MATH 825 – FALL 2014
–
ANALYSIS AND GEOMETRY IN CARNOT-CARATHÉODORY SPACES

Contents

1. Introduction

Despite its official title[1], this course will focus on the geometry and analysis of 'Carnot-Carathéodory' metrics, sometimes called 'control metrics', which are constructed from appropriate collections of first order partial differential operators or vector fields. Such metrics arise in many examples, and before dealing with the general case, we will study a number of classical examples. This will involve the development of a number of tools of classical harmonic analysis,

Compiled Monday 8th September, 2014, 2:30pm. LectureNotes-1/Math825/Desktop.

[1]Topics in Functional Analysis

such as maximal operators, and fractional and singular integral operators. We will then discuss applications of the general theory to some more recent problems arising in complex analysis in several variables. However, to introduce the subject, let us quickly introduce some general definitions, and state some theorems we will prove.

1.1. **Vector fields and flows.**

We will work in an open subset $\Omega \subset \mathbb{R}^N$. A smooth vector field on Ω is a first order partial differential operator

$$X = \sum_{j=1}^{N} a_j(\mathbf{x})\,\frac{\partial}{\partial x_j} = \sum_{j=1}^{N} a_j(\mathbf{x})\partial_{x_j} \tag{1.1}$$

where each $a_j \in \mathcal{C}^\infty(\Omega)$. This is an analytic definition, but we also think geometrically, so that at each point $\mathbf{x} \in \Omega$, X defines a vector $X(\mathbf{x}) = (a_1(\mathbf{x}), \dots, a_N(\mathbf{x}))$ having direction and length.

For each point $\mathbf{y} \in \Omega$, the *flow along* X *through* $\mathbf{y}$ is a smooth curve $\Phi : (-\epsilon, +\epsilon) \to \mathbb{R}^N$ so that $\Phi(0) = \mathbf{y}$ and $\Phi'(t) = X(\Phi(t))$ for $|t| < \epsilon$; thus the velocity vector of the curve Φ is given by the vector field X. To see that the flow along X always exists for ϵ sufficiently small, write $\mathbf{y} = (y_1, \dots, y_N)$ and $\Phi(t) = (\varphi_1(t), \dots, \varphi_N(t))$. Then $\Phi'(t) = (\varphi_1'(t), \dots, \varphi_N'(t))$, and so we need to solve the following system of first order ordinary differential equations with initial conditions:

$$\begin{aligned} \varphi_1'(t) &= a_1\big(\varphi_1(t), \dots, \varphi_N(t)\big), & \varphi_1(0) &= y_1, \\ &\vdots & &\vdots \\ \varphi_N'(t) &= a_N\big(\varphi_1(t), \dots, \varphi_N(t)\big), & \varphi_N(0) &= y_N. \end{aligned} \tag{1.2}$$

It follows from the basic theorem on ordinary differential equations that if $a_1, \dots, a_N \in \mathcal{C}^\infty(\Omega)$, there exists $\epsilon > 0$ so that the system (1.2) has a unique solution on an interval $(-\epsilon, +\epsilon) \subset \mathbb{R}$. Moreover, if $E \Subset \Omega$ is compact, the same ϵ works for all $\mathbf{y} \in E$, and the solution depends smoothly on $\mathbf{y}$.

1.2. **Metrics.**

Let $\mathcal{X} = \{X_1, \dots, X_p\}$ be a finite set of smooth vector fields on Ω. At each point $\mathbf{y} \in \Omega$, these vectors span a subspace $\mathcal{X}_\mathbf{y} \subset \mathbb{R}^N_\mathbf{y}$, where $\mathbb{R}^N_\mathbf{y}$ is the set of all (tangent) vectors at $\mathbf{y}$. (Thus $\mathcal{X}_\mathbf{y}$ is a subset of the tangent space to $\mathbb{R}^N$ at $\mathbf{y}$.) We consider a set $\mathcal{C}(\mathcal{X})$ of mappings $\Phi : [0, 1] \to \Omega$ defined as follows: $\Phi \in \mathcal{C}(\mathcal{X})$ if and only if:

a) There exist points $t_0 = 0 < t_1 < \cdots < t_{M-1} < t_M = 1$ so that Φ is of class $\mathcal{C}^\infty$ on each closed subinterval $[t_j, t_{j+1}]$, $0 \le j \le M - 1$. Thus Φ is piecewise of class $\mathcal{C}^\infty$, and at the points t_j the curve may have two distinct tangents.

b) For each $0 \le j \le M - 1$ and each $t \in [t_j, t_{j+1}]$, $\Phi'(t) \in \mathcal{X}\big(\Phi(t)\big)$. Thus the tangent vector to the curve $\Phi(t)$ is restricted to lie in the specified subspaces.

Now define $\rho_\mathcal{X} : \Omega \times \Omega \to [0, +\infty]$ as follows. Let $\mathbf{y}_0, \mathbf{y}_1 \in \Omega$, and put

$$\mathcal{C}(\mathcal{X})(\mathbf{y}_0, \mathbf{y}_1) = \big\{\Phi \in \mathcal{C}(\mathcal{X}) : \Phi(0) = \mathbf{y}_0, \Phi(1) = \mathbf{y}_1\big\}. \tag{1.3}$$

a) If $\mathcal{C}(\mathcal{X})(\mathbf{y}_0, \mathbf{y}_1) = \emptyset$, set $\rho_\mathcal{X}(\mathbf{y}_1, \mathbf{y}_1) = +\infty$.

b) If If $\mathcal{C}(\mathcal{X})(\mathbf{y}_0, \mathbf{y}_1) \neq \emptyset$, set $\rho_\mathcal{X}(\mathbf{y}_0, \mathbf{y}_1)$ equal to the infimum of the set of $\delta > 0$ such that there exists $\Phi \in \mathcal{C}(\mathcal{X})(\mathbf{y}_0, \mathbf{y}_1)$ so that for all $t \in [0, 1]$, there exist $(t_1, \dots, t_p) \in \mathbb{R}^p$ with

$$\Phi'(t) = \sum_{j=1}^{p} c_j X_j(\Phi(t)) \qquad \text{and} \qquad \sum_{j=1}^{p} |c_j|^2 < \delta^2. \tag{1.4}$$

The function $\rho_{\mathcal{X}}$ is most interesting when $\mathcal{C}(\mathcal{X})(\mathbf{y}_0, \mathbf{y}_1) \neq \emptyset$ for all $\mathbf{y}_0, \mathbf{y}_1 \in \Omega$. In this case, it is not too hard to show that $\rho_{\mathcal{X}}$ is actually a metric:

$$\begin{aligned}
&\rho_{\mathcal{X}}(\mathbf{y}_0, \mathbf{y}_1) \geq 0, \\
&\rho_{\mathcal{X}}(\mathbf{y}_0, \mathbf{y}_1) = 0 \Longleftrightarrow \mathbf{y}_0 = \mathbf{y}_1, \\
&\rho_{\mathcal{X}}(\mathbf{y}_0, \mathbf{y}_1) = \rho_{\mathcal{X}}(\mathbf{y}_1, \mathbf{y}_0), \\
&\rho_{\mathcal{X}}(\mathbf{y}_0, \mathbf{y}_2) \leq \rho_{\mathcal{X}}(\mathbf{y}_0, \mathbf{y}_1) + \rho_{\mathcal{X}}(\mathbf{y}_1, \mathbf{y}_2).
\end{aligned} \tag{1.5}$$

Also, it is fairly clear that if $\mathcal{X}_{\mathbf{y}} = \mathbb{R}^N$ for every $\mathbf{y} \in \Omega$, then $\mathcal{C}(\mathcal{X})(\mathbf{y}_0, \mathbf{y}_1) \neq \emptyset$ for all $\mathbf{y}_0, \mathbf{y}_1$. However, there are interesting and important examples where $\rho_{\mathcal{X}}$ is a metric even if $\mathcal{X}_{\mathbf{y}}$ is a proper subspace of $\mathbb{R}^N$.

Exercise 1: In $\mathbb{R}^N$ let $X_j = \frac{\partial}{\partial x_j}$, and let $\mathcal{X} = \{X_1, \ldots, X_N\}$. Show that $\rho_{\mathcal{X}}$ is the Euclidean metric.

Exercise 2: Using coordinates $(x, y) \in \mathbb{R}^2$ let

$$X = \frac{\partial}{\partial x}, \qquad Y = x\,\frac{\partial}{\partial y},$$

and let $\mathcal{X} = \{X, Y\}$. Then $\mathcal{X}_{(x,y)} = \mathbb{R}^2$ if and only if $x \neq 0$. Show that $\mathcal{C}(\mathcal{X})\big((0,0),(0,1)\big) \neq \emptyset$.

Exercise 3: Using coordinates (x, y, t) in $\mathbb{R}^3$, let

$$X = \frac{\partial}{\partial x} - \frac{y}{2}\frac{\partial}{\partial t}, \qquad Y = \frac{\partial}{\partial y} + \frac{x}{2}\frac{\partial}{\partial t},$$

and let $\mathcal{X} = \{X, Y\}$. In this case $\mathcal{X}_{(x,y,t)}$ is a two dimensional subspace of $\mathbb{R}^3$ at every point. Nevertheless, $\rho_{\mathcal{X}}$ is a metric.

1.3. **Commutators.**

If X, Y are two smooth vector fields, then XY and YX are second order partial differential operators. However, the equality of mixed derivatives shows that all the second order terms in the difference $XY - YX$ cancel, and the result is again a first order operator or vector field, called the commutator of X and Y, and written $[X, Y]$. Explicitly, if $X = \sum_{j=1}^N a_j(\mathbf{x})\partial_{x_j}$ and $Y = \sum_{k=1}^N b_k(\mathbf{x})\partial_{x_k}$ then

$$\begin{aligned}
[X, Y] &= \sum_{j,k=1}^N a_j(\mathbf{x})\partial_{x_j} b_k(\mathbf{x})\frac{\partial}{\partial x_k} - \sum_{j,k=1}^N b_k(\mathbf{x})\partial_{x_k} a_j(\mathbf{x})\frac{\partial}{\partial x_j} \\
&= \sum_{k=1}^N \left[\sum_{j=1}^N \Big[a_j(\mathbf{x})\frac{\partial b_k}{\partial x_j}(\mathbf{x}) - b_j(\mathbf{x})\frac{\partial a_k}{\partial x_j}(\mathbf{x})\Big]\right]\frac{\partial}{\partial x_k}.
\end{aligned} \tag{1.6}$$

The space of all smooth vector fields on $\Omega \subset \mathbb{R}^N$ is a module over the algebra of smooth functions $\mathcal{C}^\infty(\Omega)$. With the commutator product, this module becomes a Lie algebra. Thus for any vector fields X, Y, Z,

$$\begin{aligned}
0 &= [X, Y] + [Y, X], && \text{(anti-commutativity);} \\
0 &= [X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]], && \text{(Jacobi identity).}
\end{aligned} \tag{1.7}$$

Now if $\{X_1,\ldots,X_p\}$ are vector fields on Ω, let $\mathcal{L}(X_1,\ldots,X_p)$ denote the Lie sub-algebra they generate. Thus $\mathcal{L}(X_1,\ldots,X_p)$ is spanned over the space of smooth functions by the vector fields

$$\begin{aligned}&X_1, X_2, \ldots, X_p;\\ &[X_1,X_2],[X_1,X_3],\ldots,[X_{p-1},X_p];\\ &[X_1,[X_1,X_2]],\ldots,\text{etc.}\end{aligned}$$

Let $\mathcal{L}(\mathcal{X})_{\mathbf{y}}$ be the span of all vectors $Y(\mathbf{y})$ where $Y\in\mathcal{L}(\mathcal{X})$. Then $\mathcal{X}=\{X_1,\ldots,X_p\}$ satisfies the *Hörmander condition* at $\mathbf{y}$ if $\mathcal{L}(\mathcal{X})_{\mathbf{y}}=\mathbb{R}^N_{\mathbf{y}}$.

1.4. **Consequences of Hörmander's condition.**

The result in Exercise 3 is a consequence of what is called Chow's Theorem (see [2]).

Theorem 1.1 (Chow)**.** *Let $\mathcal{X}=\{X_1,\ldots,X_p\}$ be smooth vector fields on $\Omega\subset\mathbb{R}^N$. Suppose that $\mathcal{X}$ satisfies Hörmander's condition at $\mathbf{y}$ for all $\mathbf{y}\in\Omega$. Then $\mathcal{C}(\mathcal{X})(\mathbf{y}_0,\mathbf{y}_1)\neq\emptyset$ for every pair of points $\mathbf{y}_0,\mathbf{y}_1\in\Omega$, and so the function $\rho_{\mathcal{X}}$ is a metric on Ω, the Carnot-Carathéodory metric.*

We will also study the properties of the Carnot-Carathéodory balls

$$B_{\mathcal{X}}(\mathbf{x},\delta)=\Big\{\mathbf{y}\in\Omega:\rho_{\mathcal{X}}(\mathbf{x},\mathbf{y})<\delta\Big\}, \tag{1.8}$$

and in particular, find estimates for their volumes. We will establish the following critical 'doubling property':

Theorem 1.2. *Let $\mathcal{X}=\{X_1,\ldots,X_p\}$ be smooth vector fields on $\Omega\subset\mathbb{R}^N$, and suppose that for each $\mathbf{y}\in\Omega$, $\big\{X(\mathbf{y}):X\in\mathcal{L}(X_1,\ldots,X_n)\big\}=\mathbb{R}^N_{\mathbf{y}}$. Let $K\Subset\Omega$ be a compact subset. Then there exist constants $\eta>0$ and $C>0$ so that for all $\mathbf{x}\in K$ and all $\delta<\eta$,*

$$\big|B_{\mathcal{X}}(\mathbf{x},2\delta)\big|\leq C\,\big|B_{\mathcal{X}}(\mathbf{x},\delta)\big|.$$

1.5. **Sums of Squares and Variants.**

Let $\mathcal{X}=\big\{X_1,\ldots,X_p\big\}$ be smooth vector fields on Ω, and consider the second order partial differential operator

$$\mathcal{L}=X_1^2+\cdots+X_p^2. \tag{1.9}$$

Motivated by results of J.J. Kohn in complex analysis ([7], [8]), Hörmander [5] established the following fundamental regularity result.

Theorem 1.3 (Hörmander)**.** *Suppose that the vector fields $\mathcal{X}=\{X_1,\ldots,X_p\}$ on $\Omega\subset\Omega$ satisfy the Hörmander condition at each point of Ω. Then the sum of squares operator $\mathcal{L}$ is hypo-elliptic: if u is a distribution on Ω and if $\mathcal{L}u\in\mathcal{C}^\infty(\Omega)$ where $U\subset\Omega$ is open, the $u\in\mathcal{C}^\infty(\Omega)$.*

Hörmander established the same result for a variant of $\mathcal{L}$. This time let $\mathcal{X}=\big\{X_0,X_1,\ldots,X_p\big\}$ be smooth vector fields and consider the second order partial differential operator

$$\mathcal{H}=X_0+X_1^2+\cdots+X_p^2. \tag{1.10}$$

Theorem 1.4 (Hörmander)**.** *Suppose that the vector fields $\mathcal{X}=\{X_0,X_1,\ldots,X_p\}$ on $\Omega\subset\Omega$ satisfy the Hörmander condition at each point of Ω. Then the operator $\mathcal{H}$ is hypo-elliptic.*

Going beyond Hörmander's results, we will establish more refined regularity theorems for the operators $\mathcal{L}$ and $\mathcal{H}$ by constructing parametrices for these operators. (A parametrix for $\mathcal{L}$ is an operator $\mathcal{K}$ so that the operators $I - \mathcal{K}\mathcal{L}$ and $I - \mathcal{L}\mathcal{K}$ are either zero, in which case we have found an inverse, or are appropriately smoothing.) To describe these parametrices, one needs to use the corresponding Carnot-Carathéodory metrics. In the case of the vector fields $\mathcal{X} = \{X_1, \ldots, X_p\}$ satisfying Hörmander's condition, the operator $\mathcal{K}$ will be given as an integral operator

$$\mathcal{K}f(\mathbf{x}) = \int_{\Omega} K(\mathbf{x}, \mathbf{y}) f(\mathbf{y})\, d\mathbf{y} \tag{1.11}$$

where the kernel $K \in \mathcal{C}^{\infty}(\Omega \times \Omega \setminus \text{diagonal})$ satisfies differential inequalities

$$\left|X^{\alpha}_{\mathbf{x}} X^{\beta}_{\mathbf{y}} K(\mathbf{x}, \mathbf{y})\right| \lesssim \rho_{\mathcal{X}}(\mathbf{x}, \mathbf{y})^{2-|\alpha|-|\beta|} \left|B_{\mathcal{X}}(\mathbf{x}, \rho_{\mathcal{X}}(\mathbf{x}, \mathbf{y})\right|^{-1} \tag{1.12}$$

for $\mathbf{x}, \mathbf{y} \in E \Subset \Omega$ with $\mathbf{x} \neq \mathbf{y}$. Here if $\alpha = (\alpha_1, \ldots, \alpha_p)$ is a multi-index, $X^{\alpha} = X_1^{\alpha_1} \cdots X_p^{\alpha_p}$. There are similar results for the operator $\mathcal{H}$, but in that case one needs to modify the definition of the Carnot-Carathéodory metric with the operator X_0 treated like a second order operator. With the parametrix in hand, we will obtain regularity results for the equations $\mathcal{L}u = g$ or $\mathcal{H}u = g$ when g belongs to an L^p-Sobolev or Lipschitz space.

1.6. **Four Classical Examples.**

Before starting on the general theory, we will look at several classical examples.

1) *The Laplace operator* Δ *on* $\mathbb{R}^n$. If $X_j = \partial_{x_j}$, then

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2} = \sum_{j=1}^{n} X_j^2.$$

2) *The heat operator* $\mathcal{H}$ *on* $\mathbb{R}^{n+1}$. If $X_j = \partial_{x_j}$ for $1 \leq j \leq n$ and $X_0 = -\partial_t$, then the inhomogeneous heat equation

$$\frac{\partial u}{\partial t}(\mathbf{x}, t) = \frac{\partial^2 u}{\partial x_1^2}(\mathbf{x}, t) + \cdots + \frac{\partial^2 u}{\partial x_n^2}(\mathbf{x}, t) - g(\mathbf{x}, t)$$

can be written $X_0 u + \sum_{j=1}^{n} X_j^2 u = g$.

3) *An example of Grushin on* $\mathbb{R}^2$. If $X = \partial_x$ and $Y = x\partial_y$, consider the equation

$$\mathcal{G}u = \frac{\partial^2 u}{\partial x^2} + x^2 \frac{\partial^2 u}{\partial y^2} = (X^2 + Y^2)u = g.$$

The operator $\mathcal{G}$ is elliptic except where $x = 0$. However we will also be able to obtain sharp estimates when $x = 0$.

