{R}^n. \tag{3.3}$$

Moreover, we say that f is *locally* α-Lipschitz if

$$|f(x+h) - f(x)| \lesssim |h|^\alpha$$

holds for all $x \in \mathbb{R}^n$ and for all h in a ball which radius depends on x, i.e. $B(x, r(x))$.

Remark 3.15. A function which is α-Lipschitz with $\alpha > 1$ is constant; this is why we require α to be in $(0, 1]$.

Remark 3.16. We can also consider functions satisfying the inequality

$$|f(x+2h) - 2f(x+h) + f(x)| \lesssim |h|^\alpha. \tag{3.4}$$

In this case, $\alpha \in (0, 2]$ does not lead to a trivial definition and, for all $\alpha \in (0, 1)$ - note that $\alpha = 1$ is excluded! - it is equivalent to (3.4). It is easy to verify that

$$\alpha > 2 \implies f \text{ is affine},$$

which makes sense if we think about $f(x+2h) - 2f(x+h) + f(x)$ as a *good* approximation of the second derivative of f.

[1]**N.B.** In the literature, functions satisfying this property with $\alpha \in (0, 1)$ are called α-Hölder while the terminology Lipschitz is reserved to the special case $\alpha = 1$.

In any case, it is not hard to verify that the condition (3.3) can be also rewritten using translations as follows:

$$\sup_{h\in\mathbb{R}^n\setminus\{0\}} |h|^{-\alpha}\|\tau_h f - f\|_\infty < \infty$$

Starting from this observation, we can define the *α-Besov space* as follows:

$$\Lambda^\infty_\alpha = \{f \in L^\infty(\mathbb{R}^n) \ : \ \sup_{h\in\mathbb{R}^n\setminus\{0\}} |h|^{-\alpha}\|\tau_h f - f\|_\infty < \infty\},$$

endowed with the norm

$$\|f\|_{\Lambda^\infty_\alpha} := \|f\|_\infty + \sup_{h\in\mathbb{R}^n\setminus\{0\}} |h|^{-\alpha}\|\tau_h f - f\|_\infty.$$

Proposition 3.17. *Let $\alpha \in (0,\,1]$. A function f belongs to Λ^∞_α if and only if f is continuous, bounded and α-Lipschitz.*

Proof. Let $(\varphi_\epsilon)_{\epsilon>0}$ be an approximate identity with $\operatorname{spt}(\varphi) \subset B_1$. Then

$$\varphi_\epsilon * f(x) - f(x) = \int_{\mathbb{R}^n} [f(x-y) - f(x)]\varphi_\epsilon(y)\,\mathrm{d}y.$$

Taking the $\|\cdot\|_\infty$ norm immediately leads to the following chain of inequalities:

$$\|\varphi_\epsilon * f(x) - f(x)\|_\infty \le \int_{\mathbb{R}^n} \|\tau_y f - f\|_\infty \varphi_\epsilon(y)\,\mathrm{d}y \le$$

$$\le \|f\|_{\Lambda^\infty_\alpha} \int_{B_\epsilon} |y|^\alpha \varphi_\epsilon(y)\,\mathrm{d}y = \epsilon^\alpha \|f\|_{\Lambda^\infty_\alpha}.$$

□

We can now introduce a slightly more refined version of Besov spaces, where L^∞ is replaced by L^p and the supremum norm by the $\|\cdot\|_p$ one.

Definition 3.18. For $1 \le p \le \infty$ we define the *p-Besov space* as

$$\Lambda^p_\alpha = \{f \in L^p(\mathbb{R}^n) \ : \ \sup_{h\in\mathbb{R}^n\setminus\{0\}} |h|^{-\alpha}\|\tau_h f - f\|_{L^p(\mathbb{R}^n)} < \infty\},$$

endowed with the norm

$$\|f\|_{\Lambda^p_\alpha} := \|f\|_{L^p(\mathbb{R}^n)} + \sup_{h\in\mathbb{R}^n\setminus\{0\}} |h|^{-\alpha}\|\tau_h f - f\|_{L^p(\mathbb{R}^n)}.$$

Remark 3.19. We proved earlier that $f \in L^p(\mathbb{R}^n)$ is enough to conclude that

$$\|\tau_h f - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{|h|\to 0} 0,$$

but here we are asking for something more: that the convergence happens with "$|h|^\alpha$-speed".

Remark 3.20. The p-Besov space is nonempty since we always have that

$$f_s(x) := |x|^{-s}\chi_{B(0,\,1)}(x)$$

is an element of $\Lambda^p_{\frac{n}{p}-s}$ for $s < \frac{n}{p}$.

Definition 3.21. For $1 \leq p \leq \infty$ and $1 \leq q < \infty$ we define the $(p,\, q)$-*Besov space* as

$$\Lambda^{p,\,q}_\alpha = \left\{ f \in L^p(\mathbb{R}^n) \ : \ \left[\int_{\mathbb{R}^n} \left(|h|^{-\alpha} \|\tau_h f - f\|_{L^p(\mathbb{R}^n)} \right)^q \frac{\mathrm{d}h}{|h|^n} \right]^{\frac{1}{q}} < \infty \right\},$$

endowed with the norm

$$\|f\|_{\Lambda^{p,\,q}_\alpha} := \|f\|_{L^p(\mathbb{R}^n)} + \left[\int_{\mathbb{R}^n} \left(|h|^{-\alpha} \|\tau_h f - f\|_{L^p(\mathbb{R}^n)} \right)^q \frac{\mathrm{d}h}{|h|^n} \right]^{\frac{1}{q}}.$$

The $(p,\, \infty)$-Besov space $\Lambda^{p,\,\infty}_\alpha$ is defined in such a way that it corresponds with the Λ^p_α defined above.

Lemma 3.22. *Let $1 \leq p < \infty$. For all $h \in \mathbb{R}^n$ there results*

$$\lim_{|h| \to \infty} \|\tau_h f - f\|_{L^p(\mathbb{R}^n)} = 2^{\frac{1}{p}} \|f\|_{L^p(\mathbb{R}^n)}. \tag{3.5}$$

Proof. First, assume that f is compactly supported in a ball of radius r. Then, if $|h| > 2r$, we have

$$\operatorname{spt}(f) \cap \operatorname{spt}(\tau_h f) = \varnothing.$$

The L^p-norm is now easy to estimate since

$$\|\tau_h f - f\|^p_{L^p(\mathbb{R}^n)} = \int_{B_r} |f|^p \,\mathrm{d}x + \int_{B_r + h} |\tau_h f|^p \,\mathrm{d}x = 2\|f\|^p_{L^p(\mathbb{R}^n)},$$

and this concludes the proof since we can always approximate $f \in L^p(\mathbb{R}^n)$ by a sequence of compactly supported functions. $\square$

Now let $a > 0$ be fixed and suppose that $p < \infty$. We would like to estimate the seminorm associated to $\|\cdot\|_{\Lambda^{p,\,q}_\alpha}$ in the range $|h| > a$. From (3.5) it follows that

$$\left[\int_{|h|>a} \left(|h|^{-\alpha} \|\tau_h f - f\|_{L^p(\mathbb{R}^n)} \right)^q \frac{\mathrm{d}h}{|h|^n} \right]^{\frac{1}{q}} \leq \left(2\|f\|_{L^p(\mathbb{R}^n)} \right)^q \int_{|h|>a} \frac{\mathrm{d}h}{|h|^{n+\alpha q}} \leq$$

$$\leq C_{q,\,\alpha} \|f\|^q_{L^p(\mathbb{R}^n)}.$$

Definition 3.23 (Sobolev space)**.** Let $s > 0$. We define the *s-fractional Sobolev space* as

$$H^s(\mathbb{R}^n) := \left\{ f \in L^2(\mathbb{R}^n) \ : \ (1 + |\cdot|^2)^s \mathscr{F}(f)(\cdot) \in L^2(\mathbb{R}^n) \right\}.$$

We shall now study the connection between the fractional space $H^s(\mathbb{R}^n)$ and the Besov potential space $\Lambda^{2,\,2}_s$, when $0 < s = \alpha < 1$. First, recall that

$$\|\mathscr{F}(g)\|_{L^2(\mathbb{R}^n)} = \|g\|_{L^2(\mathbb{R}^n)} \tag{3.6}$$

as a consequence of Plancherel's theorem. It follows that

$$\|\tau_h f - f\|_{L^2(\mathbb{R}^n)} = \|\mathscr{F}(\tau_h f - f)\|_{L^2(\mathbb{R}^n)},$$

and we can easily compute the first term via a simple change of variables:

$$\mathscr{F}(\tau_h f)(\xi) = \mathrm{e}^{-\imath h\cdot\xi}\mathscr{F}f(\xi).$$

Now let $f \in \Lambda_\alpha^{2,\,2}$. It follows from the properties above that

$$\|f\|_{\Lambda_\alpha^{2,\,2}} = \|\mathscr{F}f\|_{L^2(\mathbb{R}^n)} + \left[|h|^{-2\alpha}\int_{\mathbb{R}^n} \|\mathrm{e}^{-\imath h\cdot\xi}\mathscr{F}f - \mathscr{F}f\|^2_{L^2(\mathbb{R}^n)} \frac{\mathrm{d}h}{|h|^n}\right]^{\frac{1}{2}} =$$

$$= \|\mathscr{F}f\|_{L^2(\mathbb{R}^n)} + \left[\int_{\mathbb{R}^n} |\mathscr{F}f(\xi)|^2 \,\mathrm{d}\xi \int_{\mathbb{R}^n} |\mathrm{e}^{-\imath h\cdot\xi} - 1|^2 \frac{\mathrm{d}h}{|h|^{n+2\alpha}}\right]^{\frac{1}{2}} \stackrel{\star}{=}$$

$$\stackrel{\star}{=} \|\mathscr{F}f\|_{L^2(\mathbb{R}^n)} + \left[2\int_{\mathbb{R}^n} \frac{1-\cos(h')}{|h'|^{n+2\alpha}}\,\mathrm{d}h' \int_{\mathbb{R}^n} |\xi|^{2\alpha}|\mathscr{F}f(\xi)|^2\,\mathrm{d}\xi\right]^{\frac{1}{2}}.$$

The identity $\star$ follow from the change of variables $h' := |\xi|h$ which Jacobian is $|\xi|^n$. It is easy to verify that

$$2\int_{\mathbb{R}^n} \frac{1-\cos(h')}{|h'|^{n+2\alpha}}\,\mathrm{d}h' \le C(n,\,\alpha)$$

since the integrand is asymptotically equivalent respectively to $|h'|^{2-n-2\alpha}$ for $|h'| \to 0$ and to $|h'|^{-n-2\alpha}$ for $|h'| \to \infty$. It turns out that

$$\|f\|_{\Lambda_\alpha^{2,\,2}} \le \|\mathscr{F}f\|_2 + C(n,\,\alpha)\|| \cdot |^\alpha \mathscr{F}f(\cdot)\|_{L^2(\mathbb{R}^n)} \le \|f\|_{\Lambda_\alpha^{2,\,2}},$$

so we can finally conclude that

$$\Lambda_\alpha^{2,\,2} = H^\alpha(\mathbb{R}^n).$$

Proposition 3.24. *Let $0 < \beta < \alpha < 1$. Then the following inclusions hold:*

$$\Lambda_\alpha^{2,\,2} \subset \Lambda_\alpha^{2,\,\infty} \subset \Lambda_\beta^{2,\,2}.$$

Proof. The second inclusion is left as an exercise for the reader. As for the first one, let $f \in \Lambda_\alpha^{2,\,2}$ be an arbitrary function. We claim that for each $h \in \mathbb{R}^n$ we have

$$|h|^{-2\alpha}\|\tau_h f - f\|^2_{L^2(\mathbb{R}^n)} \le C|h|^{2-2\alpha}\|f\|_{\Lambda_\alpha^{2,\,2}}.$$

However, this follows immediately from the following inequalities:

$$\|\tau_h f - f\|^2_{L^2(\mathbb{R}^n)} \le c\int_{\mathbb{R}^n} |\mathscr{F}f(\xi)|^2|\mathrm{e}^{-\imath\xi\cdot h} - 1|^2\,\mathrm{d}\xi \le$$

$$\le c'|h|^2\int_{\mathbb{R}^n} |\mathscr{F}f(\xi)|^2|\xi|^2\,\mathrm{d}\xi,$$

where the last one is obtained as before using this time the real part of $\mathrm{e}^{-\imath\xi\cdot h} - 1$. By assumption, the quantity $2-2\alpha$ is always strictly positive so, taking into account that

$$\|\tau_h f - f\|^2_{L^2(\mathbb{R}^n)} \simeq \|f\|^2_{L^2(\mathbb{R}^n)} \quad \text{as } |h| \to \infty,$$

$$|h|^{-2\alpha}\|\tau_h f - f\|^2_{L^2(\mathbb{R}^n)} \xrightarrow{|h|\to\infty} 0,$$

we readily infer that

$$\sup_{h\neq 0} |h|^{-2\alpha}\|\tau_h f - f\|^2_{L^2(\mathbb{R}^n)} \leq \widetilde{C}\|f\|_{\Lambda^{2,\,2}_{\alpha}}.$$

□

3.3.2 Besov spaces of vector fields

Let $\Omega \subset \mathbb{R}^n$ be an open set, X a smooth vector field on Ω and denote by $\Phi(x,\,t)$ the *flow* generated by X. For $\alpha \in (0,\,1]$ and $\delta > 0$ we define

$$\|f\|_{X,\,\alpha,\,\delta} := \|f\|_{L^2(\Omega)} + \sup_{|t|\leq\delta} |t|^{-\alpha}\|\exp(tX)f - f\|_{L^2(\Omega)}.$$

Notice that $\mathrm{e}^{tX}f(x) = f\circ\Phi_t(x)$, so the quantity above is well-defined provided that $\delta > 0$ is small enough for the flow to be defined **at all** $x \in \Omega$. Since it might happen that

$$\inf_{x\in\Omega}\{t\in\mathbb{R} \,:\, (x,\,t)\in \mathrm{dom}(\Phi)\} = 0,$$

we usually restrict ourselves to compact subsets of Ω, as the following remark illustrates.

Remark 3.25. If $K \subset \Omega$ is a compact set, then the flow is defined at all $x \in K$ up to a uniform time $\delta_K > 0$. Thus, the quantity

$$\|f\|_{X,\,\alpha} = \|f\|_{L^2(\Omega)} + \sup_{|t|\leq\delta_K} |t|^{-\alpha}\|\exp(tX)f - f\|_{L^2(\Omega)}$$

is well-defined for all functions $f \in L^2(\Omega)$ with compact support in K.

Lemma 3.26. *Let $K \subset \Omega$ and let $0 < \delta' < \delta_K$. Then*

$$\|f\|_{X,\,\alpha,\,\delta'} \leq \|f\|_{X,\,\alpha,\,\delta_K} \quad \textit{for all } f \in L^2(\Omega) \textit{ with } \mathrm{spt}(f)\subset K.$$

Proof. This is obvious since on the right-hand side we are simply taking the supremum over a larger value of possible t's. □

Proposition 3.27. *Let $K \subset \Omega$ be a compact subset and let*

$$\varphi : \Omega \times [0,\,\delta) \longrightarrow \Omega$$

be a function satisfying the following properties:

- *(a) For all $t \in [0,\,\delta)$ the function φ_t is $C^\infty(\Omega)$ with respect to the variable x.*
- *(b) All partial derivatives of any order with respect to x are continuous in t.*
- *(c) There exists $\mu > 0$ such that $\varphi(x,\,t) = x + \mathcal{O}(|t|^\mu)$ when $t \to 0^+$.*

Then there exists a constant $C = C(K, \varphi)$ such that for all $\alpha \in (0, 1]$ and $|t|$ small enough we have

$$\|f \circ \varphi_t - f\|_{L^2(\Omega)} \le C|t|^{\mu\alpha}\|f\|_{\Lambda^{2,\infty}_\alpha} \quad \textit{for all } f \in L^2(\Omega) \textit{ with } \operatorname{spt}(f) \subset K. \tag{3.7}$$

Proof. Start by considering all h such that $|h| < |t|^\mu$ and apply the triangular inequality to the left-hand side of (3.7). Then

$$\|f \circ \varphi_t - f\|_{L^2(\Omega)} \le \|f \circ \varphi_t - \tau_h f\|_{L^2(\Omega)} + \|\tau_h f - f\|_{L^2(\Omega)},$$

where the second addendum is already known to be estimates by $|h|^\alpha\|f\|_{\Lambda^{2,\infty}_\alpha}$. As for the first addendum, take the average with respect to $|t|^\mu$,

$$\frac{1}{|t|^{n\mu}}\int_{|h|<|t|^\mu}\|f \circ \varphi_t - \tau_h f\|^2_{L^2(\Omega)}\,\mathrm{d}h = \frac{1}{|t|^{n\mu}}\int_{|h|<|t|^\mu}\int_\Omega |f(\varphi_t(x)) - f(x-h)|^2\,\mathrm{d}x\mathrm{d}h,$$

and apply the following change of variables:

$$(y, u) = \Psi_t(x, h) := (x - h, \varphi_t(x) - (x - h)).$$

Notice that $\Psi_0(x, h) = (x - h, h)$ is a **diffeomorphism** and Ψ_t, for $|t|$ small enough, is a perturbation, small in the C^1-norm, of Ψ_0, and thus a diffeomorphism. Thus

$$|f(\varphi_t(x)) - f(x-h)|^2 = |f(u+y) - f(y)|^2,$$

and this concludes the proof using the estimate above and the fact that $|h|$ and $|u|$ are not so different. $\square$

Proposition 3.28. *Let $K \subset \Omega$ be a compact subset and let X be a vector field on Ω. There exists a constant $C = C(K, X) > 0$ such that, for all $\alpha \in (0, 1]$ and $|t|$ small enough,*

$$\|f\|_{X,\alpha} \le C\|f\|_{\Lambda^{2,\infty}_\alpha} \quad \textit{for all } f \in L^2(\Omega) \textit{ with } \operatorname{spt}(f) \subset K. \tag{3.8}$$

Exercise 3.1. Prove (3.8) exploiting the same methods proposed for the proof of (3.7).

Recall that, if X is a vector field and f a function, then there is a notion of product fX which is also a vector field. In particular, there is a map

$$C^\infty_c(\Omega) \ni \eta \longmapsto X_\eta := \eta X \in \mathfrak{X}(\Omega),$$

which sends a smooth compactly supported function to a smooth vector field that is defined by setting

$$X_\eta(p) := \eta(p)X(p).$$

Our next goal is to understand the connection between $\|\cdot\|_{X,\alpha,\delta}$ and $\|\cdot\|_{X_\eta,\alpha,\delta}$. First, it is useful to see how integral curves are related to each other. Let $\gamma^\eta_x(t)$ and $\gamma_x(t)$ be the respective integral curves originating from the same point $x \in \Omega$, and notice that

$$\gamma^\eta_x(t) = \gamma_x(\tau(t, x)), \tag{3.9}$$

where $\tau(t, x)$ is a function describing the time-discrepancy between the two curves. We differentiate (3.9) with respect to t and obtain

$$X_\eta(\gamma_x^\eta(t)) = \partial_t \tau(t,\, x) X(\gamma_x^\eta(t)),$$

which easily leads to

$$\partial_t \tau(x,\, t) = \eta(\gamma_x^\eta(t)) = \eta(\gamma_x(\tau(x,\, t))).$$

We thus obtain, for each $x \in \Omega$, a ODE problem with fixed initial value

$$\begin{cases} \partial_t \tau_x(t) = \eta(\gamma_x(\tau_x(t))), \\ \tau_x(0) = 0, \end{cases}$$

which gives local existence and also a uniform lower bound on the time of existence if we restrict x to K.

Proposition 3.29. *Let $K \subset \Omega$ be compact. There exists a constant $C = C(K,\, \eta)$ such that for all $\alpha \in (0,\, 1]$ we have*

$$\|f\|_{X_\eta,\, \alpha} \le C \|f\|_{X,\, \alpha} \quad \textit{for all } f \in L^2(\Omega) \textit{ with } \mathrm{spt}(f) \subset K. \tag{3.10}$$

Proof. Let Φ and Φ^η be the flows of, respectively, X and X_η. As before, we first apply the triangular inequality to the left-hand side of (3.10) to obtain

$$\| \exp(tX_\eta)f - f\|_{L^2(\Omega)} \le \| \exp(tX_\eta)f - \exp(sX)f\|_{L^2(\Omega)} + \| \exp(sX)f - f\|_{L^2(\Omega)}.$$

By definition, we can estimate the second addendum $\| \exp(sX)f - f\|_{L^2(\Omega)}$ with $|s|^\alpha \|f\|_{X,\, \alpha}$, so we choose s with $|s| \le |t|$. As before, take the average over t,

$$\frac{1}{2t} \int_{-|t|}^{|t|} \|\mathrm{e}^{tX_\eta} f - \mathrm{e}^{sX} f\|^2_{L^2(\Omega)} \,\mathrm{d}s = \frac{1}{2t} \int_{-|t|}^{|t|} \int_\Omega |f \circ \Phi^\eta(x,\, t) - f \circ \Phi(x,\, s)|^2 \,\mathrm{d}x \mathrm{d}s =$$

$$= \frac{1}{2t} \int_{-|t|}^{|t|} \int_\Omega |f \circ \Phi(x,\, \tau(x,\, t)) - f \circ \Phi(x,\, s)|^2 \,\mathrm{d}x \mathrm{d}s.$$

Now apply the change of variables

$$(y,\, u) = \Psi_t(x,\, s) := (\Phi(x,\, s),\, \tau(x,\, t) - s),$$

and notice that $\Psi_0(x,\, s) = (\Phi(x,\, s),\, -s)$ is a diffeomorphism since its Jacobian has determinant equal to

$$-\nabla_x \Phi(x,\, s).$$

Therefore, since we can choose $|t|$ to be as small as we want, we can apply the same argument given in (3.9) to conclude that Ψ_t is a diffeomorphism and

$$\|\mathrm{e}^{tX_\eta} f - \mathrm{e}^{sX} f\|_2 \le C |u|^\alpha \|f\|_{X,\, \alpha}.$$

Finally, from the definition of u and the fact that $|t|$ is small we find that

$$|u| \le |s| + |\tau(x,\, t)| \le |t| + |\tau(x,\, t)| \overset{|t| \ll 1}{\le} C' |t|,$$

and this concludes the proof. $\square$

3.3.3 Besov spaces of combinations of vector fields

The goal of this section is to investigate the relation between $\|\cdot\|_{X+Y,\,\alpha}$ and $\|\cdot\|_{X,\,\alpha}+\|\cdot\|_{Y,\,\alpha}$, where X and Y are two smooth vector fields. If $[X,\,Y]=0$, then

$$\|f\|_{X+Y,\,\alpha,\,\delta} = \|f\|_{L^2(\Omega)} + \sup_{|t|<\delta}\left\{|t|^{-\alpha}\|\mathrm{e}^{tX+tY}f - f\|_{L^2(\Omega)}\right\}$$

may be estimated using the triangular inequality with either e^{tX} or e^{tY}:

$$\begin{aligned}\|\mathrm{e}^{tX+tY}f - f\|_{L^2(\Omega)} &\le \|\mathrm{e}^{tX}(\mathrm{e}^{tY}f-f)\|_{L^2(\Omega)} + \|\mathrm{e}^{tX}f-f\|_{L^2(\Omega)} = \\ &= \|(\mathrm{e}^{tY}f-f)\circ\varphi_{X,\,t}\|_{L^2(\Omega)} + \underbrace{\|\mathrm{e}^{tX}f-f\|_{L^2(\Omega)}}_{\le|t|^\alpha\|f\|_{X,\,\alpha}} \le \\ &\le C|t|^\alpha\left(\|f\|_{Y,\,\alpha}+\|f\|_{X,\,\alpha}\right).\end{aligned}$$

The blue inequality follows from the fact that, if $|t|$ is small enough, the flow $\varphi_{X,\,t}$ is a diffeomorphism and hence preserves the L^2-norm:

$$\|(\mathrm{e}^{tY}f-f)\circ\varphi_{X,\,t}\|_{L^2(\Omega)} = \|\mathrm{e}^{tY}f-f\|_{L^2(\Omega)}.$$

Proposition 3.30. *Let $K\subset\Omega$ be compact and let X and Y be smooth vector fields on Ω such that*

$$[X,\,Y]=0.$$

Then there exists a constant $C=C(K,\,X,\,Y)>0$ such that for all $\alpha\in(0,\,1]$ we have

$$\|f\|_{X+Y,\,\alpha} \le C(\|f\|_{X,\,\alpha}+\|f\|_{Y,\,\alpha}) \quad \text{for all } f\in L^2(\Omega) \text{ with } \operatorname{spt}(f)\subset K. \tag{3.11}$$

If $[X,\,Y]\neq 0$, then things do not work out in the same way. We now state a lemma which strengthens the conclusions reached in Proposition B.8.

