Lectures in Analysis[1]

Rodrigo Bañuelos
Department of Mathematics
Purdue University
West Lafayette, IN 47907

In memory of my two younger brothers
José and Sergio
and older brother
Javier,
who left this world way too early.
From them I learned the true meaning of the words faith and courage.

Preface

These Lecture Notes are based on a one semester course, Mathematics 545, which I have taught at Purdue University for many years. They are more than 20 years old by now. For quite some time I thought of publishing them as a book and had them reviewed by some publishers who expressed interest modulo revisions. After learning how much they were proposing to charge for the book, I decided not to publish them and instead continue to distribute them to the students in my courses and to share them with others via email when requested. As I continued to teach this and other standard courses in probability and analysis at Purdue (both graduate and undergraduate) and using the books available, and even those that kept (keep) coming out, I realized that many (perhaps most) don't really contain any new ideas nor really present any new material. It seems that for the standard courses we are all just reshuffling the same material in different order. We teach our courses, as part of our jobs, prepare and revise lecture notes, often taking our student comments into account as is the case here, publish them as books and then turn around and ask our future students to buy them to take the same courses. Because of this, and memories from my undergraduate years of always struggling to find money to buy textbooks, the idea of cleaning and publishing the notes as a book became less and less attractive. So after many years of distributing the notes to students locally and to others by request, I am just putting them out for anyone to download. Please be aware of typos and even possible math errors. The Lecture Notes are free so not much to complain about!

The background of most of the students who typically enroll in the course consists of nothing more than the basics of measure theory. My goal for the course has been twofold. (1) To review some of those basic tools such as differentiation of monotone functions, absolute continuity, signed measures, the Radon–Nikodym theorem, Fubini's theorem, the basic theory of the Fourier transform and (2) to introduce material which is rarely part of such an introductory course. This more advanced material includes the Hardy–Littlewood maximal function, approximations to the identity, interpolation theorems, the Calderón–Zgymund decomposition and its applications to the classical singular integrals, fractional integration, and square functions. I have made an effort to present some of the basic applications of these topics to, for example, the basic boundary value problems for the Laplacian and the heat equation in the upper half space of $\mathbb{R}^n$ and to the Hörmander multiplier theorem.

One difficulty I encountered the first time I taught this course was that despite the large number of introductory analysis books on the market, I was not able to find a satisfactory text that starts at such an elementary level and develops some of the more advanced topics mentioned above. The recommended books for the course are usually Royden's *Real Analysis*, Rudin's *Real and Complex Analysis*, Torchinsky's *Real Variables* and *Fourier Analysis*, Stein's *Singular Integrals and Differentiability Properties of Functions*, and Stein and Weiss' *Introduction to Fourier Analysis on Euclidean Spaces*. Readers familiar with these references will immediately note that the material in this notes consists of pieces from each of them with additions and subtractions, keeping in mind my overall goals of the course. Thus in a certain sense these lecture notes only collects material from the above books and puts it in a single source.

I have made an effort to go through the basic topics as quickly as possible without compromising rigor and completeness in any serious way. It is up to the reader to decide if this has been achieved. There are a number of exercises but probably not as many as one would like to see in a text of this level. These are, for the most part, interwoven with the material as it is being presented. Many of the exercises are used in subsequent results and others are used for various generalizations and extensions. For these reasons, the exercises are an essential part of the Notes.

The Lecture Notes consist of ten chapters. The more elementary/introductory material is contained in Chapters 1–4. These Chapters include differentiation of monotone functions, absolute continuity, Fubini's theorem, the Radon-Nykodym theorem, applications to the duality of L^p–spaces and properties of convolutions. Chapter 6 is devoted to the Fourier transform. One item that is discussed here, which is not often taught in introductory analysis courses, is the inversion formula for regular Borel measures on the real line. Such a formula is very useful in applications to limit theorems in probability theory. Even though probability is not part of this course, I never miss the opportunity to informally point out various connections when they arise naturally. The more advanced chapters are, obviously, Chapters 5, and 7–10, and in the course I devote much of the fifteen weeks of the semester to these topics. These chapters include some of the standard tools of harmonic analysis such as the Hardy-Littlewood maximal operator, the Calderón–Zygmund decomposition, the basic theory of singular integrals and their applications to the Riesz transforms and the Beurling–Ahlfors operator (both in the plane and in $\mathbb{R}^n$), fractional integration and its connections to the inequalities of Sobolev and Nash, and a brief introduction to the classical Lusin and Littlewood–Paley square functions and their applications to the Hörmander multiplier theorem.

I wish to express my thanks to the many students who have taken this course over the years. Many made valuable contributions to the material presented here. In particular, the first draft of the notes was carefully read by Pedro Méndez many years ago and he made innumerable suggestions for improvements. My dear friend and now deceased colleague Christoph Neugebauer provided valuable comments at the start of the notes. Chris taught some the best analysis courses at Purdue and these Lecture Notes would exist if I had been able to convince him to write up his excellent lecture notes in measure theory and harmonic analysis. Earlier drafts were also read by Tom Carroll and David Applebaum who made many corrections and provided valuable comments.

It is always a great pleasure for me to see the mathematical growth that many of the students achieve as the semester progresses and as they begin to master and appreciate some of the basic material on singular integrals and Littlewood–Paley theory presented in this book. I have been fortunate to have had a number Ph.D students, many of whom began their studies from this book and just a couple of years later wrote beautiful research papers using many of the tools and ideas presented here. What more can a teacher ask for?