4) *Complex analysis and the Heisenberg group.* We use coordinates $(x_1, \ldots, x_n, y_1, \ldots, y_n, t)$ on $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{2n+1}$. Set

$$X_j = \frac{\partial}{\partial x_j} - 2y_j \frac{\partial}{\partial t},$$
$$Y_j = \frac{\partial}{\partial y_j} + 2x_j \frac{\partial}{\partial t},$$
$$T = \frac{\partial}{\partial t}.$$

It is easy to check that $[X_j, X_k] = [Y_j, Y_k] = [X_j, T] = [Y_j, T] = 0$ for $1 \le j, k \le n$ and that $[X_j, Y_k] = 4\delta_{j,k}T$. The *Kohn Laplacian* for the Siegal upper half-space in $\mathbb{C}^{n+1}$ is the operator

$$\Box_b = \sum_{j=1}^{n} \big(X_j^2 + Y_j^2\big).$$

2. The Laplace operator: fundamental solutions and hypoellipticity

2.1. Invariance properties of Δ.

The Laplace operator has invariance properties related to the geometry of $\mathbb{R}^n$ that can be expressed by the commutation properties of the operator Δ with three kinds of mappings: translations, dilations, and rotations. Let $\mathbf{y} \in \mathbb{R}^n$, let $\lambda > 0$, and let $U \in O(n)$ (the orthogonal group) so that U is a linear transformation on $\mathbb{R}^n$ satisfying $UU^* = I$. If $f \in L^1_{loc}(\mathbb{R}^n)$, set

$$\begin{aligned} \tau_{\mathbf{y}} f(\mathbf{x}) &= f(\mathbf{x} - \mathbf{y}) && \text{(translation by } \mathbf{y}); \\ D_\lambda f(\mathbf{x}) &= f(\lambda \mathbf{x}) && \text{(dilation by } \lambda); \\ R_U f(\mathbf{x}) &= f(U\mathbf{x}) && \text{(rotation by } U). \end{aligned} \tag{2.1}$$

If $f \in L^1_{loc}(\mathbb{R}^n)$ and $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ then

$$\begin{aligned}
\int_{\mathbb{R}^n} \tau_{\mathbf{y}} f(\mathbf{x})\varphi(\mathbf{x})\,d\mathbf{x} &= \int_{\mathbb{R}^n} f(\mathbf{x}-\mathbf{y})\varphi(\mathbf{x})\,d\mathbf{x} = \int_{\mathbb{R}^n} f(\mathbf{x})\varphi(\mathbf{x}+\mathbf{y})\,d\mathbf{y} = \int_{\mathbb{R}^n} f(\mathbf{x})\tau_{-\mathbf{y}}\varphi(\mathbf{x})\,d\mathbf{x}, \\
\int_{\mathbb{R}^n} D_\lambda f(\mathbf{x})\varphi(\mathbf{x})\,d\mathbf{x} &= \int_{\mathbb{R}^n} f(\lambda\mathbf{x})\varphi(\mathbf{x})\,d\mathbf{x} = \int_{\mathbb{R}^n} f(\mathbf{x})\lambda^{-n}\varphi(\lambda^{-1}\mathbf{x})\,d\mathbf{x} = \int_{\mathbb{R}^n} f(\mathbf{x})\lambda^{-n}D_{\lambda^{-1}}\varphi(\mathbf{x})\,d\mathbf{x}, \\
\int_{\mathbb{R}^n} R_U f(\mathbf{x})\varphi(\mathbf{x})\,d\mathbf{x} &= \int_{\mathbb{R}^n} f(U\mathbf{x})\varphi(\mathbf{x})\,d\mathbf{x} = \int_{\mathbb{R}^n} f(\mathbf{x})\varphi(U^*\mathbf{x})\,d\mathbf{x} = \int_{\mathbb{R}^n} f(\mathbf{x})R_{U^*}\varphi(\mathbf{x})\,d\mathbf{x}.
\end{aligned}$$

The operators $\tau_{\mathbf{y}}$, D_λ, and R_U extend mappings of the space $\mathcal{D}'(\mathbb{R}^n)$ of distributions to itself[2]: if $u \in \mathcal{D}'(R^n)$ and $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$, set

$$\begin{aligned} \big\langle \tau_{\mathbf{y}} u, \varphi \big\rangle &= \big\langle u, \tau_{-\mathbf{y}}\varphi \big\rangle; \\ \big\langle D_\lambda u, \varphi \big\rangle &= \big\langle u, \lambda^{-n} D_{\lambda^{-1}}\varphi \big\rangle; \\ \big\langle R_U u, \varphi \big\rangle &= \big\langle u, R_{U^*}\varphi \big\rangle. \end{aligned} \tag{2.2}$$

The three operators $\tau_{\mathbf{y}}$, D_λ, R_U and the Laplacian Δ thus all map the space $\mathcal{D}'(\mathbb{R}^n)$ to itself. We have the following commutation properties.

Proposition 2.1. *Let $\mathbf{y} \in \mathbb{R}^n$, $\lambda > 0$, and $U \in O(n)$. Then*

$$\begin{aligned} \Delta\tau_{\mathbf{y}} &= \tau_{\mathbf{y}}\Delta \\ \Delta D_\lambda &= \lambda^2 D_\lambda \Delta \\ \Delta R_U &= R_U \Delta. \end{aligned}$$

Proof. These identities are easy computations. The first identity holds for any constant coefficient differential operator of any order, and the second holds for any constant coefficient differential operator of order exactly two.

[2] An excellent reference for material on distributions is [6]. A brief summary of needed material is presented in Appendix A below.

The third is a special case of a more general commutation formula. Let $A = \{a_{j,k}\}$ and $M = \{m_{j,k}\}$ be real $n \times n$ matrices and for $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ set $P_A[\varphi](\mathbf{x}) = \sum_{j,k=1}^n a_{j,k}\, \partial^2_{x_j,x_k} \varphi(\mathbf{x})$ and $R_M[\varphi](\mathbf{x}) = \varphi(M\mathbf{x})$. The chain rule shows that

$$\sum_{j,k=1}^n a_{j,k} \partial^2_{x_j,x_k} R_M[\varphi](\mathbf{x}) = \sum_{r,s=1}^n \Big[\sum_{j,k=1}^n m_{r,j} a_{j,k} m_{s,k} \Big] \partial^2_{x_r,x_s} \varphi(M\mathbf{x})$$

so that $P_A R_M = R_M P_{MAM^*}$ where M^* is the transpose (or adjoint) of M.

Now if A is the identity matrix, the operator $P_A = \Delta$ is the Laplace operator, and if $M = U \in O(n)$, then $UU^* = I$. It follows that $\Delta R_U[\varphi] = R_U \Delta[\varphi]$ for every test function φ. If $u \in \mathcal{E}'(\mathbb{R}^n)$ and $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$, it follows from the third equality in (2.2) that

$$\begin{aligned} \big\langle \Delta R_U u, \varphi \big\rangle &= \big\langle R_U u, \Delta\varphi \big\rangle = \big\langle u, R_{U^*} \Delta\varphi \big\rangle \\ &= \big\langle u, \Delta R_{U^*} \varphi \big\rangle = \big\langle \Delta u, R_{U^*}\varphi \big\rangle = \big\langle R_U \Delta u, \varphi \big\rangle \end{aligned}$$

Thus $\Delta R_U = R_U \Delta$. □

2.2. **A fundamental solution: the Newtonian potential.**

Let $\delta_{\mathbf{x}} \in \mathcal{D}'(\mathbb{R}^n)$ be the Dirac delta-function at $\mathbf{x} \in \mathbb{R}^n$ given by $\langle \delta_{\mathbf{x}}, \varphi \rangle = \varphi(\mathbf{x})$. Note that $\delta_{\mathbf{x}} = \tau_{\mathbf{x}} \delta_0$. We look for a *fundamental solution* for the Laplace operator at $\mathbf{x}$; *i.e.* a tempered distribution $\mathcal{N}_{\mathbf{x}} \in \mathcal{S}'(\mathbb{R}^n)$ such that $\Delta \mathcal{N}_{\mathbf{x}} = \delta_{\mathbf{x}}$. Observe that the translation invariance of Δ means that we only need to find a fundamental solution at 0. Thus if $\Delta\mathcal{N} = \delta_0$, it follows that

$$\delta_{\mathbf{x}} = \tau_{\mathbf{x}}\delta_0 = \tau_{\mathbf{x}}\Delta\mathcal{N} = \Delta\tau_{\mathbf{x}}\mathcal{N}$$

so that $\mathcal{N}_{\mathbf{x}} \equiv \tau_{\mathbf{x}}\mathcal{N}$ is a fundamental solution at $\mathbf{x}$.

The symmetries of Δ suggest the form of a fundamental solution. Note that $D_\lambda \delta_0 = \lambda^{-n}\delta_0$ and $R_U\delta_0 = \delta_0 R_U$. If $\mathcal{N}$ is a fundamental solution for Δ, it follows from Proposition 2.1 that

$$\begin{aligned} \Delta\big[\lambda^{n-2} D_\lambda \mathcal{N}\big] &= \lambda^{n-2}\Delta D_\lambda\mathcal{N} = \lambda^n D_\lambda \Delta\mathcal{N} = \lambda^n D_\lambda \delta_0 &&= \delta_0, \\ \Delta\big[R_U \mathcal{N}\big] &= R_U \Delta\mathcal{N} = R_U\delta_0 &&= \delta_0. \end{aligned}$$

Thus $\lambda^{n-2}D_\lambda\mathcal{N}$ and $R_U\mathcal{N}$ would be fundamental solutions for Δ as well. This suggests that we look a fundamental solution such that $\mathcal{N} = R_U\mathcal{N}$ and $\mathcal{N} = \lambda^{n-2}D_\lambda\mathcal{N}$. The first identity means that $N(\mathbf{x})$ depends only on $|\mathbf{x}|$, and the second identity means that $N(\lambda\mathbf{x}) = \lambda^{2-n}N(\mathbf{x})$. This suggest that we might consider $N(\mathbf{x}) = c|\mathbf{x}|^{2-n}$ for some constant c. When $n > 2$, we see in Lemma 2.3 below that this heuristic reasoning is correct.[3]

Definition 2.2. *Let*

$$\omega_n = 2\pi^{\frac{n}{2}}\,\Gamma\left(\frac{n}{2}\right)^{-1}$$

be the surface measure of the unit sphere in $\mathbb{R}^n$. *For* $n \geq 2$ *the Newtonian potential on* $\mathbb{R}^n$ *is the function*

$$N(\mathbf{x}) = \begin{cases} \omega_2^{-1}\, \log(|\mathbf{x}|) & \text{when } n = 2; \\ \\ \omega_n^{-1}\,(2-n)^{-1}\,|\mathbf{x}|^{2-n} & \text{when } n > 2. \end{cases}$$

Using polar coordinates, it is easy to check that the function $N \in L^p_{\mathrm{loc}}(\mathbb{R}^n)$ provided that $p < \frac{n}{n-2}$, and in particiluar, N is always locally integrable. We let $\mathcal{N} \in \mathcal{S}'(\mathbb{R}^n)$ be the tempered distribution given by integration against N.

[3] Note that when $n = 2$ this formula cannot be correct.

Lemma 2.3. *The distribution $\mathcal{N}$ is a fundamental solution for Δ.*

Proof. We must show that $\Delta\mathcal{N} = \delta_0$ as distributions. Since for every $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$, $\langle \Delta\mathcal{N}, \varphi \rangle = \langle \mathcal{N}, \Delta\varphi \rangle$ so we must show that $\langle \delta_0, \varphi \rangle = \langle \mathcal{N}, \Delta\varphi \rangle$, *i.e.*

$$\varphi(0) = \int_{\mathbb{R}^n} N(\mathbf{y})\Delta\varphi(\mathbf{y})\, d\mathbf{y}.$$

Choose $R > 0$ so large that the compact support of φ is contained in the open Euclidean ball centered at the origin of radius R. For any $0 < \epsilon < R$, the function $\mathbf{x} \to N(\mathbf{x})$ has no singularities in a neighborhood of the closure of the spherical shell $B(\epsilon, R) = \{\mathbf{x} \in \mathbb{R}^n \,\big|\, \epsilon < |\mathbf{x}| < R\}$, so Green's theorem gives

$$\int\limits_{B(\epsilon,R)} \Big[N(\mathbf{x}) \,\triangle\varphi(\mathbf{x}) - \varphi(\mathbf{x}) \,\triangle N(\mathbf{x})\Big]\, d\mathbf{x} \;=\; \int\limits_{\partial B(\epsilon,R)} \Big[N(\zeta)\partial_n\varphi(\zeta) - \varphi(\zeta)\, \partial_n\varphi(\zeta)\Big]\, d\sigma(\zeta). \tag{2.3}$$

Here ∂_n denotes the outward unit normal derivative on the boundary $\partial B(\epsilon, R)$.

The function N is infinitely differentiable away from the origin, and a direct calculation shows that $\triangle[N](\mathbf{x}) = 0$ for $\mathbf{x} \neq 0$. Thus, since φ and $\Delta[\varphi]$ have compact support inside the ball of radius R, the left-hand side of equation (2.3) reduces to

$$\int\limits_{B(\epsilon,R)} N(\mathbf{x})\triangle\varphi(\mathbf{x})\, d\mathbf{x} = \int\limits_{|\mathbf{x}|>\epsilon} N(\mathbf{x})\triangle\varphi(\mathbf{x})\, d\mathbf{x}.$$

To deal with the right-hand side of equation (2.3), note that the boundary of $B(\epsilon, R)$ has two connected components: the set S_R where $|\mathbf{x}| = R$ and the set S_ϵ where $|\mathbf{x}| = \epsilon$. The function φ is identically zero in a neighborhood of S_R, so this part of the boundary gives no contribution. Thus equation (2.3) reduces to

$$\int\limits_{|\mathbf{x}|>\epsilon} N(\mathbf{x}) \,\triangle\varphi(\mathbf{x})\, d\mathbf{x} = \int\limits_{|\zeta|=\epsilon} \Big[N(\zeta)\partial_n\varphi(\zeta) - \varphi(\zeta)\, \partial_n\varphi(\zeta)\Big]\, d\sigma(\zeta).$$

On S_ϵ the outward unit normal derivative is $\partial_n = -\partial_r$ where $r = |\mathbf{x}|$ is the distance from the origin. It follows from the explicit formula for $N(\mathbf{x})$ that when $|\zeta| = \epsilon$ we have

$$N(\zeta) = \begin{cases} \omega_2^{-1}\log(\epsilon) & \text{if } n = 2, \\[2ex] \omega_n^{-1}\,(2-n)^{-1}\epsilon^{2-n} & \text{if } n > 2, \end{cases}$$

and

$$\frac{\partial N}{\partial n}(\zeta) = -\omega_n^{-1}\epsilon^{1-n} \quad \text{for } n \geq 2.$$

Since ω_n is the surface measure of the unit sphere, $\varphi(0) = \omega_n^{-1}\epsilon^{1-n}\int_{|\zeta|=\epsilon}\varphi(0)\, d\sigma(\zeta)$, so adding and subtracting $\varphi(0)$, we have

$$\int_{|\mathbf{x}|\geq\epsilon} N(\mathbf{x}) \,\triangle\varphi(\mathbf{x})\, d\mathbf{x}$$

$$= \varphi(0) + \omega_n^{-1}\epsilon^{1-n}\int_{|\zeta|=\epsilon}\big[\varphi(\zeta) - \varphi(0)\big]\, d\sigma(\zeta) - \begin{cases} \omega_2^{-1}\,\log(\epsilon)\int_{S_\epsilon}\partial_r\varphi(\zeta)\, d\sigma(\zeta) & \text{if } n = 2 \\[2ex] \omega_n^{-1}\,\epsilon^{2-n}\int_{S_\epsilon}\partial_r\varphi(\zeta)\, d\sigma(\zeta) & \text{if } n > 2 \end{cases}.$$

Now let $\epsilon \to 0$. The Lebesgue dominated convergence theorem shows that

$$\lim_{\epsilon\to 0}\int_{|x|\geq\epsilon} N(x)\,\triangle[\varphi](x)\, dx = \int_{\mathbb{R}^n} N(x)\,\triangle\varphi(x)\, dx.$$

Also, as $\epsilon \to 0$,

$$\left|\omega_n^{-1}\epsilon^{1-n}\int_{S_\epsilon}\Big[\varphi(\zeta)-\varphi(0)\Big]\,d\sigma(\zeta)\right| \le \epsilon \sup_{\mathbf{x}\in\mathbb{R}^n}\left|\nabla\varphi(\mathbf{x})\right| \to 0;$$

$$\left|\omega_2^{-1}\,\log(\epsilon)\int_{S_\epsilon}\frac{\partial\varphi}{\partial n}(\zeta)\,d\zeta\right| \le \epsilon\,\log(\epsilon)\sup_{\mathbf{x}\in\mathbb{R}^2}\left|\nabla\varphi(\mathbf{x})\right| \to 0;$$

$$\left|\omega_n^{-1}\,(2-n)^{-1}\,\epsilon^{2-n}\int_{S_\epsilon}\frac{\partial\varphi}{\partial n}(\zeta)\,d\zeta\right| \le \epsilon\,(2-n)^{-1}\sup_{\mathbf{x}\in\mathbb{R}^n}\left|\nabla\varphi(\mathbf{x})\right| \to 0.$$

Thus the limit as $\epsilon \to 0$ of the right-hand side of equation (2.3) is $\varphi(0)$, and so $\varphi(0) = \int_{\mathbb{R}^n} N(\mathbf{y})\,\triangle\varphi(\mathbf{y})\,d\mathbf{y}$. This completes the proof. □

2.3. **Solutions of the Poisson equation $\Delta u = g$.**

We first consider the case when the right hand side in this equation is a test function. Since $N \in L^1_{loc}(\mathbb{R}^n)$, if $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ the convolution $\varphi * N$ is given by the absolutely convergent integral

$$\varphi * N(\mathbf{x}) = \int_{\mathbb{R}^n}\varphi(\mathbf{x}-\mathbf{y})N(\mathbf{y})\,d\mathbf{y},$$

and differentiating under the integral sign, it easily follows that $\varphi * N \in \mathcal{C}^\infty(\mathbb{R}^n)$ with

$$\partial^\alpha[\varphi * N](\mathbf{x}) = \int_{\mathbb{R}^n}\partial^\alpha\varphi(\mathbf{x}-\mathbf{y})N(\mathbf{y})\,d\mathbf{y} = [\partial^\alpha\varphi] * N(\mathbf{x}).$$

If $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$, set $\widetilde{\varphi}(\mathbf{y}) = \varphi(-\mathbf{y})$ and observe that $\Delta\widetilde{\varphi}(\mathbf{y}) = \Delta\varphi(-\mathbf{y})$. Then

$$\begin{aligned}\Delta[\varphi * N](\mathbf{x}) &= \int_{\mathbb{R}^n}\Delta\varphi(\mathbf{x}-\mathbf{y})N(\mathbf{y})\,dy = \int_{\mathbb{R}^n}\Delta\widetilde{\varphi}(\mathbf{y}-\mathbf{x})N(\mathbf{y})\,dy = \int_{\mathbb{R}^n}\tau_{\mathbf{x}}\Delta\widetilde{\varphi}(\mathbf{y})N(\mathbf{y})\,dy\\ &= \big\langle\mathcal{N},\tau_{\mathbf{x}}\Delta\widetilde{\varphi}\big\rangle = \big\langle\mathcal{N},\Delta\tau_{\mathbf{x}}\widetilde{\varphi}\big\rangle = \big\langle\Delta\mathcal{N},\tau_{\mathbf{x}}\widetilde{\varphi}\big\rangle = \big\langle\delta_0,\tau_{\mathbf{x}}\widetilde{\varphi}\big\rangle = \tau_{\mathbf{x}}\widetilde{\varphi}(0) = \varphi(\mathbf{x}).\end{aligned}$$

There is a similar result if the right hand side in the Poisson equation is a distribution $u \in \mathcal{E}'(\mathbb{R}^n)$ with compact support. We want to define $u * N$. Note that formally

$$\begin{aligned}\int_{\mathbb{R}^n} u * N(\mathbf{x})\varphi(\mathbf{x})\,d\mathbf{x} &= \iint_{\mathbb{R}^{2n}} u(\mathbf{y})N(\mathbf{x}-\mathbf{y})\varphi(\mathbf{x})\,d\mathbf{y}\,d\mathbf{x} = \int_{\mathbb{R}^n} u(\mathbf{y})\int_{\mathbb{R}^n}\widetilde{N}(\mathbf{y}-\mathbf{x})\varphi(\mathbf{x})\,d\mathbf{x}\,d\mathbf{y}\\ &= \int_{\mathbb{R}^n} u(\mathbf{y})\,\varphi * \widetilde{N}(\mathbf{y})\,dy.\end{aligned}$$

Since $\widetilde{N} = N$, and $N * \varphi \in \mathcal{C}^\infty(\mathbb{R}^n)$ if $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$, we can rigorously define the distribution $u * N$ by the formula

$$\big\langle u * N, \varphi\big\rangle = \big\langle u, \varphi * N\big\rangle. \tag{2.4}$$

Then

$$\big\langle\Delta[u * N],\varphi\big\rangle = \big\langle u * N,\Delta\varphi\big\rangle = \big\langle u,\Delta\varphi * N\big\rangle = \big\langle u,\varphi\big\rangle$$

and

$$\big\langle\Delta u * \mathcal{N},\varphi\big\rangle = \big\langle\Delta u,\varphi * N\big\rangle = \big\langle u,\Delta[\varphi * N]\big\rangle = \big\langle u,\varphi\big\rangle$$

so

$$\Delta[u * N] = \Delta u * N = u.$$

Thus we have proved

Lemma 2.4.

a) *If* $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$, *then* $\varphi * N \in \mathcal{C}^\infty(\mathbb{R}^n)$, *and* $\Delta[\varphi * N] = \Delta\varphi * N = \varphi$.

b) *If* $u \in \mathcal{E}'(\mathbb{R}^n)$ *then* $u * N \in \mathcal{D}'(\mathbb{R}^n)$, *and* $\Delta[u * N] = \Delta u * N = u$.

