Modern Harmonic Analysis: Singular Integrals, Maximal Functions and Littlewood-Paley theory

Bernat Ramis Vich

Supervised by James R. Wright (UoE)

Cosupervised by Albert Mas Blesa (UPC)

Final Bachelor's Thesis

Bachelor's degree in Mathematics

Bachelor's degree in Engineering Physics

May, 2023

Als Matfis, en especial a en Marc.

Abstract

The tools developed in the 1950s by Calderón and Zygmund, which led to the birth of modern harmonic analysis, are studied. Some of these concepts are the Hardy-Littlewood maximal function together with its L^p estimates, interpolation theorems and, foremost, the Calderón-Zygmund decomposition. These techniques allow us to show that some singular integrals are well defined and bounded on L^p spaces. Although Euclidean space is the original setting where these ideas were developed, a main aim of the project is to understand how these estimates generalise to other measure metric spaces and to vector-valued singular integrals.

Despite being powerful, classical Calderón-Zygmund theory has its limitations. For example, the spherical maximal operator introduced by Stein in the 1970's falls outside the scope of the original theory. However, one can utilise Littlewood-Paley theory via square function estimates to prove optimal estimates for the spherical maximal operator. Nevertheless, endpoint bounds for similar singular integral and maximal operators remain as open problems.

Resum

S'estudien les eines desenvolupades en la dècada de 1950 per Calderón i Zygmund, les quals varen portar al naixement de l'anàlisi harmònica moderna. Alguns d'aquests conceptes són la funció maximal de Hardy-Littlewood juntament amb les seves propietats d'acotació en espais L^p, teoremes d'interpolació i, sobretot, la descomposició de Calderón i Zygmund. Aquestes tècniques ens permeten demostrar que algunes integrals singulars estan ben definides i fitades en els espais L^p. Tot i que l'espai euclidià fos el context original on totes aquestes idees es varen desenvolupar, un objectiu principal del projecte és entendre com aquestes propietats es generalitzen a altres espais mètrics de mesura i a integrals singulars de valors vectorials.

Tot i ser potent, la teoria clàssica de Calderón-Zygmund té les seves limitacions. Per exemple, l'operador maximal esfèric introduït per Stein en la dècada de 1970 cau fora de l'abast de la teoria original. Tot i així, un pot utilitzar la teoria de Littlewood-Paley via desigualtats de *square functions* per demostrar desigualtats òptimes per l'operador maximal esfèric. Per contra, hi ha fites en els extrems del rang de valors dels exponents p per a operadors integrals singulars i maximals que romanen com a problemes oberts.

Resumen

Se estudian las herramientas desarrolladas en la década de 1950 por Calderón y Zygmund, las cuales llevaron al nacimiento del análisis armónico moderno. Algunos de estos conceptos son la función maximal de Hardy-Littlewood junto con sus propiedades de acotación en espacios L^p, teoremas de interpolación y, sobre todo, la descomposición de Calderón y Zygmund. Estas técnicas nos permiten demostrar que algunas integrales singulares están bien definidas y acotadas en los espacios L^p. Aunque el espacio euclídeo sea el contexto original donde todas estas ideas fueron desarrolladas, un objetivo principal del proyecto es entender cómo estas propiedades se generalizan a otros espacios métricos de medida y a integrales singulares con valores vectoriales.

A pesar de ser potente, la teoría clásica de Calderón-Zygmund tiene sus limitaciones. Por ejemplo, el operador maximal esférico introducido por Stein en la década de 1970 cae fuera del alcance de la teoría original. Sin embargo, uno puede utilizar la teoría de Littlewood-Paley via desigualdades de *square functions* para demostrar desigualdades óptimas para el operador maximal esférico. Aun así, hay cotas en los extremos del rango de valores de los exponentes p para operadores integrales singulares y maximales similares que permanecen como problemas abiertos.

Keywords: Singular integrals, Calderón−Zygmund, maximal function, Littlewood−Paley

Paraules clau: Integrals singulars, Calderón−Zygmund, funció maximal, Littlewood−Paley

Palabras clave: Integrales singulares, Calderón−Zygmund, función maximal, Littlewood−Paley

AMS: 43-01

Acknowledgements

First of all, I thank my home centre CFIS for the valuable education I received, as well as for keeping track of my progress. On top of that, I thank CFIS and Fundació Privada Mir-Puig for making my mobility experience possible. I have been feeling like the whole programme was tailored to me.

In the other end, I thank the Mathematical Analysis group of the University of Edinburgh for welcoming me and taking me into account for the weekly events they organise: seminars, working groups and pub and dinner outings. Both researchers and PhD students taught me a lot and made it easy to build friendship.

In particular, I thank my supervisor Jim Wright for his outstanding labour and for the pile of hours he spent with me. I take his enthusiasm, kindness and modesty as an inspiring example. In parallel, I thank my cosupervisor Albert Mas for his attentive tracking from UPC in Barcelona and for recommending me Jim as a mentor.

On the one hand, I appreciate having shared a flat with Giorgio, who flavoured my routine with humour, and his personal advice. On the other hand, I am grateful that our landlord Chris made his flawless home ours for the academic year.

Per acabar, vull agrair a ma mare, mon pare, la meva germana i els meus padrins el seu suport al llarg de tots aquests anys i la resiliència que suposa estar lluny els uns dels altres la majoria del temps. Gracis!

Contents

CHAPTER 1
Introduction and preliminaries

1.1. Introduction to the topic

By "singular integral operators" we mean, in the first instance, convolution operators in $\mathbb{R}^n$ the kernel function of which presents a singularity, say, at the origin. Singular integrals show up in a number of problems of analytic nature. For instance, they generate solutions of some partial differential equations, either elliptic or hyperbolic; they arise in complex analysis; they underpin apparently unrelated settings in geometric measure theory, etc.

For decades, analysts felt uncomfortable when utilising such tools because there was no knowledge regarding their boundedness properties. Were they handling continuous operators on L^p spaces or not?

Harmonic analysis is the natural framework for studying singular integral operators. In the middle and end 20^{th} century, the field experienced a burst. Brilliant mathematicians contributed to the expansion of the theory concerning singular integrals. Calderón, Zygmund, Littlewood, Paley, Hardy, Bourgain and Stein are just some of the most influential driving forces in the field.

Due to its ubiquity, in the literature, theory of singular integrals is often just partially explained, because it serves to step forward at stages within problems of different natures. Therefore, this document intends to collect most of the theory of singular integrals, gathering altogether all of its pieces, putting them into context and depicting their most emblematic applications. This is, instead of regarding it an auxiliary tool, we centre them in the spotlight.

1.2. Essential initial tools and results

Throughout the chapters, the reader is going to come across a series of recurrent tools and topics that are succinctly explained in this section. It is necessary to bear them in mind.

1.2.1. The Fourier transform

The single construction of the Fourier transform on $L^p(\mathbb{R}^n)$ spaces would occupy a whole chapter, therefore we will restrict ourselves to highlighting the most relevant results for our concern, besides leaving the reference [13], Chapter 1, Section 1 and 2.

To set the convention:

Definition 1.1. *The* ***Fourier transform*** *acting on functions in $L^1(\mathbb{R}^n)$ is defined as*

$$\hat{f}(\xi) := \int_{\mathbb{R}^n} f(x)e^{-2\pi i x\xi}dx, \quad \xi \in \mathbb{R}^n.$$

Accordingly,

Definition 1.2. *The* ***inverse Fourier transform*** *acting on functions in $L^1(\mathbb{R}^n)$ is defined as*

$$\check{g}(x) := \int_{\mathbb{R}^n} g(\xi)e^{2\pi i x\xi}d\xi, \quad x \in \mathbb{R}^n.$$

It is well known that the Fourier transform operator admits a unique extension from $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ to all functions on $L^2(\mathbb{R}^n)$.

Theorem 1.3 (Plancherel's Theorem)**.** *The Fourier transform acting on functions in $L^2(\mathbb{R}^n)$ defines an isometric isomorphism from $L^2(\mathbb{R}^n)$ to itself. That is, equally denoting by $\hat{f}$ the Fourier transform of $f \in L^2(\mathbb{R}^n)$,*

$$\left\| \hat{f} \right\|_2 = \| f \|_2 .$$

1.2.2. Weak Lebesgue spaces

When L^p spaces are not large enough, the following is a candidate for a substitute.

Definition 1.4. *Let $X \equiv (X, \Sigma, \mu)$ be a measure space. Inspired by the fact*

$$\lambda^p \mu(\{x \in X : |f(x)| > \lambda\}) \leq \int_{\{x \in X: |f(x)|>\lambda\}} |f(x)|^p d\mu(x) \leq \int_X |f(x)|^p d\mu(x) = \| f \|_p^p$$

*for $f \in L^p(X)$ and $\lambda > 0$, we define the **weak L^p spaces**, with $1 \leq p < \infty$ as*

$$L^{p,\infty}(X)^{[1]} := \{f : X \to \mathbb{C} \text{ measurable } : \sup_{\lambda>0} \lambda^p \mu(\{x \in X : |f(x)| > \lambda\}) < \infty\}. \tag{1.1}$$

It is clear that $L^p(X) \subseteq L^{p,\infty}(X)$[2]. The fact that these spaces broaden the conventional $L^p(X)$ is the reason why they are called "weak L^p spaces".

It turns out that

$$\| f \|_{L^{p,\infty}(X)} = \sup_{\lambda>0} \lambda \mu(\{x \in X : |f(x)| > \lambda\})^{\frac{1}{p}}$$

is not a proper norm, since it does not verify the triangle inequality. However, for $1 < p < \infty$, it is comparable to some other actual norm, which turns $L^{p,\infty}(X)$ into a Banach space.

It is conventional to set $L^{\infty,\infty}(X) := L^\infty(X)$.

1.2.3. Convolution in Lebesgue spaces

Let us recall how convolution behaves in $L^p(X)$ spaces.

Proposition 1.5. *Let X be a measure space.*

(a) If $f, g \in L^1(X)$ then $f * g \in L^1(X)$ and

$$\| f * g \|_1 = \| f \|_1 \| g \|_1 .$$

(b) If $f \in L^p(X)$ for $1 \leq p \leq \infty$ and $g \in L^1(X)$ then $f * g \in L^p(X)$ and

$$\| f * g \|_p \leq \| f \|_p \| g \|_1 .$$

(c) If $f \in L^p(X)$ for $1 \leq p \leq \infty$, and $g \in L^{p'}(X)$, where p and p' are conjugate exponents $\frac{1}{p} + \frac{1}{p'} = 1$, then $f * g \in L^\infty(X)$ and

$$\| f * g \|_\infty \leq \| f \|_p \| g \|_{p'} .$$

The proof is straightforward: by Fubini–Tonelli theorem for (a), by Minkowski integral inequality for (b) and by Hölder inequality for (c).

[1] The notation for the weak Lebesgue spaces comes from the fact that they are a particular case of the Lorentz spaces $L^{p,q}(X)$.

[2] Moreover, in $\mathbb{R}^n$, the inclusion is strict: take $f(x) = |x|^{-\frac{n}{p}}$. It is easy to check that $f \in L^{p,\infty}(\mathbb{R}^n)$ but $f \notin L^p(\mathbb{R}^n)$.

1.2.4. Approximations to the identity

This one is a versatile tool intervening in density arguments or convergence proofs.

Definition 1.6. *Let $K \in L^1(\mathbb{R}^n)$ and let $K_r(x) := r^{-n}K(r^{-1}x)$ be their integral-preserving dilates. The family $(K_r)_{r>0}$ is an* ***approximation to the identity*** *if it verifies:*

(a) $K(x) \geq 0 \;\; \forall\, x \in \mathbb{R}^n$,

(b) $\int K = 1$.

Remark 1.7. *It follows from the previous definition that, under the same hypotheses, $\forall\, \delta > 0$*

$$\int_{|x|>\delta} K_r(x)dx \to 0 \qquad \textit{if } r \to 0.$$

1.2.5. Schwartz functions and tempered distributions

When one seeks for a space of functions which makes the Fourier transform behave nicely, one of the outstanding candidates is the space of Schwartz functions. Its definition is inspired by the two following properties.

First, introduce the multi-index notation $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ on derivatives of multivariable functions $\partial^\alpha f(x) := \partial_{x_1}^{\alpha_1} \ldots \partial_{x_n}^{\alpha_n} f(x)$ and on monomials $x^\alpha := x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}$.

(a) $\partial^\alpha \hat{f}(\xi) = ((-2\pi i x)^\alpha f(x))\hat{}(\xi)$

(b) $\widehat{\partial^\alpha f} = (2\pi i \xi)^\alpha \hat{f}(\xi)$

Thus, we need the operations of deriving and multiplying by polynomials to be well-behaved with our functions.

Definition 1.8. *The space of* ***Schwartz functions*** *$\mathscr{S}(\mathbb{R}^n)$ is the space of all smooth functions $\varphi \in C^\infty(\mathbb{R}^n)$ such that*

$$\|\, \varphi \,\|_{\alpha,\beta} := \sup_{x\in\mathbb{R}^n} |x^\alpha \partial^\beta \varphi(x)| < \infty \tag{1.2}$$

for all multi-indexes $\alpha, \beta \in \mathbb{N}^n$.

Moreover, each $\|\cdot\|_{\alpha,\beta}$ turns out to be a seminorm on $\mathscr{S}(\mathbb{R}^n)$. Combining all of them in a certain way, one can construct a metric for $\mathscr{S}(\mathbb{R}^n)$, giving it the nature of a metric vector space, see [13], Chapter 1, Section 3.

Some of the graceful properties of the space of Schwartz functions $\mathscr{S}(\mathbb{R}^n)$ are:

(a) It is a complete metric space.

(b) The Fourier transform is a homeomorphism from $\mathscr{S}(\mathbb{R}^n)$ to itself.

(c) It is separable.

(d) $\mathscr{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$ for $1 \leq p \leq \infty$ and $\mathscr{S}(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for $1 \leq p < \infty$.

Again, see [13], Chapter 1, Section 3 for futher details.

Equally important is the dual space of $\mathscr{S}(\mathbb{R}^n)$.

Definition 1.9. *The topological dual space of $\mathscr{S}(\mathbb{R}^n)$, $\mathscr{S}^*(\mathbb{R}^n)$ is called the space of* ***tempered distributions****.*

Although tempered distributions arise as a vast topic, let us just state the results that are worth mentioning here. Check [13] as before for expanded explanations.

Theorem 1.10. *Let u be a linear functional on $\mathscr{S}(\mathbb{R}^n)$. u is a tempered distribution if and only if there exist $C > 0$, $N, M \in \mathbb{N}$ such that*

$$|u(\varphi)| \leq C \sum_{\substack{\alpha \leq N \\ \beta \leq M}} \| \varphi \|_{\alpha,\beta} \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^n).$$

The proof can be found in [13], Theorem 3.11 in Chapter 1. This theorem is practical for checking whether a linear functional is in fact a tempered distribution. For example, one can check this way that we can define an inclusion map $\mathscr{S}(\mathbb{R}^n) \hookrightarrow \mathscr{S}^*(\mathbb{R}^n)$ by assigning

$$\varphi \to \int_{\mathbb{R}^n} \varphi(x)(\cdot)\, dx.$$

Definition 1.11. *We say that an equality of tempered distributions $u, v \in \mathscr{S}^*(\mathbb{R}^n)$ holds* ***in the sense of tempered distributions*** *$u = v$ if*

$$u(\varphi) = v(\varphi) \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^n).$$

One of the operations one can define involving tempered distributions is the convolution of a tempered distribution with a Schwartz function. In particular, assume $u, \varphi, \phi \in \mathscr{S}$. Then, by Fubini-Tonelli, it holds that

$$\int_{\mathbb{R}^n} (u * \varphi)(x)\phi(x)\, dx = \int_{\mathbb{R}^n} u(x)(\tilde{\varphi} * \phi)(x)\, dx$$

where $\tilde{\varphi}(x) := \varphi(-x)$ is the reflected function. The following definition is based on this equality to generalise the operation of convolution.

Definition 1.12. *Let $u \in \mathscr{S}^*(\mathbb{R}^n)$ and $\varphi \in \mathscr{S}(\mathbb{R}^n)$. Define their* ***convolution*** *as the function*

$$(u * \varphi)(x) := u(\tau_x \tilde{\varphi})$$

where $\tau_x f(y) := f(y - x)$ is the translation operator.

Theorem 1.13. *The function f resulting from the convolution of $u \in \mathscr{S}^*(\mathbb{R}^n)$ and $\varphi \in \mathscr{S}(\mathbb{R}^n)$, $f(x) = (u * \varphi)(x)$ is a smooth function with growth and growth of its derivatives at most polynomial.*

The proof is available in [13], Theorem 3.13 in Chapter 1.

It is also possible to extend the definition of the Fourier transform on tempered distributions, in similarity with the multiplication formula.

Definition 1.14. *The* ***Fourier transform of a tempered distribution*** *$u \in \mathscr{S}^*(\mathbb{R}^n)$ is defined for tempered distributions by their action as*

$$\hat{u}(\varphi) := u(\hat{\varphi}) \quad \forall\, \varphi \in \mathscr{S}(\mathbb{R}^n).$$

Note the Fourier transform of a tempered distribution is a tempered distribution. In fact, Definition 1.14 allows us to talk about the Fourier transform of $L^p(\mathbb{R}^n)$ functions with $p > 2$, by viewing these functions as tempered distributions.

We refer to [13], Chapter 1, Section 3, for the following result.

Proposition 1.15. *The Fourier transform on the space of tempered distributions defines an isomorphism of topological vector spaces from $\mathscr{S}^*(\mathbb{R}^n)$ to itself.*

1.2.6. Marcinkiewicz interpolation theorem

Sometimes, one wishes to show that an operator is bounded on $L^p(X)$ spaces for a whole range of values of p. Instead of working out the proof in the generality $f \in L^p(X)$ for any p of the desired range, there are some available interpolation theorems which allow us to conclude boundedness properties of our operator by dealing only with the endpoint values of p.

Theorem 1.16 (Marcinkiewicz interpolation theorem)**.** *Let $X_i \equiv (X_i, \Sigma_i, \mu_i)$, $i = 1, 2$ be two σ-finite measure spaces and let T be a sublinear operator mapping from $L^{p_1}(X_1) + L^{p_2}(X_1)$ to the space of measurable functions over X_2, for fixed $1 \le p_1 < p_2 \le \infty$, such that:*

(i) T is weak-type (p_1, p_1). That is, $\forall\, \lambda > 0$ and $f \in L^{p_1}(X_1)$,

$$\mu_2(\{x \in X_2 : |Tf(x)| > \lambda\}) \le \left(C_1 \frac{\| f \|_{p_1}}{\lambda} \right)^{p_1}$$

for some constant $C_1 > 0$.

(ii) T is weak-type (p_2, p_2). This means, $\forall\, \lambda > 0$ and $f \in L^{p_2}(X_1)$,

$$\mu_2(\{x \in X_2 : |Tf(x)| > \lambda\}) \leq \left(C_2 \frac{\| f \|_{p_2}}{\lambda} \right)^{p_2}$$

for some constant $C_2 > 0$.

Then, T is strong-type (r, r) $\forall\, r$ in the range $p_1 < r < p_2$ and it holds that

$$\| Tf \|_r \leq A_{p_1,p_2} \| f \|_r$$

for $f \in L^r(X_1)$ and A_{p_1,p_2} only depending on p_1, p_2, C_1 and C_2.

The proof in the case $X_1 = X_2 = \mathbb{R}^n$ and $p_1 = 1$ is available in [10], Chapter 1, Section 4. The generalization of the proof to the setting of Theorem 1.16 is straightforward though.

Indeed, Marcinkiewicz interpolation theorem is the main reason why we are interested in the weak L^p spaces.

1.2.7. The Hardy-Littlewood maximal function

Apparently, as one can read from their original paper [7], G. H. Hardy and J. E. Littlewood got inspired by cricket to define their maximal function. Both eager followers of this sport wanted to understand how to maximise a player's "satisfaction" after several innings based on the averages of their scores. Beyond its origin, the Hardy-Littlewood maximal function comes in extremely handy for the study of singular integrals, and will be the key to extend the precious Calderón-Zygmund decomposition to a broad class of measure metric spaces.

Definition 1.17. *Let $f \in L^1_{loc}(\mathbb{R}^n)$ be a locally integrable function. The* ***centred Hardy-Littlewood maximal function*** *of f is defined as*

$$\mathfrak{M}f(x) := \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)| dy. \tag{1.3}$$

The image of $\mathfrak{M}$ is not clear to determine from the early definition. This is going to be the issue of the corresponding theorem, Theorem 1.22. In fact, we are often going to restrict the domain of $\mathfrak{M}$ to smaller spaces such as $f \in L^p(\mathbb{R}^n) \subset L^1_{\text{loc}}(\mathbb{R}^n)$. In the particular case of $f \in L^1(\mathbb{R}^n)$, observe that the behaviour of the operator $\mathfrak{M}$ is local, in the sense that the supremum over r will be approached by small values of r, thus smaller balls.

Remark 1.18. *It is possible to write the Hardy-Littlewood maximal function in terms of a convolution operator. Define the convolution kernel*

$$b_r(x) := \frac{1}{|B_r(0)|} \mathbb{1}_{B_r(0)}(x), \tag{1.4}$$

and write

$$\mathfrak{M}f(x) = \sup_{r>0} (b_r * f)(x). \tag{1.5}$$

In a similar way, we define a sibling of (1.3).

Definition 1.19. *Let $f \in L^1_{loc}(\mathbb{R}^n)$ be a locally integrable function. The* ***uncentred Hardy-Littlewood maximal function*** *of f is defined as*

$$\mathfrak{M}^{unc} f(x) := \sup_{B \ni x} \frac{1}{|B|} \int_B |f(y)| dy, \tag{1.6}$$

where this time the supremum is take over all balls B that contain x.

Remark 1.20. *Clearly, $\mathfrak{M}f \leq \mathfrak{M}^{unc} f$. Furthermore, the reverse inequality holds up to a multiplicative constant.*

Given $x \in \mathbb{R}^n$ and a ball $B \ni x$ of radius δ, consider a second ball, centred at x with radius 2δ, thus $B \subseteq B_{2\delta}(x)$. Then,

$$\frac{1}{|B|} \int_B |f(y)| dy \leq \frac{2^n}{|B_{2\delta}(x)|} \int_{B_{2\delta}(x)} |f(y)| dy \leq 2^n \, \mathfrak{M}f(x).$$

Now taking the supremum over any ball B containing x, we get $\mathfrak{M}^{\text{unc}} \leq 2^n \mathfrak{M}$. Note that this has been possible thanks to the following property of balls in $\mathbb{R}^n$: $|B_{2r}(x)| = 2^n |B_r(x)|$, $\forall\, x \in \mathbb{R}^n,\ r > 0$. We are baring it in mind, since we will need a similar condition at the time of swapping the setting for a generic measure metric space.

Remark 1.21. *The Hardy-Littlewood maximal function is measurable, so it makes sense to compute its L^p norms: It is easy to check that the averages over balls $A(x,r) := \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)| dy$ are continuous functions of x, hence measurable functions of x. Likewise, $A(x,r)$ are continuous functions of r, therefore taking the supremum of the averages over $r > 0$ and over $r \in \mathbb{Q}_{>0}$ yields the same result. That the supremum of a countable collection of measurable functions is a measurable function is a basic fact from measure theory.*

One of the reasons why the Hardy-Littlewood maximal function turns out so useful in the theory of singular integrals is the following theorem. It shows the behaviour of such operator on $L^p(\mathbb{R}^n)$ spaces. Consequently, if one proves a bound for a given operator T by the Hardy-Littlewood maximal function, then boundedness of T is concluded.

Theorem 1.22 (Maximal theorem in $\mathbb{R}^n$)**.** *Let f be a measurable complex-valued function on $\mathbb{R}^n$. Then:*

(a) If $f \in L^p(\mathbb{R}^n)$ *for* $1 \le p \le \infty$, $\mathfrak{M}f(x)$ *is finite a.e.* $x \in \mathbb{R}^n$.

(b) For every $\lambda > 0$ *and* $f \in L^1(\mathbb{R}^n)$,

$$\lambda|\{x \in \mathbb{R}^n : \mathfrak{M}f(x) > \lambda\}| \le A \, \| f \|_1 \, , \tag{1.7}$$

where A *only depends on the dimension* n.

(c) If $f \in L^p(\mathbb{R}^n)$, $1 < p \le \infty$, *then* $\mathfrak{M}f \in L^p(\mathbb{R}^n)$ *and*

$$\| \mathfrak{M}f \|_p \le A_p \, \| f \|_p \, , \tag{1.8}$$

where A_p *only depends on the dimension* n *and the exponent* p.