Rodrigo Bañuelos
West Lafayette, Indiana
July 2020

Contents

Chapter 1

Differentiation

We know from elementary calculus that:

$$f(b) - f(a) = \int_a^b f'(t)\, dt \tag{1.1}$$

and that whenever

$$F(x) = \int_a^x f(t)\, dt, \tag{1.2}$$

then

$$F'(x) = f(x). \tag{1.3}$$

But, what do we really mean by 1.1 and 1.3? When are these statements really true? These and related questions will be answered precisely in this chapter.

1.1 Covering Lemmas

Covering lemmas are at the heart of differentiation theory and key in answering the questions above. We begin this section by proving two such lemmas. These lemmas are two of the most basic in a large family of similar results. They will be sufficient for our applications. More covering lemmas can be found in the classical book M. De Guzman [Gu]. Our presentation follows [Ro] and [To2].

We remind the reader that given a set $E \subset \mathbb{R}^n$, its outer measure is defined by

$$m^*(E) = \inf\{\sum_k |I_k| : E \subset \bigcup_k I_k\} \tag{1.4}$$

where the infimum is taken over families of at most countable coverings of E by closed rectangles $I = \{x = (x_1, x_2, \dots, x_n) : a_j \le x_j \le b_j, 1 \le j \le n\}$ in $\mathbb{R}^n$ of finite volume $|I|$. If the set is Lebesgue measurable, its Lebesgue measure will be denoted by $m(E)$ but often simply just by $|E|$. As usual, a property that holds except possibly on a set of measure zero will be said to hold almost everywhere and we will often just write a.e. in such case. We state and prove our first lemma in $\mathbb{R}$ which suffices for our needs here.

Definition 1.1. A family $\mathcal{F}$ of closed intervals in $\mathbb{R}$ is said to be a covering of E in the sense of *Vitali* if for every $x \in E$ and $\varepsilon > 0$, there exists an interval $I \in \mathcal{F}$ such that $x \in I$ and $|I| < \varepsilon$. That is, every point of E belongs to an interval in $\mathcal{F}$ of arbitrarily small length.

Theorem 1.2 (Vitali Covering Lemma)**.** *Suppose that E is a subset of $\mathbb{R}$, not necessarily measurable, with $m^*(E) < \infty$. Suppose that $\mathcal{F}$ is a Vitali covering for E. Then we can find a countable (finite or infinite) pairwise disjoint collection of intervals $\{I_k\}$ in $\mathcal{F}$ such that $m^*(E \setminus \cup I_k) = 0$.*

Proof. Let $\mathcal{O}$ be an open set containing E with $m(\mathcal{O}) < \infty$. We may assume, without loss of generality, that whenever $I \in \mathcal{F}$ then $I \subset O$. (If we remove from $\mathcal{F}$ all the intervals not contained in $\mathcal{O}$ the remaining intervals still form a Vitalli covering of E.) Let $I_1 \in \mathcal{F}$. If $m^*(E \setminus I_1) = 0$, we are done. If not, choose $I_2, I_3, \dots$ recursively as follows: Suppose we have chosen pairwise disjoint intervals $I_1, I_2, \dots, I_n$ in $\mathcal{F}$ such that

$$m^*\Big(E \setminus \bigcup_{k=1}^{n} I_k\Big) > 0. \tag{1.5}$$

Let

$$G_n = \mathcal{O} \setminus \bigcup_{k=1}^{n} I_k.$$

Then G_n is open and not empty, otherwise it would contradict (1.5). We now describe how to select the next interval. We simply take the largest interval in $\mathcal{F}$ which is disjoint from $I_1, \dots, I_n$. More precisely, let $k_n = \sup\{|I| \colon I \in \mathcal{F},\ I \subset G_n\}$. Then $0 < k_n < \infty$ by (1.5) and since the intervals are contained in $\mathcal{O}$. Let $I_{n+1} \in \mathcal{F}$ be such that $I_{n+1} \subset G_n$, $|I_{n+1}| > \frac{1}{2}k_n$ and disjoint from $I_1, \dots, I_n$. Either this process ends in a finite number of steps, that is

$$m^*\Big(E \setminus \bigcup_{k=1}^{m} I_k\Big) = 0 \tag{1.6}$$

for some m, or we get an infinite sequence of disjoint intervals $I_1, I_2, \ldots$ with the above properties. We claim that this collection of intervals satisfies the conclusion of the theorem. To see this observe that since $\bigcup I_n \subset \mathcal{O}$, $\Sigma|I_k| < \infty$. Thus for each $\eta > 0$ there exists an N such that

$$\sum_{k=N+1}^{\infty} |I_k| < \eta/5. \tag{1.7}$$

Let $R_N = E \setminus \bigcup_{k=1}^{N} I_k$. We will show that

$$R_N \subset \bigcup_{k=N+1}^{\infty} J_k, \tag{1.8}$$

where $J_k = 5I_k$, where $dI = [da, db]$, if $I = [a, b]$. If this is the case, then

$$m^*\Big(E \setminus \bigcup_{k=1}^{\infty} I_k\Big) \le m^*\Big(E \setminus \bigcup_{k=1}^{N} I_k\Big) \le 5 \sum_{k=N+1}^{\infty} |I_k| < \eta,$$

which completes the proof.