We denote by N the operator

$$Nf = \mathcal{N} * f$$

when the convolution is defined.

2.4. **Hypoellipticity of Δ.**

A differential operator $P[f](\mathbf{x}) = \sum_{|\alpha|\leq m} a_\alpha(\mathbf{x})\partial^\alpha f(\mathbf{x})$ on an open set $\Omega \subset \mathbb{R}^n$ is *hypoelliptic* if whenever $u \in \mathcal{D}'(\Omega)$ is a distribution such that $P[u] \in \mathcal{C}^\infty(\Omega)$, it follows that $u \in \mathcal{C}^\infty(\Omega)$.

Theorem 2.5 (Weyl's Lemma)**.** *The Laplace operator is hypoelliptic.*

Proof. Let $V \Subset U$ be any relatively compact open subset. Choose $\chi \in \mathcal{C}_0^\infty(\Omega)$ with $\chi(\mathbf{x}) \equiv 1$ in an open neighborhood of $\overline{V}$. The distribution χu then has compact support, and $\Delta[\chi u]$ is given on V by integration against a $\mathcal{C}^\infty$-function. But

$$\chi u = N * \Delta[\chi u]$$

and hence

$$\begin{aligned}\text{Singular support } (\chi u) &\subset \text{Singular support } (N) + \text{Singular support } (\chi u)\\ &= \{0\} + \text{Singular support } (\chi u)\\ &= \text{Singular support } (\chi u)\\ &\subset \mathbb{R}^n \setminus V.\end{aligned}$$

Thus χu is given on V by integration against a $\mathcal{C}^\infty$-function. □

3. Regularity of the Laplace operator

Suppose that u and g are distributions on a domain $\Omega \subset \mathbb{R}^n$, and that $\Delta[u] = g$. We can use the Newtonian potential to obtain information on the regularity of the solution u in terms of regularity of the given data g. Thus let $\chi \in \mathcal{C}_0^\infty(\Omega)$, and suppose $\chi(\mathbf{x}) \equiv 1$ for $\mathbf{x} \in \Omega_1 \Subset \Omega$. Then χg is a distribution with compact support, and according to part (2) of Lemma 2.3, we have $\Delta\big[N[\chi g]\big] = \chi g$. Hence $\Delta\big[u - N[\chi g]\big] = (1 - \chi)g$, which is identically zero on Ω_1. It follows that the distribution $u - N[\chi g]$ is given by integration against an infinitely differentiable (actually real-analytic) function. Thus modulo smooth functions, u and $N[\chi g]$ have the same regularity.

We will discuss classical regularity results in three scales of function spaces.

A. We study the behavior of the Newtonian potential N on the scale of L^p spaces on $\mathbb{R}^n$. This operator maps $L^p(\mathbb{R}^n)$ to $L^q(\mathbb{R}^n)$ with $q > p$, and it is this sense that the operator is 'improving' or 'smoothing'. Our main result is the Hardy-Littlewood Sobolev theorem on fractional integration. Since we are dealing with an operator which is smoothing of positive order, cancellation conditions are not needed.

B. We study the behavior of the Newtonian potential N on the scale of Lipschitz spaces $\Lambda^k_\alpha(\mathbb{R}^n)$. These are spaces of functions which are k-times continuously differentiable, and all k^{th}-order derivatives satisfy a Lipschitz or Hölder condition of order α. (This is defined precisely in Section 3.4 below). We show that for $0 < \alpha < 1$, the operator N gains two derivatives: if $f \in \Lambda^k_\alpha(\mathbb{R}^n)$ then $N[f] \in \Lambda^{k+2}_\alpha(\mathbb{R}^n)$. Such results are often known as *Schauder* estimates.

C. Instead of measuring smoothness with the scale of spaces $\{\Lambda^k_\alpha(\mathbb{R}^n)\}$, one can use Sobolev spaces $L^p_k(\mathbb{R}^n)$ of functions which have k-derivatives in $L^p(\mathbb{R}^n)$. The precise definition will be given in Section 3.5, and the result is that if $f \in L^p_k(\mathbb{R}^n$ then $N * f \in L^p_{k+2}(\mathbb{R}^n)$.

There are many good references for a detailed discussion of these matters[4], and it is not our objective to give an exhaustive account of this material. Rather, we want to indicate that many of these results do not depend on the explicit formula for N given in Definition 2.2, but are true for any operator $\mathcal{K}[f](\mathbf{x}) = \int K(\mathbf{x},\mathbf{y})f(\mathbf{y})\,d\mathbf{y}$ whose Schwartz kernel $K(\mathbf{x},\mathbf{y})$ satisfies appropriate differential inequalities and cancellation conditions. Our purpose here is to formulate one possible definition of such a more general class of operators. It is convenient to formulate this theory in terms of the notion of spaces of 'homogeneous type'.[5]

3.1. **Spaces of homogeneous type.**

Let $\Omega \subset \mathbb{R}^n$ be open, let ρ be a metric on Ω, and let $d\mathbf{x}$ be Lebesgue measure. We say that the triple $(\Omega, \rho, d\mathbf{x})$ is a *space of homogeneous type* if there is a constant A so that for all $\mathbf{x} \in \Omega$ and all $\delta > 0$,

$$\big|\mathbb{B}(\mathbf{x}, 2\delta)\big| \leq A\,\big|\mathbb{B}(\mathbf{x}, \delta)\big|$$

where $|E|$ is the Lebesgue measure of a subset $E \subset \Omega$, and $\mathbb{B}(\mathbf{x},\delta) = \{\mathbf{y} \in \Omega : \rho(\mathbf{x},\mathbf{y}) < \delta\}$. Note that the space $\mathbb{R}^n$ or any open subset with the standard Euclidean metric

$$d_E(x,y) = |x-y| = \Big(\sum_{j=1}^n (x_j - y_j)^2\Big)^{\frac{1}{2}}$$

is a space of homogeneous type since the volume of the Euclidean ball is

$$\big|\mathbb{B}_E(\mathbf{x},\delta)\big| = \frac{\omega_n}{n}\,\delta^n.$$

When $n \geq 3$, there is an explicit connection between the Newtonian potential N and the Euclidean metric d_E since we can write

$$N(\mathbf{x}-\mathbf{y}) = -\frac{1}{n(n-2)}\,\frac{d_E(\mathbf{x},\mathbf{y})^2}{\big|\mathbb{B}_E\big(\mathbf{x}, d_E(\mathbf{x},\mathbf{y})\big)\big|}.$$

At this stage, it may seem strange to introduce the inverse of the volume of the ball of radius $d_E(\mathbf{x},\mathbf{y})$ rather than simply write a constant times $d_E(\mathbf{x},\mathbf{y})^{2-n}$. However, in later examples, we will see that the volume of the ball is not always equal to a power of the radius, but estimates of fundamental solutions still have this form. Using just the Euclidean distance and volume, we formulate the essential estimates for derivatives of the Newtonian potential as follows.

---

[4]See, for example, books on elliptic partial differential equations such as [1] or [4], or books on pseudo-differential operators such as [11].

[5]In the late 1960's and early 1970's there were a number of authors who introduced concepts like this, but the term may have been first used in [3].

Lemma 3.1. *Suppose $n \geq 3$ and let $N(\mathbf{x},\mathbf{y}) = N(\mathbf{x}-\mathbf{y}) = c_n|\mathbf{x}-\mathbf{y}|^{-n+2}$. For any multi-indices α, β there is a constant $C_{\alpha,\beta} > 0$ so that for all $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$*

$$\big|\partial_{\mathbf{x}}^{\alpha}\,\partial_{\mathbf{y}}^{\beta}\,N(\mathbf{x},\mathbf{y})\big| \leq C_{\alpha,\beta}\,\frac{d_E(\mathbf{x},\mathbf{y})^{2-|\alpha|-|\beta|}}{\big|\mathbb{B}_E\big(\mathbf{x}, d_E(\mathbf{x},\mathbf{y})\big)\big|}.$$

The same inequality holds when $n = 2$ provided that $|\alpha| + |\beta| > 0$.

Proof. The function N is homogeneous of degree $-n+2$, and $\partial_{\mathbf{x}}^{\alpha} N$ is homogeneous of degree $-n-\alpha+2$. Thus the function $\mathbf{x} \to |\mathbf{x}|^{n+\alpha-2}\,\partial_{\mathbf{x}}^{\alpha} N(\mathbf{x})$ is continuous and homogeneous of degree zero on the unit sphere, and hence bounded there. This gives the required estimate. □

3.2. Calderón-Zygmund kernels of positive order.

Let $\Omega \subset \mathbb{R}^n$ be a connected open set and let $d : \Omega \times \Omega \to [0,\infty)$ be a metric. For $\mathbf{x} \in \Omega$, let $B(\mathbf{x},\delta) = \big\{\mathbf{y} \in \Omega : d(\mathbf{x},\mathbf{y}) < \delta\big\}$, and assume that $B(\mathbf{x},\delta)$ is an open subset of Ω for all $\mathbf{x}$ and δ. Assume also that there is a constant $A > 0$ so that for all $\mathbf{x} \in \Omega$ and all $\delta > 0$,

$$\big|B(\mathbf{x},2\delta)\big| \leq A\,\big|B(\mathbf{x},\delta)\big|.$$

Let $D_\Omega = \{(\mathbf{x},\mathbf{y}) \in \Omega \times \Omega : \mathbf{x} = \mathbf{y}\}$ be the diagonal. A *Calderón-Zygmund kernel of positive order $m > 0$* on Ω is a smooth function $K \in \mathcal{C}^\infty(\Omega \times \Omega \setminus D_\Omega)$ such that for all multi-indices α and β there is a constant $C_{\alpha,\beta}$ such that

$$\big|\partial_{\mathbf{x}}^{\alpha}\partial_{\mathbf{y}}^{\beta} K(\mathbf{x},\mathbf{y})\big| \leq C_{\alpha,\beta}\,d(\mathbf{x},\mathbf{y})^{m-|\alpha|-|\beta|}\big|B\big(\mathbf{x},d(\mathbf{x},\mathbf{y})\big)\big|^{-1}.$$

Lemma 3.2. *If K is a Calderón-Zygmund kernel on Ω of order $m > 0$, the functions $\mathbf{x} \to K(\mathbf{x},\mathbf{y})$ and $\mathbf{y} \to K(\mathbf{x},\mathbf{y})$ are locally integrable. There is a constant $C(A,m) > 0$ so that for all $\mathbf{x},\mathbf{y} \in \Omega$,*

$$\int\limits_{\mathbf{t}\in B(\mathbf{x},1)} |K(\mathbf{x},\mathbf{t})|\,d\mathbf{t} \leq C(A,m) \quad \textit{and} \quad \int\limits_{\mathbf{t}\in B1,\mathbf{y})} |K(\mathbf{t},\mathbf{y})|\,d\mathbf{t} \leq C(A,m).$$

Proof. Since K is smooth away from the diagonal, it is clearly locally integrable away from the diagonal. Thus it suffices to prove the two estimates above. If $B_j(\mathbf{x}) = B(\mathbf{x},2^{-j}) \setminus \overline{B(\mathbf{x},2^{-j-1})}$,

$$\begin{aligned}\int\limits_{\mathbf{t}\in B(\mathbf{x},1)} |K(\mathbf{x},\mathbf{t})|\,d\mathbf{t} &= \sum_{j=0}^{\infty}\int\limits_{B_j(\mathbf{x})} |K(\mathbf{x},\mathbf{t})|\,dt \leq \sum_{j=0}^{\infty}\int\limits_{B_j(\mathbf{x})} d(\mathbf{x},\mathbf{t})^m\big|B\big(\mathbf{x},d(\mathbf{x},\mathbf{t})\big)\big|^{-1}\,dt \\ &\leq \sum_{j=0}^{\infty} 2^{-mj}|B(\mathbf{x},2^{-j-1})|^{-1}\int_{B_j} d\mathbf{t} \leq \sum_{j=0}^{\infty} 2^{-mj}|B(\mathbf{x},2^{-j-1})|^{-1}\int_{B(\mathbf{x},2^{-j})} d\mathbf{t} \\ &= \sum_{j=0}^{\infty} 2^{-mj}|B(\mathbf{x},2^{-j-1})|^{-1}\big|B(\mathbf{x},2^{-j})\big| \leq A\sum_{j=0}^{\infty}[2^{-m}]^j = A[1-2^{-m}]^{-1}.\end{aligned}$$

The second estimate is proved in the same way. □

3.3. Hardy-Littlewood-Sobolev estimates.

We study the regularity of the Newtonian potential or more general Calderón-Zygmund operators of positive order on the scale of L^p-spaces. For example, if $f \in L^p(\mathbb{R}^n)$, we investigate for which q is it true that $N[f] \in L^q(\mathbb{R}^n)$ or $N[f] \in L^q_{\mathrm{loc}}(\mathbb{R}^n)$ where

$$N[f](\mathbf{x}) = N * f(\mathbf{x}) = c_n\int_{\mathbb{R}^n} |\mathbf{x}-\mathbf{y}|^{-n+2} f(\mathbf{y})\,d\mathbf{y}.$$

We start by observing that invariance properties of the function N easily provide a necessary condition for the existence of a constant $C_{p,q}$ such that for all $f \in L^p(\mathbb{R}^n)$,

$$\big|\big| N[f] \big|\big|_{L^q(\mathbb{R}^n)} \leq C_{p,q} \big|\big| f \big|\big|_{L^p(\mathbb{R}^n)}. \tag{3.1}$$

For $f \in L^p(\mathbb{R}^n)$ and $\lambda > 0$ set $D_\lambda f(\mathbf{x}) = f(\lambda \mathbf{x})$. Then $\big|\big| D_\lambda f \big|\big|_p = \lambda^{-n/p} \big|\big| f \big|\big|_p$. On the other hand,

$$\begin{aligned} N[D_\lambda f](\mathbf{x}) &= c_n \int_{\mathbb{R}^n} |\mathbf{x} - \mathbf{y}|^{-n+2} f(\lambda \mathbf{y})\, d\mathbf{y} = c_n \lambda^{-n} \int_{\mathbb{R}^n} |\mathbf{x} - \lambda^{-1}\mathbf{y}|^{-n+2} f(\mathbf{y})\, dy \\ &= \lambda^2\, c_n \int_{\mathbb{R}^n} |\lambda\mathbf{x} - \mathbf{y}|^{-n+2} f(\mathbf{y})\, d\mathbf{y} = \lambda^2 D_\lambda N[f](\mathbf{x}). \end{aligned}$$

Thus if the inequality in equation (3.1) holds for all $f \in L^p$, we can apply it to the functions $D_\lambda[f]$ and obtain

$$\lambda^{-\frac{n}{q}} ||N[f]||_q = ||D_\lambda N[f]||_q = \lambda^{-2} ||N D_\lambda[f]||_q \leq C_{p,q}\, \lambda^{-2} ||D_\lambda[f]||_p = C_{p,q}\, \lambda^{-2-\frac{n}{p}} ||f||_p.$$

Thus a consequence of inequality (3.1) is the estimate

$$\lambda^{-\frac{n}{q}+\frac{n}{p}+2}\, ||N[f]||_q \leq C_{p,q}\, ||f||_p$$

which holds for all $\lambda > 0$. If the exponent $\frac{n}{q} - \frac{n}{p} - 2 \neq 0$, we could let λ tend to zero or infinity, and conclude that $||f||_p = +\infty$. Thus a necessary condition for the boundedness of the operator N asserted in (3.1) is that $q^{-1} = p^{-1} - 2n^{-1}$.

This relationship between p, q, and n is also sufficient. In fact, we shall show the following more general result. Let $\Omega \subset \mathbb{R}^n$ be open, let d be a metric on Ω, and let $B(\mathbf{x}, \delta) = \{\mathbf{y} \in \Omega : d(\mathbf{x}, \mathbf{y}) < \delta\}$. Suppose that d satisfies

i) there is a constant $A > 0$ so that for all $\mathbf{x} \in \Omega$ and all $\delta > 0$,

$$\big|B(\mathbf{x}, 2\delta)\big| \leq A \big|B(\mathbf{x}, \delta)\big|.$$

ii) there is a constant $B > 0$ so that for all $\mathbf{x} \in \Omega$ and all $\delta > 0$,

$$\big|B(\mathbf{x}, \delta)\big| \geq B\, \delta^\nu.$$

(For Lebesgue measure, we have $\nu = n$.)