Estimates of this Chebyshev kind as in (b) are called weak-type $(1, 1)$ estimates. We shall content ourselves with the maximal function being of weak-type (1,1) and not type (1,1). Drastically, if $f \in L^1(\mathbb{R}^n)$ is not identically 0 a.e., the corresponding maximal function is *never* in $L^1(\mathbb{R}^n)$.

Remark 1.23. *If* $f \in L^1(\mathbb{R}^n)$ *is not identically* 0 *a.e., then* $\mathfrak{M}f \notin L^1(\mathbb{R}^n)$.

Proof. Start by assuming that $f \in L^1(\mathbb{R}^n)$ has compact support. Let $B_c(0)$ be a ball of radius c that fully contains the support of f, $\text{supp}(f) \subseteq B_c(0)$. For each $x \notin B_c(0)$, we have

$$\begin{aligned}\mathfrak{M}f(x) \ge \frac{1}{|B_{2|x|}(x)|} \int_{B_{2|x|}(x)} |f(y)|dy &= \frac{1}{|B_{2|x|}(x)|} \int_{B_c(0)} |f(y)|dy \\ &= \frac{1}{2^n|B_1(0)|} \int_{B_c(0)} |f(y)|dy \frac{1}{|x|^n} \equiv \frac{C(f)}{|x|^n}\end{aligned}$$

because $B_c(0) \subseteq B_{2|x|}(x)$. This shows $\mathfrak{M}f(x)$ is not integrable.

For the general case when $f \in L^1(\mathbb{R}^n)$ does not have compact support, restrict f to a compact set K such that $f|_K$ is not identically 0 a.e. in K. Hence, $|f| \ge | \, f|_K \, |$ which implies $\mathfrak{M}(f) \ge \mathfrak{M}(f|_K)$, so that the previous argument applies. ☐

Since Theorem 1.22 is key in the results we will be using to bound singular integrals, let us provide the proof, taken from [10], Chapter 1, Section 1.3.

The proof relies on a covering lemma.

Lemma 1.24 (Vitali-type covering lemma in $\mathbb{R}^n$)**.** *Let* E *be a measurable set in* $\mathbb{R}^n$ *covered by the union of a family of balls of finite diameter* $\{B_k\}_{k\in K}$. *Then, from this family we can substract a countable sequence of disjoint balls* $\{B'_k\}_{k\in\mathbb{N}} \subseteq \{B_k\}_{k\in K}$ *such that*

$$\sum_{k\in\mathbb{N}} |B'_k| \ge C|E|.$$

Here, $C > 0$ *is a constant that can be taken to be* $C = 5^{-n}$.

Proof. (*Of Theorem 1.22*) (a) follows from (b) and (c) because functions in $L^p(\mathbb{R}^n)$ are finite a.e. Since the Hardy-Littlewood maximal function is trivially bounded on $L^\infty(\mathbb{R}^n)$, if we show that it is a weak-type (1,1) operator (that is, statement (b)) then by Marcinkiewicz interpolation theorem (Theorem 1.16), (c) follows.

So, let $\lambda > 0$, $f \in L^1(\mathbb{R}^n)$ and consider $E := \{x \in \mathbb{R}^n : \mathfrak{M}f(x) > \lambda\}$. For each $x \in E$, there exists a ball centered at x, B_x such that $\int_{B_x} |f(y)|dy > \lambda |B_x|$.

We have that the family $\{B_x\}_{x\in E}$ covers E, thus by Lemma 1.24 there exists a countable subset of disjoint balls $\{B'_k\}_{k\in\mathbb{N}} \subseteq \{B_k\}_{k\in K}$ such that

$$\sum_{k\in\mathbb{N}} |B'_k| \geq C|E|,$$

reaching

$$\lambda|E| \leq C^{-1}\lambda \sum_{k\in\mathbb{N}} |B'_k| \leq C^{-1} \| f \|_1 .$$

□

The latter theorem yields the celebrated Lebesgue differentiation theorem.

Corollary 1.25 (Lebesgue differentiation theorem in $\mathbb{R}^n$)**.** *Whenever* $f \in L^1_{loc}(\mathbb{R}^n)$,

$$\lim_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} f(y)dy = f(x) \qquad a.e. \ x \in \mathbb{R}^n. \tag{1.9}$$

The shocking aspect of the Lebesgue differentiation theorem is that we do not even require regularity in a locally integrable function for its local averages around a point to tend to the value on the point, except for a set of measure zero.

The remaining proofs of Lemma 1.24 and Corollary 1.25 are available in [10], Chapter 1, Section 1.3.

1.2.8. Layer cake representation

Next, a lemma that helps figure out some computations.

Lemma 1.26 (Layer cake representation)**.** *For any* $1 \leq p < \infty$ *and* $f \in L^p(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} |f(x)|^p dx = p \int_0^\infty \alpha^{p-1} |\{x \in \mathbb{R}^n : |f(x)| > \alpha\}| \, d\alpha.$$

Proof. Start with the right hand side and $p = 1$.

$$\int_0^\infty |\{x \in \mathbb{R}^n : |f(x)| > \alpha\}| d\alpha = \int_0^\infty \int_{\mathbb{R}^n} \mathbb{1}_{\{|f(x)|>\alpha\}}(x) dx \, d\alpha$$

Since the integrand is positive, we are entitled to apply Fubini-Tonelli theorem.

$$\int_0^\infty \int_{\mathbb{R}^n} \mathbb{1}_{\{|f(x)|>\alpha\}}(x)dx \; d\alpha = \int_{\mathbb{R}^n} \int_0^\infty \mathbb{1}_{\{|f(x)|>\alpha\}}(x)d\alpha \; dx = \int_{\mathbb{R}^n} |f(x)|dx$$

This way, we prove the case $p = 1$. The remaining cases for p unlock by considering a general $f \in L^p(\mathbb{R}^n)$ and $g(x) = |f(x)|^p$, with $g \in L^1(\mathbb{R}^n)$ and applying the formula for the case $p = 1$ to g. □

1.3. Motivation for studying singular integrals

Singular integrals arise in a myriad of problems and settings both in mathematics and physics. Therefore, mathematics needed a rigorous theory for such operators. In this section, we present a rather simple yet interesting problem the solution of which demands knowledge on our central topic.

1.3.1. Dirichlet's problem for the Laplace equation on the upper half plane

Let $\mathbb{R}^+ = \{(x, y) \in \mathbb{R}^2 : y > 0\} \subset \mathbb{R}^2$ be the upper half plane. Consider the following Dirichlet problem:

$$\begin{cases} \Delta u(x, y) = 0 \,, \quad (x, y) \in \mathbb{R}^+ \\ \\ \lim\limits_{y\to 0} u(x, y) = f(x) \end{cases} \tag{1.10}$$

Here, f is a real-valued boundary function. Let its ambient function space be $L^2(\mathbb{R})$ for now. In concordance, let the limit for the boundary condition be understood in the $L^2(\mathbb{R})$ sense.

Consider

$$u(x, y) = \int_{\mathbb{R}} \hat{f}(t)e^{2\pi ixt}e^{-2\pi|t|y}dt. \tag{1.11}$$

This is an absolutely convergent integral by Cauchy-Schwarz, since by Plancherel theorem $\hat{f} \in L^2(\mathbb{R})$, and $e^{-2\pi|\cdot|y} \in L^2(\mathbb{R})$. We are also allowed to differentiate $u(x, y)$ under the integral sign with respect to either variable thanks to the rapid decay of the real exponential function. Indeed,

$$\Delta u(x, y) = \int_{\mathbb{R}} \hat{f}(t)(2\pi it)^2 e^{2\pi ixt}e^{-2\pi|t|y}dt + \int_{\mathbb{R}} \hat{f}(t)(-2\pi|t|)^2 e^{2\pi ixt}e^{-2\pi|t|y}dt = 0$$

thus u is a harmonic function. For the boundary value, using Plancherel theorem once again, we have

$$\| u(\cdot, y) - f(\cdot) \|_2 = \left\| \hat{f}(\cdot)e^{-2\pi|\cdot|y} - \hat{f}(\cdot) \right\|_2 \to 0 \qquad \text{when } y \to 0$$

by the dominated convergence theorem. This solves problem (1.10) in the setting of $L^2(\mathbb{R})$. Notice that not only does the solution (1.11) tend to the initial datum in the L^2 sense but also it is regular $C^2(\mathbb{R}^+)$.

A standard argument for uniqueness of classical solution reads as follows. Consider a conformal map from the unit complex disk $\mathbb{D}$ to the upper half complex plane $\mathbb{H}$ (identified with $\mathbb{R}^+$) like $F : \mathbb{D} \to \mathbb{H}$

$$F(z) = i\frac{1-z}{1+z}.$$

In particular, F maps the unit circle $\mathbb{S}^1 \subset \mathbb{C}$ to the real line $\{(x, y) \in \mathbb{R}^2 : y = 0\}$, except for the point $z = -1$ that "is mapped to infinity".

Complex variable functions theory grants that if u is harmonic and F is analytic, then $u \circ F$ is harmonic. But now, $u \circ F$ is a harmonic C^2 function on the open unit disk. Therefore, by the maximum principle of harmonic functions on a bounded regular domain, if two solutions of the Laplace equation agree on the boundary, then they are the same solution. Since $u \circ F$ solves uniquely Dirichlet's problem on the unit disk, and F is invertible

$$F^{-1}(z) = \frac{i-z}{i+z},$$

the solution on $\mathbb{R}^+$, u, is the unique solution that vanishes at infinity. All in all, this means that (1.11) is *the* classical solution of (1.10) that tends to 0 at infinity.

In view of solution (1.11):

Definition 1.27. *For $x \in \mathbb{R}^n$ and $y > 0$, call*

$$P_y(x) = \int_{\mathbb{R}^n} e^{2\pi i x t} e^{-2\pi|t|y} dt \tag{1.12}$$

*the **Poisson kernel** in $\mathbb{R}^n$.*

One may write the solution of (1.10) as the convolution of the boundary value function with the Poisson kernel in $\mathbb{R}$:

$$u(x, y) = (P_y * f)(x).$$

Furthermore, a computation (see Chapter 3, Section 2.1, Proposition 5 in [10]) shows that the Poisson kernel (in $\mathbb{R}^n$) can be rewritten as follows.

Proposition 1.28.

$$P_y(x) = c_n \frac{y}{(|x|^2 + y^2)^{\frac{n+1}{2}}}, \qquad c_n = \frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}} \tag{1.13}$$

From (1.13), it is straightforward to observe that the Poisson kernel is a well-defined integrable function $P_y \in L^1(\mathbb{R}^n)$ for all $y > 0$, meaning that, for $f \in L^p(\mathbb{R}^n)$, $1 \leq p \leq \infty$, the Poisson integral $u(x, y) = (P_y * f)(x)$ belongs to $L^p(\mathbb{R}^n)$ as a function of x (as remarked in the preliminars, Proposition 1.5). What is more, $P_y \in L^p(\mathbb{R}^n)$ for all $y > 0$ and $1 \leq p \leq \infty$. Besides this, it is not hard to see that (1.13) defines an approximation to the identity in the sense of Definition 1.6.1

Once a harmonic real-valued function on the plane $u(x, y)$ is obtained, a licit query is finding the harmonic conjugate function $v(x, y)$, in the framework of complex variable functions theory. This harmonic conjugate is given by the solution of the respective Cauchy-Riemann equations:

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial y} &= -\frac{\partial v}{\partial x}. \end{aligned} \tag{1.14}$$

Notice that

$$u(x, y) = \int_{\mathbb{R}} P_y(x - t) f(t) dt$$

behaves nicely in order to differentiate under integral sign, both with respect to x and y, thanks to the rapid decay of the Poisson kernel (1.13), so in essence we are interested in checking (1.14) for the Poisson kernel P_y and its harmonic conjugate kernel, say Q_y, both thought of as functions of two variables $P(x, y) := P_y(x)$ and $Q(x, y) := Q_y(x)$ for $(x, y) \in \mathbb{R}^+$. From the first Cauchy-Riemann equation:

$$\begin{aligned} \frac{\partial P(x, y)}{\partial x} &= -c_1 \frac{2xy}{(x^2 + y^2)^2} \\ Q(x, y) &= \int \frac{\partial P(x, y)}{\partial x}(x, y) dy + g(x) = c_1 \frac{x}{x^2 + y^2} + g(x) \end{aligned}$$

for a certain function g depending only on x. Imposing now the second Cauchy-Riemann equation (1.14),

$$\begin{aligned} -\frac{\partial Q(x, y)}{\partial x} &= -\left(c_1 \frac{x^2 + y^2 - 2x^2}{(x^2 + y^2)^2} + \frac{\partial g(x)}{\partial x} \right) = c_1 \frac{x^2 + y^2 - 2y^2}{(x^2 + y^2)^2} \implies \\ g(x) &\equiv 0. \end{aligned}$$

So we are left with

$$Q_y(x) = Q(x,y) = c_1 \frac{x}{x^2+y^2} = P(y,x) = P_x(y), \tag{1.15}$$

which solves the Cauchy-Riemann equations, thus $v(x,y) = (Q_y * f)(x)$ is the harmonic conjugate of u (up to an additive constant that we set to 0 to avoid integrability issues in what follows). Note that v is a bounded function of x for every $y > 0$ by Cauchy-Schwarz inequality[3].

The function u was generated by the boundary function f. This is not the case of v, obtained as the harmonic conjugate of u. This construction leads us to wondering about the boundary limit of v (see Figure 1.1). First, take the pointwise limit of (1.15).

$$\lim_{y\to 0} Q_y(x,y) = \lim_{y\to 0} c_1 \frac{x}{x^2+y^2} = \frac{1}{\pi x} \quad \text{if } x \neq 0 \tag{1.16}$$

Formally, (1.16) is the kernel of the Hilbert transform! We are going to make precise the meaning of

$$\text{“}Hf(x) = \lim_{y\to 0} v(x,y) = \lim_{y\to 0}(Q_y * f)(x)\text{”}$$

and uncover its properties later on. To be fair, the definition of this transform requires taking into account the subtlety of the singularity it presents at the origin, not to mention the lack of integrability. Here is the first time we encounter a singular kernel, which drives us to worrying about its definition and the boundedness properties of the convolution operator it defines, both in the setting of $L^2(\mathbb{R})$ and the rest of the $L^p(\mathbb{R})$ spaces.

1.3.2. The Hilbert transform

There are several reasons why the operator of the Hilbert transform deserves having its own name. Arguably, it carried the first singular kernel mathematicians worried about. Interestingly, it is a useful tool in applied sciences such as spectroscopy in chemistry

Due to the fact that the function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \frac{1}{x}$ is not integrable around the origin, it makes no sense to integrate it straightaway against any $L^p(\mathbb{R})$ function. For instance, let $g : \mathbb{R} \to \mathbb{R}$, $g(x) = \mathbb{1}_{[-1,1]}(x) \in L^p(\mathbb{R})$ for any $1 \le p \le \infty$, yet

$$\int_{\mathbb{R}} f(x)g(x)\,dx = \int_{-1}^{1} \frac{1}{x}\,dx$$

[3] What is more, if $f \in L^p(\mathbb{R})$ for a fixed $1 \le p < \infty$ and noticing that $Q_y \in L^q(\mathbb{R})$ $\forall\, 1 < q \le \infty$ and $y > 0$, by applying the well known Young convolution inequality one gets that $Q_y * f \in L^r(\mathbb{R})$ $\forall\, r > p$.

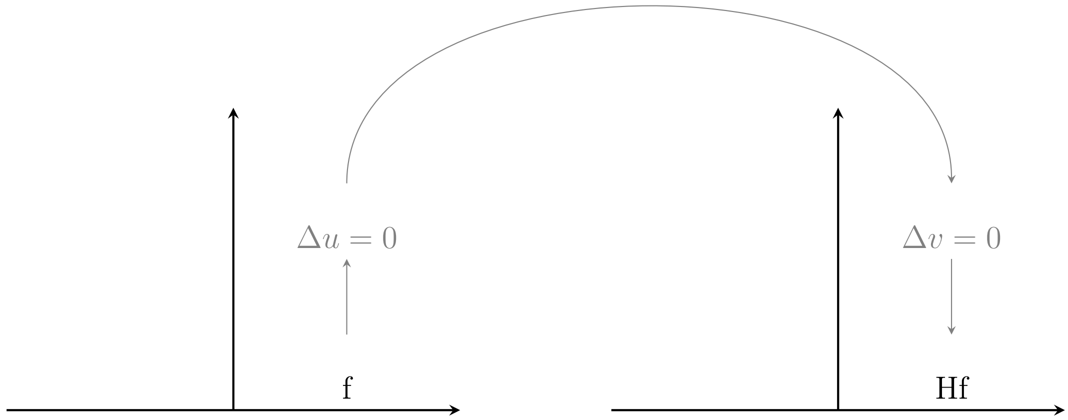


Figure 1.1: Rise of the Hilbert transform in Dirichlet's problem for Laplace's equation. First, let f be defined in the axis $y = 0$. Obtain u such that $\Delta u = 0$ and f is its boundary value. Then, obtain the conjugate harmonic function v of u (the one that turns $u(x, y) + iv(x, y)$ into a holomorphic function on the complex plane, setting the additive constant to 0). Finally, obtain the Hilbert transform of f, Hf by computing the limit $\lim_{y\to 0} v(x, y)$.

is not computable in the Lebesgue sense. However, we still have hope for a definition of the convolution of f against other functions because of the fact that f is an odd function. We would like to exploit this feature to achieve enough cancellation to overcome the effect of the singularity.

Definition 1.29. *Let $f : \mathbb{R} \to \mathbb{R}$ be a function with a singularity around the origin. Its integral is computed in the* ***principal value*** *sense as*

$$\text{p.v.} \int_{\mathbb{R}} f(x)dx := \lim_{\epsilon\to 0^+} \int_{|x|>\epsilon} f(x)dx.$$

This way, if $f \cdot \mathbb{1}_{B_\epsilon(0)^c} \in L^1(\mathbb{R})\ \forall\, \epsilon > 0$, then the limit of the above integral as $\epsilon \to 0^+$ may also yield a real number.

Proposition 1.30. *Define the linear functional acting on $\varphi \in \mathscr{S}$:*

$$\text{p.v.} \left(\frac{1}{x}\right)(\varphi) := \text{p.v.} \int_{\mathbb{R}} \frac{\varphi(x)}{x} dx$$

This yields a tempered distribution.

Proof. Inroduce a $\varphi(0)$ term taking advantage of the oddness of $\frac{1}{x}$ and split the domain of integration for a separate treatment of the integrability issues.

$$\begin{aligned}\left|\text{p.v.}\left(\frac{1}{x}\right)(\varphi)\right| &\leq \lim_{\epsilon\to 0^+}\int_{\epsilon<|x|<1}\frac{|\varphi(x)-\varphi(0)|}{|x|}dx+\int_{|x|\geq 1}\frac{|\varphi(x)|}{|x|}dx \\ &\leq \lim_{\epsilon\to 0^+}\int_{\epsilon<|x|<1}\frac{|\varphi'(c)||x|}{|x|}dx+\int_{|x|\geq 1}\frac{|x\varphi(x)|}{|x|^2}dx \\ &\leq \|\varphi\|_{0,1}\int_{|x|<1}dx+\|\varphi\|_{1,0}\int_{|x|\geq 1}\frac{1}{|x|^2}dx = C_1\|\varphi\|_{0,1}+C_2\|\varphi\|_{1,0}\end{aligned}$$

We used the mean value theorem (with some $c\in[0,x]$ appearing) and the seminorms from Definition 1.8. So the result follows from Theorem 1.10. □

The principal value is a tool that allows us to properly define the Hilbert transform.

Definition 1.31. *Let $\varphi\in\mathscr{S}(\mathbb{R}^n)$ be a Schwartz function. The* ***Hilbert transform*** *of φ, $H\varphi$ is defined as the convolution of the principal value distribution $\frac{1}{\pi}\text{p.v.}\left(\frac{1}{x}\right)$ against the Schwartz function φ (see Definition 1.12):*

$$H\varphi(x) := \frac{1}{\pi}\text{p.v.}\left(\frac{1}{x}\right)*\varphi(x)=\frac{1}{\pi}\text{p.v.}\int_{\mathbb{R}}\frac{\varphi(x-y)}{y}dy$$

Naturally, we would like to take advantage of the density of $\mathscr{S}(\mathbb{R}^n)$ in $L^p(\mathbb{R}^n)$, $1\leq p<\infty$ to extend this definition to any $L^p(\mathbb{R}^n)$ function. This is one of the goals of Chapter 2.

Alternatively, it is possible to define the Hilbert transform as a multiplier operator.

Definition 1.32. *In the setting of $L^2(\mathbb{R}^n)$, where Plancherel theorem for the Fourier transform holds, an essentially bounded function $m\in L^\infty(\mathbb{R}^n)$ is called a* ***multiplier****, when we consider an associated operator $T_m : L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)$ such that*

$$\widehat{T_m f}(\xi)=m(\xi)\hat{f}(\xi),\ \forall f\in L^2(\mathbb{R}^n).$$

The intuition one gets from the definition is that the role of a multiplier function is to modify the frequency spectrum of a function. The topic of multipliers is once again a broad world on its own. There are some examples linking this topic with the theory of Chapter 2 in Chapter 4.

Definition 1.33. *The* ***Hilbert transform*** *is the operator mapping $H : L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)$ defined via the multiplier $m(\xi)=-i\text{sgn}(\xi)$. This is,*

$$\widehat{Hf}(\xi):=-i\text{sgn}(\xi)\hat{f}(\xi).$$

Again, it is desirable to somehow extend this definition to the rest of $L^p(\mathbb{R}^n)$ spaces, relying, for example, on the fact that $L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for $1 \leq p < \infty$.

Note the following coherence:

Proposition 1.34. *Definition 1.31 and definition 1.33 are equivalent:*

$$\left(\frac{1}{\pi}\text{p.v.}\left(\frac{1}{x}\right)\right)\hat{} = -i\text{sgn}(\cdot)$$

in the sense of tempered distributions.

Proof. Let us first show that

$$\frac{1}{\pi}\text{p.v.}\left(\frac{1}{x}\right) = \lim_{y\to 0} Q_y$$

holds in the sense of tempered distributions, where Q_y denotes the conjugate Poisson kernel (understand the right hand side as in the following computation).

Take any $\varphi \in \mathscr{S}(\mathbb{R}^n)$ and combine the limit on the right-hand side with the one in the principal value distribution:

$$\begin{aligned}
&\frac{1}{\pi}\text{p.v.}\left(\frac{1}{x}\right)(\varphi) - \lim_{y\to 0} Q_y(\varphi) \\
&= \lim_{y\to 0}\int_{|x|>y}\left(\frac{1}{\pi x} - \frac{1}{\pi}\frac{x}{x^2+y^2}\right)\varphi(x)dx + \lim_{y\to 0}\int_{|x|\leq y}\left(-\frac{1}{\pi}\frac{x}{x^2+y^2}\right)\varphi(x)dx \\
&\quad = \lim_{y\to 0}\int_{|x|>1}\left(\frac{1}{\pi}\frac{1}{x(x^2+1)}\right)\varphi(yx)dx + \lim_{y\to 0}\int_{|x|\leq 1}\left(-\frac{1}{\pi}\frac{x}{x^2+1}\right)\varphi(yx)dx
\end{aligned}$$

yields after a scaling change of variables. Realise that

$$r_1(x) := \frac{1}{\pi}\frac{1}{|x|(x^2+1)}\mathbb{1}_{\{|x|>1\}}(x) \qquad r_2 = \frac{1}{\pi}\frac{|x|}{x^2+1}\mathbb{1}_{\{|x|\leq 1\}}(x)$$

are both integrable functions. This, combined with the fact that φ is a Schwartz function, allows us to invoke the dominated convergence theorem that leads to

$$\begin{aligned}
&\frac{1}{\pi}\text{p.v.}\left(\frac{1}{x}\right)(\varphi) - \lim_{y\to 0} Q_y(\varphi) \\
&\qquad = \int_{|x|>1}\left(\frac{1}{\pi}\frac{1}{x(x^2+1)}\right)\varphi(0)dx + \int_{|x|\leq 1}\left(-\frac{1}{\pi}\frac{x}{x^2+1}\right)\varphi(0)dx = 0
\end{aligned}$$

because of oddness.