Lemma 3.31. *Let $X_1,\,\dots,\,X_p$ be smooth vector fields on Ω.*

(i) *For all $m\ge1$ there exists $N(m)=N$ such that*

$$\exp\left(t(X_1+\cdots+X_p)\right)f(x) = \mathrm{e}^{\pm tX_{i_1}}\dots\mathrm{e}^{\pm tX_{i_N}}f(x)+\mathcal{O}(|t|^m), \tag{3.12}$$

where $i_k\in\{1,\,\dots,\,p\}$ for all $k=1,\,\dots,\,N$ and $\mathcal{O}(|t|^m)$ is uniform once we fix a compact subset $K\subset\Omega$.

(ii) *Let C_q be a elementary iterated commutator of multidegree q. Then for all $m\ge1$ we can find $N(m)=N$ such that*

$$\exp\left(tC_q\right)f(x) = \mathrm{e}^{\pm t^{\frac{1}{q}}X_{i_1}}\cdots\mathrm{e}^{\pm t^{\frac{1}{q}}X_{i_N}}f(x)+\mathcal{O}(|t|^m), \tag{3.13}$$

where $i_k\in\{1,\,\dots,\,p\}$ for all $k=1,\,\dots,\,N$ and $\mathcal{O}(|t|^m)$ is uniform once we fix a compact subset $K\subset\Omega$.

Remark 3.32. Multiplying by the inverses of the exponentials in the right-hand side in (3.12) yields

$$\mathrm{e}^{\mp t X_{i_N}} \cdots \mathrm{e}^{\mp t X_{i_1}} \mathrm{e}^{t(X_1+\cdots+X_p} f(x) - f(x) = \mathcal{O}(|t|^m), \tag{3.14}$$

which means that the flow Ψ associated to this composition of these exponentials satisfies

$$f \circ \Psi(x,\, t) = f(x) + \mathcal{O}(|t|^m).$$

If we choose f to be the coordinate function e_j, $j = 1, \ldots, n$, we get

$$\Psi(x,\, t) = x + \mathcal{O}(|t|^m),$$

and hence Ψ **satisfies** the assumptions of Proposition 3.27.

Theorem 3.33. *Let $X_1, \ldots, X_p$ be smooth vector fields on Ω, $K \subset \Omega$ compact and $\sigma \in (0,\, 1)$. Then for all $\alpha \geq 1$ and all $f \in L^2(\Omega)$ with support contained in K we have:*

(i) *There exists $C(X,\, p,\, \sigma) = C$ such that*

$$\|f\|_{X_1+\cdots+X_p,\, \alpha} \leq C \sum_{i=1}^{p} \|f\|_{X_i,\, \alpha} + C' \|f\|_{\Lambda^{2,\,\infty}_{\sigma}}. \tag{3.15}$$

(ii) *Let $q \in \mathbb{N}$. There exists $C(p,\, q) = C$ such that*

$$\|f\|_{C_q,\, \frac{\alpha}{q}} \leq C \sum_{i=1}^{p} \|f\|_{X_i,\, \alpha} + C' \|f\|_{\Lambda^{2,\,\infty}_{\sigma}}. \tag{3.16}$$

Proof. We only prove (**ii**). Use (3.13) to rewrite the left-hand side as

$$\|\mathrm{e}^{tC_q} f - f\|_{L^2(\Omega)} = \|\mathrm{e}^{\pm t^{\frac{1}{q}} X_{i_1}} \cdots \mathrm{e}^{\pm t^{\frac{1}{q}} X_{i_N}} (f \circ \Psi_t) - f\|_{L^2(\Omega)}.$$

This implies the following chain of inequalities

$$\begin{aligned}
(\mathbf{LHS}) &\leq \|\mathrm{e}^{\pm t^{\frac{1}{q}} X_{i_1}} \cdots \mathrm{e}^{\pm t^{\frac{1}{q}} X_{i_N}} (f \circ \Psi_t - f)\|_{L^2(\Omega)} + \|\mathrm{e}^{\pm t^{\frac{1}{q}} X_{i_1}} \cdots \mathrm{e}^{\pm t^{\frac{1}{q}} X_{i_N}} f - f\|_{L^2(\Omega)} \leq \\
&\leq C \|f \circ \Psi_t - f\|_{L^2(\Omega)} + \|\mathrm{e}^{\pm t^{\frac{1}{q}} X_{i_1}} \cdots \mathrm{e}^{\pm t^{\frac{1}{q}} X_{i_N}} f - f\|_{L^2(\Omega)} \leq \\
&\leq C \|f \circ \Psi_t - f\|_{L^2(\Omega)} + C' \sum_{i=1}^{p} \|\mathrm{e}^{t^{\frac{1}{q}} X_i} f - f\|_{L^2(\Omega)} \leq \\
&\leq C |t|^{m\sigma} \|f\|_{\Lambda^{2,\,\infty}_{\sigma}} + C' \sum_{i=1}^{p} |t|^{\frac{\alpha}{q}} \|f\|_{X_j,\, \alpha},
\end{aligned}$$

and we conclude by choosing $m\sigma \geq \frac{\alpha}{q}$.

The orange inequality follows from the fact that each $\varphi_{X_j,\, t}$ is a diffeomorphism for $|t|$ small enough, while the blue one follows from the fact that the second term can be rewritten as a telescopic sum (removing one exponential at a time.) $\square$

Lemma 3.34. *Let $0 < \beta < \alpha \leq 1$ and let $\epsilon > 0$. Then there exists a positive constant C_ϵ such that*

$$\|f\|_{\Lambda^{2,\infty}_\beta} \leq C_\epsilon \|f\|_{L^2(\Omega)} + \epsilon \|f\|_{\Lambda^{2,\infty}_\alpha} \quad \textit{for all } f \in L^2(\mathbb{R}^n). \tag{3.17}$$

Notice that the constant C_ϵ does not depend on ϵ only, but on the length of the interval over which the supremum is taken. Indeed, recall that the Besov norm is

$$\|f\|_{\Lambda^2_\alpha} = \|f\|_{L^2(\Omega)} + \sup_{0<|t|<a} |h|^{-\alpha} \|\tau_h f - f\|_2,$$

so it depends on the choice of a - although taking $a \neq a'$ only leads to equivalent norms.

Proof. Fix $a > 0$ and take $0 < \delta < a$. Then

$$\|f\|_{\Lambda^{2,\infty}_\beta} \leq \|f\|_{L^2(\Omega)} + \sup_{0<|h|\leq\delta} |h|^{-\beta} \|\tau_h f - f\|_{L^2(\Omega)} + \sup_{\delta\leq|h|<a} |h|^{-\beta} \|\tau_h f - f\|_{L^2(\Omega)} \leq$$

$$\leq \|f\|_{L^2(\Omega)} + \delta^{\alpha-\beta} \sup_{0<|h|\leq\delta} |h|^{-\alpha} \|\tau_h f - f\|_{L^2(\Omega)} + 2\delta^{-\beta} \|f\|_{L^2(\Omega)} =$$

$$\leq (1 + 2\delta^{-\beta}) \|f\|_{L^2(\Omega)} + \delta^{\alpha-\beta} \sup_{0<|h|<a} |h|^{-\alpha} \|\tau_h f - f\|_{L^2(\Omega)}.$$

Now choose δ in such a way that $\delta^{\alpha-\beta} = \epsilon$ and notice that (3.17) holds with C_ϵ given by

$$C_\epsilon = (1 + 2\delta^{-\beta} - \delta^{\alpha-\beta}).$$

□

Theorem 3.35. *Let $X_1, \ldots, X_k$ be smooth vector fields on Ω satisfying the condition*

$$\mathrm{Span}\,\langle X_1(x), \ldots, X_k(x), \ldots, \textit{commutators up to order } m \textit{ at } x\rangle = \mathbb{R}^n$$

at all $x \in K \subset \Omega$, K compact subset. Then for all $\alpha \in (0, 1]$ we have

$$\|f\|_{\Lambda^{2,\infty}_{\frac{\alpha}{m}}} \leq C_{\alpha,m,K} \sum_{j=1}^{k} \|f\|_{X_j,\alpha} \quad \textit{for all } f \in L^2(\Omega) \textit{ with } \mathrm{spt}(f) \subset K. \tag{3.18}$$

Proof. First, notice that

$$\|f\|_{\Lambda^{2,\infty}_\beta} \simeq \sum_{j=1}^{n} \|f\|_{\partial_{x_j},\beta},$$

where ∂_{x_j} denote the coordinate vector fields. Fix $x \in K$ and let

$$\{Y_1, \ldots, Y_n\} \subset \{X_1, \ldots, X_k, \ldots, \text{commutators up to order } m\}$$

be the basis of $\mathbb{R}^n$ at x which exists by assumption. Then there exists a small neighbourhood U_x of x such that

$$\mathrm{Span}\langle Y_1(y), \ldots, Y_n(y)\rangle = \mathbb{R}^n \quad \text{for all } y \in U_x.$$

Using the coordinate vector fields, we can write in a unique way

$$Y_j(y) = \sum_{i=1}^{n} \lambda_{i,\,j}(y)\partial_{x_i}$$

for some smooth functions $\lambda_{i,\,j}$ defined on U_x. Invert the metric to find smooth functions $\eta_{i,\,j}$ such that the following holds:

$$\partial_{x_j} = \sum_{i=1}^{n} \eta_{i,\,j}(y)Y_i(y).$$

The collection of open sets $\{U_x \ : \ x \in K\}$ covers K so, by compactness, we can select a finite family of point $x_1, \ldots, x_N$ such that

$$K \subseteq \bigcup_{i=1}^{N} U_{x_i}.$$

Let us consider a partition of unity relative to $\mathcal{U}$,

$$\{\varphi_i : U_{x_i} \longrightarrow \mathbb{R}\}_{1\le i\le N},$$

and glue together the vector fields ∂_{x_j} to find a global representation that holds at all points of K; namely, we have

$$\partial_{x_j} = \sum_{\ell=1}^{N} \varphi_\ell(x) \sum_{i=1}^{n} \eta^\ell_{i,\,j}(x)Y^\ell_i(x) \quad \text{for all } x \in K.$$

We added the superscript ℓ to η and Y because they depend also on the choice of the point x_i. Using (3.15), followed by (3.10), we can easily infer that

$$\|f\|_{\partial_{x_j},\,\frac{\alpha}{m}} \le C_\sigma \sum_{\ell=1}^{N}\sum_{i=1}^{n} \|f\|_{\varphi_\ell\eta_{i,\,j}Y_i,\,\frac{\alpha}{m}} \le$$

$$\le C'_\sigma \sum_{i=1}^{n} \|f\|_{Y_i,\,\frac{\alpha}{m}} + C'\|f\|_{\Lambda^{2,\,\infty}_\sigma}.$$

Now apply (3.16) with $q = m$ - since Y_k is at most a commutator of order m - and notice that

$$\|f\|_{\partial_{x_j},\,\frac{\alpha}{m}} \le C''_{m,\,\sigma} \sum_{i=1}^{k} \|f\|_{X_i,\,\alpha} + C'\|f\|_{\Lambda^{2,\,\infty}_\sigma}.$$

It turns out that

$$\|f\|_{\Lambda^{2,\,\infty}_\beta} \simeq \sum_{j=1}^{n} \|f\|_{\partial_{x_j},\,\beta} \le \widetilde{C}_{m,\,n,\,\sigma} \sum_{i=1}^{k} \|f\|_{X_i,\,\alpha} + C''_n\|f\|_{\Lambda^{2,\,\infty}_\sigma}.$$

To estimate the second term in the right-hand side, take σ smaller than $\frac{\alpha}{m}$ and notice that by (3.17) we can find a constant $C_\epsilon > 0$ such that

$$\|f\|_{\Lambda^{2,\infty}_{\sigma}} \leq \epsilon \|f\|_{\Lambda^{2,\infty}_{\frac{\alpha}{m}}} + C_\epsilon \|f\|_{L^2(\Omega)}.$$

The conclusion follows immediately if we choose ϵ in such a way that

$$C''_n \epsilon < \frac{1}{2}.$$

□

Corollary 3.36. *Let $f \in C_c^\infty(\Omega)$ with support contained in some compact set $K \subset \Omega$. Then there exists a constant $C = C(K) > 0$ such that*

$$\|f\|_{\Lambda^{2,\infty}_{\frac{1}{m}}} \leq C \left(\sum_{j=1}^{2} \|X_j f\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)} \right). \tag{3.19}$$

Proof. It suffices to estimate $\|f\|_{X_j,\alpha}$ when $\alpha = 1$. We have

$$\left| \int_\Omega \mathrm{e}^{tX} \, \mathrm{d}x \right|^2 = \left| \int_\Omega \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} \left(\mathrm{e}^{sX} \right) f(x) \, \mathrm{d}s \mathrm{d}x \right|^2 \leq$$

$$\leq a \int_\Omega \int_0^t |\mathrm{e}^{sX} X f(x)|^2 \, \mathrm{d}s \mathrm{d}x =$$

$$= a \int_0^t \|\mathrm{e}^{sX} X f \| L^2(\Omega)^2 \, \mathrm{d}s = a^2 \|X f\| L^2(\Omega)^2,$$

where a is the parameter determining the Besov norm. This shows that

$$\|f\|_{X_j,1} \simeq \|X_j f\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)},$$

and thus the conclusion follows from a straightforward application of (3.18). □

Corollary 3.37. *For every $s < \frac{1}{m}$ there results*

$$\|f\|_{H^s(\Omega)} \leq C \left(\sum_{j=1}^{2} \|X_j f\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)} \right). \tag{3.20}$$

Proof. This is a simple consequence of the inclusions

$$\Lambda^{2,\infty}_\beta \subset \Lambda^{2,2}_\alpha = H^\alpha(\Omega) \subset \Lambda^{2,\infty}_\alpha,$$

which holds true as soon as $\beta > \alpha$ - see Proposition 3.24. □

3.4 A step in Hörmander's theorem

In this section, we will show how to apply all these estimates proved so far to obtain a small step towards Hörmander's theorem.

Theorem 3.38 (Hörmander)**.** *Assume that, at each $x \in \Omega$, the vector fields*

$$\{X_j\}_{j=1,\,\dots,\,k},\ \{[X_j,\, X_i]\}_{1\le i<j\le k},\ \{[X_j,\, [X_i,\, X_\ell]]\}_{i,\,j,\,\ell},\ \dots$$

span $\mathbb{R}^n$. Then the operator $L = \sum_{j=1}^k X_j^2$ is hypoelliptic.

Lemma 3.39 (A priori estimate)**.** *Let $f \in C_c^\infty(\Omega)$ with support contained in a compact set $K \subset \Omega$. Then there exists a constant $C_s(K) = C_s > 0$ such that*

$$\|f\|_{H^s(\Omega)} \le C_s \left(\|f\|_{L^2(\Omega)} + \|Lf\|'_{\mathfrak{X}}\right)$$

for every $s < \frac{1}{m}$, where m is the order of commutators that are necessary to generate $\mathbb{R}^n$.

Proof. Let $K \subset \Omega' \Subset \Omega$. Define the norm

$$\|f\|_{\mathfrak{X}} := \left[\|f\|^2_{L^2(\Omega)} + \sum_{j=1}^k \|X_j f\|^2_{L^2(\Omega)}\right]^{\frac{1}{2}}$$

and, for $u \in \mathcal{D}'(\Omega')$, introduce the dual norm

$$\|u\|'_{\mathfrak{X}} = \sup_{\substack{f\in\mathcal{D}(\Omega')\\ \|f\|_{\mathfrak{X}}\le 1}} |\langle u,\, f\rangle|.$$

If $f \in \mathcal{D}(\Omega')$, then it is easy to check that

$$\|f\|'_{\mathfrak{X}} \le \|f\|_{L^2(\Omega)} \le \|f\|_{\mathfrak{X}}.$$

The scalar product $\langle Lf,\, f\rangle$, in terms of the vector fields X_j, is given by

$$\langle Lf,\, f\rangle = -\left\langle \sum_{j=1}^k X_j^2 f,\, f\right\rangle = -\sum_{j=1}^k \left\langle X_j f,\, X_j^* f\right\rangle,$$

so we first need to understand how the duals X_j^* behave. A simple computation shows that

$$\langle X^* f,\, g\rangle := \langle f,\, Xg\rangle = \int_\Omega f(x)\left[\sum_{i=1}^n a_i(x)\partial_{x_i} f(x)\right] \mathrm{d}x =$$

$$= -\int_\Omega \sum_{i=1}^n \partial_{x_i}(a_i f)(x) g(x)\,\mathrm{d}x =$$

$$= -\int_\Omega g(x) \sum_{i=1}^n a_i(x)\partial_{x_i} f(x)\,\mathrm{d}x - \int_\Omega g(x)f(x) \underbrace{\sum_{i=1}^n \partial_{x_i} a_i(x)}_{:=b(x)}\,\mathrm{d}x$$

so that

$$Xf(x) = \sum_{i=1}^{n} a_i(x)\partial_{x_i} f(x) \implies X^* f(x) = -\sum_{i=1}^{n} a_i(x)\partial_{x_i} f(x) - b(x)f(x).$$

It follows that

$$\langle Lf,\, f\rangle = \sum_{j=1}^{k} \langle X_j f,\, (X_j + b_j)f\rangle = \sum_{j=1}^{k} \|X_j f\|^2_{L^2(\Omega)} + \sum_{j=1}^{k} \langle X_j f,\, b_j f\rangle\,,$$

and the second term in the right-hand side can easily be dealt with using once more the formula for the dual X_j^*:

$$\sum_{j=1}^{k} \langle X_j f,\, b_j f\rangle = -\sum_{j=1}^{k} \langle f,\, (X_j + b_j) b_j f\rangle =$$

$$= -\sum_{j=1}^{k} \langle f,\, X_j(b_j f)\rangle - \left\langle f,\, b_j^2 f\right\rangle =$$

$$= -\sum_{j=1}^{k} \left[\langle f,\, X_j(b_j) f\rangle - \langle f,\, b_j X_j(f)\rangle - \left\langle f,\, b_j^2 f\right\rangle\right].$$

Since $\sum_j \langle f,\, b_j X_j(f)\rangle$ is equal to the quantity on the left-hand side, we find that

$$\sum_{j=1}^{k} \langle X_j f,\, b_j f\rangle = -\frac{1}{2}\sum_{j=1}^{k} \left[\langle f,\, X_j(b_j) f\rangle + \left\langle f,\, b_j^2 f\right\rangle\right].$$

Therefore, the scalar product between Lf and f is equal to

$$\langle Lf,\, f\rangle = \sum_{j=1}^{k} \|X_j f\|^2_{L^2(\Omega)} - \frac{1}{2}\sum_{j=1}^{k} \left[\langle f,\, X_j(b_j) f\rangle + \left\langle f,\, b_j^2 f\right\rangle\right],$$

from which we immediately deduce that

$$\|f\|_{\mathfrak{X}} = \|f\|_{L^2(\Omega)} + \langle Lf,\, f\rangle + \frac{1}{2}\sum_{j=1}^{k} \left[\langle f,\, X_j(b_j) f\rangle + \left\langle f,\, b_j^2 f\right\rangle\right] =$$

$$= \langle Lf,\, f\rangle + \frac{1}{2}\sum_{j=1}^{k} \left\langle f,\, (X_j b_j + b_j^2 + 2) f\right\rangle.$$

Now $(X_j b_j + b_j^2 + 2)$ is bounded on Ω' so there exists a positive constant C' such that

$$\frac{1}{2}\sum_{j=1}^{k} \left\langle f,\, (X_j b_j + b_j^2 + 2) f\right\rangle \le C' \|f\|^2_{L^2(\Omega)}.$$

Consequently,

$$\|f\|_{\mathfrak{X}} \le \|Lf\|'_{\mathfrak{X}}\|f\|_{\mathfrak{X}} + C'\|f\|^2_{L^2(\Omega)} \le$$

$$\le \|Lf\|'_{\mathfrak{X}}\|f\|_{\mathfrak{X}} + C'\|f\|_{L^2(\Omega)}\|f\|_{\mathfrak{X}},$$

which implies that

$$\|f\|_{\mathfrak{X}} \le \widetilde{C}(\|f\|^2_{L^2(\Omega)} + \|Lf\|'_{\mathfrak{X}}).$$

Using (3.20) we conclude the proof of this lemma. □

3.5 A few consequences of Hörmander's theorem

In this conclusive section, we show a few consequences of Hörmander's theorem such as the behaviour of integral curves of commutators and *nonisotropic lengths.* For example,

Theorem A. Let Ω be a connected subset of $\mathbb{R}^n$ and let $X_1, \ldots, X_k$ be a Hörmander system on Ω. Then any two points in Ω can be joined by a *horizontal curve.*

3.5.1 Integral curves of commutators

Let $X_1, \ldots, X_k$ be smooth vector fields on Ω satisfying the Hörmander condition with order p; in other words, for all $x \in \Omega$ we have

$$\mathbb{R}^n = \operatorname{Span}\langle X_1(x), \ldots, X_k(x), \text{commutators of order} \le p \text{ at } x\rangle .$$

Fix $x \in \Omega$ and denote by

$$\mathcal{Y}_x := \{Y_1, \ldots, Y_n\}$$

the collection of vector fields that gives a basis of $\mathbb{R}^n$ at x among the ones above. By continuity of the determinant map, the set

$$\{Y_1(y), \ldots, Y_n(y)\}$$

is still a basis of $\mathbb{R}^n$ at all y in an neighbourhood U_x of x small enough. Now consider the map

$$\Phi_x(t_1, \ldots, t_n) := \gamma_x^{\sum_{j=1}^n t_jY_j}(1) = \varphi_{\sum_{j=1}^n t_jY_j,\,1}(x),$$

where $\gamma_x^{\sum_{j=1}^n t_jY_j}(1)$ is the integral curve relative to the vector field

$$t_1Y_1 + \cdots + t_nY_n$$

starting from the point $x \in \Omega$ and evaluated at time $\tau = 1$. This map is well-defined provided that we take $|t_i|$ sufficiently small for each $i = 1, \ldots, n$. The Jacobian at the origin is given by

$$J_{\Phi_x}(0, \ldots, 0) = (Y_1(x), \ldots, Y_n(x))$$

since

$$\frac{\partial}{\partial t_j}\Big|_{\vec{t}=0}\Phi_x = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\Phi_x(0, \ldots, s, \ldots, 0) = Y_j(x)$$

follows immediately by exploiting the equation defining the integral curve, that is,

$$\gamma_x'(\tau) = Y_j(\gamma_x(\tau)).$$

The vector fields Y_i are linearly independent at x so the determinant of the Jacobian is nonzero, namely

$$\det\left(J_{\Phi_x}(0,\,\dots,\,0)\right) \neq 0,$$

and hence Φ_x, restricted to a small ball of $\mathbb{R}^n$ centred at the origin, is a diffeomorphism onto a neighbourhood of x. In a similar fashion, the map

$$\Psi_x(t_1,\,\dots,\,t_n) := \varphi_{Y_1,\,t_1} \circ \cdots \circ \varphi_{Y_n,\,t_n}(x)$$

is the map that follows the integral curve of Y_1 starting from $x \in \Omega$ for a time t_1 and then iteratively following the integral curve of Y_{j+1} starting from $Y_j(t_j)$ for a time t_{j+1}. But

$$J_{\Psi_x}(0,\,\dots,\,0) = (Y_1(x),\,\dots,\,Y_n(x))$$

so Ψ_x is also a diffeomorphism restricted to a small ball of $\mathbb{R}^n$ centred at the origin (and its differential coincides with the one given by the map Φ_x.)