To prove (1.8), let $x \in R_N$. Since G_N is open and $R_N \subset G_N$, there exists $I \in \mathcal{F}$ such that $x \in I$, and $|I|$ is as small as we like with $I \cap (I_1 \cup \cdots \cup I_N) = \emptyset$. We claim there is an $n_0 > N$ such that $I \cap I_{n_0} \neq \emptyset$. Indeed, if $I \cap I_j = \emptyset$ for all $j \le n$, then $|I| \le k_n < 2|I_{n+1}|$. Since $\lim_{n\to\infty} |I_{n+1}| = 0$, the interval I must intersect an I_{n_0} for some n_0. Clearly, $n_0 > N$ and $|I| \le k_{n_0-1} \le 2|I_{n_0}|$. Let x_0 be the midpoint of the interval I_{n_0}. Then

$$|x - x_0| \le |I| + \frac{1}{2}|I_{n_0}| \le 2|I_{n_0}| + \frac{1}{2}|I_{n_0}| = \frac{5}{2}|I_{n_0}|. \tag{1.9}$$

A simple geometric argument now shows that (1.9) implies $I \subset 5I_{n_0}$, which proves (1.8) and hence Theorem 1.2. $\square$

The following corollary is immediate from the proof of the Theorem.

Corollary 1.3. *Under the assumptions of Theorem 1.2, given $\varepsilon > 0$ there exist disjoint $I_1, \ldots, I_N$ such that $m^*(E \setminus \bigcup_{k=1}^{N} I_k) < \varepsilon$.*

The above theorem also holds for sets in $\mathbb{R}^n$. A less well known version of the above theorem, that is not as refined but that we will use in the study of the Hardy-Littlewood maximal operator in Chapter 5, is

Theorem 1.4 (Wiener Covering Lemma)**.** *Let $E \subset \mathbb{R}^n$ be measurable and suppose $E \subset \bigcup_{j \in A} B_j$, where the B_j are balls of uniformly bounded diameter. In other words, $\sup_j \operatorname{diam}(B_j) < \infty$. There exists a sequence of disjoint balls $B_1, B_2, \ldots$ (finite or infinite) from this covering such that*

$$m(E) \leq 5^n \sum_k m(B_k).$$

Proof. The proof of this theorem proceeds exactly as does the proof of Theorem 1.2. We first select a B_1 as large as possible. Let B_1 be such that $\operatorname{diam}(B_1) \geq \frac{1}{2}\sup(\operatorname{diam}(B_j))$. Assume $B_1, B_2, \ldots, B_k$ have been chosen. Let B_{k+1} be disjoint from $B_1, \ldots, B_k$ and such that

$$\operatorname{diam}(B_{k+1}) \geq \frac{1}{2}\sup\{\operatorname{diam}(B_j) : \ B_j \text{ is disjoint from } B_1, B_2, \ldots, B_k\}$$

This produces a countable (finite or infinite) collection of disjoint balls $B_1, B_2, \ldots$. That is, if there is no B_{k+1} with the above property, we stop and the collection is finite. We claim these balls have the desired property. If

$$\sum_k m(B_k) = \infty,$$

we are clearly done. Thus we may assume that

$$\sum_k m(B_k) < \infty.$$

Let $B_k^* = 5B_k$ where by cB we shall mean, here and throughout these notes, the ball concentric with B with radius c times the radius of B. We observe that with this notation, $|cB| = c^n|B|$. We claim that

$$E \subset \bigcup_k B_k^*. \tag{1.10}$$

This and the previous observation imply the lemma. To prove (1.10) assume we have an infinite collection of balls, the other case being easy. Let $x \in E \subset \bigcup_{j \in A} B_j$. Pick a $j_0 \in A$ such that $x \in B_{j_0}$. (1.10) will follow if we prove

$$B_{j_0} \subset \bigcup_k B_k^*. \tag{1.11}$$

Since $\sum_k m(B_k) < \infty$, $\mathrm{diam}(B_k) \to 0$. Let k_0 be the first integer such that $\mathrm{diam}(B_{k_0}) < (1/2)\,\mathrm{diam}(B_{j_0})$. Then $\mathrm{diam}(B_{j_0}) > \sup\{\mathrm{diam}(B_j) : B_j$ is disjoint from $B_1 \ldots B_{k_0-1}\}$ and so there is a $\tilde{k}$ satisfying $1 \le \tilde{k} < k_0$ such that

$$B_{j_0} \cap B_{\tilde{k}} \neq \emptyset. \tag{1.12}$$

By the choice of k_0,

$$\mathrm{diam}(B_{\tilde{k}}) \ge \frac{1}{2}\,\mathrm{diam}(B_{j_0}) \tag{1.13}$$

and it follows from (1.12) and (1.13) and simple geometric considerations that

$$B_{j_0} \subset B^*_{\tilde{k}},$$

proving the theorem. □

1.2 Monotone Functions

We will now apply Theorem 1.2 to study differentiation of monotone functions in $\mathbb{R}$. First, for any measurable function f on $\mathbb{R}$ we define the derivates of f by

$$\begin{aligned}
D^+f(x) &= \limsup_{h\to 0^+} \frac{f(x+h)-f(x)}{h},\\
D_+f(x) &= \liminf_{h\to 0^+} \frac{f(x+h)-f(x)}{h},\\
D^-f(x) &= \limsup_{h\to 0^+} \frac{f(x)-f(x-h)}{h},\\
D_-f(x) &= \liminf_{h\to 0^+} \frac{f(x)-f(x-h)}{h}.
\end{aligned}$$

The function is differentiable at x if $D^+f(x) = D^-f(x) = D_+f(x) = D_-f(x)$ and the common value is neither $+\infty$ nor $-\infty$. In this case we denote their common value by $f'(x)$.