Let $m > 0$ and let $K : \Omega \times \Omega \to \mathbb{C}$ be a measurable function such that for $\mathbf{x}, \mathbf{y} \in \Omega$,

$$|K(\mathbf{x}, \mathbf{y})| \leq C\, d(\mathbf{x}, \mathbf{y})^m \big|B\big(\mathbf{x}, d(\mathbf{x}, \mathbf{y})\big)\big|^{-1}.$$

Finally, let

$$\mathcal{K}[f](\mathbf{x}) = \int_\Omega K(\mathbf{x}, \mathbf{y}) f(\mathbf{y})\, d\mathbf{y}.$$

Theorem 3.3 (Hardy-Littlewood-Sobolev)**.** *With the above notation, let $1 \leq p < \frac{\nu}{m}$.*

(a) *If $f \in L^p(\Omega)$ with $1 \leq p < \frac{\nu}{m}$ then the integral defining $\mathcal{K}[f]$ converges absolutely for almost every $\mathbf{x} \in \Omega$.*

(b) *If $p > 1$ and if $\dfrac{1}{q} = \dfrac{1}{p} - \dfrac{m}{\nu} > 0$, there is a constant $C_{p,m,\nu}$ so that for every $f \in L^p(\Omega)$ we have*

$$\big|\big| \mathcal{K}[f] \big|\big|_{L^q(\mathbb{R}^n)} \leq C_{p,m,\mu} \big|\big| f \big|\big|_{L^p(\mathbb{R}^n)}.$$

Proof. For $\mathbf{x} \in \Omega$ and $\lambda > 0$

$$\begin{aligned}|\mathcal{K}[f](\mathbf{x})| &\leq \int_\Omega |K(\mathbf{x},\mathbf{y})|\,|f(\mathbf{y})|\,d\mathbf{y} = \int_{d(\mathbf{x},\mathbf{y})\leq\lambda} |K(\mathbf{x},\mathbf{y})|\,|f(\mathbf{y})|\,d\mathbf{y} \;+ \int_{d(\mathbf{x},\mathbf{y})>\lambda} |K(\mathbf{x},\mathbf{y})|\,|f(\mathbf{y})|\,d\mathbf{y}\\ &=: I_\lambda(\mathbf{x}) + II_\lambda(\mathbf{x}).\end{aligned}$$

We estimate the two terms separately.

To estiamte $I_\lambda(\mathbf{x})$ let $R_j = \{\mathbf{y} \in \Omega \,\big|\, 2^{-j-1}\lambda < d(\mathbf{x},\mathbf{y}) \leq 2^{-j}\lambda\}$. Then

$$\begin{aligned}I_\lambda(\mathbf{x}) &= \sum_{j=0}^{\infty} \int_{R_j} |K(\mathbf{x},\mathbf{y})|\,|f(\mathbf{y})|\,d\mathbf{y}\\ &\leq C\sum_{j=0}^{\infty} 2^{-mj}\lambda^m |B(\mathbf{x},2^{-j-1}\lambda)|^{-1} \int_{B(\mathbf{x},2^{-j}\lambda)} |f(\mathbf{y})|\,d\mathbf{y}\\ &= C\lambda^m \sum_{j=0}^{\infty} 2^{-mj}\frac{|B(\mathbf{x},2^{-j}\lambda)|}{|B(\mathbf{x},2^{-j-1}\lambda)|}\;\frac{1}{|B(\mathbf{x},2^{-j}\lambda)|}\int_{B(\mathbf{x},2^{-j}\lambda)} |f(\mathbf{y})|\,d\mathbf{y}\\ &\leq C\lambda^m \sum_{j=0}^{\infty} 2^{-mj}\frac{|B(\mathbf{x},2^{-j}\lambda)|}{|B(\mathbf{x},2^{-j-1}\lambda)|}\,\mathcal{M}[f](\mathbf{x})\end{aligned}$$

where

$$\mathcal{M}[f](\mathbf{x}) = \sup\left\{\frac{1}{|B(\mathbf{y},\delta)|}\int_{B(\mathbf{y},\delta)} |f(\mathbf{t})|\,d\mathbf{t} : \mathbf{x} \in B(\mathbf{y},\delta) \text{ and } \delta > 0\right\}$$

is the Hardy-Littlewood maximal operator.[6] Using the doubling property for the volumes of balls in spaces of homogeneous type, it follows that $|B(\mathbf{x},2^{-j}\lambda)| \leq A|B(\mathbf{x},2^{-j-1}\lambda)|$. Thus since $m > 0$,

$$I_\lambda(\mathbf{x}) \leq AC\lambda^m\left[\sum_{j=0}^{\infty} 2^{-mj}\right]\mathcal{M}[f](\mathbf{x}) = AC[1-2^{-m}]^{-2}\,\lambda^m\,\mathcal{M}[f](\mathbf{x}) =: C_1\lambda^m\mathcal{M}[f](\mathbf{x}). \tag{a}$$

To estimate $II_\lambda(\mathbf{x})$, we use Hölder's inequality:

$$II_\lambda(\mathbf{x}) = \int_{d(\mathbf{x},\mathbf{y})>\lambda} |K(\mathbf{x},\mathbf{y})|\,|f(\mathbf{y})|\,d\mathbf{y} \leq \big|\big|\, f \,\big|\big|_{L^p(\Omega)}\Big[\int_{d(\mathbf{x},\mathbf{y})>\lambda} |K(\mathbf{x},\mathbf{y})|^{p'}\,d\mathbf{y}\Big]^{\frac{1}{p'}}$$

[6]We will use the following result.

Theorem 3.4 (Hardy, Littlewood)**.**

(1) There is a constant $A_1 > 0$ so that if $f \in L^1(\Omega)$,

$$\Big|\Big\{\mathbf{x} \in \Omega : \mathcal{M}[f](\mathbf{x}) > \lambda\Big\}\Big| \leq A_1\lambda^{-1}||\,f\,||_{L^1(\Omega)}.$$

(2) For $1 < p \leq \infty$ there is a constant $A_p > 0$ so that if $f \in L^p(\Omega)$ then $\mathcal{M}[f] \in L^p(\Omega)$ and

$$||\,\mathcal{M}[f]\,||_{L^p(\Omega} \leq A_p||\,f\,||_{L^p(\Omega)}.$$

where $\frac{1}{p} + \frac{1}{p'} = 1$. Let $S_j = \{\mathbf{y} \in \Omega \,\big|\, 2^j\lambda < d(\mathbf{x},\mathbf{y}) \le 2^{j+1}\lambda\}$, then

$$\begin{aligned}
&\int\limits_{d(\mathbf{x},\mathbf{y})>\lambda} |K(\mathbf{x},\mathbf{y})|^{p'}\,d\mathbf{y}\\
&\qquad = \sum_{j=0}^{\infty}\int_{S_j}|K(\mathbf{x},\mathbf{y})|^{p'}\,d\mathbf{y} \le C\lambda^{mp'}\sum_{j=0}^{\infty}2^{mp'(j+1)}|B(\mathbf{x},2^j\lambda)|^{-p'}|B(\mathbf{x},2^{j+1}\lambda)|\\
&\qquad \le AC\,2^{mp'}\lambda^{mp'}\sum_{j=0}^{\infty}2^{mp'j}|B(\mathbf{x},2^j\lambda)|^{1-p'} \le ABC\,2^{mp'}\lambda^{mp'}\sum_{j=0}^{\infty}2^{mp'j}(2^j\lambda)^{\nu(1-p')}\\
&\qquad = ABC\,2^{mp'}\lambda^{mp'+\nu(1-p')}\sum_{j=0}^{\infty}2^{[mp'+\nu(1-p')]j}\\
&\qquad = ABC\,2^{mp'}\big[1-2^{mp'+\nu(1-p')}\big]^{-1}\lambda^{mp'+\nu(1-p')} =: C_2^{p'}\lambda^{mp'+\nu(1-p')}
\end{aligned}$$

since $p' = \frac{p}{p-1}$ and so $mp' + \nu(1-p') = \frac{mp-\nu}{p-1} < 0$. Since $m + \nu\frac{1-p'}{p'} = m - \frac{\nu}{p}$, it follows that

$$II_\lambda(\mathbf{x}) \le C_2\,\big|\big|\,f\,\big|\big|_{L^p(\Omega)}\lambda^{m-\frac{\nu}{p}} \tag{b}$$

It now follows from equations (a) and (b) that there is a constant C depending only on ν, p, and m so that

$$|\mathcal{K}[f](\mathbf{x})| \le C\left[\lambda^m\,\mathcal{M}[f](\mathbf{x}) + \lambda^{m-\frac{\nu}{p}}\,\big|\big|\,f\,\big|\big|_{L^p(\Omega)}\right]. \tag{c}$$

Since $\mathcal{M}[f](\mathbf{x}) < \infty$ for almost all $\mathbf{x} \in \Omega$, it follows that the integral defining $\mathcal{K}[f](\mathbf{x})$ converges absolutely for almost all $\mathbf{x}$. Choose λ so that the two terms on the right-hand side of (c) are equal; that is, let $\lambda = \big|\big|\,f\,\big|\big|_{L^p(\Omega)}^{p/\nu}\,\mathbb{M}[f](x)^{-p/\nu}$. It follows that

$$|\mathcal{K}[f](\mathbf{x})| \le C\,\big|\big|\,f\,\big|\big|_{L^p(\Omega)}^{\frac{mp}{\nu}}\,\mathcal{M}[f](\mathbf{x})^{1-\frac{mp}{\nu}} = C\,\big|\big|\,f\,\big|\big|_{L^p(\Omega)}^{1-\frac{p}{q}}\,\mathcal{M}[f](\mathbf{x})^{\frac{p}{q}}.$$

Since $||\mathcal{M}[f]||_p \le C_p||f||_p$ for $1 < p \le \infty$ we have $\big|\big|\,\mathcal{K}[f]\,\big|\big|_{L^q} \le C\,\big|\big|\,f\,\big|\big|_{L^p}$. □

We make several remarks about the theorem and its proof.

1. If $p = 1$, if $f \in L^1(\Omega)$, and if $\frac{1}{q} = 1 - \frac{m}{\nu}$, it does not follow that $\mathcal{K}[f] \in L^q(\Omega)$. The proof breaks down because the Hardy-Littlewood maximal operator is not bounded on L^1. In the case of the Newtonian potential, if we take f to be the characteristic function of the unit ball, then for large $|x|$ one can check that $N[f](\mathbf{x}) \ge c|\mathbf{x}|^{-n+2}$, and when raised to the power $\frac{2}{n-2}$ this just fails to be integrable.
2. Using the fact that the operator $\mathcal{M}$ is weak-type (1,1), it does follow that if $f \in L^1(\mathbb{R}^n)$, then $|\mathcal{K}[f](\mathbf{x})| < \infty$ for almost every x, and if $\frac{1}{q} = 1 - \frac{m}{\nu}$,
$$\left|\left\{\mathbf{x} \in \mathbb{R}^n \,\Big|\, |\mathcal{K}[f](\mathbf{x})| > \lambda\right\}\right| \le C\,\frac{||f||_1^q}{\lambda^q}. \tag{3.2}$$
3. If $p = \frac{\nu}{m}$ and $f \in L^p(\mathbb{R}^n)$, it does not follow that $\mathcal{K}[f] \in L^\infty(\mathbb{R}^n)$. The proof breaks down because $\int_{d(\mathbf{x},\mathbf{y})\ge 1}|K(\mathbf{x},\mathbf{y})|^{p'}\,d\mathbf{y}$ is now infinite. A discussion of the relevant examples which show that the conclusion is false can be found, for example, in [9], pages 159 – 160.
4. If $f \in L^p(\mathbb{R}^n)$, then $\mathcal{K}[f] \in L^q_{\text{loc}}(\mathbb{R}^n)$ if $\frac{1}{q} \ge \frac{1}{p} - \frac{m}{n}$.

3.4. **Lipschitz Estimates.**

We next turn to the study of regularity measured on the scale of Lipschitz spaces. For the standard Euclidean metric, these are defined as follows. For $0 < \alpha < 1$, $\Lambda_\alpha(\mathbb{R}^n)$ denotes the space of complex-valued functions f on $\mathbb{R}^n$ for which

$$\big|\big|\, f \,\big|\big|_{\Lambda_\alpha} = \sup_{\mathbf{x}\in\mathbb{R}^n} |f(\mathbf{x})| + \sup_{\mathbf{x}_1\neq\mathbf{x}_2\in\mathbb{R}^n} \frac{\big|f(\mathbf{x}_2) - f(\mathbf{x}_1)\big|}{\big|\mathbf{x}_2 - \mathbf{x}_1\big|^\alpha} < \infty.$$

We say that $f \in \Lambda_\alpha$ satisfies a *Lipschitz* or *Hölder condition* of order α. If k is a positive integer and $0 < \alpha < 1$, $\Lambda^k_\alpha(\mathbb{R}^n)$ is the space of k-times continuously differentiable complex-valued functions f on $\mathbb{R}^n$ such that every partial derivative of f of order k satisfies a Lipschitz condition of order α. We put

$$\big|\big|\, f \,\big|\big|_{\Lambda^k_\alpha} = \sum_{0\le|\alpha|<k} \sup_{\mathbf{x}\in\mathbb{R}^n} |\partial^\alpha f(\mathbf{x})| + \sum_{|\beta|=k} \big|\big|\, \partial^\beta f \,\big|\big|_{\Lambda_\alpha}.$$

Our objective is to show that the operators like the Newtonian potential N increases smoothness by two orders in this scale of spaces: if $0 < \alpha < 1$ and k is a non-negative integer, there is a constant $C_{k,\alpha}$ so that if $f \in \Lambda^k_\alpha(\mathbb{R}^n)$, then $N[f] \in \Lambda^{k+2}_\alpha(\mathbb{R}^n)$ and

$$\big|\big|\, N[f] \,\big|\big|_{\Lambda^{k+2}_\alpha} \le C_{k,\alpha} \,\big|\big|\, f \,\big|\big|_{\Lambda^k_\alpha}. \tag{3.3}$$

To establish estimates of this sort, we must use more information about the Schwartz kernel $N(\mathbf{x},\mathbf{y}) = c_n|\mathbf{x}-\mathbf{y}|^{-n+2}$ of the operator N than the simple size estimates used in the proof of the Hardy-Littlewood-Sobolev Theorem. There are two main points.

1. $N[\varphi] = \int_{\mathbb{R}^n} N(\mathbf{y})\varphi(\mathbf{x}-\mathbf{y})\,d\mathbf{y}$ is given by a convolution, and hence it commutes with differentiation:
$$\partial^\alpha N[\varphi](\mathbf{x}) = \int_{\mathbb{R}^n} N(\mathbf{y})\partial^\alpha\varphi(\mathbf{x}-\mathbf{y})\,d\mathbf{y} = N[\partial^\alpha\varphi](\mathbf{x})$$
for every test function φ.
2. Two derivatives of the kernel $N(\mathbf{x}) = c_n|\mathbf{x}|^{-n+2}$ has vanishing mean value over any sphere centered at the origin. One computes
$$\frac{\partial^2 N}{\partial x_j\partial x_k}(\boldsymbol{\zeta}) = \begin{cases} c_n n(n-2)\zeta_j\zeta_k|\boldsymbol{\zeta}|^{-n-2} & \text{if } j\neq k \\ \\ c_n n(n-2)\zeta_j^2|\boldsymbol{\zeta}|^{-n-2} - (n-2)|\boldsymbol{\zeta}|^{-n} & \text{if } j = k \end{cases}.$$
Then in all cases, if $R > 0$,
$$\int_{|\zeta|=R} \frac{\partial^2 N}{\partial x_j\partial x_k}(\zeta)\,d\sigma(\zeta) = 0.$$
This follows if $j \neq k$ since the integrand is odd in both ζ_j and ζ_k. On the other hand, for any $1 \le j,k \le n$, symmetry of the sphere shows that $\int_{|\zeta|=R}\zeta_j^2\,d\sigma(\zeta) = \int_{|\zeta|=R}\zeta_k^2\,d\sigma(\zeta)$, and hence
$$\begin{aligned}\int_{|\zeta|=R} n(n-2)\zeta_j^2|\boldsymbol{\zeta}|^{-n-2}d\sigma(\zeta) &= \sum_{k=1}^n \int_{|\zeta|=R}(n-2)\zeta_k^2|\boldsymbol{\zeta}|^{-n-2}d\sigma(\zeta) \\ &= \int_{|\zeta|=R}(n-2)|\boldsymbol{\zeta}|^{-n}d\sigma(\zeta).\end{aligned}$$

To formulate a result for an operator $\mathcal{K}[f](\mathbf{x}) = \int_{\mathbb{R}^n} K(\mathbf{x},\mathbf{y})f(\mathbf{y})\,d\mathbf{y}$ which generalizes the estimate in equation (3.3) we need to make assumptions about the kernel $K(\mathbf{x},\mathbf{y})$ which reflect the two facts listed above for the special kernel $N(\mathbf{x},\mathbf{y}) = c_n|\mathbf{x}-\mathbf{y}|^{-n+2}$. Since Lipschitz spaces are defined using the standard Euclidean metric, we will restrict attention to spaces of homogeneous type where the the metric is given by $d(\mathbf{x},\mathbf{y}) = |\mathbf{x}-\mathbf{y}|$.

Definition 3.5. *Let m be a non-negative integer. A function $K \in \mathcal{C}^\infty(\mathbb{R}^n \times \mathbb{R}^n \setminus D_{\mathbb{R}^n})$ belongs to the class $CZ(\mathbb{R}^n, m)$ if*

A) (Differential inequalities) There is a positive integer m and for any muti-indices α, β there is a constant $C_{\alpha,\beta}$ so that

$$\big|\partial_{\mathbf{x}}^\alpha \partial_{\mathbf{y}}^\beta K(\mathbf{x},\mathbf{y})\big| \leq C_{\alpha,\beta}|\mathbf{x}-\mathbf{y}|^{m-|\alpha|-|\beta|}\big|B(\mathbf{x},|\mathbf{x}-\mathbf{y}|)\big|^{-1}.$$

Here $B(\mathbf{x},\delta)$ is the Euclidean ball of radius δ so $\big|B(\mathbf{x},|\mathbf{x}-\mathbf{y}|)\big| = c_n|\mathbf{x}-\mathbf{y}|^n$.

B) (Cancellation condition) For any $0 < R < \infty$, if $|\alpha| + |\beta| = m$

$$\int\limits_{|\mathbf{x}-\zeta|=R} \partial_{\mathbf{x}}^\alpha \partial_{\mathbf{y}}^\beta K(\mathbf{x},\zeta)\,d\sigma(\zeta) = \int\limits_{|\mathbf{y}-\zeta|=R} \partial_{\mathbf{x}}^\alpha \partial_{\mathbf{y}}^\beta K(\zeta,\mathbf{y})\,d\sigma(\zeta) = 0.$$

C) (Commutation property) For any multi-index α there is a family of functions $\{K_\beta(\mathbf{x},\mathbf{y}) : 0 \leq |\beta| \leq |\alpha|\}$ so that for any test function φ,

$$\partial_{\mathbf{x}}^\alpha\Big[\int_{\mathbb{R}^n} K(\mathbf{x},\mathbf{y})\varphi(\mathbf{y})\,d\mathbf{y}\Big] = \sum_{|\beta|\leq|\alpha|}\int_{\mathbb{R}^n} K_\beta(\mathbf{x},\mathbf{y})\partial^\beta\varphi(\mathbf{y})\,d\mathbf{y}$$

and each function $K_\beta \in \mathcal{C}^\infty(\mathbb{R}^n \times \mathbb{R}^n \setminus D_{\mathbb{R}^n})$ also satisfies the differential inequalities and cancellation conditions in A) and B).