Remark 3.40. Let $Y_1,\,\dots,\,Y_N$ denote the vector fields $X_1,\,\dots,\,X_k$ and their elementary commutators up to order p. The corresponding map

$$\Phi_x(t_1,\,\dots,\,t_N) := \varphi_{Y_1,\,t_1} \circ \cdots \circ \varphi_{Y_N,\,t_n}(x)$$

has Jacobian with maximal rank, and hence it maps a small ball of $\mathbb{R}^n$ centred at the origin onto a neighbourhood of x. It is worth remarking that the choice of the vector fields does not depend on $x \in \Omega$ since the Hörmander condition holds.

Definition 3.41 (Horizontal Curve)**.** We say that a piecewise C^1 function $\gamma : [a,\,b] \to \Omega$ is a *horizontal curve* if

$$\gamma'(t) \in \mathrm{Span}\langle X_1(\gamma(t)),\,\dots,\,X_k(\gamma(t))\rangle \quad \text{for almost every } t \in [a,\,b].$$

Problem. Can we join any two points of Ω via a horizontal curve? If Ω is connected, it is enough to show that each $x \in \Omega$ has a neighbourhood U_x such that

$$y \in U_x \implies \exists\, \gamma : [a,\,b] \to \Omega \text{ horizontal with } \gamma(0) = x,\, \gamma(1) = y.$$

If Ψ is the map defined above, we might try to move along integral curves to connect points. However, there is no guarantee that

$$Y_j \in \{X_1,\,\dots,\,X_k\} \quad \text{for all } j = 1,\,\dots,\,n,$$

and this is, in fact, impossible when $k < n$. Therefore, our goal is to exploit the theory developed in the previous sections to "approximate" the flow when Y_j is a commutator of order ≥ 1. We start with the simplest case possible:

Lemma 3.42. *Suppose that $Y = [X_1,\,X_2]$. Then the map*

$$\widetilde{\varphi}_{Y,\,t}(x) := \begin{cases} \varphi_{X_2,\,\sqrt{t}} \circ \varphi_{X_1,\,\sqrt{t}} \circ \varphi_{X_1,\,-\sqrt{t}} \circ \varphi_{X_2,\,-\sqrt{t}}(x), & \textit{for } t \geq 0, \\ \varphi_{X_2,\,\sqrt{|t|}} \circ \varphi_{X_1,\,\sqrt{|t|}} \circ \varphi_{X_1,\,-\sqrt{|t|}} \circ \varphi_{X_2,\,-\sqrt{|t|}}(x) & \textit{for } t < 0 \end{cases}$$

is C^1 with respect to t and satisfies the identity

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\widetilde{\varphi}_{Y,\,t}(x) = Y(x).$$

Proof. Denote by $\widetilde{\exp}$ the exponential map related to the modified flow $\widetilde{\varphi}_{Y,\,t}$ so that

$$\widetilde{\exp}(tY)f(x) = \exp(\sqrt{t}X_2)\exp(\sqrt{t}X_1)\exp(-\sqrt{t}X_1)\exp(-\sqrt{t}X_2)f(x).$$

We can easily compute the derivative for $t > 0$ as follows:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\varphi}_{Y,\,t}(x) &= \frac{1}{2\sqrt{t}}\left(\mathrm{e}^{\sqrt{t}X_2}\left[(X_1+X_2),\, \mathrm{e}^{\sqrt{t}X_1}\mathrm{e}^{-\sqrt{t}X_1}\right]\mathrm{e}^{-\sqrt{t}X_2}\right)f(x) = \\ &= \frac{1}{2\sqrt{t}}\left(\mathrm{e}^{\sqrt{t}X_2}\left[(X_1+X_2),\, \mathrm{Id}+\sqrt{t}(X_2-X_1)+\mathcal{O}(t)\right]\mathrm{e}^{-\sqrt{t}X_2}\right)f(x) = \\ &= ([X_1,\, X_2]+\mathcal{O}(t))f(x).\end{aligned}$$

We can do a similar computation for $t < 0$, and this gives the continuity of the t-derivative at 0 which concludes the proof since at any other point it is trivial. $\square$

We can now find a way to replace higher-order commutators iteratively. More precisely, suppose that $Y = [X_1,\, [X_2,\, X_3]] =: [X_1,\, X']$ and write for $t > 0$

$$\widetilde{\varphi}_{Y,\,t}(x) := \varphi_{X',\,t^{\frac{2}{3}}} \circ \varphi_{X_1,\,t^{\frac{1}{3}}} \circ \varphi_{X_1,\,-t^{\frac{1}{3}}} \circ \varphi_{X',\,-t^{\frac{2}{3}}}(x).$$

Now apply the formula given in Lemma 3.42 to X' to rewrite the right-hand side as the composition of $\varphi_{X_j,\,t^{\frac{1}{3}}}$, $j = 1,\, 2,\, 3$, and iterate it to elementary commutators of any order.

Theorem 3.43. *Suppose that Ω is connected and let $X_1,\, \ldots,\, X_k$ be a Hörmander system on Ω. Then any two points in Ω can be joined by a horizontal curve.*

3.5.2 Nonisotropic lengths

Let γ be a horizontal curve associated to the Hörmander system $\mathfrak{X} := \{X_1,\, \ldots,\, X_k\}$. By definition

$$\gamma'(t) \in \mathrm{Span}\langle X_1(\gamma(t)),\, \ldots,\, X_k(\gamma(t))\rangle,$$

so we can find coefficients $a_1,\, \ldots,\, a_k$ such that

$$\gamma'(t) = \sum_{j=1}^{k} a_j X_j(\gamma(t)),$$

but the representation is not unique (because, in general, $\mathfrak{X}$ is not made up of linearly independent vectors). Therefore, to define the velocity of γ at t, we consider all possible representation and write it as the smallest possible, namely

$$|\gamma'(t)|_{\mathfrak{X}} := \inf\left\{\sum_{j=1}^{k}|a_j| \,:\, \gamma'(t) = \sum_{j=1}^{k} a_j X_j(\gamma(t))\right\}.$$

The *length* of γ is defined through the velocity as usual,

$$L_{\mathfrak{X}}(\gamma) := \int_a^b |\gamma'(t)|_{\mathfrak{X}}\,\mathrm{d}t,$$

and since it is not restrictive to assume Ω connected, we can apply Theorem 3.43 and infer that the function defined by setting

$$d_{\mathfrak{X}}(x,\, y) := \inf\left\{L_{\mathfrak{X}}(\gamma) \,:\, \gamma \text{ horizontal curve joining } x \text{ and } y\right\}$$

is a well-defined distance between the points of Ω. This is usually referred to in the literature as *control distance* associated to the vector fields X_j.

Now let $\gamma : [a,\, b] \to \Omega$ be a piecewise C^1 curve and $\mathfrak{Y} := \{Y_1,\, \dots,\, Y_N\}$ the collection of all the X_j's and commutators up to order p given by the Hörmander condition. Then

$$\gamma'(t) = \sum_{j=1}^{N} a_j Y_j(\gamma(t)),$$

but this representation is also non-unique since $\{Y_1,\, \dots,\, Y_N\}$ is in general not a minimal set of generators. If d_j is the degree of Y_j, we define

$$|\gamma'(t)|_{\mathfrak{Y}} := \inf\left\{\sum_{j=1}^{k}|a_j|^{\frac{1}{d_j}} \,:\, \gamma'(t) = \sum_{j=1}^{N} a_j Y_j(\gamma(t))\right\}$$

and, as above, the corresponding length

$$L_{\mathfrak{Y}}(\gamma) := \int_a^b |\gamma'(t)|_{\mathfrak{Y}}\,\mathrm{d}t,$$

and the distance

$$d_{\mathfrak{Y}}(x,\, y) := \inf\left\{L_{\mathfrak{Y}}(\gamma) \,:\, \gamma \text{ piecewise } C^1 \text{ curve joining } x \text{ and } y\right\}.$$

We will now make a similar construction, which is highly local but gives accurate information about the magnitude of the distance in a small neighbourhood. Fix $x \in \Omega$ and pick a basis of $\mathbb{R}^n$ following these "rules":

(a) Up to a relabeling, pick $X_1, \dots X_{\ell_1}$ in such a way that $X_1(y), \dots, X_{\ell_1}(y)$ are linearly independent for all $y \in U_x$.

(b) Pick $[X_{i_1},\, X_{j_1}], \dots [X_{i_{\ell_2}},\, X_{j_{\ell_2}}]$ in such a way that $[X_{i_1},\, X_{j_1}](y), \dots [X_{i_{\ell_2}},\, X_{j_{\ell_2}}](y)$ are linearly independent for all $y \in U_x$.

(c) Similar process up to commutators of order p. We get a basis, which we denote by

$$\{Y_1, \ldots, Y_{\ell_1}, Y_{\ell_1+1}, \ldots, Y_{\ell_2}, \ldots, Y_{\ell_{p+1}}\}$$

satisfying the following dimensional relation

$$\sum_{j=1}^{p+1} \ell_j = n, \quad \ell_j \geq 0.$$

We proved that this basis induces a diffeomorphism $\Psi\big|_{B(0,\,\epsilon)}$ onto a neighbourhood U_x of x for some $\epsilon > 0$. If $y \in U_x$ and γ joins y with x, then there exists a unique way to write

$$\gamma'(t) = \sum_{j=1}^{n} a_j Y_j(\gamma(t)),$$

and, consequently,

$$|\gamma'(t)|_{\mathfrak{Y}} := \sum_{j=1}^{n} |a_j|^{\frac{1}{d_j}}$$

is well-defined and leads to a distance, which is obviously local. Furthermore, if we write the n-uple $(t_1, \ldots, t_n)$ as

$$(\vec{t_0}, \ldots, \vec{t_p}), \quad \text{where } \vec{t_j} = (t_{\ell_j+1}, , \ldots, t_{\ell_{j+1}}),$$

then we can easily prove that

$$y = \Psi(\vec{t_0}, 0, \ldots, 0) \implies d(x, y) \simeq \epsilon,$$

$$y = \Psi(\vec{t_0}, \vec{t_1}, 0, \ldots, 0) \implies \epsilon \leq d(x, y) \leq \epsilon^{\frac{1}{2}},$$

$$y = \Psi(\vec{t_0}, \vec{t_1}, \vec{t_2}, 0, \ldots, 0) \implies \epsilon^{\frac{1}{2}} \leq d(x, y) \leq \epsilon^{\frac{1}{3}},$$

and, more in general, the distance between x and $y \in U_x$ is always $\leq \epsilon^{\frac{1}{p}}$. It follows that the corresponding ball $B_d(x, r)$ satisfies the following:

$$y \in B_d(x, r) \text{ in the direction generated by commutators of order } m \implies d(x, y) \leq r^m.$$

3.5.3 Applications to distributions

Before going through this section, the reader is encouraged to refresh some concepts in differential geometry such as distributions, foliations, etc. In Section A.5, the bare minimum necessary to understand this section is presented without any proof.

Framework. Let $M \subset \mathbb{C}^n$ be a hypersurface and let $\varphi : \mathbb{C}^n \to \mathbb{R}$ be the function with nonzero gradient $\nabla\varphi \neq 0$ such that

$$M = \{z \in \mathbb{C}^n \,:\, \varphi(z) = 0\} = \varphi^{-1}(0),$$

and, for $z \in M$, consider the real tangent space

$$T_z M = (\nabla\varphi(z))^{\perp}.$$

The symbol $\perp$ denotes the orthogonal with respect to the real scalar product $\langle \cdot, - \rangle$ and it is easy to verify that $T_z M$ is a $(2n-1)$-dimensional real vector space. We can also define the complex tangent space

$$T_z^{\mathbb{C}} M := (\nabla\varphi(z))^{\perp},$$

where the orthogonal is now taken with respect to the Hermitian inner product, which we will always denote by $\langle \cdot, - \rangle_{\mathbb{C}}$. Observe that

$$T_z^{\mathbb{C}} M = (T_z M) \cap (\imath T_z M),$$

so the complex one is, in general, different from the real one and actually its dimension is lower than $2n-1$. For each $a \in \mathbb{R}$ we can analogously define a hypersurface by setting

$$M_a := \{z \in \mathbb{C}^n \, : \, \varphi(z) = a\} = \varphi^{-1}(a).$$

Then the corresponding family of *real tangent spaces*

$$\left\{T_z M_{\varphi(z)} \, : \, z \in \mathbb{C}^n\right\}$$

is an integrable distribution since it is tangent to the foliation of $\mathbb{C}^n$ made up of the hypersurfaces $\{M_a\}_{a\in\mathbb{R}}$. On the other hand, it is easy to see that

$$\left\{T_z^{\mathbb{C}} M_{\varphi(z)} \, : \, z \in \mathbb{C}^n\right\}$$

might fail to be integrable.

Example 3.44. If $\varphi(z) := \mathfrak{Im}(z_n)$, then

$$T_z^{\mathbb{C}} M_a = \xi + \{(z', 0) \, : \, z' \in \mathbb{C}^{n-1}\},$$

where $\xi \in \mathbb{C}^n$ is such that $\mathfrak{Im}(\xi) = a$. Then the distribution of complex tangent spaces defined by φ is integrable.

Example 3.45. If we consider the sphere of radius one in $\mathbb{C}^n$, it is easy to see that the real tangent space cannot coincide with the complex one because

$$\eta \perp T_p \mathbb{S}^{2n-1} \implies \imath\eta \parallel T_p \mathbb{S}^{2n-1}.$$

Remark 3.46. Let $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ be the real unit sphere. Then a vector field X is tangent to the sphere at p if and only if

$$X(x_1^2 + \cdots + x_n^2) = 0.$$

For example, the angular derivative $\Omega_{i,j} := x_i \partial_j - x_j \partial_i$ is always tangent to the sphere (in any dimension and for all $i \neq j$).

Example 3.47. In $\mathbb{C}^2$, consider the vector fields

$$X = \bar{z}_2 \partial_{z_1} - \bar{z}_1 \partial_{z_2},$$

$$Y = z_2 \partial_{\bar{z}_1} - z_1 \partial_{\bar{z}_2}.$$

It is rather easy to prove that X and Y are tangent to the complex tangent space of the complex sphere $S^1_{\mathbb{C}}$. However, the commutator is equal to

$$[X, Y] = z_1 \partial_{z_1} + z_2 \partial_{z_2} - \bar{z}_1 \partial_{\bar{z}_1} - \bar{z}_2 \partial_{\bar{z}_2},$$

and this is not tangent to the complex tangent space; in particular, the distribution

$$\left\{ T_p^{\mathbb{C}} S^1_{\mathbb{C}} \;:\; p \in \mathbb{C}^2 \right\}$$

is not integrable.

Example 3.48 (Grushin plane)**.** Consider on $\mathbb{R}^2$ the vector fields $X = \partial_x$ and $Y = x\partial_y$. It is easy to verify that X and Y span $\mathbb{R}^2$ everywhere except for the line $\{x = 0\}$. However,

$$[X, Y] = \partial_y,$$

so they satisfy the Hörmander condition at order one. We now want to estimate the distance between points on $\mathbb{R}^2$, which we have defined in the previous section as

$$d_{\mathfrak{X}}(x, y) := \inf \left\{ L_{\mathfrak{X}}(\gamma) \;:\; \gamma \text{ horizontal curve joining } x \text{ and } y \right\}.$$

We recall that γ is a horizontal curve if and only if $\gamma'(t)$ belongs to the vector space spanned by $X(\gamma(t))$ and $Y(\gamma(t))$. Observe that

$$p \in \mathbb{R}^2 \;:\; p_x \neq 0 \implies d_{\mathfrak{X}}(p, q) \propto d(p, q)$$

since outside of $\{x = 0\}$, the vector field Y in a small neighbourhood of each point behaves similarly to ∂_y multiplied by some constant.

On the other hand, if both p and q belong to the y-axis, then we cannot move vertically so a horizontal curve joining them must be as in Figure 3.1. Moreover,

$$L_{\mathfrak{X}}(\gamma) = \inf_{\delta > 0} \left\{ 2\delta + \frac{h}{\delta} \right\},$$

which is achieved when $\delta = \frac{1}{\sqrt{2}}\sqrt{h}$. It follows that

$$d_{\mathfrak{X}}(p, q) = 2\sqrt{2}\sqrt{h} \simeq \sqrt{h},$$

so it is not comparable with the Euclidean distance but, at the same time, is coherent with the abstract theory developed in the previous sections.

Example 3.49 (Heisenberg space)**.** Consider on $\mathbb{R}^3$ the vector fields

$$X = \partial_x - \frac{y}{2}\partial_z \quad \text{and} \quad Y = \partial_y + \frac{x}{2}\partial_z.$$

As before, an easy computation shows that the commutator $[X, Y]$ equal to ∂_z and thus the system $\mathfrak{X}$ satisfies the Hörmander condition at order one. However, describing horizontal curves is not as easy as in the Grushin plane, and it can be done via a kind of constructive process which is briefly explained in Figure 3.2.

Remark 3.50. We can estimate the distance $d_{\mathfrak{X}}$ in these two cases because they both enjoy

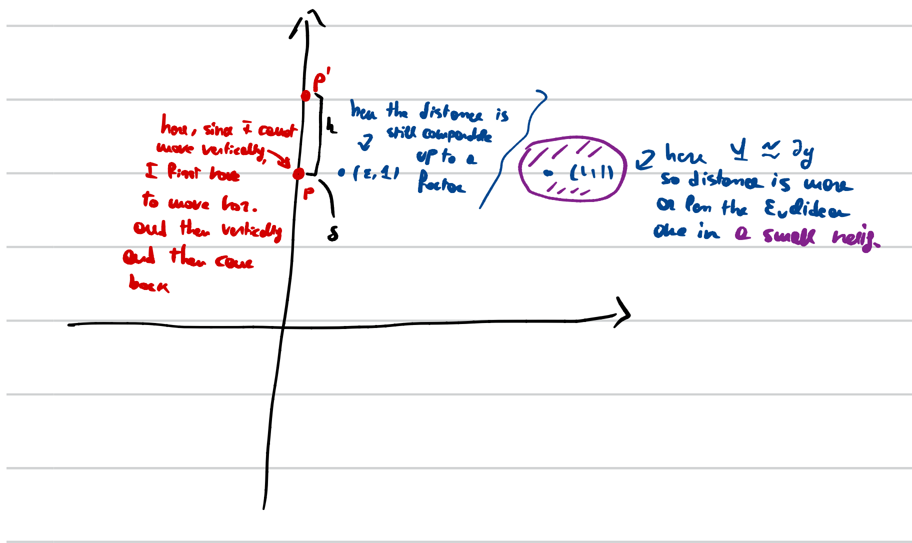

Figure 3.1: A picture of the curve joining two point on the y-axis in the Grushin plane.

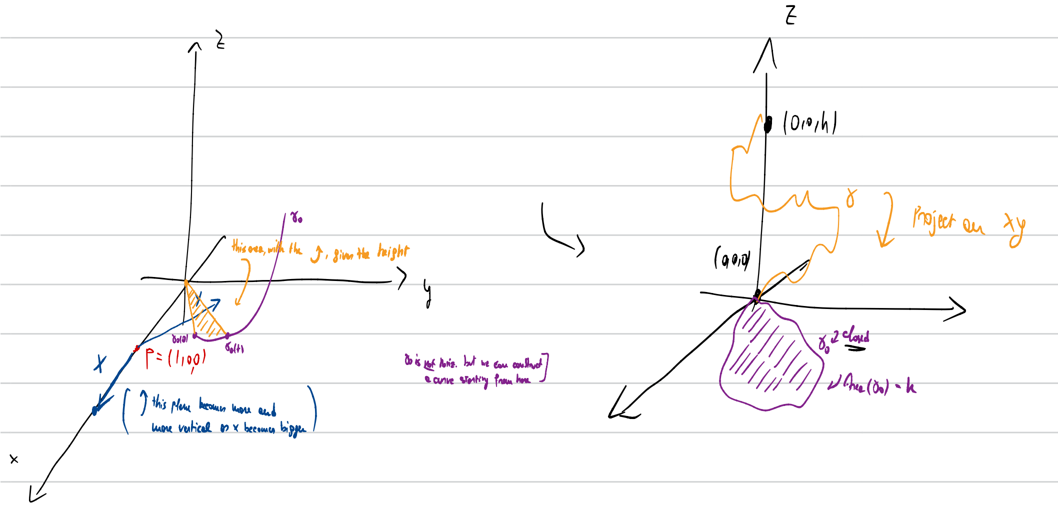

Figure 3.2: How to construct a horizontal curve in the Heisenberg group.

a very special property. Namely, it is easy to verify that

$$[X, [X, Y]] = [Y, [Y, X]] = 0$$

so the Baker-Campbell-Hausdorff formula (see Theorem B.5) only consists of three terms:

$$x \times y = x + y + \frac{1}{2}[x, y].$$

Notice that even a minor modification of the first example, namely $X = \partial_x$ and $Y = e^x \partial_y$, leads to a Lie algebra which is infinite-dimensional since

$$[X, X, \dots, [X, Y] \dots] = Y$$

for any number of the element X.

In the general case, it makes sense to put some restrictions on $\mathfrak{X} := \{X_1, \dots, X_k\}$ which are strictly related to the above examples.

(a) There is $N \in \mathbb{N}$ such that commutators of order $\geq N$ are zero. In particular,

$$\mathrm{Span}\langle X_1, \dots, X_k, \dots \rangle$$

is a finite-dimensional vector space.

(b) The dimension of the span is exactly equal to n, i.e., nonzero commutators form a Hörmander system.

The first assumption is easily justified. Indeed, let $\{X_1, \dots, X_n\}$ generate $\mathbb{R}^n$ at each point $p \in \mathbb{R}^n$ and consider the elliptic operator

$$L = \sum_{j=1}^{n} X_j^2 = \sum_{j=1}^{n} a_j(x) \partial_{x_j} \left(\sum_{i=1}^{n} a_i(x) \partial_{x_i} \right).$$

Solving the equation $Lu = 0$ is not trivial, but one could try to build an approximate solution by freezing the coefficients at some $x_0 \in \mathbb{R}^n$ so that L becomes an elliptic operator with constant coefficient. However, freezing

$$L = \partial_x^2 + x^2 \partial_y^2$$

at any point on the y-axis leads to the equation

$$\partial_x^2 u(x, y) = 0,$$

so there is a loss of information which results in a loss of regularity with respect to the actual solution.

One idea that allows performing a better solution analysis is to modify the vector fields X and Y slightly in such a way that they satisfy condition (1).

In any case, these two properties are equivalent to the fact that $\mathbb{R}^n$ can be equipped with a multiplication law $\cdot$ which is strictly related to the vector fields. We will come back to this in the next chapter, after a short introduction on Lie groups.

Part II

Fourier Analysis on Lie Groups

Chapter 4

Lie Groups and Lie Algebras

In this chapter, we introduce the notion of *Lie group* and set the ground for Fourier analysis on this class of spaces generalizing what we did on $\mathbb{R}^n$.

4.1 Introduction to Lie groups

Definition 4.1 (Lie group)**.** A *Lie group* $\mathbb{G}$ is an abstract group and a smooth manifold in which the two structures are compatible. In other words, the map

$$\mathbb{G} \times \mathbb{G} \ni (x,\, y) \longmapsto y^{-1} \cdot x \in \mathbb{G}$$

is smooth.