By an increasing function we shall mean a nondecreasing function with a similar meaning for a decreasing function. By a monotone function we mean a function which is either increasing or decreasing.

Theorem 1.5. *Let f be an increasing real–valued function on $[a,b]$. Then f is differentiable almost everywhere, f' is measurable and*

$$\int_a^b f'(t)\,dt \le f(b) - f(a).$$

Proof. We must prove that the set where any two of the derivates of f differ has Lebesgue measure zero. Let $E = \{x\colon D^+f(x) > D_-f(x)\}$, the other combinations being similar. As is usual we denote the rational numbers by $\mathbb{Q}$. For every pair r and s in $\mathbb{Q}$ with $r > s$, let

$$E_{r,s} = \{x\colon D^+f(x) > r > s > D_-f(x)\}.$$

Then $E \subset \bigcup_{r,s\in\mathbb{Q}} E_{r,s}$ and it suffices to show that $m^*(E_{r,s}) = 0$ for each pair of rational numbers $\{r, s\}$. Let $T = m^*(E_{r,s})$. Assume $T > 0$. Let $\varepsilon > 0$ and choose an open set $\mathcal{O}$ containing $E_{r,s}$ such that

$$m(\mathcal{O}) < T + \varepsilon.$$

For each $x \in E_{r,s}$, there is an h_x such that the interval $I_x = [x-h_x, x]$ is contained in $\mathcal{O}$ and with

$$f(x) - f(x - h_x) < sh_x. \tag{1.14}$$

As the reader can verify, the family $\mathcal{F} = \{I_x\colon x \in E_{r,s}\}$ is a Vitali covering for $E_{r,s}$. By Theorem 1.2, there is a collection of disjoint intervals $I_1, \dots, I_N$ such that

$$m^*\Big(E_{r,s} \setminus \bigcup_{k=1}^{N} I_k\Big) < \varepsilon. \tag{1.15}$$

That is, $\bigcup_{k=1}^{N} I_k$ covers a subset A of $E_{r,s}$ of outer measure $> T - \varepsilon$. From (1.14) we see that

$$\sum_{k=1}^{N} [f(x_n) - f(x_k - h_k)] < s\sum_{k=1}^{N} h_k = sm\Big(\bigcup_{k=1}^{N} I_k\Big) < sm(\mathcal{O}) < s(T + \varepsilon).$$

Now, since $A \subset E_{r,s}$, each $y \in A$ is the left endpoint of an arbitrary small interval $J_y = (y, y + k_y)$ contained in some I_j and such that

$$f(y + k_y) - f(y) > rk_y.$$

This intervals form a Vitali covering for the set A. Now, pick a subcollection $J_1, J_2, \dots, J_M$ from these intervals such that $\bigcup_{j=1}^{M} J_j$ contains a subset of A of outer measure $> T - 2\varepsilon$. Again we have:

$$\sum_{j=1}^{M} \big[f(y_j + k_{y_j}) - f(y_j)\big] > r\sum_{j=1}^{M} k_{y_j} > r(T - 2\varepsilon).$$

Using the fact that f is increasing and summing over all those J contained in I_k, we get

$$\sum_{\{j:\, J_j \subset I_k\}} \left[f(y_j + k_{y_j}) - f(y_j)\right] \leq f(x_k) - f(x_k - h_k)$$

and summing over k,

$$\sum_{k=1}^{N} [f(x_k) - f(x_k - h_k)] \geq \sum_{j=1}^{M} \left[f(y_i + k_{y_j}) - f(y_j)\right].$$

Thus,

$$s(T + \varepsilon) > r(T - 2\varepsilon)$$

and letting $\varepsilon \downarrow 0$ gives

$$sT \geq rT$$

which then implies that $s \geq r$, which contradicts our assumption that $r > s$, or or gives $T = 0$, which in fact must be the case. We conclude then that f is differentiable almost everywhere.

Let

$$\widetilde{f}(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

This limit exists a.e. and $\widetilde{f}(x) = f'(x)$ whenever $\widetilde{f}$ is finite. If we let

$$\widetilde{f}_n(x) = \begin{cases} n[f(x + 1/n) - f(x)] & \text{if } a \leq x \leq b - 1/n \\ 0 & \text{if } b - 1/n < x \leq b, \end{cases}$$

we obtain a sequence of measurable functions converging almost everywhere to $\widetilde{f}$ and hence this last is also measurable. In addition, since f is increasing, both $\widetilde{f}_n$ and $\widetilde{f}$ are nonnegative. By Fatou's Lemma

$$\begin{aligned}
\int_a^b \widetilde{f}(x)\, dx &\leq \liminf_{n \to \infty} \int_a^b \widetilde{f}_n(x)\, dx \\
&= \liminf_{n \to \infty} n \int_a^{b-1/n} [f(x + 1/n) - f(x)]\, dx \\
&= \liminf_{n \to \infty} \left(n \int_{b-1/n}^{b} f(x)\, dx - n \int_a^{a+1/n} f(x)\, dx \right) \\
&\leq f(b) - f(a),
\end{aligned}$$

where for the last inequality we used the fact that f is increasing. This completes the proof of Theorem 1.5. □

Exercise 1.2.1.
Prove that the Cantor function is monotone on $[0, 1]$ with derivative zero almost everywhere. Hence, strict inequality may occur in Theorem 1.5.