Remark 3.6. *If $K \in CZ(\mathbb{R}^n, m)$ and if α, β are multi-indices with $|\alpha|+|\beta| \leq m$ then $\partial_{\mathbf{x}}^\alpha \partial_{\mathbf{y}}^\beta K \in CZ(\mathbb{R}^n, m - |\alpha| - |\beta|)$.*

Theorem 3.7. *Let $K \in CZ(\mathbb{R}^n, m)$, and put $\mathcal{K}[f](\mathbf{x}) = \int_{\mathbb{R}^n} K(\mathbf{x},\mathbf{y})f(\mathbf{y})\,d\mathbf{y}$. Let $0 < \alpha < 1$ and $R > 0$. There is a constant $C_{\alpha,R}$ so that if $f \in \Lambda_\alpha$ has compact support in the ball of radius R then $\mathcal{K}[f]$ is m-times continuously differentiable, and*

$$\big|\big|\,\mathcal{K}[f]\,\big|\big|_{\Lambda_\alpha^m} \leq C_{\alpha,R}\big|\big|\,f\,\big|\big|_{\Lambda_\alpha}.$$

Moreover, there is a constant C_α independent of R so that if $|\beta| = m$ then

$$\sup_{\mathbf{x}_1\neq\mathbf{x}_2\in\mathbb{R}^n}\frac{|\partial^\beta f(\mathbf{x}_2) - \partial^\beta f(\mathbf{x}_1)|}{|\mathbf{x}_2-\mathbf{x}_1|^\alpha} \leq C_\alpha\big|\big|\,f\,\big|\big|_{\Lambda_\alpha}.$$

The proof of this result proceeds in several steps.

1. We first show that for $|\beta| \leq m-1$, we can differentiate under the integral sign to obtain

$$\partial^\beta\mathcal{K}[f](\mathbf{x}) = \int_{\mathbb{R}^n} \partial_{\mathbf{x}}^\beta K(\mathbf{x},\mathbf{y})f(\mathbf{y})\,d\mathbf{y}.$$

   This shows that $\mathcal{K}[f]$ is $(m-1)$-times continuously differentiable.

2. To show that $\mathcal{K}[f]$ is m-times continuously differentiable, we replace $\partial_{\mathbf{x}}^\beta K(\mathbf{x},\mathbf{y})$ by $L(\mathbf{x},\mathbf{y}) \in CM(\mathbb{R}^n, 1)$. It then suffices to show that if $\mathcal{L}[f](\mathbf{x}) = \int_{\mathbb{R}^n} L(\mathbf{x},\mathbf{y})f(\mathbf{y})\,d\mathbf{y}$ and if $f \in \Lambda_\alpha$ has compact support, then for $1 \leq j \leq n$ it follows that $\partial^j\mathcal{L}[f]$ is continuous. In general,

however, we cannot just differentiate one more time under the integral sign. Rather, we show that

$$\partial^j \mathcal{L}[f](\mathbf{x}) = \int\limits_{|\mathbf{x}-\mathbf{y}|\leq 1} \partial^j_{\mathbf{x}}(\mathbf{x},\mathbf{y})\big[f(\mathbf{y}) - f(\mathbf{x})\big]\, d\mathbf{y} + \int\limits_{|\mathbf{x}-\mathbf{y}|>1} \partial^j_{\mathbf{x}}(\mathbf{x},\mathbf{y}) f(\mathbf{y})\, d\mathbf{y} + A_j(\mathbf{x})\, f(\mathbf{x})$$

where A_j is a smooth function. The function $A_j f$ is then belongs to Λ_α, while the two integrals are absolutely convergent and define continuous functions.

3. Finally to show that $\partial^j \mathcal{L}[f]$ satisfies a Lipschitz condition of order α

Proposition 3.8. *Let f be continuous with compact support. Then $\mathcal{K}[f]$ is $(m-1)$-times continuously differentiable and if $0 \leq |\beta| \leq m-1$ then*

$$\partial^\beta \mathcal{K}[f](\mathbf{x}) = \int_{\mathbb{R}^n} \partial^\beta_{\mathbf{x}} K(\mathbf{x},\mathbf{y}) f(\mathbf{y})\, d\mathbf{y}.$$

Moreover, if f is supported in a ball of radius R, then

$$\sup_{\mathbf{x}\in\mathbb{R}^n} \big|\partial^\beta \mathcal{K}[f](\mathbf{x})\big| \leq C R^{m-|\beta|} \sup_{\mathbf{x}\in\mathbb{R}^n} |f(\mathbf{x})|.$$

Proof. Choose $\chi \in \mathcal{C}^\infty(\mathbb{R})$ so that $0 \leq \chi(t) \leq 1$ for all t and

$$\chi(t) = \begin{cases} 0 & \text{if } t \leq 1, \\ 1 & \text{if } t \geq 2. \end{cases}$$

For $\mathbf{x} \in \mathbb{R}^n$ and $\epsilon > 0$ put $\varphi_\epsilon(\mathbf{x}) = \chi\big(\epsilon^{-1}|\mathbf{x}|\big)$. Then φ_ϵ is supported where $|\mathbf{x}| \geq \epsilon$ and if $|\gamma| \geq 1$, $\partial^\gamma \varphi_\epsilon$ is supported where $\epsilon < |\mathbf{x}| < 2\epsilon$ and and $|\partial^\beta \varphi_\epsilon(\mathbf{x})| \leq C_\beta \epsilon^{-|\beta|} \leq 2^{|\beta|} C_\beta\, |\mathbf{x}|^{-|\beta|}$. Put

$$\mathcal{K}_\epsilon[f](\mathbf{x}) = \int_{\mathbb{R}^n} K(\mathbf{x},\mathbf{y})\, \varphi_\epsilon(\mathbf{x}-\mathbf{y})\, f(\mathbf{y})\, d\mathbf{y}.$$

Then

$$\begin{aligned}
\big|\mathcal{K}[f](\mathbf{x}) - \mathcal{K}_\epsilon[f](\mathbf{x})\big| &= \Big|\int_{\mathbb{R}^n} K(\mathbf{x},\mathbf{y})\, \big[1 - \varphi_\epsilon(\mathbf{x}-\mathbf{y})\big]\, f(\mathbf{y})\, d\mathbf{y}\Big| \\
&\leq C_n \sup_{\mathbf{x}\in\mathbb{R}^n} |f(\mathbf{x})| \int_{|\mathbf{x}-\mathbf{y}|<2\epsilon} |K(\mathbf{x},\mathbf{y})|\, d\mathbf{y} \\
&\leq C_n \sup_{\mathbf{x}\in\mathbb{R}^n} |f(\mathbf{x})| \int_{|\mathbf{x}-\mathbf{y}|<2\epsilon} |\mathbf{x}-\mathbf{y}|^{m-n}\, d\mathbf{y} \\
&\leq C_n \sup_{\mathbf{x}\in\mathbb{R}^n} |f(\mathbf{x})|\, \epsilon^m.
\end{aligned}$$

Thus $\mathcal{K}_\epsilon[f] \to \mathcal{K}[f]$ uniformly on $\mathbb{R}^n$ as $\epsilon \to 0$.

On the other hand, the function $\mathbf{x} \longrightarrow K_\epsilon(\mathbf{x},\mathbf{y}) =: K(\mathbf{x},\mathbf{y})\, \varphi_\epsilon(\mathbf{x}-\mathbf{y})$ is infinitely differentiable. Thus $\mathcal{K}_\epsilon[f]$ is infinitely differentiable, and differentiating under the integral sign, for any $|\beta| < m$ we have

$$\partial^\beta \mathcal{K}_\epsilon[f](\mathbf{x}) = \int_{\mathbb{R}^n} \partial^\beta_{\mathbf{x}} K_\epsilon(\mathbf{x},\mathbf{y})\, f(\mathbf{y})\, d\mathbf{y}.$$

Put

$$\mathcal{K}_\beta[f](x) = \int_{\mathbb{R}^n} \partial^\beta_{\mathbf{x}} K(\mathbf{x},\mathbf{y})\, f(y)\, dy.$$

Then

$$\begin{aligned}\left|\mathcal{K}_\beta[f](\mathbf{x}) - \partial^\beta \mathcal{K}_\epsilon(\mathbf{x})\right| &= \Big|\int_{\mathbb{R}^n} \partial_{\mathbf{x}}^\beta\big[K(\mathbf{x},\mathbf{y})(1-\varphi_\epsilon)(\mathbf{x}-\mathbf{y})\big]\, f(\mathbf{y})\, d\mathbf{y}\Big| \\ &\le C_n \sup_{\mathbf{x}\in\mathbb{R}^n} |f(\mathbf{x})| \int_{|\mathbf{x}-\mathbf{y}|<2\epsilon} \big(|\mathbf{x}-\mathbf{y}|^{-n+m-|\beta|} + \epsilon^{-|\beta|}|\mathbf{x}-\mathbf{y}|^{-n+m}\big)\, d\mathbf{y} \\ &\le C_n \sup_{\mathbf{x}\in\mathbb{R}^n} |f(\mathbf{x})|\, \epsilon^{m-|\beta|}.\end{aligned}$$

Thus $\partial^\beta \mathcal{K}_\epsilon[f] \longrightarrow \mathcal{K}_\beta[f]$ uniformly on $\mathbb{R}^n$ as $\epsilon \to 0$. Combined with our first observation, this shows that $\mathcal{K}[f]$ is $(m-1)$-times continuously differentiable and $\partial^\beta\mathcal{K}[f](\mathbf{x}) = \mathcal{K}_\beta[f](\mathbf{x})$ for $|\beta| < m$. The upper bound for $\partial^\beta \mathcal{K}[f](\mathbf{x})$ follows easily, completing the proof. □

To show that $\mathcal{K}[f]$ is m-times continuously differentiable when $f \in \Lambda_\alpha$, we have already observed that we cannot simply differentiate under the integral sign since if $|\beta| = m$, the kernel $\partial_{\mathbf{x}}^\beta K(\mathbf{x},\mathbf{y}) \in CZ(\mathbb{R}^n, 0)$ is not locally integrable. However if $f \in \Lambda_\alpha$, we do have

$$\left|\partial_{\mathbf{x}}^\beta K(\mathbf{x},\mathbf{y})\big[f(\mathbf{y}) - f(\mathbf{x})\big]\right| \le C_n \,\big|\big|\, f \,\big|\big|_{\Lambda_\alpha} |\mathbf{y}-\mathbf{x}|^{\alpha-n}$$

and so $\partial_{\mathbf{x}}^\beta K(\mathbf{x},\mathbf{y})\big[f(\mathbf{y}) - f(\mathbf{x})\big]$ is locally integrable as a function of $\mathbf{y}$. We use this observation in the next result.

Lemma 3.9. *Let $L(\mathbf{x},\mathbf{y}) \in CZ(\mathbb{R}^n, 1)$ and let $f \in \Lambda_\alpha$ have compact support. Then $\mathcal{L}[f](\mathbf{x}) = \int_{\mathbb{R}^n} L(\mathbf{x},\mathbf{y}) f(\mathbf{y})\, d\mathbf{y}$ is continuously differentiable, and for $1 \le j \le n$*

$$\partial^j \mathcal{L}[f](\mathbf{x}) = \int_{|\mathbf{x}-\mathbf{y}|<1} \partial_{\mathbf{x}}^j L(\mathbf{x},\mathbf{y})\big[f(\mathbf{x}) - f(\mathbf{y})\big]\, d\mathbf{y} + \int_{|\mathbf{x}-\mathbf{y}|\ge 1} \partial_{\mathbf{x}}^j L(\mathbf{x},\mathbf{y}) f(\mathbf{y})\, d\mathbf{y} + A_j(\mathbf{x}) f(\mathbf{x})$$

where

$$A_j(\mathbf{x}) = \int_{|\mathbf{x}-\zeta|=1} \partial_{\mathbf{x}}^j L(\mathbf{x},\zeta)(x_j - \zeta_j)\, d\sigma(\zeta).$$

Proof. With φ_ϵ defined as in Proposition 3.8, put

$$\mathcal{L}_\epsilon[f](\mathbf{x}) = \int_{\mathbb{R}^n} L(\mathbf{x},\mathbf{y})\varphi_\epsilon(\mathbf{x}-\mathbf{y}) f(\mathbf{y})\, d\mathbf{y}.$$

Then for $|\gamma| = m-1$, by the same argument as in the last proposition,

$$\left|\mathcal{L}[f](\mathbf{x}) - \mathcal{L}_\epsilon[f](\mathbf{x})\right| \le C\,\epsilon \sup_{\mathbf{x}\in\mathbb{R}^n} |f(\mathbf{x})|$$

and so $\mathcal{L}_\epsilon[f] \longrightarrow \mathcal{L}[f]$ uniformly on $\mathbb{R}^n$ as $\epsilon \to 0$. The function $\mathcal{L}_\epsilon[f]$ is infinitely differentiable. We will show that $\partial^j\mathcal{L}_\epsilon[f]$ converges uniformly as $\epsilon \to 0$ and hence $\mathcal{L}[f]$ is continuously differentiable. Since the kernel of $\mathcal{L}_\epsilon$ has no singularities, we have

$$\begin{aligned}\partial^j \mathcal{L}_\epsilon[f](\mathbf{x}) &= \int_{\mathbb{R}^n} \partial_{\mathbf{x}}^j[L(\mathbf{x},\mathbf{y})\varphi_\epsilon(\mathbf{x}-\mathbf{y})] f(\mathbf{y})\, d\mathbf{y} \\ &= \int_{|\mathbf{x}-\mathbf{y}|<1} \partial_{\mathbf{x}}^j[L(\mathbf{x},\mathbf{y})\varphi_\epsilon(\mathbf{x}-\mathbf{y})]\big[f(\mathbf{y}) - f(\mathbf{x})\big]\, d\mathbf{y} \\ &\qquad + f(\mathbf{x}) \int_{|\mathbf{x}-\mathbf{y}|<1} \partial_{\mathbf{x}}^j[L(\mathbf{x},\mathbf{y})\varphi_\epsilon(\mathbf{x}-\mathbf{y})]\, d\mathbf{y} \\ &\qquad\quad + \int_{|\mathbf{x}-\mathbf{y}|\ge 1} \partial_{\mathbf{x}}^j[L(\mathbf{x},\mathbf{y})\varphi_\epsilon(\mathbf{x}-\mathbf{y})] f(\mathbf{y})\, d\mathbf{y} \\ &= I_\epsilon(\mathbf{x}) + II_\epsilon(\mathbf{x}) + III_\epsilon(\mathbf{x}).\end{aligned}$$

Now

$$\left|\partial^j_{\mathbf{x}}[L(\mathbf{x},\mathbf{y})\varphi_\epsilon(\mathbf{x}-\mathbf{y})]\big[f(\mathbf{y})-f(\mathbf{x})\big]\right| \lesssim \Big[|\mathbf{x}-\mathbf{y}|^{-n}+|\mathbf{x}-\mathbf{y}|^{-n+1}\epsilon^{-1}\Big]|\mathbf{x}-\mathbf{y}|^{\alpha}$$

and since $1-\varphi_\epsilon(\mathbf{x}-\mathbf{y})$ is supported where $|\mathbf{x}-\mathbf{y}|<2\epsilon$, it follows that

$$I_\epsilon(\mathbf{x}) \longrightarrow \int_{|\mathbf{x}-\mathbf{y}|<1} \partial^j_{\mathbf{x}} L(\mathbf{x},\mathbf{y})\big[f(\mathbf{y})-f(\mathbf{x})\big]\,d\mathbf{y}$$

uniformly as $\epsilon\to 0$. Also $\varphi_\epsilon(\mathbf{x}-\mathbf{y})\equiv 1$ if $|\mathbf{x}-\mathbf{y}|>2\epsilon$, so if $\epsilon<\frac{1}{2}$,

$$III_\epsilon(\mathbf{x}) = \int_{|\mathbf{x}-\mathbf{y}|\geq 1} \partial^j_{\mathbf{x}} L(\mathbf{x},\mathbf{y}) f(\mathbf{y})\,d\mathbf{y}.$$

It remains to study the term $II_\epsilon(\mathbf{x})$. We have

$$\begin{aligned}\partial^j_{\mathbf{x}}\Big[L(\mathbf{x},\mathbf{y})\varphi_\epsilon(\mathbf{x}-\mathbf{y})\Big] &= \partial^j_{\mathbf{x}} L(\mathbf{x},\mathbf{y})\varphi_\epsilon(\mathbf{x}-\mathbf{y}) + L(\mathbf{x},\mathbf{y})\partial^j_{\mathbf{x}}\varphi_\epsilon(\mathbf{x}-\mathbf{y})\\ &= \partial^j_{\mathbf{x}} L(\mathbf{x},\mathbf{y})\varphi_\epsilon(\mathbf{x}-\mathbf{y}) - L(\mathbf{x},\mathbf{y})\partial^j_{\mathbf{y}}\varphi_\epsilon(\mathbf{x}-\mathbf{y})\\ &= \partial^j_{\mathbf{x}} L(\mathbf{x},\mathbf{y})\varphi_\epsilon(\mathbf{x}-\mathbf{y}) + \partial^j_{\mathbf{y}} L(\mathbf{x},\mathbf{y})\varphi_\epsilon(\mathbf{x}-\mathbf{y})\\ &\qquad - \partial^j_{\mathbf{y}}\big[L(\mathbf{x},\mathbf{y})\varphi_\epsilon(\mathbf{x}-\mathbf{y})\big]\end{aligned}$$

Now $\varphi_\epsilon(\mathbf{x}-\mathbf{y})$ is constant on the set of $\mathbf{y}$ where $|\mathbf{x}-\mathbf{y}|=r$, and it follows from the cancellation hypothesis on the kernel L that

$$\int_{|\mathbf{x}-\mathbf{y}|<1} \partial^j_{\mathbf{x}} L(\mathbf{x},\mathbf{y})\varphi_\epsilon(\mathbf{x}-\mathbf{y})\,d\mathbf{y} = \int_{|\mathbf{x}-\mathbf{y}|<1} \partial^j_{\mathbf{y}} L(\mathbf{x},\mathbf{y})\varphi_\epsilon(\mathbf{x}-\mathbf{y})\,d\mathbf{y} = 0.$$

Using the divergence theorem, we have for $\epsilon<\frac{1}{2}$

$$\begin{aligned}\int_{|\mathbf{x}-\mathbf{y}|<1} \partial^j_{\mathbf{y}}\big[L(\mathbf{x},\mathbf{y})\varphi_\epsilon(\mathbf{x}-\mathbf{y})\big]\,d\mathbf{y} &= \int_{|\mathbf{x}-\zeta|=1} L(\mathbf{x},\zeta)\varphi_\epsilon(\mathbf{x}-\zeta)\,(\zeta_j-x_j)\,d\sigma(\zeta)\\ &= \int_{|\mathbf{x}-\zeta|=1} L(\mathbf{x},\zeta)(\zeta_j-x_j)\,d\sigma(\zeta).\end{aligned}$$

Hence $II_\epsilon(\mathbf{x}) \longrightarrow A_j(\mathbf{x})f(\mathbf{x})$ as $\epsilon\to 0$, uniformly on $\mathbb{R}^n$. It now follows that $\mathcal{L}[f]$ is continuously differentiable and $\partial^j\mathcal{L}[f](\mathbf{x})$ is as given in the statement of the Lemma. □

To complete the proof of Theorem 3.7, it now suffices to show that if $f\in\Lambda_\alpha$ has compact support, and if

$$F(\mathbf{x}) = \int_{|\mathbf{x}-\mathbf{y}|<1} \partial^j_{\mathbf{x}} L(\mathbf{x},\mathbf{y})\big[f(\mathbf{y})-f(\mathbf{x})\big]\,d\mathbf{y} + \int_{|\mathbf{x}-\mathbf{y}|\geq 1} \partial^j_{\mathbf{x}} L(\mathbf{x},\mathbf{y})\,f(\mathbf{y})\,d\mathbf{y}$$

then $F\in\Lambda_\alpha$ and $\big|\big|\,F\,\big|\big|_{\Lambda_\alpha} \leq C\,\big|\big|\,f\,\big|\big|_{\Lambda_\alpha}$ for some constant C independent of f and its support. In fact, this holds if the kernel $\partial^j_{\mathbf{x}} L(\mathbf{x},\mathbf{y})$ is replaced by any kernel K satisfying

(1) For all $\mathbf{x}\neq\mathbf{y}$, $|K(\mathbf{x},\mathbf{y})|\leq C_0|\mathbf{x}-\mathbf{y}|^{-n}$.