Once proved this, since the Fourier transform is continuous on $\mathscr{S}^*(\mathbb{R}^n)$ (see Proposition 1.15), for any Schwartz function φ,

$$\left(\frac{1}{\pi}\text{p.v.}\left(\frac{1}{x}\right)\right)\hat{}(\varphi) = (\lim_{y\to 0} Q_y)\hat{}(\varphi) = \lim_{y\to 0}\left(\widehat{Q_y}\right)(\varphi).$$

A simple computation shows that

$$(-i\text{sgn}(\xi)e^{-2\pi y|\xi|})\check{}(x) = Q_y(x).$$

Therefore, by the dominated convergence theorem,

$$\begin{aligned}\lim_{y\to 0}\left(\widehat{Q_y}\right)(\varphi) &= \lim_{y\to 0}\int_{\mathbb{R}}\widehat{Q_y}(\xi)\varphi(\xi)d\xi \\ &= \lim_{y\to 0}\int_{\mathbb{R}} -i\text{sgn}(\xi)e^{-2\pi y|\xi|}\varphi(\xi)d\xi = \int_{\mathbb{R}} -i\text{sgn}(\xi)\varphi(\xi)d\xi\end{aligned}$$

holds.

□

1.3.3. Connection of the maximal function with the harmonic extension on the upper half plane

Here, we present the first utility of the maximal function. Our will is to transfer some of the properties of the Hardy-Littlewood maximal function to the Poisson integral. Interestingly, this is not only possible for the Poisson kernel but also for other kernels that share a common property with it.

Theorem 1.35. *Let $\varphi \in L^1(\mathbb{R}^n)$ be a kernel function and let $\varphi_y(x) = y^{-n}\varphi(y^{-1}x)$ be its integral-preserving dilates. Assume there exists a strictly decreasing function $\psi : \mathbb{R} \to \mathbb{R}$ such that $|\varphi(x)| \le \psi(|x|)$. Let $\rho : \mathbb{R}^n \to \mathbb{R}$ be the radial function defined by $\rho(x) := \psi(|x|)$ for convenience, and suppose $\rho \in L^1(\mathbb{R}^n)$. Analogously, denote their dilates by $\rho_y(x) := y^{-n}\rho(y^{-1}x)$. If $f \in L^p(\mathbb{R}^n)$, $1 \le p \le \infty$, then,*

(a)

$$\sup_{y>0}|\varphi_y * f(x)| \le C_n\mathfrak{M}f(x) \quad \forall\, x \in \mathbb{R}^n \tag{1.17}$$

where C_n is a constant depending only on the dimension and the L^1 norm of ρ.

(b) Moreover, when $\int \varphi = 1$, the convolution $\varphi_y * f(x)$ is almost everywhere convergent to $f(x)$, namely,

$$\lim_{y\to 0}\varphi_y * f(x) = f(x) \quad a.e\ x \in \mathbb{R}^n.$$

The key feature of the kernels for which this theorem applies is the possibility of approximating them by sums of centred indicator functions of concentric balls, which are, at the end of the day, the kernel present in the Hardy-Littlewood maximal function (1.4).

Proof. Note that the convolution is well defined, accounting that $f \in L^p(\mathbb{R}^n)$ and $\varphi_y \in L^1(\mathbb{R}^n)$ (Proposition 1.5). To show (a),

$$|\varphi_y * f(x)| \leq \int_{\mathbb{R}^n} |f(x-t)||\varphi_y(t)|dt \leq \int_{\mathbb{R}^n} |f(x-t)|\rho_y(t)dt.$$

Introduce an indicator function for ρ_y and, since the integrand functions are positive, apply Fubini-Tonelli theorem:

$$\int_{\mathbb{R}^n} |f(x-t)|\rho_y(t)dt = \int_{\mathbb{R}^n} |f(x-t)| \int_0^\infty \mathbb{1}_{\{u<\rho_y(t)\}}(u)du\, dt$$
$$= \int_0^\infty \int_{\mathbb{R}^n} |f(x-t)|\mathbb{1}_{\{u<\rho_y(t)\}}(u)dt\, du$$

One realises that $\{t : u < \rho_y(t)\}$ is in fact a ball on the t variable, due to ρ_y being a radial function and radially decreasing. Note it is either a ball with finite measure or the empty set, but never the whole space thanks to the fact that ψ is strictly decreasing. Denote $B_{r(y,u)}(0) := \{t : u < \rho_y(t)\}$ for $y, u > 0$. Multiplying and dividing by the measure of this ball (whenever the values of u and t yield a nonempty ball, otherwise we integrate zero), one gets

$$\int_0^\infty \int_{\mathbb{R}^n} |f(x-t)|\mathbb{1}_{\{u<\rho_y(t)\}}(u)dt\, du =$$
$$\int_0^\infty \frac{1}{|B_{r(y,u)}(0)|} \int_{B_{r(y,u)}(0)} |f(x-t)|dt\, |B_{r(y,u)}(0)|\, du$$
$$\leq \mathfrak{M}f(x) \int_0^\infty |\{t : u < \rho_y(t)\}|du = \|\rho\|_1 \mathfrak{M}f(x).$$

The last equality follows from Lemma 1.26. (1.17) yields after taking supremum on y. The result has been obtained with $C_n = \|\rho\|_1$.

The result in (b) is the analogue of the Lebesgue differentiation theorem 1.25. What (b) tells us is that not only do the averages of an $L^1_{\text{loc}}(\mathbb{R}^n)$ function tend to the central point, but also this happens for any approximation to the identity verifying the hypotheses in the theorem (notice that the kernel present in the Hardy-Littlewood maximal function, see Eq. (1.5), defines an approximation to the identity). In fact, the proof of (b) is the same as for the Lebesgue differentiation theorem. Essentially, the idea of the proof is showing uniform convergence for continuous and compactly supported functions, so that the almost everywhere convergence follows by a density argument and the result in (a). Check [10], Chapter 3, Section 2.2 for the proof. □

It is obvious that the Poisson kernel (1.13) is radial and strictly decreasing in x. Therefore, the result in part (a), in view of the maximal theorem 1.22 implies that the maximal version of the Poisson integral inherits the boundedness properties stated in the maximal theorem, i.e. that $\sup_{y>0} u(x,y) = P_y * f(x)$ belongs to $L^p(\mathbb{R}^n)$ provided $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$, and $\sup_{y>0} u(x,y)$ belongs to $L^{1,\infty}(\mathbb{R}^n)$ if $f \in L^1(\mathbb{R}^n)$.

In general, if one considers a family of linear operators $(T_y)_{y>0}$ on $L^p(X)$ and wishes to show pointwise convergence results of the kind $\lim_{y\to 0} T_y f(x) = f(x)$ (as in (b)), the usual strategy is to strive for a maximal weak-type estimate for the family $(T_y)_{y>0}$ (as result (a) implies).

1.3.4. Nontangential convergence

Despite Theorem 1.35 being a powerful result, it still has an improved version. Consider the cones with vertexes at $(x_0, 0)$, $\Gamma_\alpha(x_0) = \{(x,y) \in \mathbb{R}^+ : y > |x - x_0| \tan\alpha\}$ for any $\alpha \in (0, \frac{\pi}{2}]$. While the statement of Theorem 1.35 is based on limits for $y \to 0$ on vertical lines, the statement still holds for limits along any curve that sits in one of the cones $\Gamma_\alpha(x_0)$ and thus tends to the horizontal axis in a nontangential manner (see Figure 1.2).

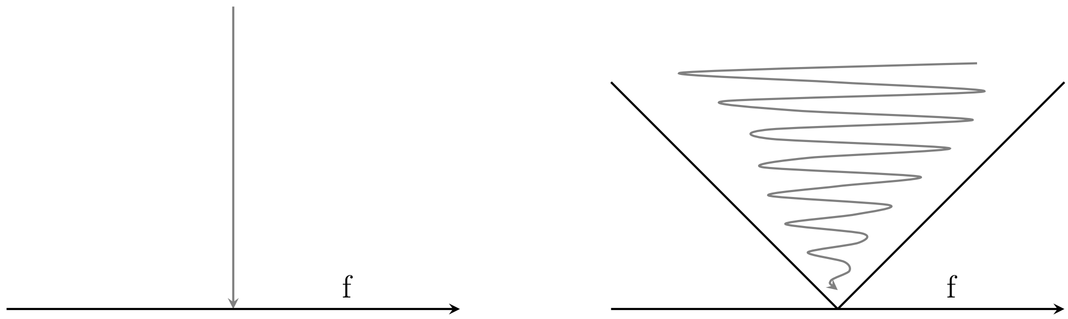


Figure 1.2: Vertical convergence as stated in Theorem 1.35 (on the left) compared to the nontangential convergence of Theorem 1.36 (on the right). The second kind of convergence generalises the result for the first kind. However, the statement turns out to be false for convergence outside any cone, in a tangential fashion.

Theorem 1.36. *Let $\varphi \in L^1(\mathbb{R}^n)$ be a kernel and let $\varphi_y(x) = y^{-n}\varphi(y^{-1}x)$ be its integral-preserving dilates. Assume there exists a strictly decreasing function $\psi : \mathbb{R} \to \mathbb{R}$ such that $|\varphi(x)| \le \psi(|x|)$. Let $\rho : \mathbb{R}^n \to \mathbb{R}$ be the radial function defined by $\rho(x) := \psi(|x|)$ for convenience, and suppose $\rho \in L^1(\mathbb{R}^n)$. Analogously, denote their dilates by $\rho_y(x) := y^{-n}\rho(y^{-1}x)$. If $f \in L^p(\mathbb{R}^n)$, $1 \le p \le \infty$, and the kernel ρ_y has the property*

$$\rho_y(x - t) \le A_\alpha \rho_y(x) \quad \textit{for } |t| < \tan(\alpha) y, \tag{1.18}$$

where A_α is a constant only depending on α. Then, for any $\alpha \in (0, \frac{\pi}{2}]$,

(a)

$$\sup_{\substack{y>0 \\ (x,y)\in\Gamma_\alpha(x_0)}} |\varphi_y * f(x)| \leq C_{n,\alpha}\mathfrak{M}f(x_0) \quad \forall\, x_0 \in \mathbb{R}^n, \tag{1.19}$$

where $C_{n,\alpha}$ is a constant depending only on the dimension, the angle α and the L^1 norm of ρ.

(b) Moreover, when $\int \varphi = 1$, the convolution $\varphi_y * f(x)$ is almost everywhere convergent to $f(x)$, namely,

$$\lim_{\substack{y\to 0 \\ (x,y)\in\Gamma_\alpha(x_0)}} \varphi_y * f(x) = f(x_0) \quad a.e\ x_0 \in \mathbb{R}^n.$$

The proof stems from a slight modification of that of Theorem 1.35, check [10], Chapter 7, Section 1 for the details.

Condition (1.18) is new with respect to Theorem 1.35. However, it is not a rare condition for usual kernels. Indeed, it is straightforward to check that the Poisson kernel verifies it.

1.4. The Hardy-Littlewood maximal function on measure metric spaces

The two main tools we used to define the Hardy-Littlewood maximal function were the balls of the metric space $\mathbb{R}^n$ and integration on the latter. Consequently, it makes sense to define an analogous version somewhere where we have balls and integration, that is, on a measure metric space $((X, d), \Sigma, \mu) \equiv X$.

Definition 1.37. *Let X be a measure metric space and let $f \in L^1_{loc}(X)$ be a locally integrable function. The* ***centred Hardy-Littlewood maximal function*** *of f is defined as*

$$\mathfrak{M}f(x) := \sup_{r>0} \frac{1}{\mu(B_r(x))} \int_{B_r(x)} |f(y)|d\mu(y). \tag{1.20}$$

This time, the balls $B_r(x)$ are built out of the distance d that arms the measure metric space X.

Similarly, one defines the uncentred version of it, as we did in Definition 1.19.

As a foothold, hereinafter we are going to provide the measure metric space with the so-called doubling property.

Definition 1.38. *A measure metric space X is said to have the* ***doubling property*** *if*

$$\mu(B_{2r}(x)) \leq C\mu(B_r(x)), \quad \forall\, r > 0, x \in X, \tag{1.21}$$

C being a universal constant for the space X.

Observe $\mathbb{R}^n$ enjoys the doubling property with $C = 2^n$. Also recall that we used such a property to see that the centred and the uncentred Hardy-Littlewood maximal functions in $\mathbb{R}^n$ are comparable, $\mathfrak{M}f \sim \mathfrak{M}^{\text{unc}}f$ in Remark 1.20. By assuming the doubling property for our measure metric space X, we keep on having $\mathfrak{M}f \sim \mathfrak{M}^{\text{unc}}f$ available. Not only do we have the comparability result but also the doubling property teams up with the generalisation of the Vitali-type covering lemma (Lemma 2.5), key to prove the forthcoming Calderón-Zygmund Lemma in our new setting, as exposed in Theorem 2.3.

Remark 1.21, Theorem 1.22 and Corollary 1.25 hold for any measure metric space enjoying the doubling property and the proofs follow the same strategy. The Vitali-type covering lemma (Lemma 1.24) can be substituted by Lemma 2.5 providing the availability of the doubling property.

Theorem 1.39 (Maximal theorem, general setting)**.** *Let f be a measurable complex-valued function on a measure metric space X enjoying the doubling property. Then:*

(a) If $f \in L^p(X)$ for $1 \leq p \leq \infty$, $\mathfrak{M}f(x)$ is finite a.e. $x \in X$.

(b) For every $\lambda > 0$ and $f \in L^1(X)$,

$$\lambda\mu(\{x \in X : \mathfrak{M}f(x) > \lambda\}) \leq A \,\|\, f \,\|_1 , \tag{1.22}$$

where A is a constant.

(c) If $f \in L^p(X)$, $1 < p \leq \infty$, then $\mathfrak{M}f \in L^p(X)$ and

$$\|\, \mathfrak{M}f \,\|_p \leq A_p \,\|\, f \,\|_p , \tag{1.23}$$

where A_p only depends on the exponent p.

Corollary 1.40 (Lebesgue differentiation theorem, general setting)**.** *Whenever $f \in L^1_{loc}(X)$,*

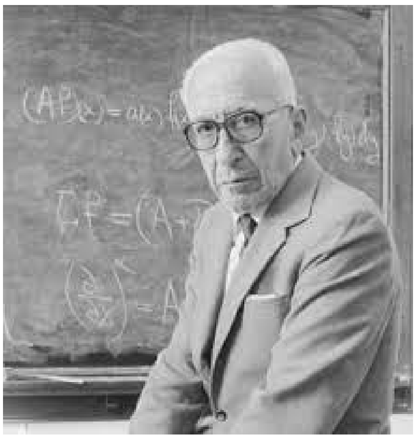
$$\lim_{r\to 0} \frac{1}{\mu(B_r(x))} \int_{B_r(x)} f(y)d\mu(y) = f(x) \qquad a.e.\ x \in X. \tag{1.24}$$

Remark 1.41. *Remark 1.21 also holds in the new setting of X: the Hardy-Littlewood maximal function is measurable.*

CHAPTER 2
Calderón-Zygmund theory

2.1. Calderón-Zygmund decomposition

The Calderón-Zygmund theory was developed originally in the setting of $\mathbb{R}^n$, in the 1950s, set off with the collaborative breakthrough paper [2] published in 1952. It aimed to prove boundedness of singular convolution-type operators on spaces of functions (mainly L^p spaces) built over $\mathbb{R}^n$. It is almost a miracle how the following, apparently disconnected idea, eventually yields the aforementioned boundedness result.



(a) Alberto Pedro Calderón. Mendoza (Argentina) 1920 - Chicago (United States of America) 1998. [Source: https://www.ams.org/notices/199809/mem-calderon.pdf]

(b) Antoni Zygmund. Warsaw (Poland) 1900 - Chicago (United States of America) 1992. [Source: https://mathshistory.st-andrews.ac.uk/Biographies/Zygmund/]

Figure 2.1: The fathers of the theory of singular integrals. They were not only close collaborators, but also Calderón was Zygmund's PhD student. Together, they revolutionised Analysis of the 20th century, and founded the Chicago School of Mathematical Analysis.

Theorem 2.1 (Calderón-Zygmund Lemma in $\mathbb{R}^n$)**.** *Let $f \in L^1(\mathbb{R}^n)$ and $\lambda > 0$.*

(a) There exists a partition $\mathbb{R}^n = F \sqcup \Omega$, such that

(b) $|f(x)| \leq \lambda$ a.e. $x \in F$, and

(c) Ω can be written as a countable union of cubes Q_k with disjoint interior, $\Omega = \bigsqcup_{k\in\mathbb{N}} Q_k$ moreover satisfying

$$\lambda \leq \frac{1}{|Q_k|}\int_{Q_k} |f(x)|dx \leq 2^n\lambda, \quad \forall\, k \in \mathbb{N}. \tag{2.1}$$

Essentially, what this theorem tells us is: for any wild function f just subject to being Lebesgue integrable, there exists a decomposition of its domain, $\mathbb{R}^n$, into two disjoint sets such that f is essentially bounded in one of them, and although it may not be in the other, the averages of f over some (almost) disjoint cubes are bounded.

Although a generalisation of this theorem is going to come up later, we leave here the proof, which consists in an elegant stopping-time argument worth explaining.

Proof. Mesh $\mathbb{R}^n$ into cubes $\{Q_k^0\}_{k\in\mathbb{N}}$ with disjoint interior and of the same size, large enough so that the averages of $|f|$ are bounded above by the given λ on all of the cubes in the mesh:

$$\frac{1}{|Q_k^0|}\int_{Q_k^0} |f(x)|dx < \lambda \quad \forall\, k \in \mathbb{N}.$$

This is possible because f is integrable,

$$\frac{1}{|Q_k^0|}\int_{Q_k^0} |f(x)|dx \leq \frac{\|\, f\,\|_1}{|Q_k^0|},$$

so choose the size of the cubes such that $|Q_k^0| > \frac{\|\, f\,\|_1}{\lambda}$.

We are going to run an algorithm. Set $\Omega = \emptyset$ and the step $s = 1$. We split each of the cubes $\{Q_k^0\}_{k\in\mathbb{N}}$ into 2^n dyadic descendent cubes of the same size $\{Q_k^1\}_{k\in\mathbb{N}}$.

<u>Case 1:</u> For each descendent cube in step s (that is, for each $k \in \mathbb{Z}$), if

$$\frac{1}{|Q_k^s|}\int_{Q_k^s} |f(x)|dx > \lambda, \tag{2.2}$$

then Q_k^s is selected to take part in the set Ω, so actualise $\Omega^{\text{new}} = \Omega^{\text{old}} \cup Q_k^s$. For such a cube Q_k^s, assume that Q_r^{s-1} is its direct ancestor cube. Then, by (2.2) and the fact that Q_r^{s-1} fell into Case 2,

$$\lambda < \frac{1}{|Q_k^s|}\int_{Q_k^s} |f(x)|dx \leq \frac{2^n}{|Q_r^{s-1}|}\int_{Q_r^{s-1}} |f(x)|dx \leq 2^n\lambda$$

which proves (2.1) for Q_k^s.

Case 2: Instead, if

$$\frac{1}{|Q_k^s|}\int_{Q_k^s}|f(x)|dx \leq \lambda,$$

then we iterate and further divide Q_k^s into 2^n identical descendent cubes, and check into which of the two cases falls each of them.

Actualise $s^{\text{new}} = s^{\text{old}} + 1$ and let the algorithm run recursively. This way, we obtain a partition like in (a), plus (c) has been verified for all cubes Q_k^s that were selected for Case 1. Fact (b) yields from the Lebesgue differentiation theorem (Corollary 1.25)[1] because

$$f(x) = \lim_{j\to\infty}\frac{1}{|Q_k^j|}\int_{Q_k^j}|f(x)|dx \leq \lambda \quad \forall\, k \in \mathbb{N}$$

since here all of the intervening cubes fall into Case 2. □

Let us now state the crucial Calderón-Zygmund decomposition of an integrable function as a corollary.

Corollary 2.2. *Let $f \in L^1(\mathbb{R}^n)$ and $\lambda > 0$. There exists a decomposition of f as sum of two functions, $f = g + b$ such that g is an essentially bounded function, and such that the support of b can in its turn be decomposed into a union of cubes with disjoint interior, in each of which b has zero average. More precisely, there exists a decomposition $f = g + b$ such that*

$$g(x) \leq 2^n\lambda \ \ a.e.\, x \in \mathbb{R}^n, \qquad \frac{1}{|Q_k|}\int_{Q_k} b(x)\,dx = 0 \ \ \forall\, k \in \mathbb{N},$$
$$\frac{1}{|Q_k|}\int_{Q_k}|b(x)|\,dx \leq 2^n\lambda, \qquad \operatorname{supp}(b) = \bigsqcup_{k\in\mathbb{N}} Q_k, \qquad b \leq f. \tag{2.3}$$

g and b are usually referred to as the "good" and the "bad" part of f. It is a worthwhile trade to gain such boundedness properties on g and b for the price of having to deal with two functions instead of only one.

Proof. Fix any $\lambda > 0$ and apply Calderón-Zygmund Lemma, Theorem 2.1, to get the decomposition $\mathbb{R}^n = F \sqcup \Omega$, as in the statement. Define

$$g(x) := \begin{cases} f(x), & x \in F \\ \frac{1}{|Q_k|}\int_{Q_k} f(x)dx, & x \in Q_k, \forall\, k \in \mathbb{N}. \end{cases}$$

[1] Being meticulous, the Lebesgue differentiation theorem was presented in Chapter 1 in the form of averages over balls, not cubes. However, this theorem still holds if the family of sets over which one averages is so-called regular. In particular, the family of all cubes in $\mathbb{R}^n$ is a regular family. Such condition of being regular resembles the doubling condition. For details, see [10], Chapter 1, Section 1.8.

Directly notice that g is essentially bounded by $2^n\lambda$. Consistently, let

$$b(x) := f(x) - g(x) = \begin{cases} 0, & x \in F \\ f(x) - \frac{1}{|Q_k|}\int_{Q_k} f(x)dx, & x \in Q_k, \forall\, k \in \mathbb{N}, \end{cases}$$

which immediately implies (2.3) and the proof is complete. □

This is the right approach to get the estimates needed in the coming-up Theorem 2.9.

In contrast, the setting of this chapter, unless otherwise specified, is a **generic σ-finite measure space over a metric space equipped with a regular measure** $((X,d),\Sigma,\mu)$ **enjoying the doubling property**.

Theorem 2.3 (Calderón-Zygmund lemma, general setting)**.** *Let $f \in L^1(X)$ and $\lambda > 0$.*

(a) There exists a partition of the space $X = F \sqcup \Omega$, F being a closed set and Ω an open set, such that

(b) $|f(x)| \leq \lambda$ a.e. $x \in F$, and

(c) Ω can be written as a countable disjoint union of smaller sets $\Omega = \bigsqcup_{k\in\mathbb{N}} \Omega_k$ moreover satisfying

$$\frac{1}{\mu(\Omega_k)}\int_{\Omega_k} |f(x)|d\mu(x) \leq C\lambda, \quad \forall\, k \in \mathbb{N} \tag{2.4}$$

for some real constant $C > 0$.

In the same way as in the case of $\mathbb{R}^n$:

Corollary 2.4. *Let $f \in L^1(X)$ and $\lambda > 0$. There exists a decomposition of f as sum of two functions, $f = g + b$ such that g is an essentially bounded function, and such that the support of b can in its turn be decomposed into a union disjoint sets $\{Q_k\}_{k\in\mathbb{N}}$, in each of which b has zero average. More precisely, there exists a decomposition $f = g + b$ such that*

$$g(x) \leq C\lambda \;\; a.e.\, x \in X, \quad \frac{1}{\mu(Q_k)}\int_{Q_k} b(x)\, d\mu(x) = 0 \;\; \forall\, k \in \mathbb{N},$$
$$\frac{1}{\mu(Q_k)}\int_{Q_k} |b(x)|\, d\mu(x) \leq C\lambda, \quad \operatorname{supp}(b) = \bigsqcup_{k\in\mathbb{N}} Q_k, \quad b \leq f \tag{2.5}$$

for some $C > 0$.

An analogous argument as that in Corollary 2.2 proves the new corollary.

Notice the slight differences with respect to the former Theorem 2.1 in the setting of $\mathbb{R}^n$. The downside of Theorem 2.3 is that it lacks the lower bound for the averages over the Ω_k. As an upside, we get a topological characterization of the sets F and Ω. These disagreements owe to the fact that the proof of Calderón-Zygmund Lemma 2.1 cannot be analogously generalised to a generic measure space, because partitioning the space $\mathbb{R}^n$ into a perfectly fitting mesh of disjoint cubes (up to a set of measure zero) is a specificity of $\mathbb{R}^n$. Accordingly, Theorem 2.3 demands a totally different proof, heavily relying on the Hardy-Littlewood maximal function.