1.3 Functions of Bounded Variation on $[a, b]$

Definition 1.6. Let $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ be a partition of the interval $[a, b]$ and let f be a function defined on $[a, b]$. We write $V(f; \mathcal{P}) = \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|$ for the variation of f over the partition $\mathcal{P}$. The function is said to be of *bounded variation* on the interval $[a, b]$, written as $f \in BV[a, b]$, if there exists a constant C independent of $\mathcal{P}$ such that $V(f; \mathcal{P}) \leq C$. The quantity

$$V(f; a, b) = \sup_{\mathcal{P}} V(f; \mathcal{P})$$

is called the total variation of f.

Example 1.1. If f is monotone on $[a, b]$, then f is of bounded variation and $V(f; a, b) = |f(b) - f(a)|$.

Example 1.2. Let $\{a_k\}$ be a strictly decreasing sequence of positive numbers with $a_1 = 1$ and $a_k \to 0$. Define the function f on $[0, 1]$ by

$$f(a_k) = f(a_{k+1}) = 0, \; f\left(\tfrac{a_k + a_{k+1}}{2}\right) = \tfrac{1}{k}$$

and linear on $(a_k, \frac{a_k + a_{k+1}}{2})$ and $(\frac{a_k + a_{k+1}}{2}, a_{k+1})$. If in addition we define $f(0) = 0$, then f is in continuous on $[0, 1]$. However, for any N,

$$V(f; 0, 1) \geq \sum_{k=1}^{N} \frac{1}{k} \to \infty.$$

Thus continuity does not imply Bounded Variation.

The following simple properties of the total variation function will be useful below.

Proposition 1.7. *Suppose $f \in BV[a, b]$. Then*

(i) $V(f; a, x)$ *is increasing,*

(ii) $|f(x + h) - f(x)| \leq V(f; x, x + h)$*,*

(iii) $V(f;a,x+h) - V(f;a,x) = V(f;x,x+h)$,

for $h > 0$.

Proof. Properties (i) and (ii) follow trivially from the definition. For (iii), let $\varepsilon > 0$ and let $\mathcal{P}_1$ and $\mathcal{P}_2$ be partitions of $[a,x]$ and $[x,x+h]$, respectively, such that

$$V(f;\mathcal{P}_1) > V(f;a,x) - \varepsilon/2$$

and

$$V(f;\mathcal{P}_2) > V(f;x,x+h) - \varepsilon/2.$$

Let $\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2$. This is a partition of the interval $[a,x+h]$ and

$$\begin{aligned} & V(f;x,x+h) + V(f;a,x) - \varepsilon \\ & \quad < \quad V(f;\mathcal{P}_1) + V(f;\mathcal{P}_2) \\ & \quad = \quad V(f;\mathcal{P}) \leq V(f;a,x+h). \end{aligned}$$

This proves half of (iii).

For the other half, let $\mathcal{P}$ be a partition of $[a,x+h]$ such that

$$V(f;a,x+h) \leq V(f;\mathcal{P}) + \varepsilon.$$

We may assume $x \in \mathcal{P}$ since the right hand side only increases if we add points to the partition. That is, we may assume

$$\mathcal{P} = \{a = x_0 < \cdots < x_n = x < x_{n+1} < \cdots < x_m = x+h\}.$$

If we let $\mathcal{P}_1 = \{a = x_0 < \cdots < x_n = x\}$ and $\mathcal{P}_2 = \{x_{n+1} < \cdots < x_m = x+h\}$, we get

$$\begin{aligned} V(f;a,x+h) &\leq V(f;\mathcal{P}_1) + V(f;\mathcal{P}_2) + \varepsilon \\ &\leq \quad V(f;a,x) + V(f;x,x+h) + \varepsilon, \end{aligned}$$

which proves the other half of (iii), since $\varepsilon > 0$ is arbitrary. □

Exercise 1.3.1.
For $f \in BV[a,b]$, define

$$\|f\| = |f(a)| + V(f;a,b).$$

(i) Prove that $BV[a,b]$ is a real vector space and that

(ii) under $\|\cdot\|$, $BV[a,b]$ is in fact a norm linear space. (See definition 2.12 below.) That is (i) $\|f+g\| \leq \|f\| + \|g\|$, (ii)$\|\beta f\| = |\beta|\|f\|$, for all $\beta \in \mathbb{R}$ and (iii) $\|f\| = 0$ implies $f = 0$.

Theorem 1.8 (Jordan Decomposition)**.** *Every function $f \in BV[a,b]$ can be written as $f = f_1 - f_2$ where both f_1 and f_2 are increasing.*

Proof. Write $f(x) = V(f;a,x) - (V(f;a,x) - f(x))$. As observed above, $V(f;a,x)$ is increasing. Also by Proposition 1.7, if $a < x < y < b$ then

$$\begin{aligned}
&(V(f;a,y) - f(y)) - (V(f;a,x) - f(x)) \\
&\quad = \quad V(f;a,y) - V(f;a,x) + f(x) - f(y) \\
&\quad = \quad V(f;x,y) - (f(y) - f(x)) \geq 0,
\end{aligned}$$

which proves the theorem with $f_1(x) = V(f;a,x)$ and $f_2(x) = V(f;a,x) - f(x)$. ☐

Theorems 1.5, 1.8 immediately imply the following

Corollary 1.9. *Let $f \in BV[a,b]$. Then f' exists almost everywhere, it is measurable and belongs to $L^1[a,b]$.*

Exercise 1.3.2.
Let $f \in L^1[a,b]$ and set

$$F(x) = \int_a^x f(t)\,dt.$$

Then $F \in BV[a,b]$ and $V(F;a,b) \leq \int_a^b |f(x)|\,dx$.