Remark 4.2. Notice that the compatibility between the two structures is also equivalent to requiring that

$$\mathbb{G} \times \mathbb{G} \ni (x,\, y) \longmapsto x \cdot y \in \mathbb{G} \quad \text{and} \quad \ni \mathbb{G}y \longmapsto y^{-1} \in \mathbb{G}$$

are both smooth maps.

Remark 4.3. In some books, a Lie group is defined as an analytic manifold rather than a smooth one. However, it can be proved that in this context smooth implies analytic - see, for example, [10].

From now on, the symbol $\mathbb{G}$ will always indicate a Lie group equipped with a multiplication law $\cdot$ compatible with the smooth manifold structure.

Definition 4.4. Fix $a \in \mathbb{G}$. The *left-translation* operator $\ell_a : \mathbb{G} \to \mathbb{G}$ is defined by setting

$$\ell_a(x) := a \cdot x$$

and, similarly, the *right-translation* operator $r_a : \mathbb{G} \to \mathbb{G}$ is defined by

$$r_a(x) := x \cdot a^{-1}.$$

Notice that we use a for ℓ_a and a^{-1} for r_a in order to make the composition laws coherent with each other, that is,

$$\ell_a \ell_b = \ell_{a\cdot b} \qquad \text{and} \qquad r_a r_b = r_{a\cdot b}.$$

Remark 4.5. It is easy to verify that both ℓ_a and r_a are diffeomorphisms of $\mathbb{G}$ with inverses $\ell_{a^{-1}}$ and $r_{a^{-1}}$ respectively.

Notation. Let $f \in C^\infty(\mathbb{G})$. We denote by $\check{f}$ the function

$$\check{f}(x) := f(x^{-1}) \in C^\infty(\mathbb{G}).$$

In a similar fashion, we introduce the translation operators L_a, $R_a : C^\infty(\mathbb{G}) \to C^\infty(\mathbb{G})$, for $a \in \mathbb{G}$, as follows:

$$L_a f(x) := f(a^{-1} \cdot x) = f \circ \ell_a^{-1}(x),$$

$$R_a f(x) := f(x \cdot a) = f \circ r_a^{-1}(x).$$

Recall that Lie group $\mathbb{G}$ is a smooth manifold and therefore the notion of tangent space at some $p \in \mathbb{G}$ is well-defined, for example, through derivations. Recall that

$$v : C^\infty(\mathbb{G}) \longrightarrow \mathbb{R}$$

is a *derivation* at $p \in \mathbb{G}$ if the following properties hold:

(a) If $f \equiv g$ in a neighbourhood of p, then $v(f) = v(g)$.

(b) The **Leibniz rule** holds, that is,

$$v(fg) = f(p)v(g) + g(p)v(f).$$

Proposition 4.6. *Let $\varphi : A \subseteq \mathbb{R}^q \to \mathbb{G}$ be a coordinate system around $p_0 \in \mathbb{G}$ and assume that $\varphi(0) = p_0$. Then $v \in T_{p_0}\mathbb{G}$ if and only if there exists $(a_1, \ldots, a_q) \in \mathbb{R}^q$ such that*

$$v(f) = \sum_{j=1}^{q} a_j \partial_{x_j}(f \circ \varphi)(0).$$

In other words, a derivation belongs to the tangent space if and only if it is a directional derivative.

Definition 4.7 (Vector field)**.** A *vector field* X defined on a Lie group $\mathbb{G}$ is a linear operator

$$X : C^\infty(\mathbb{G}) \longrightarrow C^\infty(\mathbb{G})$$

that satisfies the Leibniz rule, that is,

$$X(fg) = X(g)f + X(f)g.$$

We denote by $\mathfrak{X}(\mathbb{G})$ the space of all vector fields defined on $\mathbb{G}$ which is easily seen to be infinite-dimensional. Furthermore, for $p \in \mathbb{G}$ we have

$$f \longmapsto Xf(p) \in T_p\mathbb{G},$$

which means that a vector field X can be used to associate functions to tangent vectors.

Definition 4.8. The vector field $X \in \mathfrak{X}(\mathbb{G})$ is *left-invariant* if

$$X(L_a f) = L_a(Xf) \quad \text{for all } f \in C^\infty(\mathbb{G}) \text{ and all } p \in \mathbb{G}.$$

Example 4.9. In $\mathbb{R}^n$, the vector field defined by setting

$$X = \sum_{j=1}^n a_j(x) \frac{\partial}{\partial x_j}$$

is left-invariant if and only if $a_j(x) \equiv a_j$ is a constant function for all $j \in \{1, \ldots, n\}$.

Proposition 4.10. *Let $X \in \mathfrak{X}(\mathbb{G})$ be a left-invariant vector field and let "$v = X(e)$" be the derivation defined by setting*

$$v(f) = Xf(e).$$

Then, for all $x \in \mathbb{G}$ and for all $f \in C^\infty(\mathbb{G})$, the vector field satisfies

$$Xf(x) = v(L_x^{-1} f).$$

Conversely, given $v \in T_e\mathbb{G}$, the vector field defined by setting

$$X_v f(x) := v(L_x^{-1} f)$$

is left-invariant.

Proof. The first identity is trivial since

$$Xf(x) = L_{x^{-1}}(Xf)(e) = X(L_{x^{-1}} f)(e) = v(L_x^{-1} f).$$

To prove that X_v is a vector field, we need to check[1] that $g := X_v f \in C^\infty(\mathbb{G})$. A straightforward computation shows that

$$X_v f(x) = v(L_x^{-1} f) = \sum_{j=1}^n a_j \partial_{x_j} (L_x^{-1} f \circ \varphi)(0) = \sum_{j=1}^n a_j \partial_{x_j} f(x \cdot \varphi(t)),$$

which means that g is smooth since we can write it as the composition of smooth functions. We now show that X_v is left-invariant. Let $a \in \mathbb{G}$ and notice that

$$\begin{aligned} X_v(L_a f)(x) &= v(L_x^{-1} L_a f) = \\ &= v(L_{x \cdot a^{-1}} f) = \\ &= X_v f(a^{-1} \cdot x) = \\ &= L_a X_v(f)(x). \end{aligned}$$

$\square$

[1] It is clearly sufficient to prove that g is smooth in a neighbourhood of x.

It follows from the proposition that $v \mapsto X_v$ is a bijection between $T_e\mathbb{G}$ and the subset of $\mathfrak{X}(\mathbb{G})$ that consists of left-invariant vector fields.

Definition 4.11 (Lie Algebra)**.** The *Lie algebra* $\mathfrak{g}$ associated to a Lie group $\mathbb{G}$ is the algebra generated by left-invariant vector fields,

$$\mathfrak{g} := \{X \in \mathfrak{X}(\mathbb{G}) \ : \ X \text{ left-invariant}, \},$$

and it is a vector space of dimension q.

Remark 4.12. If $X,\, Y \in \mathfrak{g}$, then $[X,\, Y] \in \mathfrak{g}$.

The algebra isomorphism between $\mathfrak{g}$ and $T_e\mathbb{G}$ is given by the map $v \mapsto X_v$, where X_v can also be defined via the push-forward

$$X_v(x) = \ell(x)_* v.$$

4.2 Exponential map and one-parameter groups

Let $X \in \mathfrak{g}$ be a left-invariant vector field and let γ_x be the integral curve originating at some point $x \in \mathbb{G}$, namely the solution of the problem

$$\begin{cases} \gamma_x'(t) = X(\gamma_x(t)), \\ \gamma_x(0) = x. \end{cases}$$

We now claim that for all $x \in \mathbb{G}$ the following identity holds:

$$\gamma_x(t) = x \cdot \gamma_e(t). \tag{4.1}$$

One way to see this is to use the exponential map. Recall that by definition

$$\exp(tX)f(e) = f(\gamma_e(t)),$$

so for any $x \in \mathbb{G}$ it turns out that

$$\begin{aligned} \exp(tX)L_{x^{-1}}f(e) &= L_{x^{-1}}\exp(tX)f(e) \\ &= L_{x^{-1}}\left[f(\gamma_e(t))\right] \\ &= f(x \cdot \gamma_e(t)). \end{aligned}$$

The left-hand side is equal to $[L_{x^{-1}}f](\gamma_e(t))$, and thus (4.1) holds if we can prove that

$$[L_{x^{-1}}f](\gamma_e(t)) = f(\gamma_x(t)).$$

This is left as an exercise to the reader to get acquainted with the notions introduced so far.

Remark 4.13. Another way to prove (4.1) is to show that $x \cdot \gamma_e$ is the integral curve of X

starting from x. This is an immediate consequence of X left-invariant vector field:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} f(x \cdot \gamma_e(t)) &= \frac{\mathrm{d}}{\mathrm{d}t} L_{x^{-1}} f(\gamma_e(t)) \\ &= X(L_{x^{-1}} f)(\gamma_e(t)) \\ &= (Xf)(x \cdot \gamma_e(t)).\end{aligned}$$

The identity (4.1) is extremely important and carries several consequences, one of which is the existence of one-parameter groups. First, notice that

$$\gamma_e(\cdot) \text{ defined for } t \in (-\epsilon,\, \epsilon) \implies \gamma_x(\cdot) \text{ defined for } t \in (-\epsilon,\, \epsilon).$$

In particular, we have $\gamma_{\gamma_e(s)}(t) = \gamma_e(s) \cdot \gamma_e(t)$ and, since we know already (via the composition of the corresponding flows) that

$$\gamma_{\gamma_e(s)}(t) = \gamma(t+s),$$

we can infer that $\gamma_e(t+s) = \gamma_e(t) \cdot \gamma_e(s)$. This means that we can extend the domain of γ_e to coincide with the real line, and hence

$$\gamma_e \in C^\infty(\mathbb{R},\, \mathbb{G})$$

is a *smooth groups homomorphism* from $\mathbb{R}$ to $\mathbb{G}$. In this case we say that γ_e is a *one-parameter group* and we can easily define a one-to-one correspondence

$$\mathfrak{g} \longleftrightarrow \{\text{one-parameter groups}\}$$

where γ, a one-parameter group, gives a left-invariant vector field by setting

$$Xf(x) = \frac{\mathrm{d}}{\mathrm{d}t} f(x \cdot \gamma(t)).$$

Example 4.14. The one-parameter groups of $(\mathbb{R}^n,\, +)$ are all of the form

$$\gamma_v(t) := tv,$$

where $v \in \mathbb{R}^n$. Notice that from a topological point of view γ might not be a **closed** subgroup (e.g., the irrational line in $\mathbb{T}$ which is dense)

Let $\mathbb{G}$ be a Lie group and identify its Lie algebra $\mathfrak{g}$ with the tangent space.

Definition 4.15. Fix $v \in T_e\mathbb{G}$ and let γ_v be the one-parameter group satisfying the initial condition $\gamma_v'(0) = v$. We call *exponential of* $\mathbb{G}$ the mapping

$$\exp_{\mathbb{G}}(v) := \gamma_v(1).$$

Theorem 4.16. *The exponential map* $\exp_{\mathbb{G}}$ *is a local diffeomorphism from a neighbourhood* U *of* $0 \in T_e\mathbb{G}$ *to a neighbourhood* V *of the identity element* $e \in \mathbb{G}$*.*

Proof. Simply notice that $\exp_{\mathbb{G}}$ is smooth (since it is the restriction of $(t,\, v) \mapsto \gamma_v(t)$ which is smooth) and

$$d(\exp_{\mathbb{G}})_0 : T_e\mathbb{G} \longrightarrow T_e\mathbb{G}$$

is the identity map since $\gamma_v'(0) = v$ by definition. □

One might wonder whether or not $\exp_{\mathbb{G}}$ is a **global** diffeomorphism. However, it is easy to verify that any compact Lie group gives a counterexample since

$$T_e\mathbb{G} \cong \mathbb{R}^q$$

is only locally compact so it cannot be diffeomorphic to a compact manifold such as $\mathbb{G}$. For example,

$$\mathrm{SU}(2, \mathbb{C}) \cong S^3 \subset \mathbb{C}^2$$

is a compact Lie group and $\exp_{\mathbb{G}}$ is not a global diffeomorphism as one can easily see from the picture below.

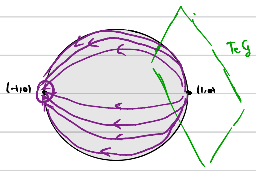

Figure 4.1: Counterexamples to $\exp_{\mathbb{G}}$ global diffeomorphism.

Furthermore, notice that for v and w in a small neighbourhood of $0 \in T_e\mathbb{G}$ we find that there exists u in the same neighbourhood such that

$$(\exp_{\mathbb{G}} v) \cdot (\exp_{\mathbb{G}} w) = \exp_{\mathbb{G}} u,$$

and u is given by the BCH formula. In particular, the series

$$v + w + \sum_{k_1+k_2 \geq 1} c_{k_1,\, k_2}, \quad c_{k_1,\, k_2} \in \mathcal{F}^{k_1,\, k_2}$$

is convergent in a small neighbourhood of the origin and this is what (more or less) leads to the definition of Lie groups with "analytic" in place of "smooth".

Example 4.17. Let $\mathbb{G} = \mathbb{R}$ and let $\mathbb{G}' = \mathbb{T}$. Then it is easy to verify that

$$\mathfrak{g} \cong \mathfrak{g}',$$

and both are isomorphic to $\mathbb{R}$.

Remark 4.18. This shows that two Lie groups with the same (up to isomorphism) Lie algebra are not necessarily isomorphic themselves; however, they are locally isomorphic.

Theorem 4.19. *Two Lie groups are locally isomorphic if and only if their Lie algebras are isomorphic.*

Remark 4.20. What we did until now only cares about the connected component of $\mathbb{G}$ and ignores all the others as if they did not exist. Therefore, assuming that $\mathbb{G}$ is connected is not restrictive in any sense.

4.3 Lie algebras and connection with Lie groups

We now introduce the notion of Lie algebra and show how is it related with the notion of Lie groups via left-invariant vector fields.

Definition 4.21. A Lie algebra $\mathfrak{g}$ is a vector space endowed with a bilinear operator $[\cdot,\, -] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ which is antisymmetric and satisfies the Jacobi identity

$$[x,\, [y,\, z]] + [z,\, [x,\, y]] + [y,\, [z,\, x]] = 0 \quad \text{for all } x,\, y,\, z \in \mathfrak{g}. \tag{4.2}$$

Example 4.22. The vector space of $n \times n$ real-valued matrices endowed with the bracket operator $[A,\, B] := AB - BA$ is a Lie algebra which is usually denoted by $\mathfrak{gl}(n,\, \mathbb{R})$. Similarly,

$$\mathfrak{sl}(n,\, \mathbb{R}) = \{A \in \mathfrak{gl}(n,\, \mathbb{R}) \,:\, \mathrm{Tr}(A) = 0\}$$

and

$$\mathfrak{so}(n,\, \mathbb{R}) = \{A \in \mathfrak{gl}(n,\, \mathbb{R}) \,:\, A \text{ is skew-symmetric}\}$$

are also Lie algebras with the same bracket operator. Notice that, in this case, the space of symmetric matrices in $\mathfrak{gl}(n,\, \mathbb{R})$ is **not** a Lie algebra.

Example 4.23. The Euclidean space $\mathbb{R}^3$ with the multiplication law given by the wedge product $\wedge$ is a Lie algebra which is isomorphic to $\mathfrak{so}(3,\, \mathbb{R})$ via

$$\mathbb{R}^3 \ni (v_1,\, v_2,\, v_3) \longmapsto \begin{pmatrix} 0 & v_3 & v_2 \\ -v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{pmatrix} \in \mathfrak{so}(3,\, \mathbb{R}).$$

Definition 4.24 (Homomorphism)**.** A Lie algebras homomorphism φ is a linear mapping that preserves the bracket, that is,

$$[x,\, y] = [\varphi(x),\, \varphi(y)].$$

Definition 4.25. We say that $\mathfrak{g}$ is a *commutative* Lie algebra if $[x,\, y] = 0$ for all $x,\, y \in \mathfrak{g}$.

We now introduce an useful notation for the commutator of linear spaces. More precisely, let $\mathfrak{h}$ be a linear subspace of a Lie algebra $\mathfrak{g}$. Then

$$[\mathfrak{g},\, \mathfrak{h}] := \mathrm{Span}\langle [x,\, y] \,:\, x \in \mathfrak{g},\, y \in \mathfrak{h} \rangle.$$

Definition 4.26. Let $\mathfrak{g}$ be a Lie algebra. A linear subspace $\mathfrak{h}$ is a *Lie subalgebra* if it is closed under the bracket, that is,

$$[\mathfrak{h},\, \mathfrak{h}] \subseteq \mathfrak{h}.$$

Moreover, we say that a Lie subalgebra $\mathfrak{h}$ is an *ideal* if

$$[\mathfrak{h},\, \mathfrak{g}] \subseteq \mathfrak{h}.$$

Remark 4.27. If $\mathfrak{h}$ is an ideal of $\mathfrak{g}$, then the quotient can be given a Lie algebra structure by setting

$$[x + \mathfrak{h},\, y + \mathfrak{h}] := [x,\, y] + \mathfrak{h}.$$

Notice that this is not well-defined (i.e., it depends on the choice of the representatives) if $\mathfrak{h}$ is a subalgebra which is not an ideal.

Definition 4.28. Let $\mathfrak{g}$ be a Lie algebra. The *centre* of $\mathfrak{g}$ is defined as

$$Z_{\mathfrak{g}} := \{x \in \mathfrak{g} \,:\, [x,\, \mathfrak{g}] = 0\}.$$

Remark 4.29. The linear subspace $Z_{\mathfrak{g}}$ is an ideal of $\mathfrak{g}$.

Definition 4.30. Let $\mathfrak{g}$ be a Lie algebra. The *derived algebra* of $\mathfrak{g}$ is defined as

$$\mathfrak{g}' := [\mathfrak{g},\, \mathfrak{g}].$$

The derived algebra of $\mathfrak{g}$ does not always coincide with $\mathfrak{g}$ itself. For example, if $\mathfrak{g} = \mathfrak{gl}(n,\, \mathbb{R})$ it is easy to verify that

$$[\mathfrak{g},\, \mathfrak{g}] \subseteq \mathfrak{sl}(n,\, \mathbb{R}),$$

and thus it is a proper ideal of $\mathfrak{g}$.

Example 4.31. Let $\mathfrak{t}(n,\, \mathbb{R})$ be the space of $n \times n$ real-valued upper-triangular matrices. Then $\mathfrak{t}(n,\, \mathbb{R})$ is a Lie algebra because $AB,\, BA \in \mathfrak{t}(n,\, \mathbb{R})$, but

$$[\mathfrak{t}(n,\, \mathbb{R}),\, \mathfrak{t}(n,\, \mathbb{R})] \subset \mathfrak{t}(n,\, \mathbb{R})$$

is a proper ideal because all its elements have the super-diagonal identically equal to zero.

The previous example suggests that we can iterate this construction a finite number of times. We obtain a sequence of derived algebras which are proper ideals of one another until

$$[\mathfrak{t}(n,\, \mathbb{R}),\, [\mathfrak{t}(n,\, \mathbb{R}),\, [\mathfrak{t}(n,\, \mathbb{R}),\, \dots,\, [\mathfrak{t}(n,\, \mathbb{R}),\, \mathfrak{t}(n,\, \mathbb{R})]] \dots] = 0.$$

This construction is what is usually called *descending central series* and can be done with any other Lie algebra, although there is no guarantee that it will eventually be zero.

Definition 4.32 (Nilpotent)**.** Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. If the descending central series stabilises at 0, then we say that $\mathfrak{g}$ is *nilpotent*.

Example 4.33. Let $\mathfrak{g}$ be the Heisenberg Lie algebra, namely the one generated as

$$\mathfrak{g} = \mathrm{Span}\langle \partial_x - \frac{y}{2}\partial_z,\, \partial_y + \frac{x}{2}\partial_z,\, \partial_z \rangle.$$

Then the derived algebra is $\mathfrak{g}' = \mathrm{Span}\langle \partial_z \rangle = \mathbb{R}$ and it is easy to verify that the $\mathfrak{g}$ is nilpotent of step two.

Theorem 4.34 (Ado)**.** *Every finite-dimensional Lie algebra over $\mathbb{R}$ (or $\mathbb{C}$) is isomorphic to a matrices Lie algebra.*

The original proof of this result due to Ado can be found in [1]. There are several others with new methods, but they are all reasonably complicated so we will just skip it.

Theorem 4.35 (Engel)**.** *Every finite-dimensional nilpotent Lie algebra over $\mathbb{R}$ (or $\mathbb{C}$) is isomorphic to a upper-triangular matrices Lie algebra.*

Example 4.36 (Heisenberg)**.** Let $\mathfrak{g}$ be the Heisenberg space, namely

$$\mathfrak{g} = \mathrm{Span}\langle \partial_x - \frac{y}{2}\partial_z,\, \partial_y + \frac{x}{2}\partial_z,\, \partial_z \rangle.$$

It is easy to see that this is isomorphic to the matrices algebra generated by

$$\begin{pmatrix} 0 & x & z \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix},$$

where

$$X = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The next theorem asserts that finite-dimensional Lie algebras are always given by $\mathrm{Lie}(\mathbb{G})$, where $\mathbb{G}$ is a Lie group which is uniquely determined if we require connected and simply connected.

Theorem 4.37 (Lie)**.** *Let $\mathfrak{g}$ be a finite-dimensional real Lie algebra. Then there exists a Lie group $\mathbb{G}$ such that $\mathfrak{g}$ i its Lie algebra. Furthermore, there exists a unique $\mathbb{G}$ which is connected and simply connected.*

One might wonder if the proof of this result is easier if we use Ado's theorem to restrict ourselves to a smaller class of Lie algebras. However, this is not the case. Let $\mathfrak{g}$ be a Lie algebra contained in $\mathfrak{gl}(n, \mathbb{R})$ and notice that

$$\mathfrak{gl}(n, \mathbb{R}) = \mathrm{Lie}(\mathrm{GL}(n, \mathbb{R})),$$

the set of all matrices with nonzero determinant. Now $\exp_{\mathbb{G}}(\mathfrak{g})$ is a "nice" submanifold close to the identity, but we get no information about the behaviour of faraway points.

Theorem 4.38. *Let $\mathfrak{g}$ be a Lie algebra and let $\mathbb{G}$ be the unique connected and simply connected Lie group with $\mathrm{Lie}(\mathbb{G}) \cong \mathfrak{g}$. If $\mathbb{H}$ is a connected Lie group with $\mathrm{Lie}(\mathbb{H}) \cong \mathfrak{g}$, then $\mathbb{H}$ is isomorphic to a quotient of $\mathbb{G}$ modulo a discrete central subgroup.*

Example 4.39. For example, $\mathbb{R}$ is the unique connected and simply connected Lie group with Lie algebra isomorphic to $\mathbb{R}$. The torus $\mathbb{T}$ is, in fact, isomorphic to a quotient of $\mathbb{R}$:

$$\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}.$$

Example 4.40. Let us consider

$$\mathrm{SU}(2, \mathbb{C}) = \left\{ \begin{pmatrix} z_1 & -\bar{z}_2 \\ z_2 & \bar{z}_1 \end{pmatrix} \,:\, |z_1|^2 + |z_2|^2 = 1 \right\}$$

and

$$\mathrm{SO}(3, \mathbb{R}) = \left\{ R_{\theta, v} \,:\, \theta \in [0, 2\pi),\, v \in \mathbb{R}^3,\, |v| = 1 \right\}.$$

It is easy to verify that $\mathrm{SU}(2, \mathbb{C})$ is diffeomorphic to $S^3 \subset \mathbb{C}^2$ so it is connected and simply connected, while $\mathrm{SO}(3, \mathbb{R})$ is connected but not simply connected. It can be proved that

$$\mathfrak{su}(2, \mathbb{C}) \cong \mathfrak{so}(3, \mathbb{R}),$$

and since the center of $\mathrm{SU}(2, \mathbb{C})$ is $\{\pm \mathrm{Id}\}$, then the unique possibility is that

$$\mathrm{SO}(3, \mathbb{R}) = \mathrm{SU}(2, \mathbb{C})/\{\pm \mathrm{Id}\},$$

which means that SU(2, $\mathbb{C}$) is a double covering of SO(3, $\mathbb{R}$).