Furthermore, we are going to be needing a convenient Vitali-type covering lemma for the proof. Let us introduce some notation: let $B = B_r(x)$ be a ball of radius r and centre x in a metric space. Denote by B^* an enlarged dilation of the ball B, sharing the same centre. That is, say $c^* > 1$ is the dilating factor, then $B^* = B_{c^* r}(x)$.

Lemma 2.5 (Vitali-type covering lemma, general setting)**.** *Let (X, d) be a metric space enjoying the the following engulfing property: there exists $c_1 > 1$ such that for all $x, y \in X$ and $\delta > 0$*

$$B_\delta(x) \cap B_\delta(y) \neq \emptyset \implies B_\delta(y) \subset B_{c_1\delta}(x).$$

Let $F \subseteq X$ be a nonempty closed set. Then, there exists a sequence of balls $(B_k)_{k\in\mathbb{N}}$ and two families of each dilations, $(B_k^*)_{k\in\mathbb{N}}$ and $(B_k^{**})_{k\in\mathbb{N}}$, such that

(a) $(B_k)_{k\in\mathbb{N}}$ are pairwise disjoint,

(b) $\bigcup_k B_k^* = F^c$, and

(c) $B_k^{**} \cap F \neq \emptyset$, $\forall\, k$.

The interest of this lemma is that each family of balls exhibits different useful properties: $(B_k)_k$ are pairwise disjoint, $(B_k^*)_k$ gather together to recover F^c and the last dilation is large enough so that any ball in $(B_k^{**})_k$ meets the border of F. We clarify that each family of dilations shares the same dilation factor. Check [12], Chapter 1, Section 3.2., for the proof of the lemma.

Remark 2.6. *It is convenient to extract another sequence of sets from Lemma 2.5. Take the first element in $(B_k^*)_{k\in\mathbb{N}}$ and define $Q_1 := B_1^*$. Next, define $Q_2 := B_2^* \setminus (Q_1)$. By an inductive process, build*

$$Q_k := B_k^* \setminus \left(\bigcup_{j=1}^{k-1} Q_j \right).$$

The sequence $(Q_k)_k$ has the properties that their sets are pairwise disjoint and $\bigcup_k Q_k = F^c$. We paid the price that Q_k are no longer balls, but other less elementary sets.

The name Q_k of such new sets is inspired by their role in the proof of Theorem 2.12, which mimics the one carried out by the cubes in the proof of the $X = \mathbb{R}^n$ case.

Proof. (Of Calderón-Zygmund Lemma, Theorem 2.3) Let $f \in L^1(X)$ and fix $\lambda > 0$. Choose $F := \{x \in X : \mathfrak{M}f(x) \leq \lambda\}$ and so $\Omega := \{x \in X : \mathfrak{M}f(x) > \lambda\}$. Accounting that averaging over balls $A(x,r) := \frac{1}{\mu(B_r(x))}\int_{B_r(x)} |f(y)|d\mu(y)$ is a continuous function of x, and that the measure μ is assumed to be regular, it is easy to see that the set F so defined is a closed set; hence, Ω open.

Working first with the set F, let us use the Lebesgue differentiation theorem 1.40 (and Theorem 1.39).

$$\begin{aligned}\lambda \geq \mathfrak{M}f(x) &= \sup_{r>0} \frac{1}{\mu(B_r(x))}\int_{B_r(x)} |f(y)|d\mu(y)\\ &\geq \lim_{r\to 0} \frac{1}{\mu(B_r(x))}\int_{B_r(x)} |f(y)|d\mu(y) = |f(x)|, \quad \text{a.e. } x \in F,\end{aligned}$$

so (b) is shown.

In order to prove (c), take into account Lemma 2.5 and Remark 2.6. For each B_k in the sequence $(B_k)_{k\in\mathbb{N}}$ given by the lemma, choose a point $p_k \in B_k^{**} \cap F$ (the lemma ensures this set is nonempty). By the definition of F,

$$\begin{aligned}\lambda \geq \mathfrak{M}f(p_k) &\geq C^{\text{unc}}\mathfrak{M}^{\text{unc}}f(p_k) \geq \frac{C^{\text{unc}}}{\mu(B_k^{**})}\int_{B_k^{**}} |f(x)|d\mu(x)\\ &\geq \frac{C^{\text{unc}}}{\mu(B_k^{**})}\int_{Q_k} |f(x)|d\mu(x) \geq \frac{C^{\text{unc}}}{C^{**}}\frac{1}{\mu(Q_k)}\int_{Q_k} |f(x)|d\mu(x).\end{aligned}$$

The two last inequalities stem from the facts that $B_k \subseteq Q_k \subseteq B_k^{**}$ and the doubling property: $\mu(Q_k) \leq \mu(B_k^{**}) \leq C^{**}\mu(B_k) \leq C^{**}\mu(Q_k)$. Since $(Q_k)_{k\in\mathbb{N}}$ partition Ω, $\Omega = \bigsqcup_k \Omega_k \equiv \bigsqcup_k Q_k$, the proof is complete. □

Note that this proof unveils the precise identity of the sets F and Ω, which are defined in terms of the Hardy-Littlewood maximal function.

In exactly the same way as in Corollary 2.2, the Calderón-Zygmund decomposition of an integrable function $f \in L^1(X)$ is deduced.

2.2. Bounding singular integral operators

2.2.1. First steps in the Euclidean space

Here is where the L^p boundedness theorem for convolution-type operators will shine. We will get to the desired theorem in the broadest setting after introducing the specific hypotheses and its corresponding version in $\mathbb{R}^n$ which serves as inspiration.

To start with, general convolution-type operators in $\mathbb{R}^n$ over $L^p(\mathbb{R}^n)$ functions looks like

$$(Tf)(x) = \int_{\mathbb{R}^n} K(x-y)f(y)dy, \quad \text{for } f \in L^p(\mathbb{R}^n), \tag{2.6}$$

being $K : \mathbb{R}^n \to \mathbb{C}$ a function called the convolution kernel. Assume for the moment the ideality that the kernel is integrable, $K \in L^1(\mathbb{R}^n)$. Then,

- if $p = 1$, $Tf \in L^1(\mathbb{R}^n)$ and moreover, $\| Tf \|_1 = \| K \|_1 \| f \|_1$ by Fubini theorem.
- if $1 < p \leq \infty$, $Tf \in L^p(\mathbb{R}^n)$ and moreover, $\| Tf \|_p \leq \| K \|_1 \| f \|_p$ by Minkowski integral inequality.

As easy as that for an integrable kernel: T is a bounded operator on $L^p(\mathbb{R}^n)$ for any $1 \leq p \leq \infty$. However, our interest relies on kernels that are not integrable due to a single singularity, say in the origin of the kernel, $K : \mathbb{R}^n \setminus \{0\} \to \mathbb{C}$. Always keep in mind the example of the Hilbert transform, (1.16).

The integrability issue of singular kernels is already a hassle for the associated operator to be defined. One can encounter many different approaches digging in the literature. The typical strategies to overcome such problems are:

(a) The following is inspired by the concept of principal value. Consider the truncations of the kernel around the singularity

$$K_\epsilon(x) := \begin{cases} K(x) & \text{if } |x| \geq \epsilon \\ 0 & \text{if } |x| < \epsilon \end{cases}$$

and so

$$T_\epsilon f(x) := \int_{\mathbb{R}^n} K_\epsilon(x-y)f(y)dy$$

This way, it is usually a simple matter to show that T_ϵ is well defined and bounded on some $L^p(\mathbb{R}^n)$ spaces for all $\epsilon > 0$. The subsequent procedure is defining $T := \lim_{\epsilon \to 0} T_\epsilon$ in a suitable way and proving it inherits the boundedness property from T_ϵ. Usually, these kind of approaches involve several uniform and $L^p(\mathbb{R}^n)$-norm convergence arguments. See [10], Chapter 2.

(b) Enter the world of tempered distributions. In this approach, one would say that initially $T \in \mathscr{S}^*(\mathbb{R}^n)$ is a tempered distribution. However, one imposes that T agrees with a measurable function K away from the origin, namely

$$\langle K, \varphi \rangle = \int_{\mathbb{R}^n} K(x)\varphi(x)dx$$

for any $\varphi \in \mathscr{S}(\mathbb{R}^n)$ with $0 \notin \operatorname{supp}(\varphi)$ (this way, the singularity is dodged). It is implicitly assumed that $K\varphi \in L^1(\mathbb{R}^n)$. Then, one defines the operator T by means of the convolution of the tempered distribution $K \in \mathscr{S}^*(\mathbb{R}^n)$ against the Schwartz function $\varphi \in \mathscr{S}(\mathbb{R}^n)$, as in Definition 1.12; the result of such a convolution is a smooth function with at most polynomial growth.

$$T\varphi(x) := K * \varphi(x)$$

For this strategy, one would hopefully prove the *a priori* version of the desired estimates for such a definition of T and then use a density argument for the class of Schwartz functions to obtain the estimates in the setting of $L^p(\mathbb{R}^n)$ spaces of functions.

(c) We are going to follow an approach aligned with the previous one, but working with another class of dense functions: $L^p(\mathbb{R}^n) \cap L^q(\mathbb{R}^n)$. It is similar to the previous one in the sense that it tries to avoid the singularity by choosing conveniently supported functions. See [3], Chapter 5.

Essentially, two hypothesis on the kernel are required to succeed in our mission. The first of them is a foothold on a particular $L^p(\mathbb{R}^n)$ space: after correctly defining the T operator, assume that $T : L^q(\mathbb{R}^n) \to L^q(\mathbb{R}^n)$ boundedly: $\| Tf \|_q \leq A \| f \|_q$. This serves as an ingredient to invoke Marcinkiewicz interpolation theorem.

The second hypothesis is a technical one.

Definition 2.7. *A convolution kernel K on $\mathbb{R}^n$ is said to satisfy the **Hörmander condition** if*

$$\sup_{|y|>0} \int_{|x|\geq 2|y|} |K(x-y) - K(x)|dx = B < \infty, \tag{2.7}$$

where $B > 0$ is a finite number.

Since the integral is computed over the region $\{x \in \mathbb{R}^n : |x| > 2|y|\}$, the singularity of the kernel is avoided both for $x-y$, $|x-y| \geq |x| - |y| \geq 2|y| - |y| = |y| > 0$ and for x, $|x| \geq 2|y| > 0$. In some sense, we are asking that the global variation of the kernel is not so wild that is not integrable. Nevertheless, the Hörmander condition is usually seen as a weakened version of the stronger condition

$$|\nabla K(x)| \leq \frac{C}{|x|^{n+1}} \tag{2.8}$$

for $K \in C^1(\mathbb{R}^n \setminus \{0\})$. Even though condition (2.8) is neater than the Hörmander condition, we are still interested in keeping the latter since some kernels fulfil the Hörmander condition, but not condition (2.8) (in this regard, we discuss the Hörmander multipliers in Chapter 4, Section 4.1).

Proposition 2.8. *If a kernel $K \in C^1(\mathbb{R}^n \setminus \{0\})$ fulfils (2.8), then it satisfies the Hörmander condition (2.7).*

Proof. Using the multidimensional mean value theorem for K,

$$\int_{|x|\geq 2|y|} |K(x-y) - K(x)|dx \leq \int_{|x|\geq 2|y|} |\nabla K(c)||y|dx$$

for some c in the segment joining $x-y$ and x. We may assume that this segment does not contain the origin, because the set of x for which the segment contains the origin is of measure zero. Applying the gradient bound,

$$\int_{|x|\geq 2|y|} |\nabla K(c)||y|dx \leq C \int_{|x|\geq 2|y|} \frac{|y|}{|c|^{n+1}} dx.$$

Now c is comparable in modulus to x, since it lies in the segment joining x and $x-y$ and both endpoints are comparable in modulus to x, ($\frac{1}{2}|x| \leq |x-y| \leq \frac{3}{2}|x|$). This means $|c| \geq A|x|$ for a universal constant A.

$$C \int_{|x|\geq 2|y|} \frac{|y|}{|c|^{n+1}} dx \leq AC \int_{|x|\geq 2|y|} \frac{|y|}{|x|^{n+1}} dx$$

With the change of variables $z = \frac{x}{|y|}$,

$$AC \int_{|x|\geq 2|y|} \frac{|y|}{|x|^{n+1}} dx = AC \int_{|z|\geq 2} \frac{1}{|z|^{n+1}} dz$$

which is now a finite number independent of y, yielding the uniform bound. □

We are in the position of stating a theorem to bound singular integrals on $L^p(\mathbb{R}^n)$ spaces in the setting of $\mathbb{R}^n$.

Theorem 2.9. *Let T be a linear operator such that there exists a measurable kernel function K such that*

$$Tf(x) = \int_{\mathbb{R}^n} K(x-y) f(y) dy$$

converges absolutely whenever $f \in L^2(\mathbb{R}^n)$ and $x \notin \operatorname{supp}(f)$. Suppose the following:

(i) T is bounded on $L^2(\mathbb{R}^n)$: $\| Tf \|_2 \leq A \| f \|_2$.

(ii) The kernel K verifies the Hörmander condition (2.7) with constant B.

Then,

(a) T is bounded on $L^p(\mathbb{R}^n)$, $1 < p < \infty$, and

$$\| Tf \|_p \leq C_{n,p} \| f \|_p$$

for $f \in L^p(\mathbb{R}^n)$ and $C_{n,p}$ only depending on n, p, A and B.

(b) T is weak-type $(1,1)$, i.e, for all $\lambda > 0$ and $f \in L^1(\mathbb{R}^n)$

$$\lambda |\{x \in \mathbb{R}^n : |Tf(x)| > \lambda\}| \leq C_n \| f \|_1$$

where C_n is a constant only depending on the dimension n, A and B.

The proofs of (some slight variants of) this theorem are available in [10], Chapter 2, Section 2 and [3], Chapter 5, Section 1. Often, condition (i) of Theorem 2.9 is encapsulated in other, perhaps more practical, hypotheses, such as the Fourier transform of the kernel being uniformly bounded, $\hat{K} \leq A$. We are skipping the proof of the theorem because a more general one is going to be discussed in detail in due time.

Corollary 2.10. *The Hilbert transform is a bounded operator on $L^p(\mathbb{R})$ for $1 < p < \infty$.*

Proof. Since the multiplier function of the Hilbert transform $m_H(\xi) = -i\text{sgn}(\xi)$ is a bounded function, condition (i) in Theorem 2.9 is fulfilled. That the gradient condition (2.8) is satisfied by the kernel $\frac{1}{x}$, and thus so is the Hörmander condition for (ii), is straightforward and completes the proof. □

2.2.2. Singular kernels on measure metric spaces

The first significant issue that we encounter when attempting to generalise what was done in the previous section is the fact that in a general $((X,d),\Sigma,\mu)$, subtracting points $x - y$ makes no sense because we lack a group structure. Thus, we can no longer understand convolution as we did in $\mathbb{R}^n$. Since we unavoidably need two variables to input into K (an integration variable and a variable for the resulting function Tf), we are going to get around this obstacle by considering 2-variable kernels, $K(x,y)$, that are assumed to blow up and be troublesome around $x = y$.

Immediately afterwards, we need to reformulate the Hörmander condition. As pointed out, we substitute $K(x-y)$ by $K(x,y)$. What do we swap $K(x)$ for, then? We should not give preference to any particular point in X; we do not even have an origin now, so instead we are going to introduce $K(x, y_0)$ and include y_0 in the supremum. This choice is going to lead to success in the corresponding proof. Finally, we shall drop the 2 factor in the integration domain and introduce a certain constant $C > 1$ for versatility and generality. All in all:

Definition 2.11. *A kernel K on the product measure space $((X, d), \Sigma, \mu) \times ((X, d), \Sigma, \mu)$ is said to satisfy the **Hörmander condition** if*

$$\sup_{y, y_0 \in X} \int_{d(x,y) \geq C d(y, y_0)} |K(x, y) - K(x, y_0)| \, d\mu(x) = B < \infty \tag{2.9}$$

for some constants $B > 0$ and $C > 1$.

Theorem 2.12. *Let T be a linear operator such that there exists a measurable kernel function K such that*

$$Tf(x) = \int_X K(x, y) f(y) d\mu(y)$$

converges absolutely whenever $f \in L^q(X)$ and $x \notin \operatorname{supp}(f)$. Suppose the following:

(i) T is bounded on $L^q(X)$ for some $1 < q \leq \infty$: $\| Tf \|_q \leq A \| f \|_q$.

(ii) The kernel measurable function K verifies the Hörmander condition (2.9) with constants B and C.

Then,

(a) T is bounded on $L^p(X) \cap L^q(X)$, $1 < p < q$, and

$$\| Tf \|_p \leq C_p \| f \|_p$$

for $f \in L^p(X) \cap L^q(X)$, and C_p only depending on p, q, A, B and C.

(b) T is weak-type $(1, 1)$ in the sense that for any $\lambda > 0$,

$$\lambda \, \mu\{x \in X : |Tf(x)| > \lambda\} \leq C_1 \| f \|_1$$

for $f \in L^1(X) \cap L^q(X)$ and some constant C_1 depending on q, A, B and C.

Proof. We aim at showing that T is weak-type $(1, 1)$ on $L^1(X) \cap L^q(X)$ so that Marcinkiewicz interpolation theorem applies. With this purpose, let $\lambda > 0$ and take $f \in L^1(X) \cap L^q(X)$. Since $f \in L^q(X)$, Tf is well defined and belongs to $L^q(X)$ by assumption (i). Also, thanks to the fact that $f \in L^1(X)$, we are allowed to invoke the Calderón-Zygmund decomposition, Corollary 2.4, on f at height λ, that is, $f = g + b$, with g being essentially bounded and b averages 0 on the sets Q_k obtained from Remark 2.6.

$$f = g + b \implies Tf = Tg + Tb \implies |Tf| \leq |Tg| + |Tb| \implies$$
$$\{x \in X : |Tf(x)| > \lambda\} \subseteq \left\{x \in X : |Tg(x)| > \frac{\lambda}{2}\right\} \cup \left\{x \in X : |Tb(x)| > \frac{\lambda}{2}\right\}$$

because if $|Tf(x)| > \lambda$ then it cannot happen that both $|Tg(x)|$ and $|Tb(x)|$ are smaller than $\frac{\lambda}{2}$. Hence, by countable subadditivity,

$$\mu(\{x \in X : |Tf(x)| > \lambda\}) \leq \mu\left(\left\{x \in X : |Tg(x)| > \frac{\lambda}{2}\right\}\right) + \mu\left(\left\{x \in X : |Tb(x)| > \frac{\lambda}{2}\right\}\right). \quad (2.10)$$

We will be done when we show the weak-type $(1,1)$ estimates for g and for b separately.

g is weak-type $(1,1)$: We use the assumption that T is bounded on $L^q(X)$[2].

$$\left(\frac{\lambda}{2}\right)^q \mu\left(\left\{x \in X : |Tg(x)| > \frac{\lambda}{2}\right\}\right) \leq \int_{\left\{x : |Tg(x)| > \frac{\lambda}{2}\right\}} |Tg(x)|^q d\mu(x)$$
$$\leq \int_X |Tg(x)|^q d\mu(x) = \| Tg \|_q^q \leq A^q \| g \|_q^q = A^q \int_X |g(x)|^{q-1} |g(x)| d\mu(x)$$

Since $q > 1$ and g is essentially bounded, $|g(x)|^{q-1} \leq (C_{CZ}\lambda)^{q-1}$ (the constant comes from Theorem 2.3, and is an absolute constant of the measure space X) a.e. $x \in X$. Thus

$$A^q \int_X |g(x)|^{q-1} |g(x)| d\mu(x) \leq A^q C_{CZ}^{q-1} \lambda^{q-1} \| g \|_1,$$

all in all yielding

$$\lambda \mu\left(\left\{x \in X : |Tg(x)| > \frac{\lambda}{2}\right\}\right) \leq (2A)^q C_{CZ}^{q-1} \| g \|_1 \leq (2A)^q C_{CZ}^{q-1} \| f \|_1.$$

b is weak-type $(1,1)$: Write $\Omega^{**} := \bigcup_k B_k^{**}$ (with the notation of Lemma 2.5) and notice that

$$\left\{x \in X : |Tb(x)| > \frac{\lambda}{2}\right\} = \left\{x \in \Omega^{**} : |Tb(x)| > \frac{\lambda}{2}\right\} \bigcup \left\{x \in \Omega^{**c} : |Tb(x)| > \frac{\lambda}{2}\right\}.$$

[2] The argument for $q = \infty$ needs a slight modification. Given an essentially bounded function g, by assumption Tg is essentially bounded. Instead of decomposing f at height λ (the parameter for the weak-type estimate), we may choose a height such that Tg is essentially bounded by $\frac{\lambda}{2}$. This way, $\mu\left(\left\{x \in X : |Tg(x)| > \frac{\lambda}{2}\right\}\right) = 0$ so we only need to control the term $\mu\left(\left\{x \in X : |Tb(x)| > \frac{\lambda}{2}\right\}\right)$.

The measure of the first set on the right-hand side can be bounded using the doubling property of the measure and the maximal theorem, Theorem 1.39 for the general setting as follows:

$$\mu\left\{x\in\Omega^{**}:|Tb(x)|>\frac{\lambda}{2}\right\}\leq\mu\{x\in\Omega^{**}\}$$
$$\leq c^{**}\mu\{x\in\Omega\}=c^{**}\mu\{x\in X:\mathfrak{M}f(x)>\lambda\}\leq\frac{A_{\mathrm{HL}}c^{**}}{\lambda}\parallel f\parallel_1. \quad (2.11)$$

Bounding the measure of the second set on the right-hand side is a bit more laborious. Accounting for the nature of b, one can write

$$b(x)=\sum_k b_k(x)$$

where b_k is supported on Q_k and it averages 0 there.

$$\frac{\lambda}{2}\mu\left(\left\{x\in\Omega^{**c}:|Tb(x)|>\frac{\lambda}{2}\right\}\right)\leq\int_{\{x\in\Omega^{**c}:|Tb(x)|>\frac{\lambda}{2}\}}|Tb(x)|d\mu(x)$$
$$\leq\int_{\Omega^{**c}}|Tb(x)|d\mu(x)=\int_{\Omega^{**c}}|\sum_k Tb_k(x)|d\mu(x)$$

We are now entitled to write explicitly T thanks to the integration domain, which is disjoint with the support of each b_k.

$$\int_{\Omega^{**c}}|\sum_k Tb_k(x)|d\mu(x)\leq\sum_k\int_{\Omega^{**c}}\left|\int_{Q_k}K(x,y)b_k(y)d\mu(y)\right|d\mu(x)$$
$$=\sum_k\int_{\Omega^{**c}}\left|\int_{Q_k}(K(x,y)-K(x,y_k))b_k(y)d\mu(y)\right|d\mu(x)$$

for fixed $y_k\in Q_k$. In the last step, we introduced a substracting term in the inner integral, which is licit thanks to the property that b_k integrates 0 on Q_k. Next, apply the triangle inequality and Fubini-Tonelli theorem, with the intention to reach the position of using the Hörmander condition.

$$\sum_k\int_{\Omega^{**c}}\int_{Q_k}|K(x,y)-K(x,y_k)||b_k(y)|d\mu(y)d\mu(x)$$
$$=\sum_k\int_{Q_k}\int_{\Omega^{**c}}|K(x,y)-K(x,y_k)||b_k(y)|d\mu(x)d\mu(y)$$
$$\leq\sum_k\int_{Q_k}\int_{B_k^{**c}}|K(x,y)-K(x,y_k)|d\mu(x)|b_k(y)|d\mu(y)$$

The last step is an overestimation of the integration domain. Because of $y,y_k\in Q_k\subseteq B_k^*\subseteq B_k^{**}$, we have $d(y,y_k)<2c^*\,r_k$, where $r_k>0$ is the radius of the ball

B_k and $c^* > 1$ is the common dilation factor for the family $(B_k^*)_{k\in\mathbb{N}}$. On the other hand, since $x \in B_k^{**c}$, $d(x,y) > (c^{**} - c^*)r_k$, where $c^{**} > 1$ is the common dilation factor for the family $(B_k^{**})_{k\in\mathbb{N}}$. It is desired that $c^{**} - c^* \geq 2Cc^*$ (here, C is the constant present in the integration domain in the Hörmander condition) so that the conclusion $d(x,y) \geq Cd(y,y_k)$ yields. This happens if $c^{**} \geq (1+2C)c^*$, which may not be the case. However, in the unfavourable case, we can enlarge the c^{**} factor so that we get $c^{**} \geq (1+2C)c^*$. This makes the family of balls $(B_k^{**})_k$ larger, but this preserves their property of intersecting $F^c = \Omega$ and each B_k^{**} still contains B_k^*. All in all, in either case the argument is still valid and may be carried on.

$$\sum_k \int_{Q_k} \int_{B_k^{**c}} |K(x,y) - K(x,y_k)| d\mu(x) |b_k(y)| d\mu(y)$$
$$\leq \sum_k \int_{Q_k} \int_{d(x,y)>Cd(y,y_k)} |K(x,y) - K(x,y_k)| d\mu(x) |b_k(y)| d\mu(y)$$
$$\leq B \sum_k \int_{Q_k} |b_k(y)| d\mu(y)$$

thanks to the Hörmander condition on each inner integral. It only remains to tackle a simple computation:

$$B \sum_k \int_{Q_k} |b_k(y)| d\mu(y) = B \int_\Omega |b(y)| d\mu(y) \leq B \| b \|_1 \leq B \| f \|_1 .$$

Putting everything together,

$$\mu(\{x \in X : |Tf(x)| > \lambda\}) \leq (2A)^q C_{CZ}^{q-1} \frac{\| f \|_1}{\lambda} + A_{\mathrm{HL}} c^{**} \frac{\| f \|_1}{\lambda} + B \frac{\| f \|_1}{\lambda}$$
$$= C_1(A,B,C,q) \frac{\| f \|_1}{\lambda}$$

which concludes the proof of (b).