Corollary 1.10. *Let $f \in L^1[a,b]$ and set $F(x) = \int_a^x f(t)\,dt$. Then F is differentiable almost everywhere.*

Exercise 1.3.3.
Suppose that the function f is Lipschitz continuous on $[a,b]$. That is, there is a constant K such that $|f(x) - f(y)| \leq K|x-y|$ for all $x, y \in [a,b]$. Then $f \in BV[a,b]$.

Exercise 1.3.4.
Decide which of the following functions are of Bounded Variation on $[0,1]$:

$$\begin{aligned}
f_1(x) &= x^2 \sin(1/x^2); \\
f_2(x) &= x^2 \sin(1/x); \\
f_3(x) &= x^2 \sin(1/x^{3/2}),
\end{aligned}$$

for $x \in (0,1]$ and $f_1(0) = f_2(0) = f_3(0) = 0$.

Exercise 1.3.5.
Suppose $f \in BV[a,b]$. Then f is right continuous at $x \in [a,b)$ if and only if the function $V(f;a,x)$ is right continuous at x.

Exercise 1.3.6.
Suppose $f \in BV[a,b]$. Prove that

$$\int_a^b |f'(x)|\,dx \leq V(f;a,b)$$

Exercise 1.3.7.
Suppose $f, f_n \in BV[a,b], n = 1,2,\ldots$.

(i) Suppose that $V(f_n - f;a,b) \to 0$. Prove that there is a subsequence of $\{f_n'\}$ which converges almost everywhere to f'.

(ii) Suppose

$$\sum_{n=1}^{\infty} V(f_n - f;a,b) < \infty.$$

Prove that f_n' converges almost everywhere to f'.

1.4 Absolute Continuity

We start this section with a general lemma which, as we will see below, plays an important role in the study of absolutely continuous functions.

Lemma 1.11. *Let $(X,\mathcal{F},\mu)$ be a measure space and let $f \in L^1(\mu)$. Then, given $\varepsilon > 0$ there exists a $\delta > 0$ such that*

$$\int_A |f|\,d\mu < \varepsilon \text{ whenever } \mu(A) < \delta.$$

Proof. Suppose this is not the case. Then there is an $\varepsilon > 0$ and a sequence of measurable sets E_n such that

$$\mu(E_n) < 2^{-n} \text{ but } \int_{E_n} |f|\,d\mu \geq \varepsilon.$$

Set

$$A_n = \bigcup_{j=n}^{\infty} E_j, \text{ and } A = \bigcap_{n=1}^{\infty} A_n$$

The set A consists of all those x which belong to infinitely many of the E_n. In analysis A is often denoted by $\limsup_{n\to\infty} E_n$ and in probability this is usually written as $A = \{E_n, i.o.\}$ where i.o. means infinitely often. Thus if we set

$$g(x) = \sum_{n=1}^{\infty} \chi_{E_n}(x),$$

then $x \in A$ if and only if $g(x) = \infty$. Since $\sum \mu(E_n) < \infty$ we see that $g(x) \in L^1(\mu)$. Therefore $g(x) < \infty$ almost everywhere and $\mu(A) = 0$. But then, since $A_{n+1} \subset A_n$,

$$\int_A |f|\, d\mu = \lim_{n\to\infty} \int_{A_n} |f|\, d\mu \geq \lim_{n\to\infty} \int_{E_n} |f|\, d\mu \geq \varepsilon,$$

where the first inequality follows from the monote convergence theorem. This contradicts the fact that $\mu(A) = 0$ and proves the Lemma. □

The proof above gives one of the so called "Borel-Cantelli Lemma" which is a basic tool in proving almost sure (a.e.) convergence results.

Theorem 1.12 (Borel-Cantelli Lemma)**.** *Suppose that* $\{E_n\}$ *is a sequence of measurable sets with the property that* $\sum_{n=1}^{\infty} \mu(E_n) < \infty$*. Then* $\mu\{E_n, i.o.\} = 0$*.*

We shall now introduce the concept of absolute continuity. First, suppose $f \in L^1[a, b]$ and let $\{(a_i, b_i)\}_{i=1}^N$ be a disjoint collection of intervals. Set

$$F(x) = \int_a^x f(t)\, dt. \tag{1.16}$$

We clearly have

$$\sum_{i=1}^N |F(b_i) - F(a_i)| = \sum_{i=1}^N \left| \int_{a_i}^{b_i} f(t)\, dt \right| \leq \sum_{i=1}^N \int_{a_i}^{b_i} |f(t)|\, dt = \int_A |f(t)|\, dt,$$

where $A = \bigcup_{i=1}^N (a_i, b_i)$. Thus by Lemma 1.12, for every $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\sum_{i=1}^N |F(b_i) - F(a_i)| < \varepsilon \text{ whenever } \sum_{i=1}^N |b_i - a_i| < \delta.$$

The function given by (1.16) is an example of an absolutely continuous function.

Definition 1.13. The real valued function f defined on $[a, b]$ is said to be *absolutely continuous*, and we write $f \in AC[a, b]$, if given $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\sum_{i=1}^{N} |f(b_i) - f(a_i)| < \varepsilon$$

for every finite collection of disjoint intervals $\{(a_i, b_i)\}_{i=1}^{N}$ with

$$\sum_{i=1}^{N} |b_i - a_i| < \delta.$$

Remark 1.1. Clearly every absolutely continuous function is continuous (uniformly continuous). The Cantor ternary function provides an example of a continuous function which is not absolutely continuous. We leave the verification of this fact to the interested reader.