Notation. So far we introduced an exponential map which sends a vector field to its flow,

$$\exp(tX)f(x) = f \circ \varphi_t(x),$$

and an exponential map associated to a Lie group $\mathbb{G}$ that sends $\mathfrak{g}$ to $\mathbb{G}$. The connection between them follows easily if we identify $X_v \in \mathfrak{g}$ with $X_v \in T_e\mathbb{G}$ since

$$\exp(tX_v)f(x) = f(x \exp_{\mathbb{G}}(tv)).$$

Since Ado's theorem asserts that every finite-dimensional Lie algebra $\mathfrak{g}$ over $\mathbb{R}$ is isomorphic to a matrices Lie algebra, it only makes sense to study properties of GL(n, $\mathbb{R}$).

Example 4.41. Recall that GL(n, $\mathbb{R}$) is a dense open set of $\mathbb{R}^{n\times n}$ so it quite clearly a Lie group. The tangent space is

$$T_e\mathrm{GL}(n,\, \mathbb{R}) = \mathrm{M}(n,\, \mathbb{R}),$$

the set of all $n \times n$ matrices, isomorphic to $\mathbb{R}^{n^2}$. Recall that for any $A \in \mathrm{M}(n,\, \mathbb{R})$ the exponential

$$\mathrm{e}^A = \sum_{n\in\mathbb{N}} \frac{A^n}{n!}$$

is well-defined and converges absolutely since

$$\left\| \sum_{n\in\mathbb{N}} \frac{A^n}{n!} \right\| \le \sum_{n\in\mathbb{N}} \frac{\|A^n\|}{n!} \le \sum_{n\in\mathbb{N}} \frac{\|A\|^n}{n!} = \mathrm{e}^{\|A\|}$$

and the exponential of a real number is well-defined. Now let $\gamma_A(t) := \mathrm{e}^{At}$ and notice that this map satisfies the relations

$$\gamma_A(s+t) = \gamma_A(t) \cdot \gamma_A(s) \quad \text{and} \quad \gamma_A(t) \cdot \gamma_A(-t) = \mathrm{id},$$

so $\gamma_A(t)$ belongs to GL(n, $\mathbb{R}$) for all $A \in \mathrm{M}(n,\, \mathbb{R})$. We can be more precise and prove that

$$\det \mathrm{e}^A = \mathrm{e}^{\mathrm{Tr}(A)},$$

which follows from the Jordan's normal form in $\mathbb{C}$ together with the obvious identity $\mathrm{e}^{PAP^{-1}} = P\mathrm{e}^A P^{-1}$, but we do not need to go as far. Anyway, γ_A is a one-parameter group and its derivative is given by

$$\gamma_A'(t) = A\mathrm{e}^{At}.$$

Now we want to prove that, given left-invariant vector fields X_A and X_B via the exponential identification, it turns out that

$$[X_A,\, X_B] = X_{[A,\, B]},$$

where $[A,\, B] = AB - BA$. Start by observing that

$$\begin{aligned}
X_A f(x) &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(x \exp_{\mathbb{G}}(A)) = \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(x\left(\mathrm{id} + tA + \mathcal{O}(t^2)\right)) = \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(x + xtA) = \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(x_1 + t(xA)_1,\, \dots,\, x_n + t(xA)_n) = \\
&= \sum_{i,\, j=1}^{n} (xA)_{i,\, j} \partial_{x_{i,\, j}} f(x)
\end{aligned}$$

from which it follows that

$$X_A(e) = \sum_{i,\, j=1}^{n} A_{i,\, j} \partial_{x_{i,\, j}}.$$

If B is another element, it is easy to see that

$$\begin{aligned}
&[X_A,\, X_B] f(e) = X_A(X_B f) - X_B(X_A f) = \\
&\quad= \sum_{i,\, j} A_{i,\, j} \partial_{i,\, j} \left(\sum_{k,\, \ell} (xB)_{k,\, \ell} \partial_{k,\, \ell} \right) f(e) - \sum_{i,\, j} B_{i,\, j} \partial_{i,\, j} \left(\sum_{k,\, \ell} (xA)_{k,\, \ell} \partial_{k,\, \ell} \right) f(e) \\
&\quad= \sum_{i,\, j} A_{i,\, j} \sum_{k,\, \ell} \left[\partial_{i,\, j} (xB)_{k,\, \ell} \partial_{k,\, \ell} \right] f(e) \\
&\quad= \sum_{i,\, j} \sum_{\ell} \left[B_{j,\, \ell} \partial_k f \right](e) \\
&\quad= \sum_{i,\, \ell} (AB)_{i,\, \ell} \partial_{i,\, \ell} f(e) = X_{[A,\, B]} f(e)
\end{aligned}$$

The Lie algebra $\mathfrak{gl}(n,\, \mathbb{R})$ can be identified with $\mathrm{M}(n,\, \mathbb{R})$ and the exponential map $A \mapsto \mathrm{e}^A$, while nice close to zero, is far from being injective everywhere. Indeed,

$$\exp\left(t \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right) = R_t,$$

the rotation of angle t, is periodic with respect to t.

4.4 Nilpotent Lie algebras

In this section, we take a closer look to nilpotent Lie algebras and Lie groups since in the final part of the course we will study the sublaplacian operator on a special subclass of nilpotent groups.

Theorem 4.42. *Let $\mathfrak{g}$ be a nilpotent Lie algebra of step m and define*

$$S(a,\, b) = a + b + \frac{1}{2}[a,\, b] + \sum_{k=3}^{m} c_k(a,\, b).$$

Then S is a multiplication law defining a Lie group structure on $\mathfrak{g}$ which Lie algebra is isomorphic to $\mathfrak{g}$.

Proof. First, notice that $S(a,\, 0) = a$ and $S(0,\, b) = b$ so the linear space 0 plays the role of the identity in $\mathbb{G}$. Moreover,

$$S(a,\, -a) = 0,$$

so each element $a \in \mathbb{G}$ has an inverse which is given by $-a$. The associative property,

$$S(S(a,\, b),\, c) = S(a,\, S(b,\, c),$$

on the other hand, is difficult to prove. We use Ado's theorem and reduce to a simpler case since working on a manifold $\cong \mathbb{R}^N$ allows one to assume the simple connectedness.

Associativity. The following result is given for granted:

Lemma 4.43. *There exists a neighbourhood of* 0 *in* $\mathfrak{gl}(n,\, \mathbb{R})$ *such that for all* $A,\, B \in U$ *the Baker-Campbell-Hausdorff formula converges.*

Let U be a neighbourhood of 0 in $\mathfrak{gl}(n,\, \mathbb{R})$ such that the restriction of the exponential map is a diffeomorphism with an open neighbourhood of the identity matrix V. Select a smaller neighbourhood V' in such a way that

$$V' \cdot V' \subset V,$$

and, similarly, select an even smaller one such that

$$V'' \cdot V'' \subset V'.$$

Let U' and U'' be the preimages of V' and V''. We can assume that U' satisfies the assumption of the lemma above (otherwise we can take a smaller one). It follows that

$$\forall a,\, b \in U',\, \exp(S(a,\, b)) \in V.$$

Now let $a,\, b,\, c \in U^{''}$. It is easy to see that $S(a,\, b)$, $S(a,\, c)$, $S(b,\, c) \in U'$ and thus

$$\mathrm{e}^{S(S(a,\,b),\,c)} = \mathrm{e}^{S(a,\,b)}\mathrm{e}^{c} = \mathrm{e}^{a}\mathrm{e}^{b}\mathrm{e}^{c} = \mathrm{e}^{a}\mathrm{e}^{S(b,\,c)} = \mathrm{e}^{S(a,\,S(b,\,c))}.$$

This shows that associativity rule holds in U'' and since this is a polynomial identity that holds for small parameters we can extend it everywhere by analytic continuation (since there is a finite number of addenda). $\square$

4.5 Homomorphisms of Lie algebras and dilations

Let $\Psi : \mathbb{G} \to \mathbb{G}'$, $\mathbb{G}$ connected, be a *homomorphism of Lie groups*, that is, a smooth[2] map which is also a group homomorphism. Then

$$\mathrm{d}\Psi_e : T_e\mathbb{G} \to T_e\mathbb{G}'$$

induces a map $\Psi_* : \mathfrak{g} \to \mathfrak{g}'$ composing with the isomorphisms described earlier in the chapter that identify the Lie algebra with the tangent at the identity element. Let $v \in T_e\mathbb{G}$ and let $\gamma_v(t)$ be the relative one-parameter group ($\gamma_v'(0) = v$). Then

$$\gamma_{v'}(t) := \Psi \circ \gamma_v(t)$$

is a one-parameter group in $\mathbb{G}'$ and hence turns out that $\Psi_*(v) = v'$.

Proposition 4.44. *The map Ψ_* is a Lie-algebra isomorphism, namely*

$$\Psi_*[u,\, v] = [\Psi_* u,\, \Psi_* v].$$

Proof. Let $v \in \mathfrak{g}$. Then $\gamma_v(t) = \exp_{\mathbb{G}}(tv)$ and

$$X_v f(x) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(x \cdot \exp_{\mathbb{G}}(tv)).$$

is the corresponding left-invariant vector field. Let $h \in C^\infty(\mathbb{G}')$, $f := h \circ \Psi \in C^\infty(\mathbb{G})$, and notice that

$$\begin{aligned} X_v f(x) &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} h\left(\Psi(x \cdot \exp_{\mathbb{G}}(tv))\right) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} h\left(\Psi(x) \cdot \exp_{\mathbb{G}'}(tv')\right) \\ &= Y_{v'} h(\psi(x)) = Y_{v'} f(x), \end{aligned}$$

which means that $X_v = Y_{\Psi_*(v)}$, where Y_w is the left-invariant vector field relative to w in the Lie algebra associated to $\mathbb{G}'$. $\square$

Proposition 4.45. *If $\mathbb{G}$ is connected, then Ψ_* uniquely determines Ψ.*

Proof. If Φ and Ψ are two Lie groups homomorphisms with $\Phi_* = \Psi_*$, then

$$\Phi(\exp_{\mathbb{G}}(tv)) = \exp_{\mathbb{G}'}(t\Phi_* v) = \exp_{\mathbb{G}'}(t\Psi_* v) = \Psi(\exp_{\mathbb{G}}(tv)).$$

This means that Φ and Ψ coincide in a neighbourhood U of the identity element $e \in \mathbb{G}$, which we can assume to be symmetric. Then

$$\Phi \equiv \Psi$$

on each power U^k of U, and hence we can consider $\mathcal{U} := \bigcup_{n\in\mathbb{N}} U^n$ subgroup of $\mathbb{G}$ since it is

[2] Requiring measurable here leads, surprisingly, to an equivalent definition.

closed under multiplication and inverse. By definition, $\mathcal{U}$ is open and, using the identity

$$\mathbb{G} \setminus \mathcal{U} = \bigcup_{p \notin \mathcal{U}} p \cdot \mathcal{U},$$

we also infer that $\mathcal{U}$ is closed. Since $\mathbb{G}$ is compact the only possibility is $\mathcal{U} = \mathbb{G}$. □

Corollary 4.46. *The map* $* : \mathrm{Hom}(\mathbb{G}, \mathbb{G}') \to \mathrm{Hom}(\mathfrak{g}, \mathfrak{g}')$ *is injective.*

Remark 4.47. In general, it is not surjective. If $\mathbb{G} = \mathbb{T}^2$ and $\mathbb{G}' = \mathbb{R}$, then the unique Lie algebra homomorphism $\mathrm{id} : \mathbb{R} \to \mathbb{R}$ cannot be lifted to a map from $\mathbb{T}^2$ to $\mathbb{R}$.

Theorem 4.48. *Let $\mathbb{G}$ be a connected and simply connected Lie group. If $\eta \in \mathrm{Hom}(\mathfrak{g}, \mathfrak{g}')$, then there exists $\Psi \in \mathrm{Hom}(\mathbb{G}, \mathbb{G}')$ such that*

$$\Psi_* = \eta.$$

Remark 4.49. If $\eta \in \mathrm{Hom}(\mathfrak{g}, \mathfrak{g}')$ and $\mathfrak{g}$ is nilpotent, then the subalgebra $\eta(\mathfrak{g}) \subset \mathfrak{g}'$ is also nilpotent.

Proof. We prove the result for $\mathbb{G}$ nilpotent. We proved that we can always assume $\mathbb{G} = \mathfrak{g}$ endowed with the Baker-Campbell-Hausdorff formula in place of the multiplication law. Notice that

$$\mathbb{G}' = \widetilde{\mathbb{G}'} / D,$$

so we can assume without loss of generality that $\mathbb{G}'$ is also connected and simply connected (up to replacing it with its universal covering). If $u,\, v \in \mathfrak{g}$, then

$$u \cdot v = u + v + \frac{1}{2}[u,\, v] + \sum_{k \geq 3}^{\iota} c_k(u,\, v) =: S(u,\, v),$$

and we can define

$$\Psi(u) := \exp_{\mathbb{G}'}(\eta(u)).$$

Then

$$\begin{aligned}
\Psi(u) \cdot \Psi(v) &= \exp_{\mathbb{G}'}(\eta(u)) \exp_{\mathbb{G}'}(\eta(v)) \\
&= \exp_{\mathbb{G}'}(S(\eta(u),\, \eta(v))) \\
&= \exp_{\mathbb{G}'}(\eta(S(u,\, v))) = \Psi(u \cdot v),
\end{aligned}$$

which means that $\Psi_* = \eta$. □

Definition 4.50 (Dilations)**.** A one-parameter family of automorphisms $\{\eta_t\}_{t>0} \subset \mathrm{Aut}(\mathfrak{g})$ is a family of *dilations* if there exists a basis $\{e_1,\, \ldots,\, e_q\}$ of $\mathfrak{g}$ and positive numbers $\lambda_1,\, \ldots,\, \lambda_q \in \mathbb{R}_+$ such that

$$\eta_t(e_j) = t^{\lambda_j} e_j.$$

Example 4.51. Consider e_1, e_2 orthonormal basis of $\mathbb{R}^2$ and define

$$\eta_t(e_j) = t^{\lambda_j} e_j,$$

where $\lambda_1 = 1$ and $\lambda_2 = 2$. The trajectories are half-parabolas as in the figure below:

Remark 4.52. We can decompose $\mathfrak{g}$ in such a way that $\lambda_1 < \cdots < \lambda_k$ and

$$\mathfrak{g} = \mathfrak{w}_1 \oplus \cdots \oplus \mathfrak{w}_k,$$

where $\eta_t \big|_{\mathfrak{w}_j} = t^{\lambda_j}\mathrm{id}$ for each $j = 1, \ldots, k$. Moreover

$$\eta_{t \cdot s} = \eta_t \circ \eta_s,$$

and since η_t is an automorphism of the Lie algebra, we also have that

$$\eta_t([u, v]) = [\eta_t u, \eta_t v].$$

This property has an immediate consequence, namely if $u \in \mathfrak{w}_j$ and $v \in \mathfrak{w}_\ell$, then the commutator $[u, v]$ belongs to $\mathfrak{w}_{j+\ell}$ (which may also coincide with $\{0\}$).

Clearly, $\lambda_j + \lambda_\ell > \max\{\lambda_j, \lambda_\ell\}$ so the existence of a family of dilations gives an upper bound on the number of commutators that one can take before it gives zero.

Proposition 4.53. *A Lie algebra with a family of dilations is nilpotent.*

Remark 4.54. The opposite assertion is not true, but finding a counterexample requires a big effort since one needs commutators sufficiently complicated.

Example 4.55. Let e_1, e_2, e_3 be vectors in $\mathbb{R}^3$ such that $[e_1, e_2] = e_3$. Then

$$\eta_t(e_1) = t^\lambda e_1 \quad \text{and} \quad \eta_t(e_2) = t^\mu e_2$$

implies $\eta_t(e_3) = t^{\lambda+\mu} e_3$. These are the unique dilations for which the e_i's are all eigenvectors.

Definition 4.56. If $\mathfrak{g}$ is a nilpotent Lie algebra with dilations δ_t, then $\mathbb{G} = \mathfrak{g}$ endowed with the BCH formula has a family of automorphisms such that

$$(\Psi_t)_* = \delta_t,$$

which is usually referred to as *dilations of a Lie group.*

4.6 Graded and stratified Lie algebras

We start with the definition of a particular class of Lie algebras which is rather important for the consequences the decomposition bears.

Definition 4.57. Let $\mathfrak{g}$ be a nilpotent group. If there exists a vector space decomposition

$$\mathfrak{g} = \mathfrak{w}_1 \oplus \cdots \oplus \mathfrak{w}_k$$

such that

$$[\mathfrak{w}_i, \mathfrak{w}_j] \subseteq \mathfrak{w}_{i+j},$$

then we say that $\mathfrak{g}$ is a *graded Lie algebra.*

Remark 4.58. Notice that a graded algebra is nilpotent, but it does not determine the step since, for example,

$$\mathbb{R}^3 = \mathbb{R} \oplus \mathbb{R} \oplus \mathbb{R} = \mathbb{R}^2 \oplus \mathbb{R}.$$

Lemma 4.59. *If $\mathfrak{g}$ is a graded Lie algebra, the maps η_t defined by setting*

$$\eta_t \Big|_{\mathfrak{w}_j} := t^j \, \mathrm{Id}_{\mathfrak{w}_j}$$

for $j = 1, \dots, k$ give a well-defined family of dilations on $\mathfrak{g}$.

Definition 4.60. If $\mathfrak{g}$ is a graded Lie algebra generated by $\mathfrak{w}_1$, we say that $\mathfrak{g}$ is *stratified*.

Example 4.61. The Heisenberg algebra described in Example 4.55 is stratified if and only if $\lambda = \mu$ since we need both vectors e_1 and e_2 in $\mathfrak{w}_1$.

Example 4.62. Let $\mathfrak{g}$ be the Lie algebra given by upper-triangular matrices. Then we can take $\mathfrak{w}_j$ to be the subalgebra generated by the matrices with **only** the jth over-diagonal nonzero, and this makes $\mathfrak{g}$ a stratified algebra.

Definition 4.63 (Sublaplacian)**.** Let $\mathfrak{g}$ be a stratified Lie algebra. If

$$\mathfrak{w}_1 = \mathrm{Span}\langle e_1, \, \dots, \, e_m \rangle,$$

then we can consider the differential operator

$$-\Delta := -\sum_{j=1}^{m} X_j^2.$$

We say that $-\Delta$ is the *sublaplacian operator* on $\mathbb{G}$ since it is homogeneous of degree two.

4.7 Homogeneous norms on Lie groups

Let $\mathfrak{g}$ be a nilpotent Lie algebra, $\{\delta_t\}_{t>0}$ a family of dilations and $\mathbb{G}$ the connected and simply connected Lie group given by $\mathfrak{g}$ equipped with the multiplication law (B.2).

Definition 4.64. A *homogeneous norm* on $\mathbb{G}$ is a map $|\cdot| : \mathbb{G} \to [0, \infty)$ satisfying the following properties:

(i) $x \mapsto |x|$ is continuous with respect to the topology on $\mathbb{G}$;

(ii) $|x| = 0$ if and only if $x = 0$, the zero vector of $\mathfrak{g}$;

(iii) if δ_t also denotes the dilations on $\mathbb{G}$, then $|\delta_t x| = t|x|$ for all $t > 0$.

We always refer to such a map as *homogeneous norm* because often, as we will see shortly, it does not satisfy the triangular inequality and thus it is not a norm.

Remark 4.65. Let $\mathfrak{g} = \mathfrak{v}_1 \oplus \cdots \oplus \mathfrak{v}_k$ be the decomposition induces by the dilations in such a way that

$$\delta_t \Big|_{\mathfrak{v}_i} = t^{\lambda_i} \mathrm{Id}_{\mathfrak{v}_i}.$$

If $\|\cdot\|$ is the norm inducing the topology on $\mathfrak{g}$, then it is easy to verify that

$$|x| := \sum_{i=1}^{k} \|x_i\|^{\frac{1}{\lambda_i}}$$

is a homogeneous norm. This shows that in (i) continuity is the "right" requirement because $|\cdot|$ is not even C^1 as soon as $\lambda_i \geq 2$ for some i.

Proposition 4.66. *A homogeneous norm satisfies a quasi-triangular inequality. In other words, there exists a constant $C \geq 1$ such that*

$$|x \cdot y| \leq C(|x| + |y|) \quad \textit{for all } x, y \in \mathbb{G}. \tag{4.3}$$

Proof. The set $B := \{x \in \mathbb{G} \,:\, |x| \leq 1\}$ is compact and so is $B^2 := B \cdot B$ because the multiplication law is continuous. Consequently, there exists a constant $C \geq 1$ such that

$$B^2 \subseteq \{x \in \mathbb{G} \,:\, |x| \leq C\},$$

which can be easily rewritten as follows:

$$|x|,\, |y| \leq 1 \implies |x \cdot y| \leq C.$$

Now take x, $y \neq 0$ and notice that, if $t := \frac{1}{|x|+|y|}$, then $|\delta_t x| \leq 1$ and $|\delta_t y| \leq 1$. Applying the inequality above shows that

$$|\delta_t x \cdot \delta_t y| \leq C \implies |x \cdot y| \leq \frac{C}{t} = C(|x| + |y|).$$

□

Remark 4.67. It follows from the definition and (4.3) that the map $d(x, y) := |x^{-1} \cdot y|$ is a left-invariant homogeneous quasi-distance (distance if $C = 1$):

$$d(x, y) = d(zx, zy) \quad \text{and} \quad d(\delta_t x, \delta_t y) = t d(x, y).$$

Notice that a quasi-distance does not induce another topology on $\mathbb{G}$ in general, and this is the reason why we require $|\cdot|$ to be continuous with respect to the topology of $\mathbb{G}$

Remark 4.68. The quasi-distance d is, in general, not smooth. However, it is possible to obtain smoothness outside of the origin by taking the Euclidean sphere $S_{\mathfrak{g}}$ in $\mathbb{G}$ and setting

$$|x| := \frac{1}{t} \quad \text{if } \delta_t x \in S_{\mathfrak{g}}.$$

The reader should check that this is a homogeneous norm which is also smooth by the implicit function theorem (except at $x = 0$).

In **stratified** groups, it is easy to construct a homogeneous norm that satisfies (4.3) with constant $C = 1$. For this, let $\mathfrak{v}_1 \oplus \cdots \oplus \mathfrak{v}_k$ be a stratification and write

$$\mathfrak{v}_1 = \operatorname{Span}\langle e_1, \ldots, e_q \rangle.$$

The corresponding vector fields $X_1, \ldots, X_q$ satisfy the Hörmander condition on $\mathbb{G}$, and hence we can consider the Carnot-Carathéodory distance

$$d(x, y) := \{L_{\mathfrak{X}}(\gamma) \,:\, \gamma \text{ horizontal curve joining } x \text{ and } y\},$$

where $L_{\mathfrak{X}}$ is the length defined in Section 3.5.2. It is easy to verify that this is a homogeneous distance which is left-invariant and satisfies (4.3) with $C = 1$.

Theorem 4.69 ([2]). *Every nilpotent connected and simply connected group with dilations has a homogeneous norm which satisfies the triangular inequality (i.e., $C = 1$) and is smooth outside of the origin.*

Proposition 4.70. *If $|\cdot|$ and $|\cdot|'$ are two homogeneous norm on a Lie group $\mathbb{G}$ equipped with the same family of dilations, then they are equivalent.*

Chapter 5

Fourier Analysis on Lie Groups

The main goal of this chapter is to develop Fourier analysis on unimodular Lie groups and investigate fundamental solutions and local solvability of differential operators.

Theorem A. Let $\mathbb{G}$ be a locally compact and separable group. Then there exists a locally finite left-invariant measure on $\mathbb{G}$, which is unique up to a multiplicative constant.