(a) follows from Marcinkiewicz interpolation theorem (Theorem 1.16) applied to the range of p between the endpoints 1 and q. □

Conclusions (a) and (b) of the theorem require f to live in $L^p(X) \cap L^q(X)$ ($p = 1$ in (b)) just to make sure T is well defined. Of course, since $L^p(X) \cap L^q(X)$ is always dense in $L^p(X)$ for all finite $1 \leq p < \infty$, T extends by continuity (equivalent to the boundedness just shown) to a unique operator acting on the whole $L^p(X)$ space (with $p = 1$ in (b)).

One further step is attempting to get the full range of boundedness $1 < p < \infty$ for our operator T despite only relying on the assumption that T is bounded on $L^q(X)$ for finite $q < \infty$ (so that Marcinkiewicz interpolation theorem does not yield the full range). In this direction, it is wise to use duality tools since the dual space of $L^p(X)$ is isomorphic to $L^{p'}(X)$ (for finite conjugate exponents such that $\frac{1}{p} + \frac{1}{p'} = 1$) and if $p \le q$ then $p' \ge q'$. Consider the adjoint operator $T^* : L^{p'}(X) \to L^{p'}(X)$, $g \in L^{p'}(X)$ and $f \in L^p(X)$. A corollary of Hahn-Banach theorem unlocks the characterization of norms by duality as follows.

$$\begin{aligned} \| T^* g \|_{p'} = \sup_{\| f \|_p \le 1} | \langle T^* g, f \rangle | &= \sup_{\| f \|_p \le 1} | \langle g, Tf \rangle | \\ &\le \sup_{\| f \|_p \le 1} \| Tf \|_p \| g \|_{p'} \le C_p \| g \|_{p'} \end{aligned} \tag{2.12}$$

The reasoning is: from the boundedness of T acting on $L^q(X)$ (and the Hörmander condition) we deduce that T is bounded on $L^p(X)$ for $1 < p < q$. From here, T^* acts boundedly on $L^{p'}(X)$ for $q' < p' < \infty$. At this point, we dream of somehow transferring the latter property to T acting on $L^{p'}(X)$ for $q' < p' < \infty$.

If we swap the roles of T and T^* in (2.12), we have that T is bounded on $L^{p'}(X)$ for $q' < p' < \infty$

$$\| Tg \|_{p'} \le C_p \| g \|_{p'}$$

if and only if T^* is bounded on $L^p(X)$ for $1 < p < q$.

We shall investigate the anatomy of the "kernel" functions of T and T^*. Let $f \in L^p(X)$ and $g \in L^{p'}(X)$ be functions such that $\mathrm{supp}(f) \cap \mathrm{supp}(g) = \emptyset$. Under this condition,

$$\langle Tf, g \rangle = \int_X g(x) \int_X K(x,y) f(y) d\mu(y)\, d\mu(x)$$

is well defined. Thanks to Theorem 2.12, we know that $Tf \in L^p(X)$, and so it is finite μ-almost everywhere. Also, the whole outer integral is absolutely convergent by Hölder inequality, which enables us to use Fubini-Tonelli theorem:

$$\begin{aligned} &\int_X g(x) \int_X K(x,y) f(y) d\mu(y)\, d\mu(x) \\ &\qquad = \int_X f(y) \int_X K(x,y) g(x) d\mu(x)\, d\mu(y) = \langle T^* g, f \rangle \end{aligned} \tag{2.13}$$

from where we see that

$$T^* g(x) = \int_X K(y,x) g(y) d\mu(y) = \int_X K^*(x,y) g(y) d\mu(y) \tag{2.14}$$

where K^* is the kernel of the adjoint operator, provided f and g had disjoint support. It is to be remarked that Eq. (2.13) does not completely define the operator T^* on all of its domain, but yet gives useful information: Whenever $x \notin \text{supp}(g)$, (2.14) holds. Realise that $K^*(x,y) = K(y,x)$. Therefore, this argument of duality will work only if the kernel behaves nicely when swapping their pair of input variables.

As seen, T^* is very close to satisfying the conditions of Theorem 2.12: it is linear, it can be written in terms of a kernel K^* providing the technicalities described and it is bounded on $L^{q'}(X)$. If only such kernel fulfilled the corresponding Hörmander condition, T^* would be bounded for exponents $1 < p' < q'$, meaning by the previous reasoning that T would be bounded for $q < p < \infty$. However, it is in general impossible to deduce a Hörmander condition for K^* based on that for K, meaning that we need to assume it as an extra.

Corollary 2.13. *Under the assumptions of Theorem 2.12 and supposing an extra Hörmander condition for the kernel K with swapped variables,*

(iii)

$$\sup_{y,y_0 \in X} \int_{d(x,y) \geq C' d(y,y_0)} |K(y,x) - K(y_0,x)| \, d\mu(x) = B' < \infty \tag{2.15}$$

for some $B' > 0$ and $C' > 1$,

then the following bonus track holds.

(c) T is bounded on $L^{p'}(X)$ for $q' < p' < \infty$ and

$$\| Tf \|_{p'} \leq C_{p'} \| f \|_{p'}$$

for all $f \in L^{p'}(X)$ and a constant $C_{p'} > 0$ depending on p, q, A, B' and C'.

For instance, if the setting is the euclidean space $X = \mathbb{R}^n$ and so the kernel is intended for convolution, $K(x,y) = K_E(x-y)$, then one can write

$$K^*(x,y) := K(y,x) = K_E(y-x) = K_E(-x-(-y)) = K(-x,-y).$$

Clearly in this case, if K satisfies the original Hörmander condition, so does K^* and no further assumption is required. It is then possible to show by duality that, in $\mathbb{R}^n$, if T is bounded on $L^q(\mathbb{R}^n)$, then T is bounded not only in $L^p(\mathbb{R}^n)$ for $1 < p \leq q$ but also for $q' \leq p < \infty$. Notice that if $q = 2$ (which is usually a convenient scenario due to the Hilbert space structure and the availability of Plancherel theorem, which unlocks the characterisation of bounded linear translation-invariant operators on $L^2(\mathbb{R}^n)$ by means of the boundedness of their Fourier multiplier functions) one obtains the full range $1 < p < \infty$.

CHAPTER 3
Vector-valued extensions

3.1. Integration of vector-valued functions

So far, we have only considered complex-valued functions and complex-valued kernels that act on such functions by product of complex numbers. Thinking of vector-valued functions bodes well for a generalization of the Calderón-Zygmund theory. It is of special interest to apply this broader theory to sophisticated operators, for instance square functions and maximal singular operators.

We follow a similar approach to the ones in [3] and [5]. To begin with, since we would like to, in some sense, integrate vector-valued functions, we need to give a precise meaning to generalised concepts of measurability of functions, Lebesgue spaces and integrability. In the context of this chapter, we are going to work in the setting of the following definition.

Definition 3.1. *Let $(X, \Sigma, \mu) \equiv X$ be a σ-finite measure space and let B be a Banach space. A function $F : X \to B$ is said to be* ***measurable*** *if both of the following conditions hold.*

(i) There is a separable subspace $B_F \subseteq B$ such that $F(x) \in B_F$ for μ-a.e. $x \in X$.

(ii) For every element of the dual space $b^* \in B^*$, the complex-valued function $x \to \langle b^*, F(x) \rangle$ is measurable in the usual sense.

The main goal of requirement *(ii)* is to ensure that norms of vector-valued functions are going to be measurable, and thus candidates to be integrated.

Namely, the function $\| F \|_B : x \to \| F(x) \|_B$ is measurable[1] as a real-valued function.

On the other hand, by assuming (i), we are enabling the forthcoming definition of integration of vector-valued functions via an extension from a dense class. Note that we could have assumed, with less generality, that B is separable straightaway.

We appreciate having an analogue of the Lebesgue spaces.

Definition 3.2. *Let (X, Σ, μ) be a measure space and let B be a Banach space. For every $1 \leq p \leq \infty$ let*

$$L^p_B(X) := \{F : X \to B \text{ measurable } : \int_X \| F(x) \|^p_B \, d\mu(x) < \infty\} \tag{3.1}$$

*(in the conventional understanding that the integral is substituted by an essential supremum when $p = \infty$) be the set of (equivalence classes of) vector-valued measurable functions the p-th power of the norm of which is integrable (in the real Lebesgue sense). Call them **vector-valued Lebesgue spaces**, or simply Lebesgue spaces if no confusion may occur.*

It turns out that for all $1 \leq p \leq \infty$ and Banach space B, $L^p_B(X)$ is also a Banach space, equipped with the expectable norm

$$\| F \|_{L^p_B(X)} = \left(\int_X \| F(x) \|^p_B \, d\mu(x) \right)^{\frac{1}{p}}$$

or

$$\| F \|_{L^\infty_B(X)} = \operatorname*{ess\,sup}_{x \in X} \| F(x) \|_B \, .$$

Similarly, one may define the weak-$L^p_B(X)$ spaces.

Definition 3.3. *Let (X, Σ, μ) be a measure space and let B be a Banach space. For every $1 \leq p \leq \infty$, let*

$$L^{p,\infty}_B(X) := \{F : X \to B \text{ measurable } : \sup_{\lambda > 0} \lambda^p \mu(\{x \in X : \| F(x) \|_B > \lambda\}) < \infty\} \tag{3.2}$$

[1] Characterise the norm by duality: $\| F(x) \|_B = \sup_{\| b^* \|_{B^*} \leq 1} | \langle b^*, F(x) \rangle |$ (we can take B^* to be the dual space of B_F instead of B, whenever $x \in X$ allows us). We invoke mighty theorems from functional analysis. For our Banach space B, we know that the dual closed unit ball $\overline{B^*_1(0)}$ is compact in the weak* topology. Not only this, but since B is separable, $\overline{B^*_1(0)}$ equipped with the weak* topology is metrizable. Now, any compact metric space is separable, thus $\overline{B^*_1(0)}$ with the weak* topology is separable. By definition of the weak* topology, the maps $b^* \to | \langle b^*, F(x) \rangle |$ are continuous in such a topology, meaning that we can take a countable dense set in $\{b^* : \| b^* \|_{B^*} \leq 1\}$, say $\{b^*_n\}_{n \in \mathbb{N}}$ to write $\| F(x) \|_B = \sup_{n \in \mathbb{N}} | \langle b^*_n, F(x) \rangle |$ and conclude that $\| F(x) \|_B$ is a measurable function of x owing to the supremum over a countable set and the assumption that $| \langle b^*_n, F(x) \rangle |$ are measurable functions of x μ-a.e.

(the convention for $p = \infty$ is $L_B^{\infty,\infty}(X) = L_B^\infty(X)$) be the set of (equivalence classes of) vector-valued measurable functions such that such supremum is finite. Name them ***vector-valued weak Lebesgue spaces****, or simply weak Lebesgue spaces if there is no risk of confusion.*

As an example, take $f \in L^p(X)$ (a classic complex-valued Lebesgue function) and $b \in B$ an element of some complex Banach space. Consider, for every $x \in X$, the product $f(x)b \in B$ a.e. $x \in X$. Let us check Definition 3.2 for this element.

$$\begin{aligned}\int_X \| f(x)b \|_B^p \, d\mu(x) &= \int_X |f(x)|^p \| b \|_B^p \, d\mu(x) \\ &= \int_X |f(x)|^p d\mu(x) \| b \|_B^p = \| f \|_p^p \| b \|_B^p < \infty\end{aligned}$$

From here, we conclude that $f(x)b \in L_B^p(X)$. Moreover,

$$\| fb \|_{L_B^p(X)} = \| f \|_p \| b \|_B .$$

Let $L^p(X) \otimes B := \operatorname{span}\{fb : f \in L^p(X), b \in B\}$ be the set of all finite linear combinations of elements of the form just treated.

Proposition 3.4. *$L^p(X) \otimes B$ is dense in $L_B^p(X)$ for $1 \leq p < \infty$.*

Proof. We find help in the σ-finiteness of the measure space X. Let $F \in L_B^p(X)$ and let $X_j \nearrow X$ be a countable sequence of increasing measurable sets tending to the total space such that $0 < \mu(X_j) < \infty \ \forall j \in \mathbb{N}$. By assumption (i) in Definition 3.1, we can take a countable set $\{b_n\}_{n\in\mathbb{N}} \subset B$ which is dense in the image of F, except for a set of measure 0 which we may ignore.

Fix a global $\epsilon > 0$. By density, for each $j \in \mathbb{N}$ and $\epsilon_j > 0$ (that we are going to choose later on) $\forall x \in X_j \ \exists n \in \mathbb{N}$ such that $\| F(x) - b_n \|_B < \epsilon_j$. Next, we construct a measurable function to associate with each x and $F(x)$ a $b_n \in \{b_n\}_{n\in\mathbb{N}}$ that is close enough to $F(x)$. Define[2]

$$\begin{aligned}\Phi_j : \ X_j &\longrightarrow \mathbb{N} \\ x &\longrightarrow \min\{n \in \mathbb{N} : \| F(x) - b_n \| < \epsilon_j\}.\end{aligned} \tag{3.3}$$

At this point, we have grouped all $x \in X_j$ into the measurable disjoint sets $X_j^n := \{x \in X_j : \Phi_j(x) = n\}$, $X_j = \bigsqcup_{n\geq 1} X_j^n$. With the will of constructing an

[2]The functions Φ_j are measurable because one can inductively express the preimages of singletons as $\Phi_j^{-1}(n) = \{x \in X_j : \| F(x) - b_n \|_B < \epsilon_j\} \bigcap_{i=1}^{n-1} (\Phi_j^{-1}(i))^c \ \forall n \geq 1$. Of course, the maps $x \to \| F(x) - b_n \|_B$ are measurable in the restricted measure space (X_j, Σ_j, μ), in alignment with Definition 3.1.

element of $L^p(X)\otimes B$ which resembles F, define, for each $n\in\mathbb{N}$, the measurable functions

$$f_n(x) := \begin{cases} 1 & x\in X_j^n \\ 0 & x\notin X_j^n. \end{cases}$$

In the remaining of the proof, we check that F is approximated by the sequence of elements $\sum_{m=1}^N f_m b_m \in L^p(X)\otimes B$ indexed by $N\in\mathbb{N}$. First, estimate the integrals over each X_j^n with $n\le N$.

$$\begin{aligned}\int_{X_j^n}\left\| F(x)-\sum_{m=1}^N f_m(x)b_m\right\|_B^p d\mu(x) &= \int_{X_j^n} \| F(x)-b_n \|_B^p \, d\mu(x) \\ &\le \epsilon_j^p \mu(X_j^n) = 2^{-j}\frac{\mu(X_j^n)}{\mu(X_j)}\epsilon\end{aligned}$$

once we have chosen $\epsilon_j^p := \frac{2^{-j}}{\mu(X_j)}\epsilon$. Splitting the global integral according to j and N, we reach

$$\begin{aligned}\int_X \left\| F(x)-\sum_{m=1}^N f_m(x)b_m\right\|_B^p d\mu(x) &\le \sum_{j\ge1}\int_{X_j}\left\| F(x)-\sum_{m=1}^N f_m(x)b_m\right\|_B^p d\mu(x) \\ &= \sum_{j\ge1}\int_{\cup_{n=1}^N X_j^n}\left\| F(x)-\sum_{m=1}^N f_m(x)b_m\right\|_B^p d\mu(x) \\ &\quad + \underbrace{\sum_{j\ge1}\int_{\cup_{n=N+1}^\infty X_j^n}\left\| F(x)-\sum_{m=1}^N f_m(x)b_m\right\|_B^p d\mu(x)}_{E(N)} \\ &\le \sum_{j\ge1}\sum_{n=1}^N 2^{-j}\frac{\mu(X_j^n)}{\mu(X_j)}\epsilon + E(N) \le \sum_{j\ge1} 2^{-j}\epsilon + E(N) = \epsilon + E(N).\end{aligned}$$

Proving that the error term $E(N)$ tends to zero as $N\to\infty$ would conclude the proof. After realizing that the inner sum vanishes in the integration domain of $E(N)$, that

$$E(N) = \int_{\cup_{j\ge1}\cup_{n=N+1}^\infty X_j^n} \| F(x) \|_B^p \, d\mu(x) \to 0 \text{ as } N\to\infty$$

is true can be shown by the dominated convergence theorem and the fact that $F\in L_B^p(X)$. □

There is a reasonable way of defining integration of elements in $L^p(X)\otimes B$ making the most of the $L^p(X)$ integration structure.

Definition 3.5. *Let fb be a tensor product of a function $f \in L^1(X)$ and some Banach space element $b \in B$. Define its* ***integral*** *over the measure space X:*

$$\int_X f(x)b \, d\mu(x) := \left(\int_X f(x) d\mu \right) \forall \, b \in B.$$

This definition extends to all $L^1(X) \otimes B$ by imposing linearity to this integral operation.

Resembling the triangle inequality, we get the following inequality for the integral in the previous definition. It is going to become essential for the proof of Theorem 3.9.

Proposition 3.6. *If $F_t \in L^1(X) \otimes B$, then $\| \int_X F_t d\mu \|_B \leq \| F_t \|_{L^1_B(X)}$.*

Proof. The statement is pretty obvious for simple functions: Let $F_t(x) = \sum_{n=1}^N \sum_{k=1}^K a_n^k \mathbb{1}_{X^k}(x) b_n$ be a linear combination of integrable simple functions multiplied by some elements in B , were a_n^k are their complex values and X^k are measurable sets. Then,

$$\begin{aligned} \left\| \int_X F_t d\mu \right\|_B &= \left\| \int_X \sum_{n=1}^N \sum_{k=1}^K a_n^k \mathbb{1}_{X^k}(x) b_n d\mu(x) \right\|_B = \left\| \sum_{k=1}^K \sum_{n=1}^N a_n^k \mu(X^k) b_n \right\|_B \\ &\leq \sum_{k=1}^K \left\| \sum_{n=1}^N a_n^k b_n \right\|_B \mu(X^k) = \sum_{k=1}^K \int_{X^k} \left\| \sum_{n=1}^N a_n^k b_n \right\|_B d\mu(x) \\ &= \int_X \left\| \sum_{n=1}^N \sum_{k=1}^K a_n^k \mathbb{1}_{X^k}(x) b_n \right\|_B d\mu(x) = \| F_t \|_{L^1_B(X)} . \end{aligned}$$

The last step is legit due to the indicator functions of disjoint sets. If we denote by $S(X)$ the space of integrable simple functions over X, taking into account that the inclusion $S(X) \otimes B \subset L^1(X) \otimes B$ is a dense inclusion, we are done by continuity.

□

What Proposition 3.6 tells us is that the linear map of Banach spaces

$$\begin{aligned} \int_X \cdot \, d\mu : \; L^1(X) \otimes B \subset L^1_B(X) &\longrightarrow \quad B \\ F_t \quad &\longrightarrow \int_X F_t d\mu \end{aligned} \tag{3.4}$$

is continuous (equivalently bounded). Therefore, by Proposition 3.4, we can extend the operator (3.4) by continuity uniquely to an operator acting now on all of the $L^1_B(X)$ space. We shall use the same notation for the extended operator

$$\begin{aligned} \int_X \cdot \, d\mu : \; L^1_B(X) &\longrightarrow \quad B \\ F \quad &\longrightarrow \int_X F d\mu. \end{aligned} \tag{3.5}$$

Proposition 3.7. *For $F \in L^1_B(X)$, the element $\int_X F d\mu \in B$ is characterised by*

$$\left\langle b^*, \int_X F \, d\mu \right\rangle = \int_X \langle b^*, F(x) \rangle \, d\mu(x) \quad \forall \, b^* \in B^*. \tag{3.6}$$

Proof. Checking that $\int_X F d\mu$ is the only element in B that satisfies (3.6) is immediate by the customary way of assuming there exists a different element $b \in B$ that also verifies (3.6) and reaching the contradiction that $\int_X F d\mu - b = 0 \in B$.

To prove that $\int_X F d\mu$ built as in (3.5) satisfies (3.6), let $F \in L^1_B(X)$ and write $F = F_t + G$, where $F_t \in L^1(X) \otimes B$ and $G \in L^1_B(X)$ is such that, fixed $\epsilon > 0$, $\| G \|_{L^1_B(X)} < \epsilon$ (which is permitted thanks to Proposition 3.4).

Checking (3.6) for F_t is straightforward by linearity. The density argument reads as follows.

$$\begin{aligned} &\left\langle b^*, \int_X F \, d\mu \right\rangle = \left\langle b^*, \int_X F_t + \int_X G \, d\mu \right\rangle \\ &= \left\langle b^*, \int_X F_t \, d\mu \right\rangle + \left\langle b^*, \int_X G \, d\mu \right\rangle = \int_X \langle b^*, F_t(x) \rangle \, d\mu(x) + \left\langle b^*, \int_X G \, d\mu \right\rangle \\ &\qquad = \int_X \langle b^*, F(x) \rangle \, d\mu(x) - \int_X \langle b^*, G(x) \rangle \, d\mu(x) + \left\langle b^*, \int_X G \, d\mu \right\rangle \end{aligned}$$

Eventually, it remains to show that the second and third terms are arbitrarily small. For the second term, let us use duality of norms.

$$\begin{aligned} \left| \int_X \langle b^*, G(x) \rangle \, d\mu(x) \right| \leq \int_X \| b^* \|_{B^*} \| G(x) \|_B \, d\mu(x) \\ = \| b^* \|_{B^*} \| G \|_{L^1_B(X)} \leq \| b^* \|_{B^*} \, \epsilon \end{aligned}$$

For the third term, we also rely on the inequality in Proposition 3.6.

$$\left| \left\langle b^*, \int_X G \, d\mu \right\rangle \right| \leq \| b^* \|_{B^*} \left\| \int_X G \, d\mu \right\|_B \leq \| b^* \|_{B^*} \| G \|_{L^1_B(X)} < \| b^* \|_{B^*} \, \epsilon$$

This route led to (3.6). □

Having developed a grounding for integrating vector-valued functions, we are in a position to shoot for an extension of our main theorem, Theorem 2.12, now in this new vaster setting. We are going to showcase how to make the most of the Calderón-Zygmund theory developed in Chapter 2.

The problem of, given a linear $L^p(\mathbb{R}^n)$-bounded operator, attempting to show that it has a vector-valued extension was faced a couple decades before the birth of the Calderón-Zygmund theory. Efforts were made to try to show estimates of the kind: Given a particular known linear bounded operator on $L^p(\mathbb{R}^n)$, H (say, for instance, the Hilbert transform, if $n = 1$), choose $B = \ell^q(\mathbb{R})$ and prove that the extended linear operator $\tilde{H} : L^p_{\ell^q(\mathbb{R})}(\mathbb{R}) \to L^p_{\ell^q(\mathbb{R})}(\mathbb{R})$ given by $\tilde{H}(\{f_j\}_{j\in\mathbb{N}}) := \{Hf_j\}_{j\in\mathbb{N}}$ is bounded, namely

$$\left\| \left(\sum_{j\geq 1} |Hf_j|^q \right)^{\frac{1}{q}} \right\|_p \leq C_{p,q} \left\| \left(\sum_{j\geq 1} |f_j|^q \right)^{\frac{1}{q}} \right\|_p . \tag{3.7}$$

Even though, with enough imagination, one can come up with a myriad of operators acting on vector-valued functions, it was wildly difficult to shelter a complete amount of them under the same theory. It was not until the machinery of Calderón-Zygmund was discovered that we got a general enough theory for akin operators.