(2) For all $\mathbf{x}\neq\mathbf{y}$, $|\nabla_{\mathbf{x}} K(\mathbf{x},\mathbf{y})|+|\nabla_{\mathbf{y}} K(\mathbf{x},\mathbf{y})|\leq C_1|\mathbf{x}-\mathbf{y}|^{-n-1}$.

(3) For all $0<R_1<R_2$ we have

$$\int_{R_1<|\mathbf{x}-\mathbf{y}|<R_2} K(\mathbf{x},\mathbf{y})\,d\mathbf{x} = \int_{R_1<|\mathbf{x}-\mathbf{y}|<R_2} K(\mathbf{x},\mathbf{y})\,d\mathbf{y} = 0.$$

If $f\in\Lambda_\alpha$ has compact support, we can define

$$\mathcal{K}[f](\mathbf{x}) = \int_{|\mathbf{x}-\mathbf{y}|<1} K(\mathbf{x},\mathbf{y})\big[f(\mathbf{y})-f(\mathbf{x})\big]\,d\mathbf{y} + \int_{|\mathbf{x}-\mathbf{y}|\geq 1} K(\mathbf{x},\mathbf{y})\,f(\mathbf{y})\,d\mathbf{y}.$$

and both integrals converge absolutely.

Theorem 3.10. *Let $0 < \alpha < 1$. There is a constant C_α depending only on C_0 and C_1 so that for all $f \in \Lambda_\alpha$ with compact support we have*

$$\sup_{\mathbf{x}_1 \neq \mathbf{x}_2 \in \mathbb{R}^n} \frac{|\mathcal{K}[f](\mathbf{x}_2) - \mathcal{K}[f(\mathbf{x}_1)|}{|\mathbf{x}_2 - \mathbf{x}_1|^\alpha} \leq C_\alpha \,||\, f \,||_{\Lambda^\alpha}.$$

Proof. Let $\mathbf{x}_1 \neq \mathbf{x}_2 \in \mathbb{R}^n$ and put $\delta = |\mathbf{x}_2 - \mathbf{x}_1|$. Since f has compact support, there exists $R > 3\delta$ (depending on f) so that $|\mathbf{x}_j - \mathbf{y}| \geq R$ implies $f(\mathbf{y}) = 0$. Using the fact that the function $\mathbf{y} \to K(\mathbf{x}, \mathbf{y})$ has mean value zero on the set $1 \leq |\mathbf{x} - \mathbf{y}| \leq R$, it follows that

$$\mathcal{K}[f](\mathbf{x}_j) = \int_{|\mathbf{x}_j - \mathbf{y}| < R} K(\mathbf{x}_j, \mathbf{y}) \left[f(\mathbf{y}) - f(\mathbf{x}_j)\right] d\mathbf{y}.$$

Let $\psi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ be a radial function such that $\psi(\mathbf{x}) = 1$ if $|\mathbf{x}| \leq R$ and $0 \leq \psi(\mathbf{x}) \leq 1$ and $|\nabla\psi(\mathbf{x})| \leq |\mathbf{x}|^{-1}$ for all $\mathbf{x} \in \mathbb{R}^n$. Then we can write

$$\mathcal{K}[f](\mathbf{x}_j) = \int_{R^n} K(\mathbf{x}_j, \mathbf{y})\, \psi(\mathbf{x}_j - \mathbf{y}) \left[f(\mathbf{y}) - f(\mathbf{x}_j)\right] d\mathbf{y}.$$

Then

$$\begin{aligned}
&\mathcal{K}[f](\mathbf{x}_2) - \mathcal{K}[f](\mathbf{x}_1) \\
&\quad = \int_{\mathbb{R}^n} (K\,\psi)(\mathbf{x}_2, \mathbf{y}) \big(f(\mathbf{y}) - f(\mathbf{x}_2)\big) - (K\,\psi)(\mathbf{x}_1, \mathbf{y}) \big(f(\mathbf{y}) - f(\mathbf{x}_1)\big)\, d\mathbf{y} \\
&\quad = \int_{\mathbb{B}(\mathbf{x}_2, 2\delta)} K(\mathbf{x}_2, \mathbf{y}) \big(f(\mathbf{y}) - f(\mathbf{x}_2)\big) - K(\mathbf{x}_1, \mathbf{y}) \big(f(\mathbf{y}) - f(\mathbf{x}_1)\big)\, d\mathbf{y} \\
&\qquad + \int_{\mathbb{B}(\mathbf{x}_2, 2\delta)^c} (K\,\psi)(\mathbf{x}_2, \mathbf{y}) \big(f(\mathbf{y}) - f(\mathbf{x}_2)\big) - (K\,\psi)(\mathbf{x}_1, \mathbf{y}) \big(f(\mathbf{y}) - f(\mathbf{x}_1)\big)\, d\mathbf{y} \\
&\quad = I + II.
\end{aligned}$$

Now $B(\mathbf{x}_2, 2\delta) \subset B(\mathbf{x}_1, 3\delta)$. Thus using the size estimates (1) for K we have

$$\begin{aligned}
\big|I\big| &\leq C_0 ||\, f \,||_{\Lambda_\alpha} \int_{|\mathbf{y} - \mathbf{x}_2| < 2\delta} |\mathbf{y} - \mathbf{x}_2|^\alpha |K(\mathbf{x}_2, \mathbf{y})|\, d\mathbf{y} \\
&\qquad + C_0 ||\, f \,||_{\Lambda_\alpha} \int_{|\mathbf{y} - \mathbf{x}_1| < 3\delta} |\mathbf{y} - \mathbf{x}_1|^\alpha |K(\mathbf{x}_1, \mathbf{y})|\, d\mathbf{y} \\
&\leq \frac{C_0 \omega_n}{\alpha} ||\, f \,||_{\Lambda_\alpha} \left[(2\delta)^\alpha + (3\delta)^\alpha\right] = C(n, \alpha)\, ||\, f \,||_{\Lambda_\alpha} \delta^\alpha.
\end{aligned}$$

To deal with II, we rewrite the integrand as

$$\big(f(\mathbf{y}) - f(\mathbf{x}_1)\big)\big[(K\,\psi)(\mathbf{x}_2, \mathbf{y}) - (K\,\psi)(\mathbf{x}_1, \mathbf{y})\big] + \big[f(\mathbf{x}_1) - f(\mathbf{x}_2)\big](K\,\psi)(\mathbf{x}_2, \mathbf{y}).$$

Since ψ is radial, we can use the cancellation condition (3) to conclude that

$$\int_{B(\mathbf{x}_2 2\delta)^c} \big[f(\mathbf{x}_1) - f(\mathbf{x}_2)\big](K\,\psi)(\mathbf{x}_2, \mathbf{y})\, d\mathbf{y} = 0.$$

Thus

$$\big|II\big| \leq ||\, f \,||_{\Lambda^\alpha} \int_{|\mathbf{y} - \mathbf{x}_2| > 2\delta} |\mathbf{y} - \mathbf{x}_1|^\alpha |(K\,\psi)(\mathbf{x}_2, \mathbf{y}) - (K\,\psi)(\mathbf{x}_1, \mathbf{y})|\, d\mathbf{y}$$

If $|\mathbf{y} - \mathbf{x}_2| > 2\delta$, it follows that

$$|\mathbf{y} - \mathbf{x}_1| \leq |\mathbf{y} - \mathbf{x}_2| + |\mathbf{x}_2 - \mathbf{x}_1| = |\mathbf{y} - \mathbf{x}_2| + \delta < |\mathbf{y} - \mathbf{x}_2| + \frac{1}{2}|\mathbf{y} - \mathbf{x}_2|,$$

and so

$$|\mathbf{y} - \mathbf{x}_1| < \frac{3}{2}|\mathbf{y} - \mathbf{x}_2|.$$

Also, using the mean value theorem, it follows that

$$\begin{aligned}|(K\,\psi)(\mathbf{x}_2,\mathbf{y})-(K\,\psi)(\mathbf{x}_1,\mathbf{y})| &= |\mathbf{x}_2-\mathbf{x}_1|\,|\nabla_{\mathbf{x}}(K\,\psi)(\lambda\mathbf{x}_2+(1-\lambda)\mathbf{x}_1,\mathbf{y})|\\ &\leq \frac{3}{2}\,(C_0+C_1)\,\delta\,|\mathbf{y}-\mathbf{x}_2|^{-n-1}.\end{aligned}$$

Thus

$$\begin{aligned}\big|II\big| &\leq \left(\frac{3}{2}\right)^{1+\alpha}(C_0+C_1)\,\big|\big|\,f\,\big|\big|_{\Lambda_\alpha}\,\delta\int_{|\mathbf{y}-\mathbf{x}_2|>2\delta}|\mathbf{y}-\mathbf{x}_2|^{-n-1+\alpha}\\ &= \left(\frac{3}{2}\right)^{1+\alpha}(C_0+C_1)\,\frac{\omega_n}{1-\alpha}\,\big|\big|\,f\,\big|\big|_{\Lambda_\alpha}\,\delta^\alpha = C(n,\alpha)\,\big|\big|\,f\,\big|\big|_{\Lambda_\alpha}\,\delta^\alpha.\end{aligned}$$

This completes the proof. □

3.5. **Sobolev estimates.**

Instead of measuring the smoothness of a k^{th}-derivatrive with a Lipschitz norm, we can also measure it with an L^p-norm. Let X be a space of functions on $\mathbb{R}^n$ contained in $L^1_{\text{loc}}(\mathbb{R}^n)$. Then if $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ and $\alpha \in \mathbb{Z}^n$ is a multi-index, we say that $\partial^\alpha_{\mathbf{x}} f \in X$ in the sense of distributions if there exists a function $g_\alpha \in X$ so that for every $\varphi \in \mathcal{D}(\mathbb{R}^n)$ we have

$$\int_{\mathbb{R}^n} f(\mathbf{x})\partial^\alpha_{\mathbf{x}}\varphi(\mathbf{x})\,d\mathbf{x} = (-1)^{|\alpha|}\int_{\mathbb{R}^n} g(\mathbf{x})\varphi(\mathbf{x})\,d\mathbf{x}. \tag{3.4}$$

If g is a locally integrable function and if $\int_{\mathbb{R}^n} g(\mathbf{x})\varphi(\mathbf{x})\,d\mathbf{x} = 0$ for every $\varphi \in \mathcal{C}^\infty_0(\mathbb{R}^n)$, then $g(\mathbf{x}) = 0$ for almost every $\mathbf{x}$. Thus if $\partial^\alpha_{\mathbf{x}} f \in X$ exists, it is unique. Also, if f is smooth, then a derivative in the sense of distributions equals the classical derivative since equation (3.4) is then a consequence of integration by parts.

Definition 3.11. *If k is a non-negative integer and $1 \leq p \leq \infty$,*

$$L^p_k(\mathbb{R}^n) = \left\{ f \in L^p(\mathbb{R}^n)\,\middle|\,\partial^\alpha_{\mathbf{x}} f \in L^p(\mathbb{R}^n) \text{ in the sense of distributions for all } |\alpha| \leq k \right\}.$$

If $f \in L^p_k(\mathbb{R}^n)$, set

$$||f||_{L^p_k} = \sum_{0\leq|\alpha|\leq k} ||\partial^\alpha_{\mathbf{x}} f||_{L^p(\mathbb{R}^n)}.$$

We shall use the following properties of these Sobolev spaces.

Lemma 3.12.

(1) *The space $L^p_k(\mathbb{R}^n)$ is complete with the norm introduced in Definition 3.11, and so is a Banach space. $L^2_k(\mathbb{R}^n)$ is a Hilbert space with the inner product*

$$(f,g) = \sum_{0\leq|\alpha|\leq k} (\partial^\alpha_{\mathbf{x}} f, \partial^\alpha_{\mathbf{x}} g)_{L^2(\mathbb{R}^n)}.$$

(2) *A function $f \in L^p_k(\mathbb{R}^n)$ if and only if there is a sequence of functions $f_n \in L^p(\mathbb{R}^n)\cap\mathcal{C}^\infty(\mathbb{R}^n)$ such that for every $|\alpha| \leq k$, the sequence $\{\partial^\alpha_{\mathbf{x}} f_n\}$ is Cauchy in $L^p(\mathbb{R}^n)$.*

(3) *A function $f \in L^2_k(\mathbb{R}^n)$ if and only if $\int_{\mathbb{R}^n} |\widehat{f}(\xi)|^2(1+|\xi|^2)^k\,d\xi < +\infty$ where $\widehat{f}$ is the Fourier transform of f.*

(4) *If $f \in L^p_k(\mathbb{R}^n)$ and if $\psi \in \mathcal{C}^\infty_0(\mathbb{R}^n)$, then $\psi f \in L^p_k(\mathbb{R}^n)$. Thus membership in $L^p_k(\mathbb{R}^n)$ depends on the local behavior of f.*

We have seen that the operator which takes a function f to $\partial^2_{j,k}\big[N[f]\big]$ can be written as an operator of the form

$$\mathcal{K}[\varphi](x) = \int_{|y|<R} K(y)\big[\varphi(x-y)-\varphi(x)\big]\,dy + \int_{|y|\geq R} K(y)\,\varphi(x-y)\,dy$$

where $K \in \mathcal{C}^1(\mathbb{R}^n - \{0\})$ and we assume that

(1) For all $x \neq 0$, $|K(x)| \leq C_0|x|^{-n}$.

(2) For all $x \neq 0$, $|\nabla K(x)| \leq C_1|x|^{-n-1}$.

(3) For all $0 < R_1 < R_2$ we have $\displaystyle\int_{R_1<|x|<R_2} K(x)\,dx = 0$.

We want to show that the operator $\mathcal{K}$, defined for $\varphi \in \mathcal{C}^\infty_0(\mathbb{R}^n)$, has a bounded extension to $L^2(\mathbb{R}^n)$. Since this operator is given by convolution with a distribution K, it is natural to use the Fourier transform and the Plancherel theorem to reduce the problem to showing that the Fourier transform of the distribution is uniformly bounded. This can certainly be done, but since we will not have the Fourier transform available in later examples, we prefer to use a method with wider application. This is based on an 'almost orthogonality' argument, and the key result is the following.

Theorem 3.13 (Cotlar-Stein)**.** *Let $\{T_j\}$, $j \in \mathcal{Z}$, be bounded operators on a Hilbert space $\mathcal{H}$. Assume there are constants C and $\epsilon > 0$ so that for all $j,k \in \mathcal{Z}$ we have*

(1) $\big|\big|\,T_j\,\big|\big| \leq C$.

(2) $\big|\big|\,T_j^*\,T_k\,\big|\big| \leq C\,2^{-\epsilon|j-k|}$.

(3) $\big|\big|\,T_j\,T_k^*\,\big|\big| \leq C\,2^{-\epsilon|j-k|}$.

There is a constant A so that for all N

$$\Big|\Big|\sum_{j=-N}^{N} T_j\Big|\Big| \leq A.$$

This is proved, for example, in [10], pages 279-281, so we do not reproduce the argument here.

Now let $\chi \in \mathcal{C}^\infty_0(\mathbb{R})$ satisfy

(i) $0 \leq \chi(t) \leq 1$ for all $t \in \mathbb{R}$;

(ii) We have

$$\chi(t) = \begin{cases} 0 & \text{if } t \leq \frac{1}{8}, \\ 1 & \text{if } \frac{1}{4} \leq t \leq \frac{1}{2}, \\ 0 & \text{if } 1 \leq t. \end{cases}$$

Put $\chi_j(t) = \chi(2^{-j}t)$. Then each $t > 0$ is in the support of at most 4 of the functions $\{\chi_j\}$. Put

$$\Psi_j(t) = \big[\sum_k \chi_k(t)\big]^{-1}\,\chi_j(t).$$

Then Ψ_j is supported on $2^{j-3} \leq t \leq 2^j$, and

(iii) $\displaystyle\sum_{j=-\infty}^{\infty} \Psi_j(t) \equiv 1$ for all $t > 0$.

(iv) $\displaystyle\sum_{j=-N}^{N} \Psi_j$ is supported on $2^{-N-3} \leq t \leq 2^N$, and is identically 1 for $2^{-N} \leq t \leq 2^{N-3}$.

Now put

$$K_j(x) = \Psi_j(|x|)\, K(x).$$

Then it is not difficult to check that $\{K_j\}$, $j = 0, \pm 1, \pm 2, \dots$ is a doubly infinite sequence of continuously differentiable functions on $\mathbb{R}^n$, and there is a constant C so that

(1) For all $j \in \mathcal{Z}$ and all $x \in \mathbb{R}^n$ we have $|K_j(x)| \le C\, 2^{-nj}$.

(2) For all $j \in \mathcal{Z}$ and all $x \in \mathbb{R}^n$ we have $|\nabla K_j(x)| \le C\, 2^{-(n+1)j}$.

(3) For all $j \in \mathcal{Z}$, K_j is supported in the ball $\mathbb{B}_E(0; 2^j)$.

(4) For all $j \in \mathcal{Z}$, $\displaystyle\int_{\mathbb{R}^n} K_j(x)\, dx = 0$.

Moreover

(5) $\displaystyle\sum_{j=-\infty}^{\infty} K_j(x) = K(x)$ for $x \neq 0$.

(6) $K^{[N]} = \displaystyle\sum_{j=-N}^{N} K_j$ is supported $2^{-N-3} \le |x| \le 2^N$, and is identically equal to K for $2^{-N} \le |x| \le 2^{N-3}$. Moreover, the kernel $K^{[N]}$ satisfies the same conditions (1), (2), and (3) as K with constants that are independent of N.