If $\mathbb{G}$ is a unimodular group, then the convolution of two functions (or a function and a distribution) can easily be defined as in the Euclidean setting replacing $\mathrm{d}x$ with the Haar measure. Namely, if $\Phi \in \mathcal{D}'(\mathbb{G})$ and $f \in C_c(\mathbb{G})$, one can set

$$\langle \Phi * f,\, g\rangle := \langle \Phi, g * \check{f}\rangle.$$

We show that several properties valid in the Euclidean setting can be extended to unimodular Lie groups with little effort. Next, we introduce left-invariant differential operators

$$D : C_c^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$$

and prove two characterizations in terms of polynomials defined on the corresponding Lie algebra. In conclusion, we show that some operators have fundamental solutions:

Theorem B. Let Q be the homogeneous dimension of $\mathbb{G}$ and let D be a left-invariant operator of homogeneity order $\beta \le Q$, that is,

$$D(f \circ \delta_t) = t^\beta (Tf) \circ \delta_t.$$

Assume also that D and ${}^{\mathrm{t}}D$ are hypoelliptic. Then D has a fundamental solution with homogeneity order $-Q + \beta$.

5.1 Existence of Haar measures

In this section, we recall a few properties concerning the existence and uniqueness of left-invariant measures defined on a **topological** group $\mathbb{G}$.

Definition 5.1 (Radon measure)**.** Let X be a Hausdorff topological space. A *Radon measure* is a measure on $\mathcal{B}(X)$ that satisfies the following properties:

(i) It is finite on all compact sets;

(ii) outer regular on all Borel sets; and

(iii) inner regular on all open sets.

Definition 5.2 (Push-forward)**.** Let μ be a positive Radon measure on $\mathfrak{X}$, and let $f : \mathfrak{X} \to \mathfrak{Y}$ be a Borel function. The *push-forward* measure of μ via f is defined by setting

$$f_{\#}\mu(E) := \mu\left(f^{-1}(E)\right) \quad \text{for all } E \in \mathcal{B}(\mathfrak{Y}).$$

Remark 5.3. Let $(\mathfrak{X}, \mathcal{B}(\mathfrak{X}))$ and $(\mathfrak{Y}, \mathcal{B}(\mathfrak{Y}))$ be two measure spaces, and let μ be a positive Radon measure on $\mathfrak{X}$. Then the push-forward $f_{\#}\mu$ is a well-defined measure on $\mathcal{B}(\mathfrak{Y})$.

Let $\mathbb{G}$ be a topological group. For any $y \in \mathbb{G}$, we denote by ℓ_y the left-multiplication map $(x \mapsto y \cdot x)$ and by r_y the right-multiplication map $(x \mapsto x \cdot y^{-1})$.

Definition 5.4 (Invariant Measure)**.** Let μ be a measure defined on $\mathbb{G}$. We say that μ is a *left-invariant* measure if

$$(\ell_y)_{\#}\,\mu = \mu \quad \text{for all } y \in \mathbb{G},$$

and a *right-invariant* one if

$$(r_y)_{\#}\,\mu = \mu \quad \text{for all } y \in \mathbb{G}.$$

A measure μ which is both left-invariant and right-invariant is simply called *invariant*.

Theorem 5.5. *Let $\mathbb{G}$ be a compact group. Then there exists a unique left-invariant (or right-invariant) probability measure on $\mathbb{G}$, called* **Haar measure***.*

We present the proof of this result under the additional assumption $\mathbb{G}$ commutative. The reader will find in this paper the proof of the result in the general case.

Idea of the proof. Let $\mathbb{G}$ be a commutative group, and let $\mathcal{P}(\mathbb{G})$ be the space of probability measures defined on $\mathbb{G}$. For a given $g \in \mathbb{G}$, consider

$$\mathcal{P}_g := \left\{\mu \in \mathcal{P}(\mathbb{G}) \;\middle|\; (\ell_g)_{\#}\,\mu = \mu\right\},$$

the subset of $\mathcal{P}(\mathbb{G})$ that contains all the g-invariant probability measures.

Step 1. We claim that, for any $g \in \mathbb{G}$, the set $\mathcal{P}_g$ is nonempty. Fix $\mu_0 \in \mathcal{P}$ and define, for every $n \in \mathbb{N}$, the probability measure

$$\mu_n := \frac{\mu_0 + (\ell_g)_{\#}\,\mu_0 + \cdots + (\ell_{g^n})_{\#}\,\mu_0}{n+1} \in \mathcal{P}(\mathbb{G}),$$

where $g^n := g \cdots g$ is the product of n copies of g. By compactness, there exists a subsequence μ_{n_k} weakly$\star$ converging to a measure μ_∞. It is easy to see that

$$(\ell_g)_{\#}\,\mu_{n_k} \to \mu_\infty \implies (\ell_g)_{\#}\,\mu_\infty = \mu_\infty,$$

which means that μ_∞ is g-invariant or, in other words, $\mu_\infty \in \mathcal{P}_g$.

Step 2. Let g, $h \in \mathbb{G}$ be two elements, let $\mu_0 \in \mathcal{P}_g$ be an invariant measure, and let μ_∞ be the weak$\star$ limit of the sequence

$$\mu_n := \frac{\mu_0 + (\ell_h)_\# \mu_0 + \cdots + (\ell_{h^n})_\# \mu_0}{n+1} \in \mathcal{P}(\mathbb{G}).$$

Since $\mathcal{P}_g$ is weakly$\star$ closed, we conclude that $\mu_\infty \in \mathcal{P}_g \cap \mathcal{P}_h$. An inductive argument proves that the family $\{\mathcal{P}_g\}_{g\in\mathbb{G}}$ has the finite intersection property which, by compactness, implies

$$\bigcap_{g\in\mathbb{G}} \mathcal{P}_g \neq \varnothing.$$

Step 3. We claim that the intersection above only consists of one element. In order to prove that, we define the "convolution of two measures" by setting

$$\mu_1 * \mu_2(E) := (\mu_1 \times \mu_2)\left(\{(x_1, x_2) \mid x_1 + x_2 \in E\}\right).$$

The reader may prove that the convolution is commutative, and also that

$$\mu_1 * \mu_2 = \mu_1, \tag{5.1}$$

whenever μ_1 is a left-invariant measure. Now, if λ, $\mu \in \cap_{g\in\mathbb{G}}\mathcal{P}_g$ are two invariant measures, then property (5.1) implies the uniqueness:

$$\mu = \mu * \lambda = \lambda * \mu = \lambda \implies \mu = \lambda.$$

$\square$

Theorem 5.6. *Let $\mathbb{G}$ be a locally compact and separable group. Then there exists a locally finite left-invariant measure on $\mathbb{G}$, unique up to a multiplicative constant.*

A proof of this theorem can be found in most geometric measure theory books, but for a quick overview the reader may consult this paper.

Remark 5.7. Notice that Theorem 5.6 gives the existence of a left-invariant measure μ and, with a similar proof, of a right-invariant measure λ. However, in general $\lambda \neq \mu$.

Definition 5.8 (Unimodular)**.** A locally compact and separable group $\mathbb{G}$ whose left-invariant measure is right-invariant is called *unimodular*.

We can immediately give a characterization of unimodular Lie groups in terms of the determinant of the adjoint representation; we first start off with a few definitions.

Definition 5.9 (Adjoint)**.** Let $\mathbb{G}$ be a Lie group and let

$$\Psi : \mathbb{G} \to \mathrm{Aut}(\mathbb{G})$$

be the mapping that sends g to Ψ_g, where Ψ_g is the inner automorphism $h \mapsto ghg^{-1}$. The *adjoint map* is defined as

$$\mathfrak{Ad}_g = \mathrm{d}(\Psi_g)_e : \mathfrak{g} \to \mathfrak{g}$$

using the identification $T_e\mathbb{G} \cong \mathfrak{g}$. The corresponding mapping

$$\mathfrak{Ad} : \mathbb{G} \ni g \longmapsto \mathfrak{Ad}_g \in \mathrm{Aut}(\mathfrak{g})$$

is called *adjoint representation* of the group $\mathbb{G}$.

Definition 5.10 (Adjoint)**.** Let $\mathfrak{g}$ be a Lie algebra over some field. Given $x \in \mathfrak{g}$ we can define the *adjoint action* as

$$\mathfrak{ad}_x : \mathfrak{g} \ni y \longmapsto [x, y] \in \mathfrak{g}.$$

Since the bracket is bilinear, this determines a linear mapping

$$\mathfrak{ad} : \mathfrak{g} \ni x \longmapsto \mathfrak{ad}_x \in \mathrm{End}(\mathfrak{g}),$$

which is usually called *adjoint representation* of the Lie algebra $\mathfrak{g}$.

Remark 5.11. A Lie group $\mathbb{G}$ is unimodular if and only if

$$|\det(\mathfrak{Ad}_g)| = 1 \quad \text{for all } g \in \mathbb{G}.$$

For a connected Lie group $\mathbb{G}$, this is equivalent to requiring that the trace of $\mathfrak{ad}_g$ is zero for all $g \in \mathfrak{g}$, as the reader might show to practice.

Remark 5.12. The following classes of groups are unimodular:

(a) compact groups;

(b) discrete groups;

(c) commutative locally compact groups;

(d) connected reductive Lie groups;

(e) locally compact nilpotent groups (in particular, nilpotent Lie groups).

Corollary 5.13. *If* $\mathbb{G}$ *is unimodular and* $\mathrm{d}x$ *is the invariant measure, then*

$$\int_{\mathbb{G}} f(x)\,\mathrm{d}x = \int_{\mathbb{G}} f(x^{-1})\,\mathrm{d}x.$$

Now let $\mathbb{G}$ be a nilpotent, connected and simply connected Lie group endowed with the BCH-coordinates (also known as *canonical coordinates of the first kind*) and let

$$\mathfrak{g} = \mathfrak{g}_0 \supset \mathfrak{g}_1 \supset \cdots \supset \mathfrak{g}_m = \{0\}$$

be the *central descent sequence* given by the derivate groups. Let $\mathfrak{v}_1, \ldots, \mathfrak{v}_m$ be linear spaces such that the following decomposition holds for all $0 \leq k \leq m-1$:

$$\mathfrak{g}_k = \mathfrak{g}_{k+1} \oplus \mathfrak{v}_{k+1}.$$

Then $\mathfrak{g} = \mathfrak{v}_1 \oplus \cdots \oplus \mathfrak{v}_m$ and $x \in \mathfrak{g}$ can be written as a m-tuple $(x_1, \ldots, x_m)$ with $x_k \in \mathfrak{v}_k$ for all $k = 1, \ldots, m$. Consequently, the product between any two elements is given by

$$x \cdot y = \Bigg(x_1 + y_1, x_2 + y_2 + \frac{1}{2}[x_1, y_1], \ldots, x_k + y_k + P_k(x_1, \ldots, x_{k-1}, y_1, \ldots, y_{k-1}), \\ \ldots, x_m + y_m + P_m(x_1, \ldots, x_{m-1}, y_1, \ldots, y_{m-1}) \Bigg),$$

where each P_j is a polynomial of order j. Let $\mathrm{d}x$ be the Lebesgue measure on $\mathfrak{g}$ (which is a linear space) and notice that the change of variables formula implies

$$\int_{\mathbb{G}} f(a \cdot x)\,\mathrm{d}x = \int_{\mathbb{G}} f(x)\,\mathrm{d}x.$$

This is easy to check because we can change one variable at a time starting from the last one (x_m). The reason is that, if we look at the formula above we have

$$x_m + y_m + P_m(x_1, \ldots, x_{m-1}, y_1, \ldots, y_{m-1}),$$

and this depends **linearly** on x_m (that is, equal to $x_m + c$). We now summarize the result we obtained here in the following proposition:

Proposition 5.14. *Let $\mathbb{G}$ be a nilpotent, connected and simply connected Lie group. Then the Lebesgue measure is invariant (both left and right).*

Remark 5.15. We can use the exponential map to introduce the so-called *canonical coordinates of the second kind*. Let $\mathfrak{g} = \mathfrak{v}_1 \oplus \cdots \oplus \mathfrak{v}_m$ and consider the mapping

$$\Phi(x_1, \ldots, x_m) := \exp_{\mathbb{G}}(x_1) \cdots \exp_{\mathbb{G}}(x_m).$$

Then Φ is a local diffeomorphism, which can be used to parametrize $\mathbb{G}$ via coordinates which are the ones we just mentioned.

5.1.1 Homogeneous dimension

Let $\mathbb{G}$ be a homogeneous group, let $\mathfrak{g} = \mathfrak{v}_1 \oplus \cdots \oplus \mathfrak{v}_m$ be the decomposition induced by the family of dilations $\{\delta_t\}_{t>0}$ and $\mathrm{d}x$ the Haar measure. Then it is easy to verify that

$$\int_{\mathbb{G}} f(\delta_t x)\,\mathrm{d}x = t^{-Q} \int_{\mathbb{G}} f(x)\,\mathrm{d}x, \tag{5.2}$$

where Q is the so-called *homogeneous dimension of* $\mathbb{G}$ given by the sum

$$Q = \sum_{i=1}^{m} \lambda_i \dim(\mathfrak{v}_i),$$

where $\delta_t \big|_{\mathfrak{v}_i} \equiv t^{\lambda_i} \, \mathrm{id}_{\mathfrak{v}_i}$. Using $f = \chi_A$ as a test function in (5.2) gives the identity

$$\mu(\delta_t A) = t^Q \mu(A) \quad \text{for all } A \subseteq \mathbb{G}.$$

Furthermore, the integral

$$\int_{|x|<1} |x|^{-\alpha}\,\mathrm{d}x$$

converges if and only if $\alpha < Q$. This “shows” that the homogeneous dimension plays the role of the topological dimension in Lie groups equipped with a Haar measure.

5.2 Convolution on Lie groups

Let $\mathbb{G}$ be a **unimodular** Lie group, $\mathrm{d}y$ the invariant measure and $f, g \in C_c(\mathbb{G})$. The *convolution* of f and g is the function defined by setting

$$f * g(x) := \int_{\mathbb{G}} f(x \cdot y^{-1}) g(y) \,\mathrm{d}y.$$

A simple change of variables shows that

$$f * g(x) = \int_{\mathbb{G}} f(y) g(y^{-1} \cdot x) \,\mathrm{d}y,$$

but this is different from the convolution

$$g * f(x) = \int_{\mathbb{G}} f(y) g(x \cdot y^{-1}) \,\mathrm{d}y.$$

The reason is that the product $y^{-1} \cdot x$ is not equal anymore to $x \cdot y^{-1}$ for all x, $y \in \mathbb{G}$ unless the group is commutative. We can actually prove a stronger characterization:

Proposition 5.16. *The convolution is commutative on $C_c(\mathbb{G})$ if and only if $\mathbb{G}$ is abelian.*

Proof. One implication is trivial. Therefore, assume that that $\mathbb{G}$ is not abelian and take any two points $x, y \in \mathbb{G}$ such that

$$x \cdot y \neq y \cdot x.$$

Let U_1 and U_2 be neighborhoods of x and y respectively which are disjoint. By continuity of the product, we can always find $V_1 \ni x$ and $V_2 \ni y$ such that

$$V_1 \cdot V_2 \subseteq U_1 \quad \text{and} \quad V_2 \cdot V_1 \subseteq U_2.$$

If $f \in C_c(V_1)$ and $g \in C_c(V_2)$, the well-known supports inclusions

$$\mathrm{spt}(f * g) \subseteq \mathrm{spt}(f) \cdot \mathrm{spt}(g) \subseteq U_1 \quad \text{and} \quad \mathrm{spt}(g * f) \subseteq \mathrm{spt}(g) \cdot \mathrm{spt}(f) \subseteq U_2$$

show that $f * g \neq g * f$ as they are supported on disjoint sets. □

Remark 5.17. If $f \in C_c^\infty(\mathbb{G})$ and $g \in C_c(\mathbb{G})$, then the convolution belongs to $C_c^\infty(\mathbb{G})$ and

$$X_j(f * g) = (X_j f) * g.$$

Definition 5.18. Let μ be a Radon measure on $\mathbb{G}$ and let $f \in C_c(\mathbb{G})$. The convolution between these two objects is defined as follows:

$$f * \mu(x) := \int_{\mathbb{G}} f(x \cdot y^{-1}) \,\mathrm{d}\mu(y).$$

In a similar fashion, one can define

$$\mu * f(x) := \int_{\mathbb{G}} f(y^{-1} \cdot x) \,\mathrm{d}\mu(y).$$

Remark 5.19. If $\mu := \delta_a$ is the Dirac delta, then it is easy to verify that

$$f * \delta_a(x) = \int_{\mathbb{G}} f(x \cdot y^{-1})\delta_a(y)\,\mathrm{d}y = f(x \cdot a^{-1}) = R_{a^{-1}} f(x),$$

and, similarly,

$$\delta_a * f(x) = \int_{\mathbb{G}} f(y^{-1} \cdot x)\delta_a(y)\,\mathrm{d}y = f(a^{-1} \cdot x) = L_a f(x).$$

This means that for finite Radon measures μ and ν we can define the convolution as

$$\int_{\mathbb{G}} f(x)\,\mathrm{d}(\mu * \nu)(x) = \int_{\mathbb{G}} \int_{\mathbb{G}} f(x \cdot y)\,\mathrm{d}\mu(x)\mathrm{d}\nu(y)$$

in such a way that the identity $\delta_a * \delta_b = \delta_{a \cdot b}$ holds.

Definition 5.20. Let $\Phi \in \mathcal{D}'(\mathbb{G})$ be a distribution. Then we can define the left and right translations by setting

$$\langle L_a \Phi,\, f\rangle := \langle \Phi, L_{a^{-1}} f\rangle,$$

$$\langle R_a \Phi,\, f\rangle := \langle \Phi, R_{a^{-1}} f\rangle.$$

The convolution between a distribution $\Phi \in \mathcal{D}'(\mathbb{G})$ and a function $f \in C_c(\mathbb{G})$ is defined in the usual way by setting

$$\langle \Phi * f,\, g\rangle := \langle \Phi, g * \check{f}\rangle,$$

where $\check{f}(x) := f(x^{-1})$. Notice that the order of $g * \check{f}$ is important since $\mathbb{G}$ might not be commutative. It is easy to verify that $\Phi * f$ is actually a function defined by

$$\Phi * f(x) = \langle \Phi, L_x \check{f}\rangle$$

and, similarly,

$$f * \Phi(x) = \langle \Phi, R_{x^{-1}} \check{f}\rangle.$$

Remark 5.21. We can define the convolution between two distributions in a similar fashion, but there is something to take into account now. Since $\Phi * f$ is a function, to define

$$\langle \Psi, \Phi * f\rangle,$$

we need $\Phi * f$ to be **compactly supported**. Therefore,

$$\langle \Psi * \Phi, f\rangle := \langle \Psi, \check{\Phi} * f\rangle$$

is well-defined provided that, for example, $\Psi \in \mathcal{D}'(\mathbb{G})$, $f \in C_c(\mathbb{G})$ and Φ is a compactly supported distribution.

5.2.1 Schwartz spaces

Defining $\mathscr{S}(\mathbb{G})$ for a general Lie group $\mathbb{G}$ is hard, but if we require nilpotent (connected and simply connected), then we can identify it with $\mathfrak{g}$ via (B.2) and write

$$\mathscr{S}(\mathbb{G}) := \mathscr{S}(\mathfrak{g}),$$

where $\mathfrak{g} \cong \mathbb{R}^N$ for some $N \in \mathbb{N}$. That said, there are two problems we need to take care of:

(i) The derivatives are rapidly decreasing, but what can we say about $X_j f$?

(ii) The estimate asserts that
$$|\partial^\alpha f(x)| = o(|x|^{-M}),$$
however, the norm $|\cdot|$ is the Euclidean one. What is the relation with the homogeneous norm on $\mathbb{G}$, if $\mathbb{G}$ is homogeneous?

The problem (i) is easy to solve since we can always write

$$X_j f(x) = \partial_{x_j} f(x) + \sum a_{jk}(x) \partial_{x_k} f(x),$$

where a_{jk} is a **polynomial function** that vanishes at the origin for each k. Since a_{jk} is a polynomial, we easily infer that

$$|\partial^\alpha f(x)| = o(|x|^{-M}) \implies |X_{j_1} \cdots X_{j_k} f(x)| = o(|x|^{-M}).$$

The problem (ii), on the other hand, is harder to deal with and outside the scopes of this course.

5.3 Left-invariant differential operators

The goal of this section is to characterize left-invariant differential operators on Lie groups, but we first need to recollect some definitions on manifolds.

Definition 5.22. Let $\mathcal{M}$ be a manifold. A *differential operator* on $\mathcal{M}$ is a map

$$D : C_c^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$$

such that we can write
$$Df(\varphi(t)) = \sum_{|\alpha| \le m} a_\alpha(t) \partial_t^\alpha (f \circ \varphi)(t)$$

for all $f \in C_c^\infty(\mathcal{M})$ and all local coordinates φ, where a_α are $C^\infty(\mathcal{M})$ coefficients.

Definition 5.23. Let D be a differential operator on a manifold $\mathcal{M}$ and fix $x_0 = \varphi(t_0)$ for some local coordinate φ. Then the *order of D at x_0* is

$$\operatorname{ord}(D, x_0) := \inf \{k \in \mathbb{N} \, : \, a_\alpha \equiv 0 \text{ for all } \alpha \, : \, |\alpha| = k+1\}.$$

Exercise 5.1. Show that the order $\operatorname{ord}(D, x_0)$ does not depend on the choice of the local coordinate around x_0.

Theorem 5.24 (Peetre)**.** *Let $\mathcal{M}$ be a manifold. A linear operator $D : C_c^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$ is a differential operator if and only if*

$$\operatorname{spt}(Df) \subseteq \operatorname{spt}(f) \quad \text{for all } f \in C_c^\infty(\mathcal{M}).$$

The proof of this result is far beyond the purpose of this course. For more information, see the original articles [8], [9] and the related works.

Definition 5.25. Let $\mathbb{G}$ be a Lie group and D a differential operator. We say that D is *left-invariant*, and we write $D \in \mathscr{D}(\mathbb{G})$, if

$$D(L_x f) = L_x(Df).$$

Theorem 5.26. *Let $\mathfrak{g}$ be a Lie algebra and let $p(t)$ be a polynomial on $\mathfrak{g}$. Define*

$$D_p f(x) := p(\partial_t)\Big|_{t=0} f(x \exp_{\mathbb{G}}(t))$$

for $x \in \mathbb{G}$. Then D_p is a left-invariant differential operator and, conversely, given $D \in \mathscr{D}(\mathbb{G})$ there exists a unique polynomial p such that $D = D_p$.

Proof. First, we notice that D_p is linear so by Theorem 5.24 to prove that D_p is a differential operator it suffices to show that

$$x \notin \operatorname{spt}(f) \implies x \notin \operatorname{spt}(D_p f).$$

This is trivial because there exists $\delta > 0$ such that for all $|t| < \delta$ one has $f(x \exp_{\mathbb{G}}(t)) = 0$ and, given that $p(\partial_t)\Big|_{t=0} f(x \exp_{\mathbb{G}}(t))$ depends only on small values of $|t|$, we conclude that

$$D_p f(x') = 0 \quad \text{for all } x' \text{ in a neighbourhood of } x.$$

To check that D_p is left invariant we simply notice that

$$\begin{aligned} D_p(L_y f)(x) &= p(\partial_t)\Big|_{t=0} \left(L_{(y^{-1}\cdot x)^{-1}} f \circ \exp_{\mathbb{G}}(t)\right)(0) \\ &= L_y(D_p f)(x), \end{aligned}$$

and this concludes the proof of the first part. Now let $D \in \mathscr{D}(\mathbb{G})$ and take the local chart around the origin $\varphi = \exp_{\mathbb{G}}$ so that the identity

$$Df(\varphi(t)) = \sum_{|\alpha| \le k} a_\alpha(t) \partial^\alpha (f \circ \varphi)(t)$$

implies, by setting $t = 0$, that

$$Df(e) = D_p f(e) \quad \text{where } p(t) = \sum_{|\alpha| \le k} a_\alpha(0) t^\alpha.$$

This is enough to conclude because both D and D_p are left-invariant and hence

$$Df(x) = D_p f(x) \quad \text{for all } x \in \mathbb{G}.$$

□

Remark 5.27. This shows that the order of D does not depend on the point $x \in \mathbb{G}$ and is equal to the degree of the polynomial p such that $D = D_p$.