Retrieving the idea of operators given by integration against, in some sense, a kernel, since we would like to work with vector-valued functions, the kernel $K(x, y)$ is no longer going to be a function, but a linear operator mapping Banach spaces into Banach spaces. Let $\mathcal{L}(A, B)$ denote the Banach space of bounded linear operators mapping the Banach space A into the Banach space B. In order to carefully select the operators we are homing in on, we consider K with domain in the product measure space $X\times X$, $K(x, y)$ being ill-defined along the diagonal $x = y$, and taking values $K(x, y) \in \mathcal{L}(A, B)$. In similarity with the hypotheses of Theorem 2.12, let us assume K to be measurable (in the sense of Definition 3.1) and locally integrable away from the diagonal. Moreover, whenever $F \in L^\infty_A(X)$ has compact support and $x \notin \operatorname{supp}(F)$, the object of study T is given by

$$TF(x) = \int_X K(x, y)F(y)d\mu(y).$$

Under such assumptions, for $x \notin \operatorname{supp}(F)$,

$$\begin{aligned} \| Tf(x) \|_B &\leq \int_{\operatorname{supp}(F)\subseteq K'\subseteq X} \| K(x, y)F(y) \|_B \, d\mu(y) \\ &\leq \| F \|_{L^\infty_A(X)} \int_{\operatorname{supp}(F)\subseteq K'\subseteq X} \| K(x, y) \|_{\mathcal{L}(A,B)} \, d\mu(y) < \infty \end{aligned}$$

meaning that $Tf(x)$ is a well defined element of B if $x \notin \operatorname{supp}(F)$.

It is not to be forgotten that we should impose some kind of Hörmander condition. Analogously to Definition 2.11:

Definition 3.8. *A kernel K on the product measure space $((X,d),\Sigma,\mu) \times ((X,d),\Sigma,\mu)$ taking values in $\mathcal{L}(A,B)$ is said to satisfy the **Hörmander condition** if*

$$\sup_{y,y_0 \in X} \int_{d(x,y) \geq Cd(y,y_0)} \| K(x,y) - K(x,y_0) \|_{\mathcal{L}(A,B)} \, d\mu(x) = D < \infty \tag{3.8}$$

for some constants $C > 1$ and $D > 0$.

Astonishingly, the natural generalization of Theorem 2.12 turns out to work in this setting as well!

Theorem 3.9. *Let $((X,d),\Sigma,\mu)$ be a measure metric space with the doubling property. Let A, B be Banach spaces and let T be a linear operator which is represented by*

$$TF(x) = \int_X K(x,y)F(y)d\mu(y)$$

whenever $F \in L^\infty_A(X)$ with compact support and $x \notin \operatorname{supp}(F)$, where the vector-valued kernel $K \in \mathcal{L}(A,B)$ is measurable in $X \times X$ and locally integrable away from the diagonal. Assume that

(i) T is bounded from $L^q_A(X)$ to $L^q_B(X)$ for a fixed $1 < q \leq \infty$, $\| TF \|_{L^q_B(X)} \leq C_q \| F \|_{L^q_A(X)}$, and

(ii) the operator kernel K satisfies the Hörmander condition in (3.8) with constants C and D.

Then,

(a) the operator T has a bounded extension mapping $L^p_A(X)$ to $L^p_B(X)$, with $1 < p < q$. Furthermore,

$$\| TF \|_{L^p_B(X)} \leq C_p \| F \|_{L^p_A(X)}, \qquad 1 < p < \infty$$

for $F \in L^p_A(X)$ and C_p only depending on p, q, C_q, C and D.

(b) The operator T has a bounded weak-type (1,1) extension that satisfies

$$\lambda\mu(\{x \in X : \| TF(x) \|_B > \lambda\}) \leq C_1 \| F \|_{L^1_A(X)} \quad \forall\, \lambda > 0 \tag{3.9}$$

for $F \in L^1_A(X)$ and C_1 only depending on q, C_q, C and D.

Proof. The key new remark, in comparison with the proof of Theorem 2.12 is that, when letting $F \in L^1_A(X)$ in order to prove the weak-type (1,1) estimate, we can apply the Calderón-Zygmund decomposition to the function $x \to \| F(x) \|_A$, which lies in $L^1(X)$. We sketch the proof emphasizing the slight differences.

The goal is, once again, showing T is weak-type (1,1) so that Marcinkiewicz interpolation applies. Let $\lambda > 0$ and let $F \in L^1_A(X) \cap L^\infty_A(X)$.

By Theorem 2.3 and Remark 2.6, there exists a partition of the space $X = F_X \sqcup \Omega$, $\Omega = \bigsqcup_{k\in\mathbb{N}} Q_k$ such that $\| F \|_A \leq \lambda$ a.e. $x \in \Omega$ and

$$\frac{1}{\mu(Q_k)} \int_{Q_k} \| F(x) \|_A \, d\mu(x) \leq C_{CZ}\lambda, \qquad \forall\, k \in \mathbb{N}.$$

It is suitable to define

$$G(x) := \begin{cases} F(x) & x \in F_X \\ \frac{1}{\mu(Q_k)} \int_{Q_k} F(x) d\mu(x) & x \in Q_k \subset \Omega,\ \forall\, k \in \mathbb{N}. \end{cases}$$

Consequently, $B_F(x) := F(x) - G(x)$,$\forall\, x \in X$[3].

As in the sibling proof, we reduce matters to showing the weak-type (1,1) property for each of the two terms (see (2.10)):

$$\mu(\{x \in X : \| TF(x) \|_B > \lambda\}) \leq \mu\left(\left\{x \in X : \| TG(x) \|_B > \frac{\lambda}{2}\right\}\right) + \mu\left(\left\{x \in X : \| TB_F(x) \|_B > \frac{\lambda}{2}\right\}\right). \quad (3.10)$$

<u>G is weak-type $(1,1)$</u>: This is carried out in the same way, using hypothesis (a) and the fact that, this time, $\| G(x) \|^{q-1} \leq (C_{CZ}\lambda)^{q-1}$ a.e. $x \in X$. Note that, in the case $q = \infty$, we would proceed as remarked in the former proof.

<u>B is weak-type $(1,1)$</u>: Let us check how the new Hörmander condition helps us find the way out this time. Step (2.11) works the same way using the maximal theorem in the scalar setting:

$$\mu\left\{x \in \Omega^{**} : \| TB_F(x) \|_B > \frac{\lambda}{2}\right\} \leq \frac{A_{\mathrm{HL}} c^{**}}{\lambda} \| F \|_{L^1_B(X)} .$$

[3] We keep on using the notation G and B_F standing for "good" and "bad" to preserve the tradition, but please do not mistake the function $B_F(x)$ for the Banach space B. Similarly, make a distinction for the vector-valued function F and the set F_X from the Calderón-Zygmund decomposition.

To deal with the second term in (3.10), proceed as in the scalar-valued case to reach

$$\frac{\lambda}{2}\mu\left(\left\{x \in \Omega^{**c} : \| TB_F(x) \|_B > \frac{\lambda}{2}\right\}\right) \leq \sum_k \int_{\Omega^{**c}} \left\| \int_{Q_k} K(x,y)B_k(y)d\mu(y) \right\|_B d\mu(x)$$

and then introduce a second constant kernel operator term relying on the fact that B_k (which is the restriction of B_F to Q_k) integrates zero on Q_k. After using triangle inequality, Fubini-Tonelli, the operator norm inequality, and an overestimation of the integration domain, we reach the step where to use the Hörmander condition (3.8).

$$\begin{aligned}\sum_k \int_{\Omega^{**c}} &\left\| \int_{Q_k} K(x,y)B_k(y)d\mu(y) \right\|_B d\mu(x) \\ &\leq \sum_k \int_{Q_k} \int_{d(x,y)>Cd(y,y_k)} \| K(x,y) - K(x,y_k) \|_{\mathcal{L}(A,B)} d\mu(x) \| B_k(y) \|_A d\mu(y) \\ &\leq D \| B_F \|_{L^1_B(X)} \leq D \| F \|_{L^1_F(X)}\end{aligned}$$

Once we know T is weak-type (q,q) (in fact, type (q,q)) and weak-type $(1,1)$, both of course in the vector valued setting, we want to use Macinkiewicz interpolation. Even though Theorem 1.16 is exposed in the scalar-valued nature, it turns out to be true for vector-valued functions (with the obvious modifications, in particular replacing absolute values by norms of the Banach spaces). In fact, the proof in both cases follows exactly the same route, see Theorem 1.18 in [5], Chapter 5, Section 1.

Finally, the fact that we worked with $F \in L^1_A(X) \cap L^\infty_A(X)$ due to the technical issues surrounding the kernel is not an obstacle since this space is dense in $L^1_A(X)$, thus once the weak-type estimate is proved, a density argument yields the estimate for the whole space. □

All these machinery we built turns out useful to bound vector-valued operators, or even scalar-valued ones, through a variety of different techniques. We are stating some more crucial results that take part in those techniques. Later on in this chapter, we are going to experiment with them.

Here is an extension result for operators acting on scalar-valued functions to operators acting on vector-valued functions.

Theorem 3.10 (Marcinkiewicz-Zygmund). *Let (X_i, Σ_i, μ_i) be two σ-finite measure spaces, $i = 1, 2$. Let $T : L^p(X_1) \to L^q(X_2)$, $1 \leq p, q < \infty$, be a linear* *bounded operator with norm $\| T \|$. Then,*

$$\left\| \left(\sum_{j\in\mathbb{N}} |Tf_j(x)|^2 \right)^{\frac{1}{2}} \right\|_q \leq C_{p,q} \| T \| \left\| \left(\sum_{j\in\mathbb{N}} |f_j(x)|^2 \right)^{\frac{1}{2}} \right\|_p \tag{3.11}$$

for $(f_j)_{j\in\mathbb{N}} \in L^p_{\ell^2}(\mathbb{R}^n)$ and some constant $C_{p,q} > 0$ depending on p and q.

What underpins the corresponding proof are randomisation techniques, that we are going to discuss later on. The proof can be found in [5], Chapter 5, Theorem 2.7.

This results partially answers the question we mentioned about extending the Hilbert transform to ℓ^p. Theorem 3.10 solves the problem for $p = 2$. In general, it is not possible to find extensions to functions taking values in Banach spaces that are not Hilbert spaces; such a structure is really helpful.

Likewise, if instead of a single linear bounded operator we have a collection of them, the corresponding vector-valued operator is in general unlikely to be tameable, unless in very nice situations like the following one.

Theorem 3.11. *Let $\{I_j\}_{j\in N}$ be an arbitrary countable collection of intervals in $\mathbb{R}^n$ with sides parallel to the coordinate axis, and denote by T_{I_j} the multiplier operators associated to the indicator function of each interval on the frequency side. Then, for all $1 < p < \infty$ and $(f_j)_{j\in\mathbb{N}} \in L^p_{\ell^2}(\mathbb{R}^n)$,*

$$\left\| \left(\sum_{j\in\mathbb{N}} |T_{I_j} f_j(x)|^2 \right)^{\frac{1}{2}} \right\|_p \leq C_p \left\| \left(\sum_{j\in\mathbb{N}} |f_j(x)|^2 \right)^{\frac{1}{2}} \right\|_p \tag{3.12}$$

where $C_p > 0$ is a constant depending on p and n.

The details of the proof can be found in [5], Chapter 5, Corollary 2.13. Let us explain the main idea though. Since the intervals have sides parallel to the coordinate axis, it is possible to give a tensor product structure to the operator that reduces matters to the one dimensional case. In this position, one realises that the indicator function of an interval $\mathbb{1}_{[a,b]}$ can be written as the sum of two sign functions, with the intention to make the multiplier of the Hilbert transform show up (use Definition 1.33). The proof concludes by an application of the Marcinkiewicz-Zygmund theorem (Theorem 3.10) and the fact that the Hilbert transform is a bounded operator on $L^p(\mathbb{R}^n)$, for $1 < p < \infty$.

3.2. Littlewood-Paley theory

The Littlewood-Paley theory has its origin in the 1930's, firstly developed by J. E. Littlewood and R. Paley with the main intention of studying Fourier series. Although the theory was in the beginning developed in the setting of the real line, many brilliant mathematicians like Zygmund, Marcinkiewicz and Stein made further contributions all along the century. Nowadays, the theory is powerful enough to be capable of extending properties of $L^2(\mathbb{R}^n)$ functions to $L^p(\mathbb{R}^n)$ functions.

(a) John Edensor Littlewood. Rochester (United Kingdom) 1885 - Cambridge (United Kingdom) 1977. [Source: https://londmathsoc.onlinelibrary.wiley.com/doi/pdf/10.1112/blms/11.1.59]

(b) Raymond Paley. Bournemouth (United Kingdom) 1907 - Canadian Rockies (Canada) 1933. [Source: https://mathshistory.st-andrews.ac.uk/Biographies/Paley/]

Figure 3.1: The fathers of the Littlewood-Paley theory. They could not collaborate long because Paley died young in a skiing accident in the Canadian Rockies.

Accordingly, the starting point is the framework of $L^2(\mathbb{R}^n)$. The central idea is splitting the Fourier domain $\mathbb{R}^n$ into concentric dyadic annuli in sort of a partition of unity: the Littlewood-Paley decomposition. We turn to the details.

Lemma 3.12 (Smooth Littlewood-Paley decomposition in $\mathbb{R}$)**.** *One can write*

$$1 = \sum_{j\in\mathbb{Z}} \psi(2^{-j}\xi) \quad \forall\, \xi \in \mathbb{R} \setminus \{0\} \tag{3.13}$$

where $\psi(2^{-j}\xi) \in C_c^\infty(\mathbb{R})$ is a smooth function compactly supported in $[-2^{j+1}, -2^{j-1}] \cup [2^{j-1}, 2^{j+1}]$ for every $j \in \mathbb{Z}$.

Proof. Let $\phi_0 \in C_c^\infty(\mathbb{R})$ be a bump function such that $\operatorname{supp}(\phi_0) \subseteq [-\frac{1}{2}, \frac{1}{2}]$ and $\int_{\mathbb{R}} \phi_0(\xi)d\xi = 1$.

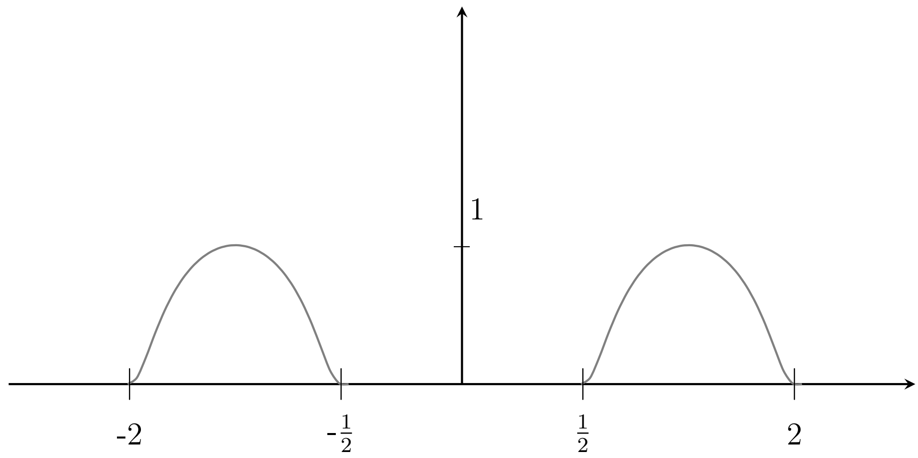


Figure 3.2: The function ψ we build in Lemma 3.12 looks like this.

Consider $\phi_1(\xi) := \mathbb{1}_{[-\frac{3}{2},\frac{3}{2}]} * \phi_0(\xi)$. We have that $\phi_1 \in C_c^\infty(\mathbb{R})$ and the support is the Minkowski sum of the supports of the convolving functions: $\mathrm{supp}(\phi_1) \subseteq [-2,2]$. Moreover, $\phi_1(\xi) = 1$ for $\xi \in [-1,1]$.

Define now $\psi(\xi) := \phi_1(\xi) - \phi_1(2\xi)$. This function satisfies the requirements of the statement. Obviously, ψ is a smooth function. Since $\phi_1(\xi) = \phi_1(2\xi) = 1$ on $|\xi| \leq \frac{1}{2}$, and $\phi_1(\xi) = \phi_1(2\xi) = 0$ on $|\xi| \geq 2$, then $\mathrm{supp}(\psi) \subseteq [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$. Finally,

$$\begin{aligned}\sum_{j\in\mathbb{Z}} \psi(2^{-j}\xi) &= \sum_{j\in\mathbb{Z}} \phi_1(2^{-j}\xi) - \phi_1(2^{-j+1}\xi) = \lim_{J\to\infty} \sum_{|j|\leq J} \phi_1(2^{-j}\xi) - \phi_1(2^{-j+1}\xi) \\ &= \lim_{J\to\infty} (\phi_1(2^{-J}\xi) - \phi_1(2^{J+1}\xi)) = \lim_{\xi\to 0} \phi_1(\xi) - \lim_{\xi\to\infty} \phi_1(\xi) = 1 \quad \forall\, \xi \neq 0,\end{aligned}$$

by telescope summation (notice that, for every $\xi \in \mathbb{R}$, at most 3 terms are contributing to the sum). □

Without much effort, the Littlewood-Paley decomposition in $\mathbb{R}$ extends to $\mathbb{R}^n$ by imposing spherical symmetry. Let $A_j := \{\xi \in \mathbb{R}^n : 2^{j-1} \leq |\xi| \leq 2^{j+1}\}$ $\forall\, j \in \mathbb{Z}$ denote the dyadic annuli in $\mathbb{R}^n$ of those frequencies that are about 2^j.

Corollary 3.13 (Smooth Littlewood-Paley decomposition in $\mathbb{R}^n$)**.** *One can write*

$$1 = \sum_{j\in\mathbb{Z}} \psi(2^{-j}\xi) \quad \forall\, \xi \in \mathbb{R}^n \setminus \{0\} \tag{3.14}$$

where $\psi(2^{-j}\xi) \in C_c^\infty(\mathbb{R}^n)$ is a smooth function compactly supported in A_j for every $j \in \mathbb{Z}$.

Proof. Just take the function $\psi_{\mathbb{R}}$ from Lemma 3.12 and set $\psi_{\mathbb{R}^n}(\xi) := \psi_{\mathbb{R}}(|\xi|)$. □

Imagine for the moment that we wish to study a certain multiplier m. Such decomposition of the frequency space is useful for exploiting the properties of m for frequencies around comparable scales. By defining $m_j(\xi) := m(\xi)\psi(2^{-j}\xi)$ and carrying out the decomposition

$$m(\xi) = \sum_{j\in\mathbb{Z}} m_j(\xi) = \sum_{j\in\mathbb{Z}} m(\xi)\psi(2^{-j}\xi)$$

we may separately treat each m_j, although we will afterwards have to face the problem of bringing back the properties to m. We do not care about m not being defined in $\xi = 0$ since m is only defined a.e.

Carrying on the previous idea, surely we can think of the multiplier operator associated to m,

$$\widehat{T_m f}(\xi) = m(\xi)\hat{f}(\xi) = \sum_{j\in\mathbb{Z}} m(\xi)\psi(2^{-j}\xi)\hat{f}(\xi).$$

This computation invites us to define a family of multiplier operators indexed by $j \in \mathbb{Z}$:

$$\widehat{T_{m_j} f}(\xi) := m_j(\xi)\hat{f}(\xi) = m(\xi)\psi(2^{-j}\xi)\hat{f}(\xi),$$

so that

$$\widehat{T_m f}(\xi) = \sum_{j\in\mathbb{Z}} \widehat{T_{m_j} f}(\xi). \tag{3.15}$$

In the spacial domain,

$$T_{m_j}(x) = (m(\xi)\psi(2^{-j}\xi))\check{\ } * f(x).$$

Accounting that $m(\xi)\psi(2^{-j}\xi) \in L^1(\mathbb{R}^n)$, its inverse Fourier transform is well defined as a continuous function, in concordance with Definition 1.2. Nevertheless, the tempting step

$$(m(\xi)\psi(2^{-j}\xi))\check{\ } = \check{m} * 2^{nj}\check{\psi}(2^j\cdot)(x)$$

will not make sense in the framework of functions unless stronger assumptions on m are made.

Besides, in contrast with the clearly absolutely convergent sum in (3.15), one should also worry about the sum in the spacial side

$$T_m f(x) = \sum_{j\in\mathbb{Z}} (m(\xi)\psi(2^{-j}\xi))\check{\ } * f(x) \tag{3.16}$$

being absolutely convergent so that the decomposition makes sense as a function. At least, the sum in (3.16) does converge in the L^2 sense thanks to Plancherel, as can be easily checked.

3.2.1. Square functions

Forget about the multiplier function now; set $m \equiv 1$. In such a case, the identity (3.16) rephrases as

$$f(x) = \sum_{j\in\mathbb{Z}} 2^{nj}\check{\psi}(2^j\cdot) * f(x) \tag{3.17}$$

understood in the sense of L^2 convergence. Let us introduce some notation for this particular case:

$$\widehat{P_j f}(\xi) := \psi(2^{-j}\xi)\hat{f}(\xi) \quad \forall\, j \in \mathbb{Z}.$$

With this notation, rewrite (3.17) as

$$f(x) = \sum_{j\in\mathbb{Z}} P_j f(x).$$

Let us here promote the linear sum to a square sum, eventually reaching the definition of the smooth Littlewood-Paley square function.

Definition 3.14 (Smooth Littlewood-Paley square function)**.**

$$Sf(x) := \left(\sum_{j\in\mathbb{Z}} |P_j f(x)|^2\right)^{\frac{1}{2}} \tag{3.18}$$

Of course this is not the identity operator. In fact, it is not even linear, which is an obstacle to overcome. The truth is that Sf comes in handy in many situations, creating the need to understand this particular operator.

Let us shoot for proving that (3.18) is bounded on L^p. The main ingredients we are going to be using are the vector-valued theory and a probabilistic trick.

First of all, we like to think of the square function as the norm of an operator acting on vector-valued functions $S : L^p(\mathbb{R}^n) \to L^p_{\ell^2}(\mathbb{R}^n)$: Define

$$\begin{aligned} P(f) &:= (P_j f)_{j\in\mathbb{Z}} = (\dots, P_{-1}f, P_0 f, P_1 f, \dots) \\ &= (\dots,\ 2^{-n}\check{\psi}(2^{-1}\xi) * f(x),\ 2^0\check{\psi}(2^0\xi) * f(x),\ 2^n\check{\psi}(2^1\xi) * f(x),\ \dots) \end{aligned}$$

which is a **linear** operator mapping functions to sequences of functions[4]. Accordingly,

$$Sf(x) = \|\, Pf(x)\,\|_{\ell^2(\mathbb{Z})}\,.$$

Next, we introduce the Rademacher functions that will provide us with some useful L^p estimates.

[4]One can play the same trick with maximal functions, for instance, $\mathfrak{M}f(x) = \|\, A(x,\cdot)\,\|_{L^\infty(\mathbb{R}_{>0})}$.

Definition 3.15. *Let*

$$r_0(t) := \begin{cases} 1 & t \in [0, \frac{1}{2}) \\ -1 & t \in [\frac{1}{2}, 1) \end{cases} \tag{3.19}$$

*and define $r_0(t)$ on all of the real line by its periodic extension, $r_0(t) = r_0(\mathrm{frac}(t))$. For any $m \in \mathbb{N}$, let $r_m(t) := r_0(2^m t)$. The family $\{r_m\}_{m\in\mathbb{N}}$ of functions are called the **Rademacher functions**.*

Remark 3.16. *The Rademacher functions are an orthonormal system in $L^2([0,1])$. However, we highlight that they are not a basis of such space.*

A satisfactory result that this family unlocks is the following.