We showed above that if a function is given as the indefinite integral of an L^1 function then such a function is absolutely continuous. The next theorem shows that this property in fact characterizes absolutely continuous. Thus absolutely continuous functions are exactly those for which the "the fundamental theorem of calculus" holds.

Theorem 1.14. *Let $f \in AC[a, b]$. Then f is differentiable almost everywhere, $f'(x) \in L^1[a, b]$ and*

$$f(x) = f(a) + \int_a^x f'(t)\, dt, \text{ for } x \in [a, b].$$

We begin by proving that functions in $AC[a, b]$ are in $BV[a, b]$ and hence differentiable almost everywhere.

Lemma 1.15. *If $f \in AC[a, b]$, then $f \in BV[a, b]$.*

Proof. We take $\varepsilon = 1$ and let δ be as guaranteed by the property of absolutely continuous. Let $\mathcal{P}$ be a partition of $[a, b]$. By inserting new points in $\mathcal{P}$ we can divide $\mathcal{P}$ into K groups of disjoint intervals such that the sum of the lengths of the intervals in each group is no more than δ. Thus $V(f; a, b) \le K$. A simple geometric argument shows K is no more than $1 + \frac{b-a}{\delta}$. This completes the proof. $\square$

Corollary 1.16. *If $f \in AC[a,b]$, then f is differentiable almost everywhere and $f' \in L^1[a,b]$. In particular, the function $F(x)$ defined by (1.16) is differentiable almost everywhere and $F' \in L^1[a,b]$.*

Lemma 1.17. *Suppose $f \in L^1[a,b]$ and that*

$$\int_a^c f(t)\,dt = 0$$

for all $c \in [a,b]$. Then $f(x) = 0$ for almost every $x \in [a,b]$.

Proof. It suffices to show that $\int_E f(x)\,dx = 0$ for all measurable subsets E of $[a,b]$. Our assumption implies that $\int_{a_1}^{b_1} f(t)\,dt = 0$ for any $(a_1,b_1) \subset [a,b]$. Given any open set $\mathcal{O} \subset [a,b]$ we can write it as the disjoint union of such open intervals and we get

$$\int_{\mathcal{O}} f(x)\,dx = 0$$

for any such open set $\mathcal{O}$. Let $\varepsilon > 0$ be given and let $\delta > 0$ be the δ given by Lemma 1.11. Let E be any measurable subset of $[a,b]$. Choose an open set $\mathcal{O}$ with $E \subset \mathcal{O}$ and such that $m(\mathcal{O} \setminus E) < \delta$. Then

$$\begin{aligned}\left|\int_E f(x)\,dx\right| &= \left|\int_{\mathcal{O}} f(x)\,dx - \int_{\mathcal{O}\setminus E} f(x)\,dx\right| \\ &= \left|\int_{\mathcal{O}\setminus E} f(x)\,dx\right| \le \int_{\mathcal{O}\setminus E} |f(x)|\,dx < \varepsilon,\end{aligned}$$

where in the last inequality we applied Lemma 1.11. Since E and ε are arbitrary we conclude that $f(x) = 0$ almost everywhere. This proves the lemma. □

Lemma 1.18. *Let $f \in L^1[a,b]$ and set*

$$F(x) = f(a) + \int_a^x f(t)\,dt.$$

Then $F'(x) = f(x)$ for almost every $x \in [a,b]$.

Proof. We have already shown in Corollary 1.16 above that F is differentiable a.e. and that its derivative is in $L^1[a,b]$. What remains to be proved is that $F'(x) = f(x)$ a.e. Assume first that $f \ge 0$ is bounded, say $|f| \le M$. Set

$$f_n(x) = n\Big[F\big(x + \tfrac{1}{n}\big) - F(x)\Big],$$

for $x \in [a, b - 1/n]$ and $f_n(x) = 0$ for $x \in [b - 1/n, b]$. With $h = 1/n$ we have

$$f_n(x) = \frac{1}{h} \int_x^{x+h} f(t)\, dt,$$

$|f_n(x)| \leq M$ and $f_n(x) \to F'(x)$ for almost every $x \in [a, b]$. By the bounded convergence theorem and the continuity of the function F, for every $c \in (a, b)$

$$\begin{aligned}
\int_a^c F'(x) dx &= \lim_{n\to\infty} \int_a^c f_n(x)\, dx \\
&= \lim_{h\to 0} \frac{1}{h} \int_a^c [F(x+h) - F(x)]\, dx \\
&= \lim_{h\to 0} \left(\frac{1}{h} \int_c^{c+h} F(x)\, dx - \frac{1}{h} \int_a^{a+h} F(x)\, dx \right) \\
&= F(c) - F(a) = \int_a^c f(t)\, dt.
\end{aligned}$$

Thus

$$\int_a^c \left(F'(x) - f(x)\right)\, dx = 0,$$

for every $c \in [a, b]$. Since the function $F'(x) - f(x)$ is in $L^1[a, b]$, we conclude from this and Lemma 1.17 that $F'(x) = f(x)$ for almost every x, provided f is bounded.