There is a different characterisation of these differential operators that requires the introduction of left-invariant vector fields. Let $(X_1, \ldots, X_n)$ be a basis of $\mathfrak{g}$ and set

$$X^\alpha := X_1^{\alpha_1} \cdots X_n^{\alpha_n}.$$

Theorem 5.28 (Poincaré-Birkhoff-Witt)**.** *Let $D \in \mathscr{D}(\mathbb{G})$. Then D admits a unique decomposition of the form*

$$D = \sum_{|\alpha| \le k} c_\alpha X^\alpha =: \widetilde{D}_p,$$

where $p(t) = \sum_{|\alpha| \le k} c_\alpha t^\alpha$.

Proof. The reader may consult [3] or the references therein. □

Remark 5.29. Let $p = \sum_{|\alpha| \le k} c_\alpha t^\alpha$. The operator $\widetilde{D}_p$ introduced above can also be written as follows:

$$\widetilde{D}_p = p(\partial_{t'}) \Big|_{t'=0} f(x \exp_{\mathbb{G}}(t_1' X_1) \cdots \exp_{\mathbb{G}}(t_n' X_n)).$$

Therefore, using the chart $\psi(t') := \exp_{\mathbb{G}}(t_1' X_1) \cdots \exp_{\mathbb{G}}(t_n' X_n)$, we have

$$\widetilde{D}_p = p(\partial_{t'}) \Big|_{t'=0} (L_{x^{-1}} f) \circ \psi(t').$$

Similarly, if we consider the chart $\varphi(t) = \exp_{\mathbb{G}}(t_1 X_1 + \cdots + t_n X_n)$, we can write

$$D_p f(x) = p(\partial_t) \Big|_{t=0} (L_{x^{-1}} f) \circ \varphi(t).$$

The change of coordinates is $t = \varphi^{-1} \circ \psi(t') = t' + \mathcal{O}(|t'|^2)$, and thus we can write

$$\partial_{t'}^\alpha (g \circ u)(0) = \partial_t^\alpha (g \circ u)(0) + \cdots,$$

which means that $\widetilde{D}_p$ and D_p coincide up to lower order terms.

Proof of Theorem 5.28. By induction on the degree of p. The base case is trivial, so we can assume that it holds for all $j < k$. By Theorem 5.26 we can write

$$D = D_p$$

for some polynomial of order k. Then

$$D_p - \widetilde{D}_p = D_{<k},$$

where $D_{<k}$ is a differential operator of order strictly less than k. By inductive hypothesis, there exists a polynomial q of degree $< k$ such that $D'_{<k} = \widetilde{D}_q$. It follows that

$$D = D_p = \widetilde{D}_p + \widetilde{D}_q = \widetilde{D}_{p+q},$$

and this concludes because $\deg(p+q) = k$. □

Definition 5.30. Let $\mathfrak{g}$ be a Lie algebra and $T(\mathfrak{g})$ be the tensor algebra. The *universal enveloping algebra* $U(\mathfrak{g})$ is defined by the quotient

$$T(\mathfrak{g}) \big/ \langle x \otimes y - y \otimes x - [x, y] \rangle.$$

Theorem 5.31. *Let $\mathfrak{g}$ be a Lie algebra. Then $U(\mathfrak{g})$ is isomorphic to $\mathscr{D}(\mathbb{G})$, where $\mathbb{G}$ is the unique connected and simply connected Lie group.*

Proof. See [3] and the references therein. $\square$

5.4 Solvability of left-invariant differential operators

We first recall the definition of locally solvable operator (at some point) in a manifold and then specialise it to Lie groups.

Definition 5.32. Let D be a differential operator on a manifold $\mathcal{M}$. We say that D is *locally solvable* at $x \in \mathcal{M}$ if for all $k \in \mathbb{N}$ there exists $V_k \ni x$ neighbourhood such that

$$\forall \psi \in \mathcal{D}'(\mathcal{M}), \exists\, u \in D'_k(V_k) \;:\; Du = \psi \quad \text{on } V_k.$$

We now recall a critical result that holds when $\mathcal{M} = \mathbb{R}^n$ is the Euclidean space, which asserts that fundamental solutions exist for certain operators.

Theorem 5.33 (Malgrange–Ehrenpreis)**.** *Let L be a linear differential operator with constant coefficients. Then there exists $\Phi \in \mathcal{D}'(\mathbb{R}^n)$ satisfying*

$$L\psi = \delta_0.$$

Corollary 5.34. *Linear differential operators with constant coefficients are locally solvable.*

Notice that linear differential operators with constant coefficients can be replaced with $\mathscr{D}(\mathbb{R}^n)$, which means that one might expect the following result to be true:

Proposition 5.35. *Every $D \in \mathscr{D}(\mathbb{G})$ is locally solvable.*

Unfortunately, this statement is **false** for a general Lie group $\mathbb{G}$. Consider, for example, the Haus Levy's operator in $\mathbb{R}^3$ with coordinates (x, y, t) given by

$$L = \left(\frac{1}{2}\partial_x + y\partial_t\right) + \imath\left(\frac{1}{2}\partial_y - x\partial_t\right) =: X + \imath Y$$

Notice that $[X, Y] = -\partial_t =: T$ and $\{X, Y, T\}$ span $\mathbb{R}^3$ at all points (x, y, t). Then $L \in \mathscr{D}(H)$, where H is the Heisenberg group, but we know that L is not locally solvable.

Definition 5.36. Let $\mathbb{G}$ be a Lie group and D a differential operator. A fundamental solution is $K \in \mathcal{D}'(\mathbb{G})$ such that

$$DK = \delta_e,$$

where e is the identity element of $\mathbb{G}$.

Proposition 5.37. *If there is a fundamental solution K for $D \in \mathscr{D}(\mathbb{G})$, then it is locally solvable.*

Proof. Let $\varphi \in \mathcal{D}'(\mathbb{G})$ with compact support. Let $u = \varphi * K$ (the order here is important) and notice that

$$\begin{aligned}
Du(x) &= D(\varphi * K)(x) \\
&= \int \varphi(y) D_x K(y^{-1} \cdot x) \, \mathrm{d}y \\
&= \int \varphi(y) D_x (L_y K)(x) \, \mathrm{d}y \\
&= \int \varphi(y) L_y (D_x K)(x) \, \mathrm{d}y = \varphi * DK(x).
\end{aligned}$$

$\square$

Proposition 5.38. *A differential operator $D \in \mathscr{D}(\mathbb{G})$ is locally solvable if and only if there exists a local fundamental solution, namely there exists $V_0 \ni e$ neighbourhood and $K \in D_0'(V_0)$ such that*

$$DK = \delta_e \quad \text{on } V_0.$$

Proposition 5.39. *If $D \in \mathscr{D}(\mathbb{G})$ is hypoelliptic, then ${}^{\mathrm{t}}D \in \mathscr{D}(\mathbb{G})$ and it admits a local fundamental solution.*

Now let $\mathbb{G}$ be a homogeneous Lie group with a family of dilations $\{\delta_s\}_{s>0}$ and write $\mathfrak{g} = \sum_\lambda \mathfrak{g}_\lambda$ in such a way that

$$\delta_s \big|_{\mathfrak{g}_\lambda} \equiv s^\lambda \mathrm{Id}.$$

We would like to find sufficient conditions on a left-invariant differential operator D for a fundamental solution to exist.

Definition 5.40. A distribution $\phi \in \mathcal{D}'(\mathbb{G})$ is *homogeneous* of order α if

$$\phi \circ \delta_s = s^\alpha \phi,$$

where the left-hand side is defined by

$$\langle \phi \circ \delta_s, f \rangle = \langle \phi, s^{-Q} f \circ \delta_{s^{-1}} \rangle.$$

Remark 5.41.

(i) The Dirac delta at the origin $0 \in \mathbb{G}$ (using the BCH coordinates), δ_0, is homogeneous of order $-Q$.

(ii) If ϕ is homogeneous of order α and $X \in \mathfrak{g}_\lambda$, then $X\phi$ has homogeneity order $\alpha - \lambda$.

Definition 5.42. Let $Tf := f * K$ be the convolution operator. We say that T is homogeneous of order β if

$$T(f \circ \delta_s) = s^\beta (Tf) \circ \delta_s$$

so that, if $T = \mathrm{Id}$, then the order is zero.

Remark 5.43.

(i) If $T = X \in \mathfrak{g}_\lambda$, then the homogeneity order of T is equal to λ.

(ii) T is homogeneous of order β if and only if K is homogeneous of order $-Q-\beta$.

(iii) If T is homogeneous of order β and U is homogeneous of order α, the composition TU has homogeneity order $\alpha+\beta$.

Definition 5.44. A differential operator D is homogeneous of order β if

$$D(f\circ\delta_s)=s^\beta(Df)\circ\delta_s.$$

Example 5.45. Let D be a differential operator homogeneous of order β and let K be a homogeneous fundamental solution for D. Since

$$DK=\delta_0,$$

we can consider the convolution operator $Tf=f*K$ and find that

$$D(Tf)=f \implies DT=\mathrm{Id}.$$

It is easy to verify that, at this point, the unique possibility is that K is homogeneous of order $-Q+\beta$.

Example 5.46. If $\mathbb{G}=\mathbb{R}^n$ and $D=\Delta$, then a fundamental solution must have homogeneity order $-n+2$ and hence it must be

$$\frac{c}{|x|^{n-2}}$$

for $n\geq 3$. If $n=2$, then there is no homogeneity and it can be proved that a fundamental solution for the Laplace operator is

$$c\log|x|.$$

Example 5.47. Let $D=\partial_t-\Delta_x$ be the heat kernel on $\mathbb{R}\times\mathbb{R}^n$ and consider the dilations

$$\delta_s(t,x)=(s^2t,sx).$$

Then the operator has homogeneity order equal to two and

$$K(t,x):=\begin{cases}\frac{1}{(2\pi t)^{\frac{n}{2}}}\mathrm{e}^{-\frac{|x|^2}{4t}} & \text{if } t>0,\\ 0 & \text{otherwise,}\end{cases}$$

is a fundamental solution. Since $\delta_s(t,x)=(s^2t,sx)$ we have that

$$K(s^2t,sx)=s^{-n}K(t,x) \implies \text{the homogeneity order of } K \text{ is } n.$$

Theorem 5.48. *Let Q be the homogeneous dimension of $\mathbb{G}$ and suppose that*

(i) D is a left-invariant operator of homogeneity order $\beta\leq Q$;

(ii) D and ${}^{\mathrm{t}}D$ are hypoelliptic.

Then D has a fundamental solution with homogeneity order $-Q+\beta$.

Proof. The operator D has a fundamental solution K defined on some neighborhood V of the origin $0\in\mathbb{G}$. Since D is hypoelliptic, from

$$DK=0 \quad \text{on } V\setminus\{0\}$$

we infer that K must be smooth away from the origin. Let $\eta \in \mathcal{D}(V)$ be a cutoff function taking values in $[0,1]$ and set

$$K_1 := \eta K.$$

Then K_1 is defined on all $\mathbb{G}$, smooth away from 0, and satisfies the equation

$$DK_1 = \delta_0 + \Phi,$$

where Φ is compactly supported in a set that does not contain the origin. Notice that, once again, the hypoellipticity of D implies $\Phi \in \mathcal{D}(\mathbb{G})$.

Step 1. Fix a homogeneous norm $|\cdot|$ on $\mathbb{G}$. Then

$$\operatorname{spt}(\Phi) \subset \{x \,:\, a < |x| < b\}$$

and, if we define $K_t := t^{\beta-Q} K_1 \circ \delta_{t^{-1}}$ and $\Phi_t := t^{-Q}\Phi \circ \delta_{t^{-1}}$, we also find that

$$DK_t = t^{-Q}(DK_1) \circ \delta_{t^{-1}} = \delta_0 + \Phi_t$$

because δ_0 has degree of homogeneity equal to $-Q$ (as verified above). Since

$$\operatorname{spt}(\Phi_t) = \delta_t\left(\operatorname{spt}(\Phi)\right) \subset \{x \,:\, ta < |x| < tb\}$$

we easily infer that K_t is another local fundamental solution of D. Moreover, given that

$$\lim_{t\to\infty} \Phi_t = 0$$

in the sense of distributions, it is sufficient to prove that $\lim_{t\to\infty} K_t = K$ exists as a distribution to conclude that K is a fundamental solution for D. If such K exists, then

$$s^{Q-\beta} K \circ \delta_s = \lim_{t\to\infty} s^{Q-\beta} K_t \circ \delta_s = \lim_{t\to\infty} K_{t/s} = K$$

shows that it is homogeneous of degree $-Q+\beta$ as claimed.

Step 2. Let $t > 1$. We would like to write K_t as

$$K_t = K_1 + \int_1^t \frac{\mathrm{d}K_s}{\mathrm{d}s}\,\mathrm{d}s,$$

but, a priori, the integrand might not be a well-defined distribution. To prove it, we start with noticing that for $\varphi \in \mathcal{D}(\mathbb{G})$ we have the identity

$$\begin{aligned}\lim_{s\to 1} \frac{1}{s-1}\langle K_s - K_1, \varphi\rangle &= \lim_{s\to 1} \frac{1}{s-1}\langle K_1, s^{\beta}\varphi\circ\delta_s - \varphi\rangle \\ &= \left\langle K_1, \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=1} (s^{\beta}\varphi\circ\delta_s)\right\rangle.\end{aligned}$$

Let $(x_1, \dots, x_n)$ be coordinates of $\mathbb{G}$ such that $\delta_s x = (s^{\lambda_1}x_1, \dots, s^{\lambda_n}x_n)$. Then

$$\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=1} \varphi(\delta_s x) = \sum_{j=1}^n \left(\lambda_j \partial_{x_j}\varphi(x)\right) x_j = E\varphi(x),$$

where E is called *modified Euler operator*. It follows that

$$\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=1} \langle K_s, \varphi\rangle = \langle K_1, \beta\varphi + E\varphi\rangle,$$

which immediately gives the identity

$$K_1' = \beta K_1 + {}^{\mathrm{t}}EK_1 = (\beta - Q)K_1 - EK_1.$$

To prove that K_1' is smooth (i.e., it belongs to $\mathcal{D}(\mathbb{G})$) simply notice that

$$DK_1' = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=1} (DK_s) = -Q\Phi - E\Phi,$$

which is a smooth function. Since D is hypoelliptic, K_1' is smooth on $\mathbb{G}$ and it has compact support.

Step 3. For any $s > 1$ we have

$$\begin{aligned}\frac{\mathrm{d}K_s}{\mathrm{d}s} &= s^{-1}\frac{\mathrm{d}}{\mathrm{d}u}\bigg|_{u=1} K_{su} \\ &= s^{-1}s^{\beta-Q}\frac{\mathrm{d}}{\mathrm{d}u}\bigg|_{u=1} K_u \circ \delta_{s^{-1}} \\ &= s^{\beta-Q-1}K_1' \circ \delta_{s^{-1}}.\end{aligned}$$

If $\varphi \in \mathcal{D}(\mathbb{G})$, then

$$\int_1^t \langle s^{\beta-Q-1}K_1' \circ \delta_{s^{-1}}, \varphi\rangle = \int_1^t s^{\beta-Q-1}\int_{\mathbb{G}} K_1'(\delta_{s^{-1}}x)\varphi(x)\,\mathrm{d}x\mathrm{d}s.$$

Since $\beta < Q$, we have

$$\int_1^t s^{\beta-Q-1}\int_{\mathbb{G}} |K_1'(\delta_{s^{-1}}x)||\varphi(x)|\,\mathrm{d}x\mathrm{d}s \le C\|\varphi\|_{L^1(\mathbb{G})},$$

showing that the integral is actually well-defined in the sense of distributions as $t \to \infty$.

Step 4. The smoothness of K away from zero follows from the hypoellipticity of D, or alternatively, from the fact that for all $x \neq 0$ we have

$$K(x) = K_1(x) + \int_1^\infty s^{\beta-Q-1}K_1'(\delta_{s^{-1}}x)\,\mathrm{d}s.$$

Step 5. For the uniqueness, let K and H be two $(-Q+\beta)$-homogeneous fundamental solutions for D. Then

$$D(K - H) = 0$$

gives, by hypoellipticity, that $K - H \in C^\infty(\mathbb{G})$, which is impossible because it is homogeneous of order $-Q + \beta$ which is always negative. ☐

Remark 5.49. If $D = X + \imath Y$ is the Levi's operator on the Heisenberg group, it is easy to see that D is not hypoelliptic and hence it is not a counterexample to Theorem 5.48.

Example 5.50. If $\mathbb{G}$ is stratified and $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_k$ with $\{X_1, \ldots, X_k\}$ basis, then

$$D = \sum_{j=1}^{k} X_j^2$$

is homogeneous of order two. Therefore, for all $Q \geq 3$ we can apply the previous theorem to infer the existence of a fundamental solution for D.

Remark 5.51. In general $K(x) = |x|^{-Q+2}$, but $|\cdot|$ is the homogeneous norm on $\mathbb{G}$ which is in most cases a rather mysterious object.

Example 5.52. In the Heisenberg group with $X_j = \partial_{x_j} - \frac{y_j}{2}\partial_t$ and $Y_j = \partial_{y_j} - \frac{x_j}{2}\partial_t$ we have the fundamental solution

$$\Phi_k(\mathbf{z}, t) = \frac{2^k \Gamma(k/2)^2}{\pi^{k+1}} \left(|\mathbf{z}|^4 + t^2\right)^{-\frac{k}{2}},$$

which is, accordingly to a previous result, smooth away from the origin and homogeneous of degree $2 - Q$.

Part III

Appendix

Appendix A

Topics in Differential Geometry

The goal of this chapter is to recollect some useful and well-known facts in differential geometry which are needed throughout the course. The reader who is totally unfamiliar with the topics discussed here may refer to [7].

A.1 Introduction

Recall that a *smooth manifold* is a topological manifold which is equipped with an equivalence class of atlases whose transition maps are all smooth. Namely, a manifold is a couple

$$(M, \{\varphi_i : U_i \to V_i \subset \mathbb{R}^q\}_{i\in I}),$$

where U_i and V_i are open sets and the transition maps

$$\varphi_{i,j} : \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j),$$

defined by setting $\varphi_{i,j} = \varphi_j^{-1} \circ \varphi_i$, are smooth.

Definition A.1. We say that a map between smooth manifolds $f : M \to N$ is smooth if it is so along some charts, that is,

$$\psi_j \circ f \circ \varphi_i^{-1} : V_i \subset \mathbb{R}^q \to Z_i \subset \mathbb{R}^p \in C^\infty.$$

The tangent space at p to a smooth manifold M is easy to define when M is embedded in $\mathbb{R}^q$, but it requires an abstract definition if M is not a submanifold of the Euclidean space.

Definition A.2. Let M be a smooth manifold and fix $p \in M$. A *derivation* v at p is an operator that assigns a number $v(f)$ to any smooth real-valued function f defined in a neighbourhood of p, that satisfies the following assumptions:

(a) If there exists $U \ni p$ such that $f\big|_U \equiv g\big|_U$, then $v(f) = v(g)$.

(b) The operator v is linear.

(c) The operator v satisfies the Leibniz rule, that is,

$$v(fg) = f(p)v(g) + g(p)v(f).$$

The set of all derivations at p is called *tangent space at p to M* and it will always be denoted with the symbol $\mathrm{Tan}_p M$.

Remark A.3. The tangent space $\mathrm{Tan}_p M$ is a vector space.

Immersion, embeddings and submanifolds

Let $f : M \to N$ be a smooth map between two smooth manifolds. We say that f is a *immersion* at $p \in M$ if the differential

$$\mathrm{d}f_p : \mathrm{Tan}_p M \longrightarrow \mathrm{Tan}_{f(p)} N$$

is injective as a linear map between vector spaces. It is worth remarking that an immersion is locally injective, but it might fail to be so globally: for example, the 8-knot in $\mathbb{R}^2$ has a self-intersection at the origin, but it is an immersion. The natural "upgrade" of this notion is that of *embedding*.

Definition A.4. We say that a smooth map $f : M \to N$ is an embedding if it is an immersion and a homeomorphism onto its image.

The homeomorphism condition implies that an embedding is globally injective, but the vice versa is not always[1] true although the counterexample is a little bit more subtle than the previous one. In any case, for us embeddings will take on a fundamental role thanks to the following property:

Lemma A.5. *If $f : M \to N$ is an embedding, then $f(M)$ is a smooth submanifold of N.*

A.2 Bundles

Work in progress...

A.3 Vector fields and flows

In this section, we introduce the notion of a *smooth vector field* and we investigate the properties of the associated flow.

Lemma A.6. *Let Ω be an open connected subset of $\mathbb{R}^n$ and let $X : C^\infty(\Omega) \to C^\infty(\Omega)$ be a linear operator. Then the following assertions are equivalent:*

[1] A proper injective immersion is an embedding, so if M is compact globally injective together with immersion is enough to get an embedding.

(a) *The operator X is a smooth vector field satisfying*

$$Xf(x) = \sum_{j=1}^{n} a_j(x)\partial_j,$$

where $a_j \in C^\infty(\Omega)$ and $\partial_j := \partial_{x_j}$.

(b) *The operator X satisfies the Leibniz rule.*

Proof. Suppose that X satisfies the Leibniz rule and let $f \in C^\infty(\Omega)$. For a fixed point $\bar{x} \in \Omega$ we can exploit the smoothness of f to write

$$\begin{aligned} f(x) = f(\bar{x}) + \frac{\mathrm{d}}{\mathrm{d}t}\int_0^1 f(tx + (1-t)\bar{x})\,\mathrm{d}t = \\ = f(\bar{x}) + \sum_{j=1}^{n}(x_j - \bar{x}_j)\int_0^1 \partial_j f(tx + (1-t)\bar{x})\,\mathrm{d}t. \end{aligned}$$

Observe[2] that X vanishes on constant functions, i.e. $X(c) = 0$ for any $c \in \mathbb{R}$ and thus

$$X(f)(x) = \underbrace{X(f(\bar{x}))}_{=0} + \sum_{j=1}^{n} X\left((x_j - \bar{x}_j)h_j(x)\right).$$

We evaluate the expression at $x = \bar{x}$ and conclude that

$$X(f)(\bar{x}) = \sum_{j=1}^{n} X(x_j)(\bar{x})h_j(\bar{x}).$$

Set $a_j(\bar{x}) := X(x_j)(\bar{x})$ and observe that $h_j(\bar{x})$ is nothing but the partial derivative ∂_j. This concludes the proof of the lemma. □

To define the flow, we first need the notion of *integral curve* of a vector field X. The idea is to look for solutions $\gamma : (-\delta, \delta) \to \Omega$ of the following Cauchy problem:

$$\begin{cases} \gamma'(t) = X(\gamma(t)), \\ \gamma(0) = x_0. \end{cases}$$

The local existence theorem tells us that we can find a unique maximal solution, say γ_{x_0}, that is defined in the (maximal) interval I_{x_0}. We say that γ_{x_0} is the integral curve of X originating at x_0.

Remark A.7. If $K \subset \Omega$ is compact, then there exists $\epsilon_K > 0$ such that

$$(-\epsilon_K, \epsilon_K) \subset I_x \text{ for all } x \in K.$$

In particular, integral curves are always defined up to a time $\delta > 0$ which is uniform with respect to the choice of the originating point.

[2]Use the Leibniz rule with $f = g$ constant.

The *flow* of a vector field X is defined by setting

$$\Phi_X(x,\, t) := \gamma_x(t) \quad (\text{or } \varphi_{X,\,t}(x)),$$

where γ_x is the maximal solution of the Cauchy problem above. It is easy to verify that Φ_X is smooth (C^∞) with respect to the couple $(x,\, t)$.