Theorem 3.17. *Let $F(t) := \sum_{m\in\mathbb{N}} a_m r_m(t)$ for any sequence of square summable complex coefficients $(a_m)_{m\in\mathbb{N}} \in \ell^2(\mathbb{C})$. Then, for every $1 \le p < \infty$, $F \in L^p([0,1])$ and*

$$A_p \, \| \, F \, \|_{L^p([0,1])} \le \| \, (a_m)_m \, \|_{\ell^2(\mathbb{C})} = \left(\sum_{m\in\mathbb{N}} |a_m|^2 \right)^{\frac{1}{2}} \le B_p \, \| \, F \, \|_{L^p([0,1])} \tag{3.20}$$

for some constants $A_p, B_p > 0$ depending on p.

The proof is available in [10], Appendix D[5].

In some sense, the Rademacher functions introduce a random component in the function F that makes such apparently unrelated norms comparable. We intend to exploit this fact to extend L^2 bounds for some operator to L^p bounds for the square function.

Theorem 3.18. *The smooth Littlewood-Paley square function (3.18) is bounded in $L^p(\mathbb{R}^n)$ for every $1 < p < \infty$.*

Proof. We cannot apply the Calderón-Zygmund theory directly to S, for instance, because it is not a linear operator. Let the Rademacher functions perturbate the Littlewood-Paley decomposition:

$$P_t f(x) := \sum_{j\in\mathbb{Z}} r_j(t) P_j f(x). \tag{3.21}$$

It is true that we only defined r_m for $m \in N$ although we are considering $m \in \mathbb{Z}$. Just reorder the Rademacher functions so that they are indexed by the integers; the only thing that matters is the orthogonality.

[5]Some authors prefer alternative results to this one, in order to tackle the goal of showing L^p boundedness of the square function. Using sequences of gaussian random variables or using Khintchine inequality are two alternatives. However, what they all have in common is the philosophy of randomness: they are all set in the framework of probability theory in order to exploit the so-called square root cancellation phenomenon.

Notice that $\| P_{(\cdot)}(x) \|_{L^2([0,1])} = Sf(x)$, recovering the square function.

Unboxing (3.21), we find that the operator P_t is given by convolution against some kernel:

$$P_t f(x) = \left(\sum_{j\in\mathbb{Z}} r_j(t) 2^{nj} \check{\psi}(2^j x) \right) * f(x). \tag{3.22}$$

We are willing to apply the Calderón-Zygmund theory to P_t. If we call K_t the kernel in (3.22),

$$|\hat{K}_t(\xi)| = \left| \sum_{j\in\mathbb{Z}} r_j(t)\psi(2^{-j}\xi) \right| \leq \sum_{j\in\mathbb{Z}} \psi(2^{-j}\xi) = 1 \quad \forall\, \xi \neq 0,$$

which shows that the corresponding multiplier is an essentially bounded function, with bound independent of t. Thus, P_t defines a bounded operator on $L^2(\mathbb{R}^n)$. Let us check that K_t verifies the gradient decay condition (2.8).

$$|\nabla K_t(x)| = \left| \sum_{j\in\mathbb{Z}} r_j(t) 2^{n(j+1)} \nabla\check{\psi}(2^j x) \right| \leq \sum_{j\in\mathbb{Z}} 2^{n(j+1)} |\nabla\check{\psi}(2^j x)|$$

Thanks to $\check{\psi} \in \mathscr{S}(\mathbb{R}^n)$, we have both $|\nabla\check{\psi}(x)| \leq \frac{C}{|x|^{n+2}}$ and $|\nabla\check{\psi}(x)| \leq \frac{C}{|x|^n}$. Let us split the sum according to x:

$$\begin{aligned}\sum_{j\in\mathbb{Z}} 2^{n(j+1)} |\nabla\check{\psi}(2^j x)| &\leq \sum_{2^j\geq|x|} 2^{n(j+1)} \frac{C}{|2x|^{n+2}} + \sum_{2^j<|x|} 2^{n(j+1)} \frac{C}{|2x|^n} \\ &= \frac{C}{|x|^{n+1}} \left(\sum_{2^j\geq|x|} \frac{1}{2^j|x|} + \sum_{2^j<|x|} 2^j|x| \right) \leq \frac{C}{|x|^{n+1}}(2+2) = \frac{4C}{|x|^{n+1}}.\end{aligned}$$

The latter is also an estimate independent of t. Consequently, because of Theorem 2.9, for all $t \in \mathbb{R}$, P_t is bounded on $L^p(\mathbb{R}^n)$:

$$\| P_t f \|_p \leq C_p \| f \|_p \tag{3.23}$$

for some constant C_p independent of t.

Having developed this, by Theorem 3.17, we have

$$Sf(x) = \left(\sum_{j\in\mathbb{Z}} |P_j f(x)|^2 \right)^{\frac{1}{2}} \leq B_p \| P_{(\cdot)}(x) \|_{L^p([0,1])} \quad \forall\, 1 < p < \infty.$$

If we then take L^p norms in x,

$$\| Sf(x) \|_p \leq B_p \Big\| \| P_{(\cdot)}(x) \|_{L^p([0,1])} \Big\|_p = B_p \left(\int_{\mathbb{R}^n} \int_{[0,1]} |P_t f(x)|^p dt\, dx \right)^{\frac{1}{p}}$$

and by Fubini-Tonelli,

$$B_p\left(\int_{\mathbb{R}^n}\int_{[0,1]}|P_tf(x)|^p dt\,dx\right)^{\frac{1}{p}} = B_p\left(\int_{[0,1]}\int_{\mathbb{R}^n}|P_tf(x)|^p dx\,dt\right)^{\frac{1}{p}}$$
$$= B_p\left(\int_{[0,1]}\|\,P_tf\,\|_p^p\,dt\right)^{\frac{1}{p}}.$$

Finally, we make use of estimate (3.23) to reach

$$B_p\left(\int_{[0,1]}\|\,P_tf\,\|_p^p\,dt\right)^{\frac{1}{p}} \leq B_p\left(\int_{[0,1]}C_p^p\,\|\,f\,\|_p^p\,dt\right)^{\frac{1}{p}} = B_pC_p\,\|\,f\,\|_p\,.$$

□

One could wonder why we made an effort to build a smooth Littlewood-Paley square function instead of chopping the frequency space with rough cutoffs. Well, the answer is in what comes next.

Let $\overline{A_j} := \{\xi \in \mathbb{R}^n : 2^{j-1} \leq |\xi| \leq 2^j\}$ be the annuli containing the frequencies with magnitude between 2^{j-1} and 2^j, $j \in \mathbb{Z}$. Define the following rough multipliers.

$$\widehat{\overline{P_j}f}(\xi) := \mathbb{1}_{\overline{A_j}}(\xi)\hat{f}(\xi) \quad \forall\, j \in \mathbb{Z}$$

As in the smooth case, one has a decomposition of the identity,

$$f(x) = \sum_{j\in\mathbb{Z}} \overline{P_j}f(x),$$

as well as a square function.

Definition 3.19 (Rough Littlewood-Paley square function)**.**

$$\overline{S}f(x) := \left(\sum_{j\in\mathbb{Z}}|\overline{P_j}f(x)|^2\right)^{\frac{1}{2}} \tag{3.24}$$

By its construction, it is natural to start the study in $L^2(\mathbb{R}^n)$. By Plancherel and Fubini-Tonelli,

$$\|\,\overline{S}f\,\|_2^2 = \int_{\mathbb{R}^n}\sum_{j\in\mathbb{Z}}|\overline{P_j}f(x)|^2dx = \sum_{j\in\mathbb{Z}}\int_{\mathbb{R}^n}|\overline{P_j}f(x)|^2dx$$
$$= \sum_{j\in\mathbb{Z}}\int_{\mathbb{R}^n}|\mathbb{1}_{\overline{A_j}}(\xi)\hat{f}(\xi)|^2d\xi = \int_{\mathbb{R}^n}\sum_{j\in\mathbb{Z}}\mathbb{1}_{\overline{A_j}}(\xi)|\hat{f}(\xi)|^2d\xi = \int_{\mathbb{R}^n}|\hat{f}(\xi)|^2d\xi = \|\,f\,\|_2^2\,.$$

Not only is the rough square function bounded in $L^2(\mathbb{R}^n)$ but also it is an isometry. From here, with $L^2(\mathbb{R}^n)$ as a foothold, one would expect some argument that proves $L^p(\mathbb{R}^n)$ boundedness. Well, with great surprise, it turns out that the behaviour of the square function depends on the dimension of $\mathbb{R}^n$, as the following results depict.

Theorem 3.20. *Let $n = 1$. The rough Littlewood-Paley square function* (3.24) *is bounded on $L^p(\mathbb{R})$ for $1 < p < \infty$.*

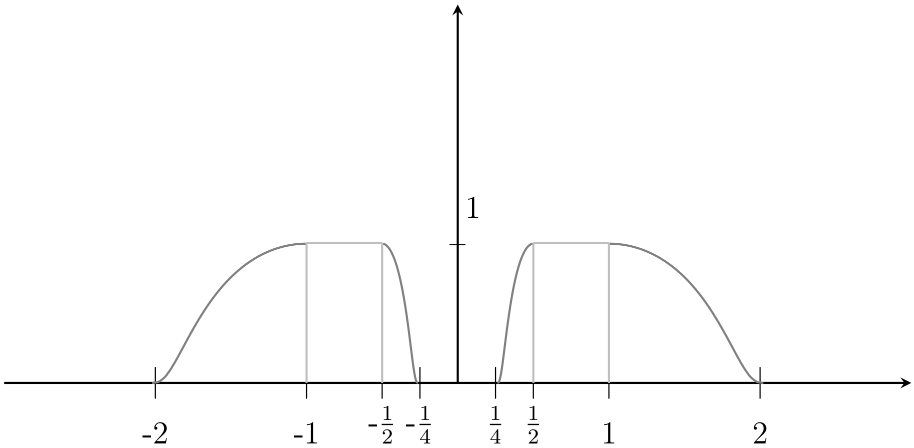


Figure 3.3: The multiplier functions $\mathbb{1}_{\overline{A_0}}$ and $\overline{\psi_0}$ in the proof of Theorem 3.20 compare like this.

Proof. If $\mathbb{1}_{\overline{A_j}}(\xi)$, $\overline{A_j} := \{\xi \in \mathbb{R}^n : 2^{j-1} \leq |\xi| \leq 2^j\}\ \forall\, j \in \mathbb{Z}$, are the multipliers corresponding to the rough square function, let us cover them by multipliers resembling the smooth ones. Let $\overline{\psi_j}(\xi) \in C_c^\infty(\mathbb{R})$ such that $\overline{\psi_j}(\xi) = 1$ on $2^{j-1} \leq |\xi| \leq 2^j$ and $\overline{\psi_j}(\xi) = 0$ on $|\xi| \leq 2^{j-2}$ and $|\xi| \geq 2^{j+1}$. Not happy with this, require

$$2 = \sum_{j\in\mathbb{Z}} \overline{\psi_j}(\xi) \quad \forall\, \xi \in \mathbb{R} \setminus \{0\}.$$

Such functions can be constructed similarly as we did in Lemma 3.12.

As a consequence we can relate the smooth and rough multipliers by

$$\mathbb{1}_{\overline{A_j}} \overline{\psi_j} = \mathbb{1}_{\overline{A_j}} \implies \overline{P_j} \circ P'_j = \overline{P_j},$$

in the understanding that

$$\widehat{P'_j f}(\xi) := \overline{\psi_j}(\xi) \hat{f}(\xi) \quad \forall\, j \in \mathbb{Z}.$$

This allows us to write

$$\left\| \overline{S} f \right\|_p = \left\| \left(\sum_{j\in\mathbb{Z}} |\overline{P_j} P'_j f(x)|^2 \right)^{\frac{1}{2}} \right\|_p \leq C_p \left\| \left(\sum_{j\in\mathbb{Z}} |P'_j f(x)|^2 \right)^{\frac{1}{2}} \right\|_p = C_p \left\| S' f \right\|_p .$$

Here, we used Theorem 3.11, but notice that it only works because in $\mathbb{R}$ the multipliers $\mathbb{1}_{\overline{A_j}}$ can be seen as indicator function of *intervals*, not only annuli. We reduced the problem to bounding the smooth Littlewood Paley square function, so the proof is complete by Theorem 3.18 (valid both for S and S', it does not matter the variant of the construction of the smooth square function). □

Theorem 3.21. *Let $n \geq 2$. The rough Littlewood-Paley square function* (3.24) *is bounded on $L^p(\mathbb{R}^n)$ if and only if $p = 2$.*

Proof. We have already seen that $\overline{S}$ defines an isometry on $L^2(\mathbb{R}^n)$. From the simple bound

$$|\overline{P_0} f(x)| \leq \overline{S} f(x) = \left(\sum_{j \in \mathbb{Z}} |\overline{P_j} f(x)|^2 \right)^{\frac{1}{2}}$$

we deduce that if $\overline{S}$ is bounded on $L^p(\mathbb{R}^n)$, then so is $|\overline{P_0} f(x)|$. But $|\overline{P_0} f(x)|$ is the rest of two ball multiplier operators, so by the ball multiplier theorem[6], for $n \geq 2$ this can only happen if $p = 2$. □

In what precedes, we have exposed a clear manifestation of an ubiquitous phenomenon in harmonic analysis: In the multidimensional case $n \geq 2$, a multiplier whose support has smooth curved boundary is going to behave nastier the rougher the decay of the multiplier function is near the boundary. Such a boundary is degenerate in the $n = 1$ case, so these heuristics do not apply.

[6] The ball multiplier theorem is a deep result in harmonic analysis, often linked to the study of convergence of partial Fourier integrals and partial Fourier sums. It was a major open problem in the 1960's; in fact, many people believed the ball multiplier would be bounded in a larger range of values of p. The breaking paper [4] of Charles Fefferman solved the problem in 1971. The proof is based, in its turn, on the existence of Kakeya sets.

CHAPTER 4

Applications and examples

4.1. Hörmander multipliers

In this section, we are dealing with a particular family of multipliers that arise, for instance, in partial differential equations. We present them as a setting where to apply the Calderón-Zygmund theory. In the understanding of the multi-index notation $\alpha \in \mathbb{N}^n$ as in Chapter 1, here is how we define them.

Definition 4.1. *A multiplier function $m \in L^\infty(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n \setminus \{0\})$ is called a* ***Hörmander multiplier*** *if*

$$|\partial^\alpha m(\xi)| \leq \frac{C_\alpha}{|\xi|^{|\alpha|}} \qquad \forall\, \xi \in \mathbb{R}^n \setminus \{0\},\ \forall\, \alpha \in \mathbb{N}^n. \tag{4.1}$$

We will also consider the case when these multipliers do not enjoy full regularity but just up to a certain order.

Theorem 4.2. *Let m be a Hörmander multiplier. Then, its associated distributional kernel K agrees with a smooth function away from the origin that verifies*

$$|\partial^\alpha K(x)| \leq \frac{A_\alpha}{|x|^{n+|\alpha|}} \quad \forall\, \alpha \in \mathbb{N}^n \tag{4.2}$$

for some constants A_α.

Proof. We start with applying the smooth Littlewood-Paley decomposition (3.14) to the multiplier. Write

$$m(\xi) = \sum_{j\in\mathbb{Z}} m_j(\xi) \quad \forall\, \xi \neq 0$$

where each m_j contains the information of the multiplier m around the frequencies 2^j. That is,

$$m_j(\xi) := m(\xi)\psi\left(2^{-j}\xi\right) \quad \forall\, j \in \mathbb{Z}.$$

Now define their corresponding kernels using the Fourier inversion formula,

$$K_j(x) := \int_{\mathbb{R}^n} m_j(\xi) e^{2\pi i x \xi} d\xi.$$

Since m is a smooth bounded function, so is m_j as well as compactly supported. Thus, each m_j is Lebesgue integrable (as well as their derivatives), from where K_j is a well defined smooth function. Note that this conclusion would be desirable for m, but a priori this may not be the case.

Let us shoot for bounds for each piece of the kernel. Begin with the basic identity:

$$(-2\pi i x)^\gamma \partial^\alpha K_j(x) = \int_{\mathbb{R}^n} \partial^\gamma (m_j(\xi)(2\pi i \xi)^\alpha) e^{2\pi i x \xi} d\xi$$

The following crude inequality follows from the assumption (4.1) and the fact that m_j is supported on the annulus of radii 2^{j-1} and 2^{j+1}.

$$\begin{aligned}(2\pi|x|)^\gamma |\partial^\alpha K_j(x)| &\le (2\pi)^\alpha \int_{\mathbb{R}^n} |\partial^\gamma (m_j(\xi)\xi^\alpha)| d\xi \\ &\le (2\pi)^\alpha |B_1(0)| 2^{n(j+1)} C_1(\alpha, \gamma, C_\alpha) 2^{j(|\alpha|-|\gamma|)} \le C_2(n, \alpha, \gamma, C_\alpha) 2^{j(n+|\alpha|-|\gamma|)}\end{aligned}$$

After this computation, γ as a parameter is freed. Set $M := |\gamma|$ and conclude that

$$|\partial^\alpha K_j(x)| \le C_2(n, \alpha, M, C_\alpha) 2^{j(n+|\alpha|-M)} |x|^{-M} \quad \forall M \in \mathbb{N}. \tag{4.3}$$

Next, we would like to estimate the complete sum $\sum_{j\in\mathbb{Z}} |\partial^\alpha K_j(x)|$ by strategically splitting it into two sums and using the inequality (4.3) with two different values of M.

With $M = 0$ and by a geometric summation,

$$\sum_{2^j \le |x|^{-1}} |\partial^\alpha K_j(x)| \le C_2(n, \alpha, C_\alpha) \sum_{2^j \le |x|^{-1}} 2^{j(n+|\alpha|)} \le C_3(n, \alpha, C_\alpha) |x|^{-n-|\alpha|}.$$

Analogously, select $M = n + |\alpha| + 1$:

$$\sum_{2^j > |x|^{-1}} |\partial^\alpha K_j(x)| \le C_2(n, \alpha, C_\alpha) |x|^{-(n+|\alpha|+1)} \sum_{2^j > |x|^{-1}} 2^{-j} \le C_4(n, \alpha, C_\alpha) |x|^{-n-|\alpha|}.$$

All in all, we have just shown

$$\sum_{j\in\mathbb{Z}} |\partial^\alpha K_j(x)| \le C_5(n, \alpha, C_\alpha) |x|^{-n-|\alpha|} \tag{4.4}$$

which resembles what we are aiming to show. Here is the limiting argument that leads to the end of the proof.

We see that $\sum_{|j|\leq N} K_j$ tends to K as $N \to \infty$ in the sense of distributions. Take any $\varphi \in \mathscr{S}(\mathbb{R}^n)$ and compute the following limit.

$$\lim_{N\to\infty}\left(\sum_{|j|\leq N} K_j - K\right)(\varphi) = \lim_{N\to\infty}\int_{\mathbb{R}^n}\left(\sum_{|j|\leq N} m_j(\xi) - m(\xi)\right)\check{\varphi}(\xi)d\xi = 0$$

In the last step, we used the facts that, by construction, $\sum_{|j|\leq N} m_j(\xi) \to m(\xi)$ pointwise as $N \to \infty$ (providing $\xi \neq 0$); and that their difference is bounded by $2m(\xi)$ which is a bounded function. dominated convergence theorem applies since $\check{\varphi}$ is a Schwartz function.

Recalling that we showed (4.4) (focus on $\alpha = 0$), we know that $\sum_{j\in\mathbb{Z}} K_j(x)$ is pointwise and absolutely convergent away from the origin. Since it converges to K in the sense of distributions, we conclude that K is a distribution which agrees with a function away from the origin. Moreover, carrying out an analogous computation for the derivatives $\partial^\alpha K_j$, we deduce that the function with which K agrees is smooth.

Eventually, once we know K can be viewed as a smooth function away from the origin, (4.2) stems from (4.4) by triangle inequality. □

From a wider perspective, what Theorem 4.2 tells us is that a Hörmander multiplier is not only a bounded function (thus giving an associated bounded operator on $L^2(\mathbb{R}^n)$) but also it verifies the gradient condition (2.8) (which, recall, is stronger than the Hörmander condition). By the main Theorem 2.9 (say, in the setting of $\mathbb{R}^n$), a Hörmander multiplier defines, on the spacial side, an operator that is bounded on $L^p(\mathbb{R}^n)$ and is weak-type (1,1)!

Corollary 4.3. *Hörmander multipliers define bounded operators on $L^p(\mathbb{R}^n)$ for $1 < p < \infty$.*

Note that in the proof of Theorem 4.2, we heavily relied on the hypothesis $m \in C^\infty(\mathbb{R}^n)$ at the time of deducing the crude estimate (4.4). As soon as we drop the full smoothness hypothesis, we loose Theorem 4.2. Nevertheless, if we assume m to be sufficiently regular (although not smooth), we can show that the associated kernel still satisfies the Hörmander condition! This is one of the reasons why we care about such a cumbersome condition in theorems like Theorem 2.9 instead of just assuming a more handy condition like the gradient one: The Hörmander multipliers without full regularity, but with enough, are still embraced by the Hörmander condition (not by the gradient one though), as the following theorem showcases.

Theorem 4.4. *Let $m \in L^\infty(\mathbb{R}^n) \cap C^\ell(\mathbb{R}^n \setminus \{0\})$ be a multiplier function such that*

$$|\partial^\alpha m(\xi)| \leq \frac{C_\alpha}{|\xi|^{|\alpha|}} \qquad \forall\, \xi \in \mathbb{R}^n \setminus \{0\},\ \forall\, \alpha \in \mathbb{N}^n,\ 0 \leq |\alpha| \leq \ell. \tag{4.5}$$

with $\ell := \lceil \frac{n}{2} \rceil$ the smallest integer greater than or equal to $\frac{n}{2}$. Then, its associated distributional kernel K agrees with a function away from the origin that verifies the Hörmander condition:

$$\sup_{|y|>0} \int_{|x|\geq 2|y|} |K(x-y) - K(x)|dx = B < \infty.$$

for some constant $B > 0$.

As the proof of Theorem 4.4 involves similar technicalities as the one of Theorem 4.2, we just note that the proof can be found in [12], Chapter 6, Section 4.4. Essentially, instead of deducing a crude estimate as (4.4), a neater application of Plancherel leads to a similar estimate contemplating less regularity yet turning out useful.

Corollary 4.5. *Hörmander multipliers with limited regularity, as in the previous theorem, define bounded operators on $L^p(\mathbb{R}^n)$ for $1 < p < \infty$.*

4.1.1. Elliptic differential operators

We are now focusing on partial differential operators with constant coefficients of the kind

$$Du = \sum_{|\alpha|\leq k} c_\alpha \partial^\alpha u \tag{4.6}$$

with $k, n \in \mathbb{N}$ and understanding the multi-index notation $\alpha \in \mathbb{N}^n$ as in Chapter 1.

Definition 4.6. *We say that an operator of the type (4.6) is **elliptic** if the characteristic polynomial of the associated homogeneous differential operator of degree k*

$$Du = \sum_{|\alpha|=k} c_\alpha \partial^\alpha u \quad \xrightarrow{FT} \quad P(\xi) = \sum_{|\alpha|=k} c_\alpha \xi^\alpha \quad \xi \in \mathbb{R}^n \tag{4.7}$$

only vanishes at $\xi = 0$: $P(\xi) \neq 0 \ \forall \xi \neq 0$.

Let us first work directly with a homogeneous differential operator of degree k, as in Eq. (4.7). Consider the PDE

$$Du(x) = f(x)$$

for a given function $f : \mathbb{R}^n \to \mathbb{R}$ (let us work out formal computations for the moment). Equivalently, in the frequency domain,

$$P(\xi)\hat{u}(\xi) = \hat{f}(\xi)$$

with $\xi \in \mathbb{R}^n$ and P being the associated characteristic polynomial to the differential operator D. The computation

$$(\partial^\alpha u)\hat{}(\xi) = (2\pi i \xi)^\alpha \hat{u}(\xi) = \frac{(2\pi i \xi)^\alpha}{P(\xi)} P(\xi)\hat{u}(\xi) = \frac{(2\pi i \xi)^\alpha}{P(\xi)} \hat{f}(\xi) \tag{4.8}$$

is valid whenever $\xi \neq 0$. Hence, (4.8) wishes to be interpreted in the sense of multipliers.