Set

$$\mathcal{K}_j[f](x) = K_j * \varphi(x) = \int_{\mathbb{R}^n} K_j(y)\, f(x-y)\, dy.$$

It follows that if $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$, we have

$$\begin{aligned}
\mathcal{K}[\varphi](x) &= \int_{|y|<R} K(y)\,\big[\varphi(x-y) - \varphi(x)\big]\, dy + \int_{|y|\ge 1} K(y)\,\varphi(x-y)\, dy \\
&= \lim_{N\to\infty} \int_{\mathbb{R}^n} \sum_{j=-N}^{N} K_j(y)\,\varphi(x-y)\, dy \\
&= \lim_{N\to\infty} \sum_{j=-N}^{N} \mathcal{K}_j[\varphi](x).
\end{aligned}$$

Thus if we can show that the operators $\{\mathcal{K}_j\}$ satisfy the almost orthogonality conditions of Theorem 3.13, it follows from Fatou's lemma that for $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ we have

$$\begin{aligned}
\big|\big|\, \mathcal{K}[\varphi]\, \big|\big|^2_{L^2(\mathbb{R}^n} &= \int_{\mathbb{R}^n} \lim_{N\to\infty} \Big| \sum_{j=-N}^{+N} \mathcal{K}_j[\varphi](x)\Big|^2\, dx \\
&\le \limsup_{N\to\infty} \int_{\mathbb{R}^n} \Big| \sum_{j=-N}^{+N} \mathcal{K}_j[\varphi](x)\Big|^2\, dx \\
&\le A^2 \big|\big|\, \varphi\, \big|\big|^2_{L^2(\mathbb{R}^n)}.
\end{aligned}$$

It follows from (1) and (3) that

$$\big|\big|\, K_j\, \big|\big|_{L^1(\mathbb{R}^n)} = \int_{|x|<2^j} \big|K_j(x)\big|\, dx \le C\, 2^{-nj}\, \big|\mathbb{B}_E(0; 2^j)\big| \le C\, n^{-1}\, \omega_n$$

since $n^{-1}\omega_n$ is the volume of the Euclidean unit ball in $\mathbb{R}^n$. Now if $f, g \in L^1(\mathbb{R}^n)$, we always have

$$\big|\big|\, f * g\, \big|\big|_{L^1} = \int_{\mathbb{R}^n} |f * g(x)|\, dx \le \iint_{\mathbb{R}^n \times \mathbb{R}^n} |f(x-y)|\, |g(y)|\, dy\, dx = \big|\big|\, f\, \big|\big|_{L^1} \big|\big|\, g\, \big|\big|_{L^1},$$

and consequently $\big|\big| K_j * K_k \big|\big|_{L^1} \leq (C\, n^{-1}\, \omega_n)^2$. However, the differential inequality (2) and the cancellation property (4) allow us to get a better estimate when $j \neq k$.

Proposition 3.14. *For all $j, k \in \mathcal{Z}$ we have*

$$\big|\big| K_j * K_k \big|\big|_{L^1} \leq 2^n \,(C\, n^{-1}\, \omega_n)^2\, 2^{-|j-k|}.$$

Proof. Without loss of generality, assume that $j < k$. Then

$$\begin{aligned}
\big|K_j * K_k(x)\big| &= \Big|\int_{\mathbb{R}^n} K_k(x-y)\, K_j(y)\, dy\Big| \\
&= \Big|\int_{\mathbb{R}^n} \big[K_k(x-y) - K_k(x)\big]\, K_j(y)\, dy\Big| && \text{(using assumption (4))} \\
&\leq \int_{\mathbb{B}_E(0;2^j)} \big|K_k(x-y) - K_k(x)\big|\, \big|K_j(y)\big|\, dy \\
&\leq \int_{\mathbb{B}_E(0;2^j)} |y| \sup_{z \in \mathbb{R}^n} \big|\nabla K_k(z)\big|\, \big|K_j(y)\big|\, dy && \text{(Mean Value Theorem)} \\
&\leq C\, 2^{j-(n+1)k} \big|\big| K_j \big|\big|_{L^1} && \text{(using estimate (3))}.
\end{aligned}$$

On the other hand, if $K_j * K_k(x) \neq 0$ we must have $|y| \leq 2^j$ and $|x-y| \leq 2^k$, so $|x| \leq |x-y| + |y| \leq 2 \cdot 2^k$. Thus

$$\big|\big| K_j * K_k \big|\big|_{L^1} \leq C\, 2^{j-(n+1)k} \big|\big| K_j \big|\big|_{L^1} \big|\mathbb{B}(0; 2\cdot 2^k)\big| \leq 2^n\, (C\, n^{-1}\, \omega_n)^2\, 2^{j-k}.$$

This completes the proof. □

Lemma 3.15. *Let $\{K_j\}$ be functions satisfying conditions (1) through (4), and for each $j \in \mathcal{Z}$ define an operator T_j by setting*

$$T_j[f](x) = K_j * f(x) = \int_{\mathbb{R}^n} K_j(x-y)\, f(y)\, dy.$$

There exists a constant C so that for any integer N we have

$$\Big|\Big| \sum_{j=-N}^{N} T_j[f] \Big|\Big|_{L^2(\mathbb{R}^n)} \leq C\, \big|\big| f \big|\big|_{L^2(\mathbb{R}^n)}.$$

3.6. L^1-estimates.

It is not true that the operator $\mathcal{K}$ from section 3.5, defined on the space $\mathcal{C}_0^\infty(\mathbb{R}^n)$, extends to a bounded operator on $L^1(\mathbb{R}^n)$, but we do have the following replacement, which is called a *weak type (1,1)* estimate. Let us write $\mathcal{K}^{[N]} = \sum_{j=-N}^{+N} \mathcal{K}_j$. We will need the following estimate, which is sometimes called the Calderón-Zygmund estimate.

Lemma 3.16. *Let $\eta > A_2$, let $B = \mathbb{B}_\rho(x_0;\delta_0)$, and let $B^* = \mathbb{B}_\rho(x_0; \eta\, \delta_0)$. Suppose $x_1, x_2 \in B$. Then*

$$\int_{\mathbb{R}^n - B^*} \big|K^{[N]}(y - x_1) - K^{[N]}(y - x_2)\big|\, dy \leq C.$$

Theorem 3.17. *There is a constant A independent of N so that if $f \in L^1(\mathbb{R}^n)$, then*

$$\Big|\Big\{x \in \mathbb{R}^n \,\Big|\, |\mathcal{K}^{[N]}[f](x) > \alpha\Big\}\Big| \leq \frac{A}{\lambda}\, \big|\big| f \big|\big|_{L^1(\mathbb{R}^n)}.$$

Proof. We apply the Calderón-Zygmund decomposition of Theorem to the function f. Let $E_\alpha = \Big\{x \in \mathbb{R}^n \,\Big|\, \mathbb{M}[f](x) > \alpha\Big\}$. Then we can write $f = g + \sum_j b_j = g + b$ where g and $\{b_j\}$ satisfy

(1) b_j is supported on a ball $\mathbb{B}_E(x_j;\delta_j) \subset E_\alpha$ and $\int_{\mathbb{R}^n} b_j(x)\,dx = 0$ while $\big|\big|\, b_j \,\big|\big|_{L^1(\mathbb{R}^n)} \leq C\,\alpha\,\big|B_j\big|$;

(2) If $x \notin E_\alpha$, then $|x - x_j| \geq 2\delta_j$;

(3) $\sum_j \big|\mathbb{B}_E(x_j;\delta_j)\big| \leq C\,\big|E_\alpha\big| \leq \dfrac{C}{\alpha}\,\big|\big|\, f \,\big|\big|_{L^1(\mathbb{R}^n)}$;

(4) $|g(x)| \leq C\,\alpha$ for almost all $x \in \mathbb{R}^n$.

Since $|\mathcal{K}^{[N]}[f](x)| \leq |\mathcal{K}^{[N]}[g](x)| + |\mathcal{K}^{[N]}[b](x)|$, we have

$$\begin{aligned}
&\Big|\Big\{x \in \mathbb{R}^n \,\Big|\, |\mathcal{K}^{[N]}[f](x) > \alpha\Big\}\Big| \\
&\qquad \leq \Big|\Big\{x \in \mathbb{R}^n \,\Big|\, |\mathcal{K}^{[N]}[g](x) > \frac{\alpha}{2}\Big\}\Big| \cup \Big|\Big\{x \in \mathbb{R}^n \,\Big|\, |\mathcal{K}^{[N]}[b](x) > \frac{\alpha}{2}\Big\}\Big|.
\end{aligned}$$

Now $\mathcal{K}^{[N]}$ is a bounded operator on $L^2(\mathbb{R}^n)$ with norm A independent of N. Thus

$$\begin{aligned}
\frac{\alpha^2}{4}\Big|\Big\{x \in \mathbb{R}^n \,\Big|\, |\mathcal{K}^{[N]}[g](x) > \frac{\alpha}{2}\Big\}\Big| &\leq \int_{\mathbb{R}^n} |\mathcal{K}^{[N]}[g](x)|^2\,dx \\
&\leq A\big|\big|\, g \,\big|\big|_{L^2}^2 \\
&\leq A\,C\,\alpha\,\big|\big|\, f \,\big|\big|_{L^1}
\end{aligned}$$

and so

$$\Big|\Big\{x \in \mathbb{R}^n \,\Big|\, |\mathcal{K}^{[N]}[g](x) > \frac{\alpha}{2}\Big\}\Big| \leq \frac{4AC}{\alpha}\,\big|\big|\, f \,\big|\big|_{L^1}.$$

Now suppose we can show that

$$\int_{\mathbb{R}^n - E_\alpha} |\mathcal{K}^{[N]}[b](x)|\,dx \leq C\big|\big|\, f \,\big|\big|_{L^1}. \tag{3.5}$$

Then

$$\begin{aligned}
&\Big|\Big\{x \in \mathbb{R}^n \,\Big|\, |\mathcal{K}^{[N]}[b](x) > \frac{\alpha}{2}\Big\}\Big| \\
&\qquad \leq \Big|E_\alpha\Big| + \Big|\Big\{x \in \mathbb{R}^n - E_\alpha \,\Big|\, |\mathcal{K}^{[N]}[b](x) > \frac{\alpha}{2}\Big\}\Big| \\
&\qquad \leq \frac{C_1}{\alpha}\big|\big|\, f \,\big|\big|_{L^1} + \frac{2}{\alpha}\int_{\mathbb{R}^n - E_\alpha} |\mathcal{K}^{[N]}[b](x)|\,dx \\
&\qquad \leq \frac{C'}{\alpha}\,\big|\big|\, f \,\big|\big|_{L^1},
\end{aligned}$$

which would complete the proof. Thus the key is to prove the estimate is equation (3.5).

Suppose $x \notin E_\alpha$. Since the integral of b_j is zero, we have

$$\begin{aligned}
\big|\mathcal{K}^{[N]}[b_j](x)\big| &= \Big|\int_{\mathbb{R}^n} K^{[N]}(x-y)\,b_j(y)\,dy\Big| \\
&= \Big|\int_{\mathbb{R}^n} \big[K^{[N]}(x-y) - K(x-x_j)\big]\,b_j(y)\,dy\Big| \\
&\leq \int_{\mathbb{R}^n} \big|K^{[N]}(x-y) - K^{[N]}(x-x_j)\big|\,|b_j(y)|\,dy
\end{aligned}$$

and hence

$$\begin{aligned}
&\int_{\mathbb{R}^n - E_\alpha} \big|\mathcal{K}^{[N]}[b_j](x)\big|\,dx \\
&\qquad \le \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n - E_\alpha} \big|K^{[N]}(x-y) - K^{[N]}(x-x_j)\big|\,dx\right] |b_j(y)|\,dy \\
&\qquad \le C \int_{\mathbb{R}^n} |b_j(y)|\,dy \\
&\qquad \le C'\,\alpha\,\big|B_j\big|.
\end{aligned}$$

It follows that

$$\begin{aligned}
\int_{\mathbb{R}^n - E_\alpha} \big|\mathcal{K}^{[N]}[b](x)\big|\,dx &\le \sum_j \int_{\mathbb{R}^n - E_\alpha} \big|\mathcal{K}^{[N]}[b_j](x)\big|\,dx \\
&\le C'\,\alpha \sum_j \big|B_j\big| \\
&\le C''\big|\big|\, f\,\big|\big|_{L^1}.
\end{aligned}$$

□

Appendix A. Test functions and distributions

A.1. **Some spaces of smooth functions.**

Let $\Omega \subset \mathbb{R}^n$ be open. If m is a non-negative integer, $\mathcal{C}^m(\Omega)$ denotes the space of m-times continuously differentiable functions on Ω and $\mathcal{C}^\infty(\Omega) = \bigcap_{m=1}^\infty \mathcal{C}^m(\Omega)$ the space of infinitely differentiable functions on Ω. Note that $\mathcal{C}^m(\Omega) \supset \mathcal{C}^{m+1}(\Omega) \supset \mathcal{C}^\infty(\Omega)$. If $K \Subset \Omega$, we put

$$\big|\big|\, f\,\big|\big|_{K,m} = \sup_{|\alpha| \le m} \sup_{\mathbf{x} \in K} \big|\partial^\alpha \varphi(\mathbf{x})\big|. \tag{A.1}$$

Then $\big|\big| \cdot \big|\big|_{K,m}$ is a semi-norm on the space $\mathcal{C}^m(\Omega)$ and also on the space $\mathcal{C}^\infty(\Omega)$. Define a topology on $\mathcal{C}^\infty(\Omega)$ so that a sequence $\{f_n\} \subset \mathcal{C}^\infty(\Omega)$ converges to a limit function $f_0 \in \mathcal{C}^\infty(\Omega)$ if and only if $\lim_{n\to\infty} \big|\big|\, f_0 - f_n\,\big|\big|_{K,m} = 0$ for every compact subset $K \subset \Omega$ and for every non-negative integer m. The space $\mathcal{C}^\infty(\Omega)$ equipped with this topology is denoted by $\mathcal{E}(\Omega)$. If $\{f_n\} \subset \mathcal{E}(\Omega)$ is a sequence which is Cauchy in every semi-norm $\big|\big| \cdot \big|\big|_{K,m}$, then the sequence converges to a limit $f_0 \in \mathcal{E}(\Omega)$.

If $f \in \mathcal{C}^m(\Omega)$, the *support* of f is the closure of the set of points $\mathbf{x} \in \Omega$ such that $f(\mathbf{x}) \neq 0$, and is denoted by $\mathrm{suppt}(f)$. Then $\mathcal{C}_0^\infty(\Omega)$ is the subspace of $\mathcal{C}^\infty(\Omega)$ consisting of functions whose support is a compact subset of Ω. There is a topology on $\mathcal{C}_0^\infty(\Omega)$ such that a sequence $\{f_n\} \subset \mathcal{C}_0^\infty(\Omega)$ converges to a limit function $f_0 \in \mathcal{C}_0^\infty(\Omega)$ if and only there is a compact set $K \subset \Omega$ so that $\mathrm{suppt}(f_n) \subset K$ for all $n \ge 0$ and $\lim_{n\to\infty} \big|\big|\, f_0 - f_n\,\big|\big|_{K,m} = 0$ for every non-negative integer m. The space $\mathcal{C}_0^\infty(\Omega)$ with this topology is denoted by $\mathcal{D}(\Omega)$.[7]

The *Schwartz space* $\mathcal{S}(\mathbb{R}^n)$ consists of all $\varphi \in \mathcal{C}^\infty(\mathbb{R}^n)$ such that for all multi-indices α and β,

$$\sup_{\mathbf{x} \in \mathbb{R}^n} \big|\mathbf{x}^\alpha \partial^\beta \varphi(\mathbf{x})\big| = \big|\big|\, \varphi\,\big|\big|_{\alpha,\beta} < +\infty.$$

The natural topology on $\mathcal{S}(\mathbb{R}^n)$ is given by the collection of these semi-norms. Thus a sequence $\{\varphi_1, \varphi_2, \ldots\}$ in $\mathcal{S}(\mathbb{R}^n)$ converges to φ_0 if and only if $\lim_{n\to\infty} \big|\big|\, \varphi_n - \varphi_0\,\big|\big|_{\alpha,\beta} = 0$.

In case $\Omega = \mathbb{R}^n$, the inclusions of spaces

$$\mathcal{D}(\mathbb{R}^n) \hookrightarrow \mathcal{S}(\mathbb{R}^n) \hookrightarrow \mathcal{E}(\mathbb{R}^n)$$

[7] Note that this topology is *not* the same as the topology that $\mathcal{C}_0^\infty(\Omega)$ inherits as a subspace of $\mathcal{E}(\Omega)$.

are all continuous.

A.2. **The space $\mathcal{D}'(\Omega)$ of distributions on Ω.**

The space of *distributions* on Ω is the dual space $\mathcal{D}'(\Omega)$; that is, the set of continuous real- or complex-valued continuous linear functionals on $\mathcal{D}(\Omega)$. The pairing between a distribution $u \in \mathcal{D}'(\Omega)$ and a smooth, compactly supported function $\varphi \in \mathcal{D}(\Omega)$ is denoted by $\big\langle u, \varphi \big\rangle$. Continuity of u means that if $\{\varphi_n\}$ is a sequence of test functions in $\mathcal{D}(\Omega)$, if there is a compact set K containing suppt(φ_n) for every n, and if $\lim_{n\to\infty} ||\varphi_n - \varphi_0||_{K,m} = 0$ for every m, then

$$\lim_{n\to\infty} \langle u, \varphi_n \rangle = \langle u, \varphi_0 \rangle.$$

If $f \in L^1_{loc}(\Omega)$, then f induces a distribution u_f whose action is given by

$$\big\langle u_f, \varphi \big\rangle = \int_\Omega f(\mathbf{x})\, \varphi(\mathbf{x})\, d\mathbf{x}.$$

Abusing notation, we shall frequently denote the distribution u_f simply by f. We shall sometimes engage in another, perhaps more serious abuse by writing the action of a distribution $u \in \mathcal{D}'(\Omega)$ on a function $\varphi \in \mathcal{D}(\Omega)$ by $\int_\Omega u(\mathbf{x})\, \varphi(\mathbf{x})\, d\mathbf{x}$.