Remark A.8. If $\{X_y\}_{y\in\Theta}$ is a family of smooth vector fields, then the map

$$(x,\, y,\, t) \longmapsto \Phi_{X_y}(x,\, t)$$

is smooth with respect to the triplet $(x,\, y,\, t)$.

Now, let us denote by D_X the domain of Φ_X (which happens to be vertically connected) and define $\varphi_{X,\,t}(\cdot)$ as the flow at a fixed time t, that is,

$$\varphi_{X,\,t}(x) := \gamma_x(t).$$

It is easy to verify that $\varphi_{X,\,0}$ is the identity map and also that

$$\varphi_{X,\,t} \circ \varphi_{X,\,s} = \varphi_{X,\,s+t}$$

by the local uniqueness of the solution of the Cauchy problem. This immediately translates to the analogous property for the flow:

$$\Phi_X(x,\, t+s) = \Phi(\Phi(x,\, t),\, s). \tag{A.1}$$

Proposition A.9. *Let $A \subseteq \Omega \times \mathbb{R}$ be an open neighbourhood of $\Omega \times \{0\}$ that is vertically connected. Let $\Phi \in C^\infty(A,\, \Omega)$ satisfy (A.1) and $\Phi(x,\, 0) = x$. Then for any $f \in C^\infty(\Omega)$ it turns out that*

$$Xf(x) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(\Phi(x,\, t))$$

is a vector field and $\Phi = \Phi_X\Big|_A$.

A.4 Exponential map

Fix a positive time t. We can consider[3] the operator

$$f \longmapsto f \circ \varphi_{X,\,t},$$

which sends $C^\infty(\Omega)$ into $C^\infty(\Omega_t)$, where Ω_t is the set of all x such that the flow survives until (at least) time t. We denote it by

$$\exp(tX)f := f \circ \varphi_{X,\,t},$$

and we call it exponential. It can easily be proved via the Cauchy problem above that

$$f \circ \varphi_{X,\,t} = f \circ \varphi_{tX,\,1}$$

[3] Whenever it makes sense, of course, but in this section we will purposely ignore this issue.

since, roughly speaking, we can interpret X as the speed at which we are running across the curve.

Lemma A.10 (Properties of the exponential)**.**

(1) *The exponential* $\exp(0X)$ *coincides with the identity map.*

(2) *The composition is given by* $\exp((s+t)X) = \exp(sX)\exp(tX)$.

(3) *The inverse is given by* $\exp(-tX) = [\exp(tX)]^{-1}$.

(4) *The derivative is given by*

$$\frac{\mathrm{d}}{\mathrm{d}t}\exp(tX) = X\exp(tX) = \exp(tX)X,$$

and therefore for all $k \in \mathbb{N}$ *and* $f \in C^\infty$ *we have the following Taylor formula:*

$$\exp(tX)f(x) = \sum_{j=0}^{k} \frac{t^j}{j!} X^j f(x) + \mathcal{O}(t^{k+1}). \tag{A.2}$$

If $x \in K$*, then this identity makes sense for all time in* $(-\epsilon_K, \epsilon_K)$ *and the big-*$\mathcal{O}$ *is uniform.*

(5) *If* $u : \Omega \to \Omega'$ *is a diffeomorphism, then*

$$u \circ \Phi_X(\cdot, t) = \Psi(\cdot, t)$$

is the flow of the pushforward vector field $X' = u_* X$.

We are now interested in composition of exponential maps generated by different vector fields that may also not commute. In particular, we will compare the following three terms:

$$\mathrm{e}^{tX}\mathrm{e}^{sY}, \quad \mathrm{e}^{sY}\mathrm{e}^{tX} \quad \text{and} \quad \mathrm{e}^{tX+sY}.$$

Example A.11. In $\mathbb{R}^2$ consider the vector fields $X = \partial_x$ and $Y = \partial_y$. Then

$$\varphi_{X,\,t}(x,\,y) = (x+t,\,y) \quad \text{and} \quad \varphi_{Y,\,s}(x,\,y) = (x,\,y+s),$$

so it is immediate to verify that these two flows commute, that is,

$$\varphi_{X,\,t} \circ \varphi_{Y,\,s} \equiv \varphi_{Y,\,s} \circ \varphi_{X,\,t}.$$

Still in $\mathbb{R}^2$, we can consider X as above and $Y = x\partial_y$. In Figure **??** we show that

$$\varphi_{X,\,t} \circ \varphi_{Y,\,s}(x,\,y) \neq \varphi_{Y,\,s} \circ \varphi_{X,\,t}(x,\,y)$$

for all points $(x,\,y) \in \mathbb{R}^2$, which means that the flows do not commute and something bad happens (as we will see later on).

Now apply the Taylor expansion (up to the order two) to both e^{tX} and e^{sY}. Assuming that s and t are admissible, we can write them as

$$\mathrm{e}^{tX} f(x) = f(x) + tXf(x) + \frac{t^2}{2} X^2 f(x) + \mathcal{O}(t^3),$$

$$\mathrm{e}^{sY} g(x) = g(x) + sYg(x) + \frac{s^2}{2} Y^2 g(x) + \mathcal{O}(s^3).$$

We can now use these formulas to compute a second-order approximation of $\mathrm{e}^{tX}\mathrm{e}^{sY}$ and $\mathrm{e}^{sY}\mathrm{e}^{tX}$ to better understand how the difference behaves. Namely, we have

$$\mathrm{e}^{sY}\mathrm{e}^{tX} f(x) = f(x) + [tX + sY]\, f(x) + \left[\frac{t^2}{2} X^2 + stYX + \frac{s^2}{2} Y^2\right] f(x) + \dots,$$

and, by symmetry, also

$$\mathrm{e}^{tX}\mathrm{e}^{sY} f(x) = f(x) + [tX + sY]\, f(x) + \left[\frac{t^2}{2} X^2 + stXY + \frac{s^2}{2} Y^2\right] f(x) + \dots.$$

It turns out that the difference at the second order is given by

$$\left[\mathrm{e}^{sY}, \mathrm{e}^{tX}\right] f(x) = st[Y, X]f(x),$$

which means that $[Y, X] = 0$ and having the flows commute are strictly related problems. As a matter of fact, even the exponential e^{tX+sY} differs from the previous ones (at the second-order approximation) by something multiplied by $[Y, X]$.

Theorem A.12. *Let X and Y be smooth vector fields on Ω. The following assertions are equivalent:*

(1) *There exists $\delta > 0$ such that for all s, t admissible with $\max\{|t|, |s|\} < \delta$ we have*

$$\mathrm{e}^{tX}\mathrm{e}^{sY} = \mathrm{e}^{sY}\mathrm{e}^{tX}.$$

(2) *There exists $\delta > 0$ such that for all t admissible with $|t| < \delta$ we have*

$$\mathrm{e}^{tX} Y = Y\mathrm{e}^{tX}.$$

(3) *There exists $\delta > 0$ such that for all s admissible with $|s| < \delta$ we have*

$$\mathrm{e}^{sY} X = X\mathrm{e}^{sY}.$$

(4) *The vector fields X and Y commute, that is, $[X, Y] = 0$.*

(5) *There exists $\delta > 0$ such that for all s, t admissible with $\max\{|t|, |s|\} < \delta$ we have*

$$\mathrm{e}^{tX}\mathrm{e}^{sY} = \mathrm{e}^{tX+sY}.$$

Furthermore, if one of these assertions hold, we can extend δ in such a way that t and s can be chosen among all the admissible ones.

Proof. Suppose that (**1**) holds and notice that we can write

$$\mathrm{e}^{tX}\mathrm{e}^{sY}\mathrm{e}^{-tX}f = f \circ \varphi_{X,\,-t} \circ \varphi_{Y,\,s} \circ \varphi_{X,\,t}.$$

Denote by $\Phi_t(x,\,s)$ the composition on the right-hand side except for f and notice that it has the property of a flow, which means that there exists X_t smooth vector field on Ω such that its flow coincide, that is,

$$\Phi_t(x,\,s) = \varphi_{X_t,\,s}(x).$$

It is now easy to verify that

$$\mathrm{e}^{tX}\mathrm{e}^{sY}\mathrm{e}^{-tX}f(x) = \mathrm{e}^{sX_t}f(x),$$

and therefore we can compute $X_t f(x)$ by differentiating the formula above at $s = 0$ with respect to s. More precisely, we have that

$$X_t = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\mathrm{e}^{sX_t} = \mathrm{e}^{tX}Y\mathrm{e}^{-tX}.$$

Since (**1**) holds, we also have that X_t is identically (w.r.t. t) equal to Y. The formula above now tells us that (**2**) (and, equivalently, (**3**)) holds. Now notice that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}X_t = \mathrm{e}^{tX}[Y,\,X]\mathrm{e}^{-sX},$$

and therefore if we assume that (**2**) (and (**3**)) holds, we find that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}X_t = [Y,\,X].$$

But Y_t is identically Y so the derivative must be equal to zero and hence the commutator $[Y,\,X]$ must be zero well. This proves that (**2**) implies (**4**), but the reverse argument proves the reverse implication (although we will not need it.) Now introduce the function

$$T(s,\,t) := \mathrm{e}^{-tX}\mathrm{e}^{tX+sY}\mathrm{e}^{-sY}.$$

If we can prove that T is constant (in both variables) and equal to the identity, then we would be able to conclude that (**5**) holds. For this, let

$$T_\alpha(s) := \mathrm{e}^{-\alpha sX}\mathrm{e}^{\alpha sX+sY}\mathrm{e}^{-sY},$$

where α is an arbitrary parameter which allows us to reduce the derivative problem to a single-variable function. Assume that (**4**) holds and notice that

$$[\alpha X + Y,\,Y] = 0.$$

An easy computation, together with this commutator relation, shows that $T'_\alpha(s)$ is identically equal to zero, and therefore we can infer (**5**). Finally (**5**) trivially implies (**1**), so the chain of implications is complete. $\square$

A.5 Foliations and distributions

In this section, we introduce some higher-dimensional analogues of vector fields and integral curves, replacing vectors with k-dimensional subspaces and integral curves with k-dimensional submanifolds.

Definition A.13. Let M be a smooth manifold. An *immersed submanifold* in M is the image of an immersion $S \hookrightarrow M$.

Definition A.14 (Foliation)**.** Let M be a smooth manifold. A *k-dimensional foliation* is a partition of M,

$$M = \bigcup_{F \in \mathscr{F}} F,$$

where each F is an injectively immersed connected submanifold, such that the following holds: for every $p \in M$ there is a chart

$$\Phi : U \longrightarrow \mathbb{R}^n,$$

with $p \in U$, that sends the intersection of every $F \in \mathscr{F}$ with U into a collection of countably many parallel affine k-planes of the type

$$\{x_{k+1} = c_{k+1}, \ldots, x_n = c_n\}.$$

The immersed submanifolds in the definition are usually referred to as *leaves* of the foliation $\mathscr{F}$. Moreover, any chart φ that satisfies the property above is said to be *compatible* with the foliation.

Remark A.15. Any foliation is made up of uncountably many leaves. Indeed, the countable union of immersed manifolds of dimension strictly smaller than M has measure zero.

There is an equivalent definition of foliation which translates everything to local conditions on the transition maps. More precisely, we have:

Definition A.16 (Foliation)**.** Let M be a smooth manifold. A *k-dimensional foliation* is an atlas $\{\varphi_i : U_i \to \mathbb{R}^n\}$, compatible with the smooth structure of M, which transition maps are locally given by

$$\varphi_{i,j}(x, y) = (\varphi^1_{i,j}(x, y), \varphi^2_{i,j}(y)),$$

where $(x, y) \in \mathbb{R}^k \times \mathbb{R}^{\mathrm{m}-k}$.

Definition A.17 (Distribution)**.** Let M be a smooth manifold. We say that D is a k-distribution in M if D is a rank-k subbundle D of the tangent bundle TM, where

$$TM = \bigcup_{p \in M} T_pM.$$

In other words, a distribution is a collection of k-subspaces $D_p \subset T_pM$ that varies smoothly (in a smooth manifold!) with $p \in M$.

Definition A.18 (Integrable)**.** We say that a distribution D on M is *integrable* if it is tangent to some foliation $\mathscr{F}$. Namely, for each $p \in M$ there is $F \in \mathscr{F}$ such that

$$D_p = T_pF.$$

We say that a distribution D is *involutive* if whenever X and Y are tangent vector fields to D (i.e., $X(p), Y(p) \in D_p$), then also $[X, Y]$ is a tangent vector field.

Theorem A.19 (Frobenius)**.** *A distribution D is integrable if and only if it is involutive.*

Appendix B

Baker–Campbell–Hausdorff Formula

Let X and Y be vector fields. If $[X, Y] = 0$, then it is easy to verify that the exponential map of the sum is given by

$$\mathrm{e}^{tX+sY} = \mathrm{e}^{tX}\mathrm{e}^{sY},$$

but this is not the case if X and Y do not commute. The main goal of this chapter is to prove the **Baker–Campbell–Hausdorff formula**, which gives us a way to express

$$\mathrm{e}^{tX+sY}$$

in terms of the so-called elementary commutators. The central role taken on by this formula will be more clear when we will talk about Lie groups and, more specifically, Lie algebras.

B.1 Free associative algebra

Let R be a commutative ring. The free (associative, unital) algebra on n variables, say $\{a_1, \ldots, a_n\}$ is the free R-module with a basis consisting of all words over the alphabet

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

including the empty word, which is the unit of the free algebra. It is easy to see that endowed with the concatenation of words as multiplication, we obtain an R-algebra

Let $R := \mathbb{R}$ and consider the *associative free algebra generated* generated by two elements, say x and y, over R. We recall that an element (or word) of length L can be written as

$$u = x^{i_1}y^{j_1}\ldots x^{i_\ell}y^{j_\ell},$$

where $i_k,\, j_k \geq 0$ and $\sum_{k=1}^{\ell}(i_k + j_k) = L$. If $[x, y] = 0$, then L-words are of the form

$$u = x^{\ell}y^{L-\ell},$$

but, in general, we do not expect to be in such a simple situation. It makes sense to introduce the notation $a[x, y]$ for the free algebra generated by x and y. Recall that

$$a[\![x, y]\!]$$

indicates the set of formal power series with elements in $a[x, y]$. Namely, $u \in a[\![x, y]\!]$ if

$$u = \sum_{k=0}^{\infty} c_k,$$

where c_k is a linear combination of words of length k. The exponential of u,

$$\mathrm{e}^u = \sum_{k=0}^{\infty} \frac{1}{k!} u^k,$$

is well-defined provided that $c_0 = 0$ (so we can apply Taylor's theorem) so we define

$$a_0[\![x, y]\!] := \{u \in a_0[\![x, y]\!] \ : \ c_0 = 0\}.$$

Remark B.1. Let $\mathcal{A}$ be an associative algebra. The bracket operator

$$[u, v] := uv - vu$$

satisfies the following properties:

(a) $[\cdot, *]$ is bilinear;

(b) $[\cdot, *]$ is skew-symmetric;

(c) $[\cdot, *]$ satisfies the Jacobi identity, that is,

$$[[u, v], w] + [[w, u], v] + [[v, w], u] = 0. \tag{B.1}$$

In particular, any associative algebra $\mathcal{A}$ endowed with this bracket operator is a Lie algebra[1]. The interested reader can find a more precise definition in Chapter 4, but we still recommend to read it after this one.

B.2 Baker–Campbell–Hausdorff formula

Let $\mathcal{F}$ be the Lie sub-algebra generated by x and y in $a[\![x, y]\!]$. Then $\mathcal{F}$ is closed under the bracket operator and it is easy to verify that

$$x, y \in \mathcal{F}, \quad [x, y] \in \mathcal{F}, \quad [[x, y], x] \in \mathcal{F}.$$

One might also check that elements such as x^2 or y^2 are not in $\mathcal{F}$ proving that it is not possible to obtain them via the bracket operator starting from x and y only.

[1] Recall that a Lie algebra only requires the existence of a bracket operator, but it does not have to be associative. In any case, we say that when $[\cdot, *] \equiv 0$ the associated Lie algebra is abelian.

Definition B.2. A commutator of the form

$$[a_1, \ldots [a_{k-2}, [a_{k-1}, a_k]] \ldots]$$

is called *elementary iterated commutator.*

Remark B.3. The Lie sub-algebra $\mathcal{F}$ is not made up entirely of elementary iterated commutators. For example,

$$[[[x, y], x], [x, y]] \in \mathcal{F}$$

is not a elementary commutator. However, using (B.1), one can prove that it can also be written as the sum of elementary commutators.

Proposition B.4. *The linear space $\mathcal{F}$ is spanned by x, y and elementary commutators.*

As a corollary of this result, we can always write $\mathcal{F}$ as the sum of linear subspaces in the following form:

$$\mathcal{F} = \sum_{k_1+k_2 \geq 1} \mathcal{F}^{k_1, k_2},$$

where (k_1, k_2), called *bidegree*, indicates the number of times x and y appear in the elementary iterated commutator.

Theorem B.5 (Baker–Campbell–Hausdorff)**.** *Let x and y be as above. Then*

$$\mathrm{e}^x \mathrm{e}^y = \mathrm{e}^u,$$

and we can express u explicitly with the following formula:

$$u = x + y + \sum_{k_1+k_2 \geq 1} c_{k_1, k_2},$$

where $c_{k_1, k_2} \in \mathcal{F}^{k_1, k_2}$ for all (k_1, k_2) admissible. In particular, if X and Y are smooth vector fields on Ω that do not commute, we have that

$$\mathrm{e}^{tX} \mathrm{e}^{sY} = \mathrm{e}^{tX + sY + \sum_{k_1+k_2 \geq 1} c_{k_1, k_2}(tX, sY)}. \tag{B.2}$$

Remark B.6. There are iterative methods to determine the coefficients c_{k_1, k_2}, but we usually care only about the abstract formula. In any case, the first two terms - which is always useful to remember - are given by

$$X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}\left([X, [X, Y]] - [Y, [X, Y]]\right).$$

Remark B.7. The BCH formula allows one to write

$$\mathrm{e}^u \mathrm{e}^v = \mathrm{e}^w$$

even if u and v are elements of $\mathcal{F}$ rather than x or y. Hence w is not, in general, in $\mathcal{F}$ but rather a formal sum of such elements.

B.2.1 Consequences of the BCH formula

Proposition B.8.

(i) *For each $n \geq 1$ there exists a unique decomposition*

$$e^{x+y} = e^x e^y e^{w_2} \dots e^{w_n} e^{r_{n+1}},$$

where $w_j \in \mathcal{F}^j$ and $r_{n+1} \in \sum_{k \geq n+1} \mathcal{F}^k$.

(ii) *If $c_p \in \mathcal{F}^p$, then there exists $q(p) := q > 0$ such that*

$$e^{c_p} = e^{a_1} \dots e^{a_q} e^{r_{p+1}}, \tag{B.3}$$

where $a_j \in \{\pm x, \pm y\}$ and $r_{p+1} \in \sum_{k \geq p+1} \mathcal{F}^k$.

Proof. We can prove both assertions with $n = 1$, $p = 2$ and then apply the induction principle to conclude that it holds for all n and all p.

(i) Set

$$\star := e^{-y} e^{-x} e^{x+y},$$

and use the BCH formula to rewrite the first two terms as follows:

$$\star = e^{-(x+y)+r_2(x,\,y)} e^{x+y}.$$

Now use the formula again to obtain the identity

$$\star = e^{r_2(x,\,y)+\frac{1}{2}[-(x+y)+r_2(x,\,y),\,x+y]} = e^{\widetilde{r_2}(x,\,y)},$$

and this proves the formula for $n = 1$ since

$$e^{x+y} = e^x e^y e^{\widetilde{r_2}(x,\,y)}.$$

(ii) Let $p = 2$. Then

$$\begin{aligned} e^x e^y e^{-x} e^{-y} &= e^{x+y+\frac{1}{2}[x,\,y]+r_3(x,\,y)} e^{-(x+y)+\frac{1}{2}[x,\,y]+r_3(-x,\,-y)} = \\ &= e^{[x,\,y]+\frac{1}{2}\left[x+y+\frac{1}{2}[x,\,y]+r_3(x,\,y),\,-(x+y)+\frac{1}{2}[x,\,y]+r_3(-x,\,-y)\right]+r_3'} = \\ &= e^{[x,\,y]+r_3''} \stackrel{\star}{=} \\ &\stackrel{\star}{=} e^{[x,\,y]} e^{r_3''} e^{r_2([x,\,y],\,r_3'')} = \\ &= e^{[x,\,y]} e^{\widetilde{r_3}(x,\,y)}, \end{aligned}$$

where the identity $\star$ follows from the previous point and all the others from applications of the BCH formula. ☐

Remark B.9. The formula (B.3) for e^{x+y} can be generalized to more than two elements. Namely, one can prove that

$$e^{x_1+\dots+x_n} = e^{a_1} \dots e^{a_q} e^{r_{p+1}}, \tag{B.4}$$

where $a_j \in \{\pm x_1, \dots, \pm x_n\}$ and $r_{p+1} \in \sum_{k \geq p+1} \mathcal{F}^k$ and $\mathcal{F}^k$ is the space of elementary commutators with n-degree equal to k.

Remark B.10. If X and Y are smooth vector fields, then

$$\mathrm{e}^{t(X+Y)} = \mathrm{e}^{tX}\mathrm{e}^{tY}\mathrm{e}^{\frac{t^2}{2}[X,\,Y]}(1+\mathcal{O}(t^3)).$$

Therefore

$$\mathrm{e}^{-\frac{t^2}{2}[X,\,Y]}\mathrm{e}^{-tY}\mathrm{e}^{-tX}\mathrm{e}^{t(X+Y)}f(x) - f(x) = \mathcal{O}(t^3),$$

so we find that the derivatives of first and second order of the left-hand side with respect to t is zero at $t = 0$. A straightforward computation shows that

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(x) = -Yf(x) - Xf(x) + (X+Y)f(x) = 0,$$

while, with a little bit more efforts, the second derivative equal to zero gives us

$$0 = \frac{\mathrm{d}}{\mathrm{d}t^2}\Big|_{t=0} f(x) = [...].$$

Index

Bibliography

[1] I. Ado, *The representation of lie algebras by matrices*, Akademiya Nauk SSSR i Moskovskoe Matematicheskoe Obshchestvo, **2** (1947), pp. 159–173.

[2] W. Hebisch and A. Sikora, *A smooth subadditive homogeneous norm on a homogeneous group*, Studia Mathematica, **XCVI** (1990), pp. 231–236.

[3] G. Hochschild, *The Theory of Lie Groups*, Holden-Day, 1965.

[4] L. Hörmander, *On the theory of general partial differential operators*, Acta Math, **94** (1955), pp. 161–248.

[5] S. G. Krantz and H. R. Parks, *A Primer of Real Analytic Functions*, Birkhäuser Advanced Texts Basler Lehrbücher, Birkhäuser Basel, 2nd ed., 2002.

[6] H. Lewy, *An example of a smooth linear partial differential equation without solution*, Annals of Mathematics, **66** (1957), pp. 155–158.

[7] B. Martelli, *Manifolds*, e-Print, 2019.

[8] J. Peetre, *Une caractérisation abstraite des opérateurs différentiels*, Math. Scand., **7** (1959), pp. 211–218.

[9] ——, *Rectification à l'article "une caractérisation abstraite des opérateurs différentiels"*, Math. Scand, **8** (1960), pp. 116–120.

[10] J. Serre, *Lie Algebras and Lie Groups*, Lectures Notes in Mathematics, Springer-Verlag Berlin Heidelberg, 2nd ed., 1992.

[11] E. M. Stein and R. Shakarchi, *Fourier Analysis: An Introduction*, Princeton Lectures in Analysis I, Princeton University Press, 2003.

[12] T. O. Thorin, *Convexity Theorems*, PhD thesis, Meddelanden fran Lunds Universitets Matematiska Seminarium, 1948.