Indeed, neglecting the constant $2\pi i$, define the multiplier

$$m_h(\xi) := \frac{\xi^\alpha}{P(\xi)} \qquad |\alpha| = k. \tag{4.9}$$

Since we assumed D to be homogeneous, P is a homogeneous polynomial of degree k, meaning that Eq. (4.9) defines a singular homogeneous multiplier of degree 0:

$$m_h \in L^\infty(\mathbb{R}^n) \qquad m_h(\lambda\xi) = m_h(\xi) \ \ \forall\, \lambda > 0, \xi \neq 0.$$

Not happy with this, we realise that a Hörmander multiplier has arisen! m_h in (4.9) is clearly smooth and bounded away from the origin. Furthermore, because of the anatomy of the derivatives of quocients of polynomials, the partial derivatives of m_h satisfy (4.1). Thanks to this, by taking inverse Fourier transforms and L^p norms in (4.8) and accounting for Corollary 4.3, we get

$$\| \partial^\alpha u \|_p \leq \| T_{m_h} \| \, \| Du \|_p \,, \qquad |\alpha| = k, 1 < p < \infty. \tag{4.10}$$

Precisely, if D is an elliptic differential operator with constant coefficients, then any monomial partial differential operator of the same order as D is bounded, in the L^p norm, by D. We leave the task of explaining the meaning of (4.10) for after discussing the nonhomogeneous case.

The next step is considering a nonhomogeneous elliptic differential operator of the kind (4.6). In such scenario, let us split its characteristic polynomial into a sum of the homogeneous terms of degree k, P_h, and the remaining lower order terms, P_l: $P(\xi) = P_h(\xi) + P_l(\xi)$. We show that the polynomial P is bounded below away from the origin.

$$|P(\xi)| = \left| \sum_{\alpha \leq k} c_\alpha \xi^\alpha \right| = \left| |\xi|^k P_h\left(\frac{\xi}{|\xi|}\right) + P_l(\xi) \right|$$

Note that the restriction to the unit sphere of the homogeneous part P_h appears, but since $P_h(\xi)$ does not vanish for $\xi \neq 0$, such a restriction is bounded below in modulus by some constant $c_h > 0$. Furthermore, P_l is a polynomial of degree at most $k-1$, thus there exists another constant $C_l > 0$ such that $|P_l(\xi)| \leq C_l |\xi|^{k-1}$ for $|\xi| \geq 1$. We then reach the bound

$$\left| |\xi|^k P_h\left(\frac{\xi}{|\xi|}\right) + P_l(\xi) \right| \geq c_h |\xi|^k - C_l |\xi|^{k-1} \geq C_l |\xi|^k$$

for $|\xi| \geq R$, where $R := \max\{1, r\}$ and $r > 0$ is such that $c_h r^k = 2C_l r^{k-1}$, so $r = \frac{2C_l}{c_h}$. All in all, P is bounded below away from the origin: $|P(\xi)| \geq C_l R^k$ for $|\xi| > R$.

Having remarked this, let $\phi \in C_c^\infty(\mathbb{R}^n)$ be a bump function such that $\phi(\xi) = 1$ for $\xi \in B_R(0)$ and $\phi(\xi) = 0$ for $\xi \notin B_{2R}(0)$. The adaptation of the computation in (4.8) is as follows.

$$\begin{aligned}(\partial^\alpha u)\hat{}(\xi) &= (2\pi i\xi)^\alpha \hat{u}(\xi) = (2\pi i\xi)^\alpha \hat{u}(\xi)\phi(\xi) + (2\pi i\xi)^\alpha \hat{u}(\xi)(1-\phi(\xi)) \\ &= (2\pi i\xi)^\alpha \phi(\xi)\hat{u}(\xi) + (2\pi i)^\alpha \frac{\xi^\alpha}{P(\xi)}(1-\phi(\xi))P(\xi)\hat{u}(\xi) \quad (4.11)\end{aligned}$$

It is safe to divide by $P(\xi)$ in the second term because $(1-\phi(\xi)) = 0$ in $B_R(0)$ and we saw that $|P|$ is bounded below elsewhere.

Again, neglecting $2\pi i$ factors from now on, define the multiplier

$$m_{nh}(\xi) := \frac{\xi^\alpha}{P(\xi)}(1-\phi(\xi)) \quad |\alpha| = k \tag{4.12}$$

which is *not* homogeneous at all this time. However, it clearly lives in $L^\infty(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n \setminus \{0\})$ and, thankfully, it is at the end of the day a Hörmander multiplier: for $|\xi| < R$, $m_{nh}(\xi) = 0$; for $R \leq |\xi| \leq 2R$, $m_{nh}(\xi)$ is bounded; and for $|\xi| > 2R$, $m_{nh}(\xi) = \frac{\xi^\alpha}{P(\xi)}$. Even though P is not homogeneous this time, since in the latter case we are away from the origin, the bound (4.1) holds for m_{nh}. To sum up, m_{nh} unmasks as a Hörmander multiplier. Thus, by Corollary 4.3, the multiplier defines an operator $T_{m_{nh}}$ which is bounded on $L^p(\mathbb{R}^n)$ for $1 < p < \infty$.

Take inverse Fourier transforms and then L^p norms in Eq. (4.11) to obtain

$$\| \partial^\alpha u \|_p \leq \| (\phi(\xi)\xi^\alpha)\check{} * u \|_p + \| T_{m_{nh}}(Du) \|_p, \quad |\alpha| = k. \tag{4.13}$$

Do not be afraid of the function $(\phi(\xi)\xi^\alpha)\check{}$. It is a Schwartz function if we take into account that $\phi \in \mathscr{S}(\mathbb{R}^n)$ also is. Utilising Proposition 1.5 for the first term and the boundedness of $T_{m_{nh}}$ (Corollary 4.3) for the second one, we reach

$$\| \partial^\alpha u \|_p \leq \| (\phi(\xi)\xi^\alpha)\check{} \|_1 \| u \|_p + \| T_{m_{nh}} \| \| Du \|_p, \quad |\alpha| = k, 1 < p < \infty. \tag{4.14}$$

This kind of estimates ((4.10) and (4.14)) are useful in the context of studying regularity of solutions of partial differential equations. Imagine that we take the initial data $f = Du$ to live in $L^p(\mathbb{R}^n)$, and suppose that we were capable to show that the solution to the equation u is also an $L^p(\mathbb{R}^n)$ function. Then, what estimate (4.14) tells us is that, in fact, the solution u belongs to the Sobolev space $W^{p,k}(\mathbb{R}^n)$.

4.2. Marcinkiewicz multipliers

Here, we present another interesting family of multipliers. Even though one can define Marcinkiewicz multipliers in the setting of $\mathbb{R}^n$, it will be more illustrative and practical to work in the real line $\mathbb{R}$. Let $I = [2^{j+1}, 2^j]$ or $I = [-2^j, -2^{j+1}]$, $j \in \mathbb{Z}$, be a dyadic interval.

Definition 4.7. *A multiplier function m on the real line is a* ***Marcinkiewicz multiplier*** *if $m \in L^\infty(\mathbb{R}) \cap C^1(\mathbb{R} \setminus \{0\})$ and there exists $B \geq 0$ such that*

$$\int_I |m'(\xi)| d\xi \leq B < \infty \tag{4.15}$$

for all dyadic intervals $I = [2^j, 2^{j+1}]$ and $I = [-2^{j+1}, -2^j]$, $j \in \mathbb{Z}$.

Remark 4.8. *The C^1 regularity condition may be relaxed so that instead of requiring (4.15), more generally we demand that m has uniformly bounded variation*[1] *on dyadic intervals.*

Remark 4.9. *The following straightforward computation shows that a Hörmander multiplier in $\mathbb{R}$ is a Marcikiewicz multiplier in $\mathbb{R}$ as well.*

$$\int_{2^j}^{2^{j+1}} |m'(\xi)| d\xi \leq C_1 \int_{2^j}^{2^{j+1}} \frac{d\xi}{|\xi|} = C_1 \log\left(\frac{2^{j+1}}{2^j}\right) = C_1 \log(2) \quad \forall j \in \mathbb{Z}$$

Indeed, Marcinkiewicz multipliers generalise Hörmander multipliers. Our natural concern now is the L^p boundedness of this broader class of multipliers.

Theorem 4.10. *A Marcinkiewicz multiplier m, as in Definition 4.7 defines an operator T_m which is bounded on $L^p(\mathbb{R})$ for $1 < p < \infty$.*

So far, every time we have stated a theorem the thesis of which is that a certain operator is bounded on $L^p(X)$ for $1 < p < \infty$, the strategy for the proof has been invoking the Calderón-Zygmund theory. This time, however, Marcinkiewick multipliers do *not* verify the Hörmander condition. To be accurate, the inverse Fourier transform of a Marcinkiewick multiplier may not even be a function, meaning that the Hörmander condition makes no sense in this context. Therefore, such a strategy is automatically discarded. In spite of this, we are going to succeed by making wise use of the vector-valued theory and the Littlewood-Paley theory. The following is a sketch of a proof; we have not introduced enough machinery to explain the full proof.

Proof. We denote by T_I the multiplier operator whose multiplier function is the indicator function of the interval I. Likewise, denote by T_{m_I} the multiplier

[1] To expand knowledge on this space of functions, see [14], Chapter 2, Section 1.

operators with multipliers $m\mathbb{1}_I$. Let us work with $\varphi \in \mathscr{S}(\mathbb{R})$ (and extend by continuity in the end) and dyadic intervals I. Clearly, by the fundamental theorem of calculus, assuming $I = [2^j, 2^{j+1}]$,

$$\widehat{T_{m_I}\varphi}(\xi) = m(2^j)\mathbb{1}_I(\xi)\hat{\varphi}(\xi) + \int_I \mathbb{1}_{[s,\infty)}(\xi)\mathbb{1}_I(\xi)\hat{\varphi}(\xi)m'(s)ds$$

or, on the spacial side,

$$T_{m_I}\varphi(x) = m(2^j)T_I\varphi(x) + \int_I T_{[s,\infty)}T_I\varphi(x)m'(s)ds.$$

Introducing absolute values and using Cauchy-Schwarz, one reaches

$$\begin{aligned} |T_{m_I}\varphi(x)| &\leq |m(2^j)||T_I\varphi(x)| + \int_I |T_{[s,\infty)}T_I\varphi(x)(m'(s))^{\frac{1}{2}}||m'(s)|^{\frac{1}{2}}ds \\ &\leq \| m \|_\infty |T_I\varphi(x)| + B^{\frac{1}{2}} \left(\int_I |T_{[s,\infty)}T_I\varphi(x)|^2|m'(s)|ds \right)^{\frac{1}{2}}, \end{aligned}$$

where B is the constant bounding the variation of the derivative of the Marcinkiewicz multiplier, as in (4.15). Now apply ℓ^2 and L^p norms.

$$\begin{aligned} &\left\| \left(\sum_I |T_{m_I}\varphi(x)|^2 \right)^{\frac{1}{2}} \right\|_p \\ &\leq \| m \|_\infty \left\| \left(\sum_I |T_I\varphi(x)|^2 \right)^{\frac{1}{2}} \right\|_p + B^{\frac{1}{2}} \left\| \left(\sum_I \int_I |T_{[s,\infty)}T_I\varphi(x)|^2|m'(s)|ds \right)^{\frac{1}{2}} \right\|_p \\ &= \| m \|_\infty \| \overline{S}\varphi \|_p + B^{\frac{1}{2}} \left\| \left(\sum_I \int_I |T_{[s,\infty)}T_I\varphi(x)|^2|m'(s)|ds \right)^{\frac{1}{2}} \right\|_p \end{aligned}$$

The sum is over all dyadic intervals, both in the negative and positive sides of the real line. Also, we identified the one dimensional rough Littlewood-Paley square function $\overline{S}$ in the first term. The following step deals with the second and cumbersome term. The main problem is tackling $T_{[s,\infty)}$. The truth is that this document does not reach a theoretical stage that enables us to justify the move. Therefore, consult [5], Chapter 5, Theorem 5.13 for the rigorous proof and higher technology. Just for illustrative purposes, let us sweep $T_{[s,\infty)}$ under the rug.

$$\begin{aligned} \left\| \left(\sum_I \int_I |T_I\varphi(x)|^2|m'(s)|ds \right)^{\frac{1}{2}} \right\|_p &= \left\| \left(\sum_I |T_I\varphi(x)|^2 \int_I |m'(s)|ds \right)^{\frac{1}{2}} \right\|_p \\ &\leq B^{\frac{1}{2}} \left\| \left(\sum_I |T_I\varphi(x)|^2 \right)^{\frac{1}{2}} \right\|_p = B^{\frac{1}{2}} \| \overline{S}\varphi \|_p \end{aligned}$$

Winding back, we reached

$$\left\| \left(\sum_I |T_{m_I}\varphi(x)|^2 \right)^{\frac{1}{2}} \right\|_p \leq (\| m \|_\infty + B) \left\| \overline{S}\varphi \right\|_p$$

and since by Theorem 3.20, the one dimensional rough square function is $L^p(\mathbb{R})$-bounded, so is such a vector-valued extension of the Marcinkiewicz multiplier T_m. Again, we refer to [5], Chapter 5, Corollary 5.11, to find out why the bounded vector valued extension implies that the simple original multiplier T_m is bounded as well. □

CHAPTER 5
Beyond the paradigm

The gift of Calderón and Zygmund illuminated the mathematical analysis community. Eventually, mathematicians could understand the behaviour of iconic operators such as the Hilbert transform or square functions. Later, the theory broadened to generic measure metric spaces and vector-valued functions. Nevertheless, together with this progress, new problems arose in the field. In particular, interest was shown in singular measure operators.

The reason for this interest relies on the thirst for understanding other appealing problems like the Kakeya problem, the Bochner-Riesz conjecture or the Fourier restriction problem, which still remain mysterious and open.

In this chapter, we would like to give a flavour of the difficulties and state-of-the-art challenges by an exposition of a particular example: Stein's spherical maximal operator.

5.1. The spherical maximal operator

In Chapter 1, we introduced the Hardy-Littlewood maximal function:

$$\mathfrak{M}f(x) := \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)| dy$$

It is to be noted that for $r \to 0$, the averages over balls of radius r tend to resemble the convolution with the Dirac delta (as a feature of approximations to the identity). By this, we want to highlight that the Hardy-Littlewood maximal operator is certainly a singular operator, in the sense that the kernels tend to generate a singularity.

If light of the latter, why not considering an analogous construction but averaging over spheres, instead of balls? This is precisely what Elias Stein wondered in the 1970's.

Definition 5.1. *Let $f : \mathbb{R}^n \to \mathbb{C}$ be a measurable function in $\mathbb{R}^n$* [1]*. Define the **spherical maximal function** as follows.*

$$\mathcal{S}f(x) := \sup_{r>0} \int_{\mathbb{S}^{n-1}} f(x - r\omega) d\sigma(\omega) \quad \forall x \in \mathbb{R}^n \tag{5.1}$$

where $\mathbb{S}^{n-1}$ is the $(n-1)$-dimensional unit sphere in $\mathbb{R}^n$, σ is the surface measure in $\mathbb{R}^n$ and $\omega \in \mathbb{S}^{n-1}$ is a unit vector.

The first remark we make noticeable is that it is necessary to use the surface measure σ instead of the Lebesgue measure. However, since we are going to be considering functions $f \in L^p(\mathbb{R}^n)$ in Lebesgue spaces over $\mathbb{R}^n$ equipped with the Lebesgue measure, the integral (5.1) is clearly singular for those functions, because an $L^p(\mathbb{R}^n)$ function may take ∞ as a value on a sphere, yet have p-th power integrable on all of $\mathbb{R}^n$ [2].

Secondly, the surface measure σ is *not* absolutely continuous at all with respect to the Lebesgue measure, meaning that Radon-Nikodym theorem[3] does not apply, thus we cannot interpret (5.1) as convolution against a kernel function. Instead, we like to view it as convolution against a measure (different to the surface measure σ).

Definition 5.2. *Let μ be a measure defined on the Borelians in $\mathbb{R}^n$, $(\mathbb{R}^n, \mathcal{B}, \mu)$. Let f be a Borel-measurable function. Define the **convolution of a function against a measure** as the function of $x \in \mathbb{R}^n$*

$$(f * \mu)(x) := \int_{\mathbb{R}^n} f(x - y) d\mu(y). \tag{5.2}$$

In the case that concerns us, the measure is

$$\mu_r(f) := \int_{\mathbb{S}^{n-1}} f(r\omega) d\sigma(\omega). \tag{5.3}$$

Here, we have defined the measure μ_r making use of Riesz representation theorem[4], a powerful result that in essence tells us that the dual space of the continuous functions that tend to 0 at infinity are the finite measures.

If f is a continuous function, it is very clear that (5.3) defines a linear and bounded functional, because f is bounded on the compact set $r\,\mathbb{S}^{n-1}$. Therefore,

[1] If $\mathcal{B}$ is the σ-algebra of the Borelians, λ is the Lebesgue measure and σ is the surface measure, measurable functions on $(X, \mathcal{B}, \lambda)$ are the same as the ones on $(X, \mathcal{B}, \sigma)$. It does not matter the measure function.

[2] It is in fact a subtle matter to correctly define the spherical average operator on functions $f \in L^p(\mathbb{R}^n)$. Check [12], Chapter 11 for more on this.

[3] Check [9], Chaper 6, Theorem 6.10.

[4] See [9], Chapter 6, Theorem 6.19.

μ_r turns out to be a finite measure, for every $r > 0$. All in all, we are now comfortable with

$$\mathcal{S}f(x) = \sup_{r>0}(f * \mu_r)(x).$$

Surely, by dealing with convolution with finite measures, we have to forget about shooting for the Hörmander condition, which requires the kernel to be an actual function. Therefore, the spherical maximal operator definitely falls outside the scope of the Calderón-Zygmund theory.

Against all odds, it is desirable to study the properties of the spherical maximal function because, as remarked in Chapter 1, maximal estimates lead to almost everywhere convergence results. It turns out that the spherical average operator is intimately related to the wave equation in $\mathbb{R}^n$. In fact, it is present in the analytic solution of it. Therefore, with maximal estimates for such operator one may deduce almost everywhere convergence for the solution of the wave equation to the initial datum.

To put even more stress in how untameable $\mathcal{S}f$ is, here is a comparison. Let us set us back to the Hardy-Littlewood maximal function, the supremum of which we took over the continuum $r > 0$,

$$\mathfrak{M}f(x) := \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)|dy.$$

It is sometimes useful in the study of maximal operators to place a supremum over a countable set of dyadic numbers, like

$$\tilde{\mathfrak{M}}f(x) := \sup_{k\in\mathbb{Z}} \frac{1}{|B_{2^k}(x)|} \int_{B_{2^k}(x)} |f(y)|dy.$$

Clearly, $\tilde{\mathfrak{M}}f \leq \mathfrak{M}f$. Moreover, the reverse inequality also holds up to a multiplicative constant. To show it, let $k \in \mathbb{Z}$ and let $2^{k-1} \leq r \leq 2^k$.

$$\begin{aligned}\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)|dy &\leq \frac{1}{|B_{2^{k-1}}(x)|} \int_{B_{2^k}(x)} |f(y)|dy \\ &= \frac{2^n}{|B_{2^k}(x)|} \int_{B_{2^k}(x)} |f(y)|dy \leq 2^n \tilde{\mathfrak{M}}f(x)\end{aligned}$$

Now taking supremum over all $k \in \mathbb{Z}$ and $r > 0$, one obtains $\mathfrak{M}f \leq 2^n \tilde{\mathfrak{M}}f$ as desired. All in all, what $\mathfrak{M}f \sim 2^n \tilde{\mathfrak{M}}f$ means is that these two manifestations of the operator are bounded on the same spaces; in this sense, they are equivalent.

The shown feature is not specific of the Hardy-Littlewood maximal function. For instance, the maximal version of the parabolic measure

$$\mu_r^{\text{par}}(f) := \int_1^2 f(rt, r^2t^2)dt$$

also presents this property, as it can be similarly shown.

Unfortunately, the dyadic version of the spherical maximal operator

$$\tilde{\mathcal{S}}f(x) := \sup_{k\in\mathbb{Z}} \int_{\mathbb{S}^{n-1}} f(x-2^k\omega)d\sigma(\omega)$$

is *not* comparable to the continuous supremum version $\mathcal{S}f$. The only true inequality is the trivial $\tilde{\mathcal{S}}f \leq \mathcal{S}f$. The reason for this property to fail is that, in the case of the Hardy-Littlewood maximal function, the set of centred solid balls are a nested collection of sets, whereas the hollow centred spheres of different radii are not nested by any means. Perhaps surprisingly, this issue leads to different behaviours of $\mathcal{S}f$ and $\tilde{\mathcal{S}}f$. Here is what we can say about this pair of operators.

Theorem 5.3. *The dyadic spherical maximal operator $\tilde{\mathcal{S}}f$ is bounded in $L^p(\mathbb{R}^n)$ for $1 < p \leq \infty$. This is, for $f \in L^p(\mathbb{R}^n)$,*

$$\left\| \tilde{\mathcal{S}}f \right\|_p \leq C_p \, \| f \|_p \,,$$

for some constant $C_p > 0$ depending on p and n.

Conjecture 5.4. *The dyadic spherical maximal operator $\tilde{\mathcal{S}}f$ is weak-type (1,1). So for any $\lambda > 0$ and $f \in L^1(\mathbb{R}^n)$,*

$$\lambda|\{x\in\mathbb{R}^n : \tilde{\mathcal{S}}f(x) > \lambda\}| \leq C_1 \, \| f \|_1 \,,$$

for some constant $C_1 > 0$ depending on n.

Theorem 5.5. *The spherical maximal operator $\mathcal{S}$ is bounded in the following cases:*

(a) For any dimension $n \geq 2$, $\frac{n}{n-1} < p \leq \infty$ and $f \in L^p(\mathbb{R}^n)$,

$$\| \mathcal{S}f \|_p \leq C_p \, \| f \|_p \,,$$

for some constant $C_p > 0$ depending on p and n.

(b) For $n \geq 3$, $\mathcal{S}$ is of restricted type[5] *at the endpoint. By this, we mean:*

$$\| \mathcal{S}f \|_{L^{\frac{n}{n-1},\infty}(\mathbb{R}^n)} \leq C_n \, \| f \|_{L^{\frac{n}{n-1},1}(\mathbb{R}^n)} \,,$$

where C_n is a constant depending on n.

However, (b) is false for $n = 2$:

(c) $\mathcal{S}$ does not map $L^{2,1}(\mathbb{R}^2)$ to $L^{2,\infty}(\mathbb{R}^2)$.

See Stein and Bourgain's work in this regard, in [11] and [1].

[5] Learn more about Lorentz spaces and their interpolating role in [6], Chapter 1.

Bibliography

[1] Bourgain, J. “Estimations de certaines fonctions maximales”. In: *C. R. Acad. Sci. Paris Sér. I Math.* 10 (1985), pp. 499–502.

[2] Calderón, A. P. and Zygmund, A. “On the existence of certain singular integrals”. In: *Acta Mathematica* 88 (1952), pp. 85–139. DOI: 10.1007/BF02392130.

[3] Duoandikoetxea, J. *Fourier Analysis*. Providence, Rhode Island: Graduate Studies in Mathematics 29, 2001.

[4] Fefferman, C. “The Multiplier Problem for the Ball”. In: *Annals of Mathematics* 94 (1971), pp. 330–336. DOI: https://doi.org/10.2307/1970864.

[5] García-Cueva, J. and Francia, J.L. Rubio de. *Weighted Norm Inequalities and Related Topics*. Amsterdam, The Netherlands: Elsevier Science Publishers B.V., 1985.

[6] Grafakos, L. *Classical Fourier Analysis*. New York City: Springer, 2008.

[7] Hardy, G. H. and Littlewood, J. E. “A Maximal Theorem With Function-Theoretic Applications”. In: *Acta Mathematica* 54 (1930), pp. 81–116. DOI: 10.1007/BF02547518.

[8] Hickman, Jonathan. *Fourier and Harmonic Analysis*. University of Edinburgh. SMSTC lecture notes. 2021.

[9] Rudin, W. *Real and Complex Analysis*. Singapore, Singapore: McGraw-Hill; 3rd edition, 1987.

[10] Stein, E. M. *Singular Integrals and Differentiability Properties of Functions*. Princeton, New Jersey: Princeton University Press, 1970.

[11] Stein, E. M. “Maximal Functions: Spherical Means”. In: *PNAS* 73 (1976), pp. 2174–2175. DOI: https://www.pnas.org/doi/abs/10.1073/pnas.73.7.2174.

[12] Stein, E. M. *Harmonic Analysis: Real-Variable Methods, Orthogonality and Oscillatory Integrals*. Princeton, New Jersey: Princeton University Press, 1993.

[13] Stein, E. M. and Weiss, G. *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton, New Jersey: Princeton University Press, 1971.

[14] Wheeden, R. L. and Zygmund, A. *Measure and Integral, An Introduction to Real Analysis*. New York, United States of America: Marcel Dekker Inc., 1977.