The space $\mathcal{D}(\Omega)$ is closed under the operation of multiplication by a smooth function in $\mathcal{E}(\Omega)$, and also under differentiation. By duality, this allows us to define the product of a distribution and a smooth function, and to define the derivative of a distribution. Thus let $u \in \mathcal{D}'(\Omega)$ be a distribution. If $a \in \mathcal{C}^\infty(\Omega)$ is a smooth function, the product au is the distribution defined by the requirement that

$$\big\langle au, \varphi \big\rangle = \big\langle u, a\varphi \big\rangle.$$

For every multi-index $\alpha \in \mathbb{Z}^n_+$, the distribution $\partial^\alpha u$ is defined by the formula

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

In particular, if

$$P[\varphi](\mathbf{x}) = \sum_{|\alpha| \le M} a_\alpha(\mathbf{x}) \partial^\alpha[\varphi](\mathbf{x})$$

is a linear partial differential operator of order M with coefficients $\{a_\alpha\} \subset \mathcal{C}^\infty(\Omega)$ and if $u \in \mathcal{D}'(\Omega)$ is a distribution, then $P[u]$ is also a distribution on Ω whose action on a test function $\varphi \in \mathcal{E}(\Omega)$ is given by

$$\big\langle P[u], \varphi \big\rangle = \big\langle u, \sum_{|\alpha| \le M} (-1)^{|\alpha|} \partial^\alpha [a_\alpha \varphi] \big\rangle. \tag{A.2}$$

Note that the operator

$$P^*[\varphi](\mathbf{x}) = \sum_{|\alpha| \le M} (-1)^{|\alpha|} \partial^\alpha \big[a_\alpha\, \varphi\big](\mathbf{x})$$

is also a linear partial differential operator of order M. It is called the *formal adjoint* of P. Thus equation (A.2) is equivalent to the statement

$$\big\langle P[u], \varphi \big\rangle = \big\langle u, P^*[\varphi] \big\rangle.$$

Suppose that $\Omega = \bigcup_{\alpha \in \mathcal{A}} U_\alpha$ where each U_α is a relatively compact open subset of Ω. Then there is a partition of unity subordinate to this collection of open sets; for each $\alpha \in \mathcal{A}$ there exists $\varphi_\alpha \in \mathcal{C}^\infty_0(U_\alpha) \subset \mathcal{C}^\infty_0(\mathbb{R}^n)$ so that

(a) for each $\mathbf{x} \in \Omega$ there is an open neighborhood $V_\mathbf{x}$ so that $\varphi_\alpha(\mathbf{y}) = 0$ for all $\mathbf{y} \in V_\mathbf{x}$ and for all except for finitely many indices $\alpha \in \mathcal{A}$;

(b) $\sum_{\alpha\in\mathcal{A}} \varphi_\alpha(\mathbf{x}) \equiv 1$ for all $\mathbf{x} \in \Omega$.

Then if $\psi \in \mathcal{C}_0^\infty(\Omega)$ there is a finite collection of indices $\mathcal{A}_\psi$ so that for all $\mathbf{x} \in \Omega$,

$$\psi(\mathbf{x}) = \sum_{\alpha\in\mathcal{A}_\psi} (\varphi_\alpha\psi)(\mathbf{x}).$$

Note that $\varphi_\alpha\psi \in \mathcal{C}_0^\infty(U_\alpha)$. Suppose that $u, v \in \mathcal{D}'(\Omega)$, and that for every $\alpha \in \mathcal{A}$, $\langle u, \theta\rangle = \langle v, \theta\rangle$ for every $\theta \in \mathcal{C}_0^\infty(U_\alpha)$. Then if $\psi \in \mathcal{C}_0^\infty(\Omega)$,

$$\langle u, \psi\rangle = \sum_{\alpha\in\mathcal{A}_\psi} \langle u, \varphi_\alpha\psi\rangle = \sum_{\alpha\in\mathcal{A}_\psi} \langle v, \varphi_\alpha\psi\rangle = \langle v, \psi\rangle,$$

which shows that distributions are determined by their local behavior. It thus makes sense to define the support of a distribution. If $u \in \mathcal{D}'(\Omega)$, there is a smallest closed set $E \subset \Omega$ such that if $\varphi \in \mathcal{C}_0^\infty(\Omega - E)$, then $\langle u, \varphi\rangle = 0$. This set is called the *support* of u, and we again denote the support of a distribution u by $\mathrm{suppt}(u)$.

The continuity of a linear functional u is needed in order to allow "differentiation under the integral sign".

Lemma A.1. *Let $U, V \subset \mathbb{R}^n$ be open sets, let $\varphi \in \mathcal{C}^\infty(U \times V)$, and let $u \in \mathcal{D}'(\Omega)$. For $\mathbf{y} \in V$ set $\varphi_{\mathbf{y}}(\mathbf{x}) = \varphi(\mathbf{x}, \mathbf{y})$, and set $\psi(\mathbf{y}) = \langle u, \varphi_y\rangle$. If there is a compact set $E \Subset U$ so that $\varphi(\mathbf{x}, \mathbf{y}) = 0$ if $\mathbf{x} \notin E$, then $\psi \in \mathcal{C}^\infty(V)$ and $\partial^\alpha\psi(\mathbf{y}) = \langle u, \partial_{\mathbf{y}}^\alpha\varphi(\,\cdot\,, \mathbf{y})\rangle$.*

Formally, we have

$$\psi(\mathbf{y}) = \langle u, \varphi_{\mathbf{y}}\rangle = \int_U u(\mathbf{x})\varphi(\mathbf{x}, \mathbf{y})\, d\mathbf{x},$$

and the Lemma says that

$$\partial^\alpha\psi(\mathbf{y}) = \int_U u(\mathbf{x})\partial_{\mathbf{y}}^\alpha\varphi(\mathbf{x}, \mathbf{y})\, d\mathbf{x}.$$

For the proof, see Theorem 2.1.3 in [6]. The main idea is to show that ψ is differentiable by using Taylor's formula to write

$$\varphi(\mathbf{x}, \mathbf{y} + \mathbf{h}) = \varphi(\mathbf{x}, \mathbf{y}) + \sum_{j=1}^n h_j\partial_{y_j}\varphi(\mathbf{x}, \mathbf{y}) + O(|\mathbf{h}|^2).$$

A.3. **Other spaces of distributions.**

Instead of using smooth, compactly supported functions as test functions, one can use larger spaces. In general this gives a restricted class of distributions.

The space $\mathcal{E}'(\Omega)$ is the space of continuous linear functionals on $\mathcal{E}(\Omega)$, the space of *all* smooth functions on Ω. Since the inclusion map $\mathcal{D}(\Omega) \hookrightarrow \mathcal{E}(\Omega)$ is continuous, every continuous linear functional on $\mathcal{E}(\Omega)$ restricts to a continuous linear functional on $\mathcal{D}(\Omega)$, so there is an induced mapping $\mathcal{E}'(\Omega) \to \mathcal{D}'(\Omega)$. It is not hard to see that this mapping is one-to-one, so $\mathcal{E}'(\Omega)$ is a subspace of distributions on Ω. In fact, $\mathcal{E}'(\Omega)$ is exactly the space of distributions on Ω with compact support.[8]

The space $\mathcal{S}'(\mathbb{R}^n)$ is the space of continuous linear functionals on the Schwartz spaces $\mathcal{S}(\mathbb{R}^n)$. Since $\mathcal{D}(\mathbb{R}^n) \hookrightarrow \mathcal{S}(\mathbb{R}^n)$ is continuous, there is an induced mapping $\mathcal{S}'(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^n)$ which is also one-to-one since $\mathcal{D}(\mathbb{R}^n)$ is dense in $\mathcal{S}(\mathbb{R}^n)$. Elements of $\mathcal{S}'(\mathbb{R}^n)$ are called *tempered distributions*. $\mathcal{S}'(\mathbb{R}^n)$ is a natural space to use when studying the Fourier transform.

---

[8] See, for example, Theorem 2.3.1 in [6].

A.4. **Convolutions of distributions and test functions.**

If $f, g \in L^1(\mathbb{R}^n)$, the *convolution* $f * g$ is the function

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

Fubini's theorem shows that $||f*g||_1 \leq ||f||_1\,||g||_1$, and in particular, the integral defining $f*g$ converges absolutely for almost every $\mathbf{x} \in \mathbb{R}^n$. If we set $\widetilde{g}(\mathbf{x}) = g(-\mathbf{x})$, then convolution can be written

$$f * g(\mathbf{x}) = \int_{\mathbb{R}^n} f(\mathbf{y})\widetilde{g}(\mathbf{y}-\mathbf{x})\,d\mathbf{y} = \int_{\mathbb{R}^n} f(\mathbf{y})\tau_{\mathbf{x}}\widetilde{g}(\mathbf{y})\,d\mathbf{y}. \tag{A.3}$$

We can now define the convolution of a distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ and a test function $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$. Motivated by equation (A.3), set

$$u * \varphi(\mathbf{x}) = \langle u, \tau_{\mathbf{x}}\widetilde{\varphi}\rangle.$$

Lemma A.2. *If $u \in \mathcal{D}'(\mathbb{R}^n)$ and $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ then:*

(a) *$u * \varphi \in \mathcal{C}^\infty(\mathbb{R}^n)$;*

(b) *For any multi-index α, $\partial^\alpha\big[u*\varphi\big] = (\partial^\alpha u)*\varphi = u*(\partial^\alpha\varphi)$.*

(c) *Suppt $(u*\varphi) \subset$ Suppt (u) + Suppt (φ), and in particular, if u has compact support then so does $u*\varphi$.*

Proof. The mapping $\mathbb{R} \to \mathcal{D}(\mathbb{R}^n)$ given by $\mathbf{x} \to \tau_{\mathbf{x}}\widetilde{\varphi}$ is continuous, and it follows that $u*\varphi$ is a continuous function. Also, if $\mathbf{e}_j$ is the j^{th} standard basis element in $\mathbb{R}^n$,

$$\lim_{h\to 0} h^{-1}\left[\tau_{\mathbf{x}+h\mathbf{e}_j}[\widetilde{\varphi}] - \tau_{\mathbf{x}}[\widetilde{\varphi}]\right] = \tau_{\mathbf{x}}\left[\widetilde{\partial_{x_j}\varphi}\right]$$

with convergence in $\mathcal{D}(\mathbb{R}^n)$. It follows easily that $u*\varphi$ is continuously differentiable with $\partial_{x_j}[u*\varphi] = u*\partial_{x_j}\varphi = \partial_{x_j}u*\varphi$. Then by induction it follows that $u*\varphi \in \mathcal{C}^\infty(\mathbb{R}^n)$ with $\partial^\alpha[u*\varphi] = u*\partial^\alpha\varphi = \partial^\alpha u*\varphi$. This establishes (a) and (b). If $u*\varphi(\mathbf{x}) \neq 0$ then $\mathrm{Suppt}(u)\cap\mathrm{Suppt}(\tau_{\mathbf{x}}\varphi) \neq \emptyset$, and so $\mathrm{Suppt}(u*\varphi) \subset \mathrm{Suppt}(u)+\mathrm{Suppt}(\varphi)$, and this completes the proof. □

A.5. **Convolution of two distributions.**

We will need to extend the definition of convolution to the case where both factors are distributions. This is not always possible unless at least one of the distributions has compact support. In this section we summarize results from [6] that we will need in order to prove that are needed.

Lemma A.3. *If $u \in \mathcal{D}'(\mathbb{R}^n)$ and $\varphi, \psi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ then*

$$(u*\varphi)*\psi = u*(\varphi*\psi).$$

For the proof, see Theorem 4.1.2 and Lemma 4.1.3 in [6]. The idea is to use the fact that the Riemann sum $\sum_{\mathbf{m}\in\mathbb{Z}^n}\varphi(\mathbf{x}-t\mathbf{m})t^n\psi(t\mathbf{m})$ converges to $\varphi*\psi(\mathbf{x})$ as $t\to 0$, and then use the continuity of u.

We will use the following important result (Theorem 4.2.1 in [6]).

Lemma A.4. *Let $A : \mathcal{C}_0^\infty(\mathbb{R}^n) \to \mathcal{C}^\infty(\mathbb{R}^n)$ be a linear map. Assume*

(a) *A commutes with translations: $\tau_{\mathbf{x}}A = A\tau_{\mathbf{x}}$ for all $\mathbf{x} \in \mathbb{R}^n$ where $\tau_{\mathbf{x}}$ is translation by $\mathbf{x}$ defined earlier.*

(b) *A is continuous in the sense that $A(\varphi_j)(0) \to 0$ if $\varphi_j \to 0$ in $\mathcal{C}_0^\infty(\mathbb{R}^n)$.*

Then there exists a unique distribution $\mathcal{A} \in \mathcal{D}'(\mathbb{R}^n)$ so that

$$A[\varphi](\mathbf{x}) = \mathcal{A} * \varphi(\mathbf{x}).$$

Proof of Lemma A.4. Note that if such a distribution exists, we must have $A[\varphi](0) = \mathcal{A} * \varphi(0) = \big\langle \mathcal{A}, \widetilde{\varphi} \big\rangle$, or $\big\langle \mathcal{A}, \varphi \big\rangle = A[\widetilde{\varphi}](0)$, so $\mathcal{A}$ is certainly uniquely determined by A. On the other hand, the linear functional

$$\mathcal{C}_0^\infty(\mathbb{R}^n) \ni \varphi \longrightarrow A[\widetilde{\varphi}](0) \in \mathbb{C}$$

is continuous and so defined a distribution $\mathcal{A} \in \mathcal{D}'(\mathbb{R}^n)$, and

$$\mathcal{A} * \varphi(0) = \big\langle \mathcal{A}, \widetilde{\varphi}_0 \big\rangle = A[\varphi](0).$$

But then

$$\mathcal{A} * \varphi(\mathbf{x}) = \big\langle \mathcal{A}, \widetilde{\varphi}_{\mathbf{x}} \big\rangle = A[\varphi_{\mathbf{x}}](0) = A[T_{-\mathbf{x}}\varphi](0) = T_{-\mathbf{x}}A[\varphi](0) = A[\varphi](\mathbf{x})$$

since A commutes with translation. □

Now let $\mathcal{S}, u \in \mathcal{D}'(\mathbb{R}^n)$ and assume that at least one of them has compact support. Note:

(a) if $\mathcal{S}$ has compact support, then $u * \varphi \in \mathcal{C}^\infty(\mathbb{R}^n)$ and so $\mathcal{S} * [u * \varphi] \in \mathcal{C}^\infty(\mathbb{R}^n)$;
(b) if u has compact support, then $u * \varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ and so $\mathcal{S} * [u * \varphi] \in \mathcal{C}^\infty(\mathbb{R}^n)$.

Thus in either case,

$$\mathcal{C}_0^\infty(\mathbb{R}^n) \ni \varphi \longrightarrow \mathcal{S} * [u * \varphi] \in \mathcal{C}^\infty(\mathbb{R}^n)$$

and the mapping is continuous. Thus by Lemma A.4 there exists a unique distribution $\mathcal{S} * u \in \mathcal{D}'(\mathbb{R}^n)$ so that

$$[\mathcal{S} * u] * \varphi = \mathcal{S} * [u * \varphi].$$

Lemma A.5. *Let $\mathcal{S}, u \in \mathcal{D}(\mathbb{R}^n)$ and suppose that at least one has compact support.*

(a) $\mathcal{S} * u = u * \mathcal{S}$.

(b) *For any multi-index α, $\partial^\alpha[\mathcal{S} * u] = [\partial^\alpha u] * \mathcal{S} = u * [\partial^\alpha \mathcal{S}]$.*

(c) *Support $(\mathcal{S} * u) \subset$ Support $(\mathcal{S})$ + Support (u).*

For the proof, see Theorem 4.2.4 and the discussion following its proof in [6].

The crucial result we need to establish hypoellipticity uses the notion of *singular support*. Thus if $u \in \mathcal{D}'(\mathbb{R}^n)$ the singular support of u is the smallest closed subset of $\mathbb{R}^n$ outside of which u is given by integration against an infinitely differentiable function. The following is Theorem 4.2.5 in [6].

Lemma A.6. *Let $\mathcal{S}, u \in \mathcal{D}'(\mathbb{R}^n)$ and suppose that at least one has compact support. Then*

$$\text{Singular support } (\mathcal{S} * u) \subset \text{Singular support } (\mathcal{S}) + \text{Singular support } (u).$$

Proof. Let U and V be open neighborhoods of Support($\mathcal{S}$) and Support(u). Let $\varphi, \psi \in \mathcal{C}^\infty(\mathbb{R}^n)$ such that Support(φ) $\subset U$, $\varphi(\mathbf{x}) \equiv 1$ in a neighborhood of Support($\mathcal{S}$), Support(ψ) $\subset V$, and $\psi(\mathbf{x}) \equiv 1$ in a neighborhood of Support(u). Put $\mathcal{S}_1 = \varphi\mathcal{S}$, $\mathcal{S}_2 = (1-\varphi)\mathcal{S}$, $u_1 = \psi u$, and $u_2 = (1-\psi)u$ so that $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2$ and $u = u_1 + u_2$. Note that $\mathcal{S}_2, u_2 \in \mathcal{C}^\infty(\mathbb{R}^n)$ and we can assume that at least one of them has compact support. Then

$$\mathcal{S} * u = \mathcal{S}_1 * u_1 + \mathcal{S}_1 * u_2 + \mathcal{S}_2 * u_1 + \mathcal{S}_2 * u_2.$$

Then the last three terms on the right hand side belong to $\mathcal{C}^\infty(\mathbb{R}^n)$ and so

$$\begin{aligned}\text{Singular support }(\mathcal{S} * u) &\subset \text{Singular support }(\mathcal{S}_1 * u_1)\\ &\subset \text{Support }(\mathcal{S}_1 * u_1)\\ &\subset \text{Support }(\mathcal{S}_1) + \text{Support }(u_1) \subset U + V.\end{aligned}$$

Since U and V are arbitrary open neighborhoods of the closed sets Singular support $(\mathcal{S})$ and Singular support (u), the Lemma follows. □

References

1. Lipman Bers, Fritz John, and Martin Schechter, *Partial differential equations*, American Mathematical Society, Providence, R.I., 1979, With supplements by Lars Gårding and A. N. Milgram, With a preface by A. S. Householder, Reprint of the 1964 original, Lectures in Applied Mathematics, 3A. MR 598466 (82c:35001)
2. Wei-Liang Chow, *Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung*, Math. Ann. **117** (1939), 98–105. MR 0001880 (1,313d)
3. Ronald R. Coifman and Guido Weiss, *Analyse harmonique non-commutative sur certains espaces homogènes*, Lecture Notes in Mathematics, Vol. 242, Springer-Verlag, Berlin-New York, 1971, Étude de certaines intégrales singulières. MR 0499948 (58 #17690)
4. David Gilbarg and Neil S. Trudinger, *Elliptic partial differential equations of second order*, Classics in Mathematics, Springer-Verlag, Berlin, 2001, Reprint of the 1998 edition. MR 1814364 (2001k:35004)
5. Lars Hörmander, *Hypoelliptic second order differential equations*, Acta Mathematica **119** (1967), 147–171.
6. ______, *The analysis of linear partial differential operators. I*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 256, Springer-Verlag, Berlin, 1983, Distribution theory and Fourier analysis. MR 717035 (85g:35002a)
7. J. J. Kohn, *Harmonic integrals on strongly pseudo-convex manifolds. I*, Ann. of Math. (2) **78** (1963), 112–148. MR 0153030 (27 #2999)
8. ______, *Harmonic integrals on strongly pseudo-convex manifolds. II*, Ann. of Math. (2) **79** (1964), 450–472. MR 0208200 (34 #8010)
9. Elias M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970. MR 0290095 (44 #7280)
10. ______, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993, With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III. MR 1232192 (95c:42002)
11. Francois Trèves, *Introduction to pseudodifferential and Fourier integral operators. Vol. 1*, Plenum Press, New York-London, 1980, Pseudodifferential operators, The University Series in Mathematics. MR 597144 (82i:35173)