For general f we may assume that f is nonnegative and set $f_n(x) = f(x) \wedge n$, the minimum of $f(x)$ and n. Then $f - f_n \geq 0$ and hence the functions defined by

$$F_n(x) = \int_a^x \big(f(t) - f_n(t)\big)\, dt$$

are increasing and their derivatives are nonnegative. Since the functions f_n are bounded, we have

$$\frac{d}{dx} \int_a^x f_n(t)\, dt = f_n(x) \text{ a.e.}$$

and

$$F'(x) = F_n'(x) + \frac{d}{dx} \int_a^x f_n(t)\, dt \geq f_n(x),$$

for all n and almost every x. Hence $F'(x) \geq f(x)$ a.e. Integrating this we obtain

$$\int_a^b F'(x)\, dx \geq \int_a^b f(x)\, dt = F(b) - F(a).$$

Combining this inequality with Theorem 1.5 gives

$$\int_a^b F'(x)\,dx = F(b) - F(a).$$

This implies that

$$\int_a^b [F'(x) - f(x)]\,dx = 0$$

and since $F'(x) \geq f(x)$ almost everywhere, $f = F'$ almost everywhere, finishing the proof of Lemma 1.18. $\square$

Lemma 1.19. *Suppose $f \in AC[a,b]$ and $f'(x) = 0$ for almost every x. Then $f =$ constant.*

Proof. Let $c \in (a,b)$. We will prove that $f(a) = f(c)$. Let $E \subset (a,c)$ be such that $f'(x) = 0$ for all $x \in E$ and such that $|E| = |c-a|$. Given $\varepsilon > 0$ take $\delta > 0$ satisfying the properties in the definition of absolute continuity. Let $\eta > 0$. Since $f'(x) = 0$ for every $x \in E$, there is an $h_x > 0$ such that $I_x = [x, x+h_x] \subset [a,c]$ and $|f(x+h_x) - f(x)| < \eta h_x$. The collection $\{I_x, x \in E\}$ is a Vitali covering for E. By Corollary 1.3, we can find a finite subcollection of such intervals, $\{[x_k, x_k + h_{x_k}]\}_{k=1}^N$ such that $|E \setminus \bigcup I_k| < \delta$. If we set $y_k = x_k + h_k$ and label $x_k < x_{k+1}$, this gives

$$\sum_{k=1}^N |y_k - x_{k+1}| < \delta. \tag{1.17}$$

By the absolute continuity property,

$$\sum_{k=1}^N |f(y_k) - f(x_{k+1})| < \varepsilon. \tag{1.18}$$

Also,

$$\sum_{k=1}^N |f(y_k) - f(x_k)| \leq \eta \sum_{k=1}^N h_{x_k} \leq \eta(c-a). \tag{1.19}$$

The triangle inequality, (1.18) and (1.19) now gives that $|f(a) - f(c)| \leq \varepsilon + \eta(c-a)$ with ε and η as small as we like. This completes the proof of Lemma 1.19. $\square$

Proof of Theorem 1.14. By Lemma 1.15, $f \in BV$ and therefore the Jordan decomposition (Theorem 1.8) gives $f(x) = f_1(x) - f_2(x)$ where the functions f_1 and f_2 are nondecreasing. We have by Theorem 1.5,

$$\int_a^b |f'(x)|\,dx \le \big(f_1(b) - f_1(a)\big) + \big(f_2(b) - f_2(a)\big) < \infty.$$

So, $f' \in L^1[a,b]$. Let $g(x) = \int_a^x f'(x)\,dx$. Then $g \in AC[a,b]$ and $g'(x) = f'(x)$. Therefore by Lemma 1.19, $g(x) - f(x) = C$, a constant. Since $g(a) = 0$, we have that $f(a) = C$ and so

$$f(x) = \int_a^x f'(t)\,dt + f(a),$$

which completes the proof. □

A function h defined on $[a,b]$ is said to be singular if $h'(x) = 0$ for almost every $x \in [a,b]$. The Cantor ternary function provides an example of a singular function. We have

Theorem 1.20 (Lebesgue Decomposition)**.** *Suppose $f \in BV[a,b]$. We can write $f(x) = g(x) + h(x)$ where $g \in AC[a,b]$ and h is singular. Up to constants, this decomposition is unique.*

Proof. Since f is of bounded variation, f' exists almost everywhere and it is in $L^1[a,b]$. If we set

$$g(x) = \int_a^x f'(t)\,dt$$

and

$$h(x) = f(x) - g(x),$$

then $g \in [a,b]$ and by Lemma 1.18, h is singular. This gives the desired decomposition. If $f = g_1 + h_1$ is another decomposition we get that $g - g_1 = h_1 - h$ and hence the function $g - g_1$ is both absolutely continuous and singular and hence constant, by Lemma 1.19. This completes the proof of the theorem. □

Exercise 1.4.1.
Prove that every Lipschitz continuous function is absolutely continuous.

Exercise 1.4.2.
Suppose $f \in AC[a,b]$. Then

$$V(f;a,b) = \int_a^b |f'(x)|\,dx.$$

Exercise 1.4.3.
Suppose $f \in AC[a,b]$ and $E \subset [a,b]$ has Lebesgue measure zero. Then $f(E)$ also has Lebesgue measure zero.

Exercise 1.4.4.
Suppose $1 \le p < \infty$, $f \in L^p(\mathbb{R})$ and

$$F(x) = \int_x^{x+1} f(t)\,dt.$$

Prove that F vanishes at infinity. Is this also the case if $p = \infty$?

Exercise 1.4.5.
Suppose $f \in AC[-k,k]$ for any k and $f' \in L^2(\mathbb{R})$. Find

$$\lim_{x \to \pm\infty} |f(x+1) - f(x)|$$

Is it true that f vanishes at ∞?

Exercise 1.4.6.
Let f be a continuous function in $BV[a,b]$. Prove that $f \in AC[a,b]$ if and only if f takes sets of measure zero to sets of measure zero.

Chapter 2

The Radon-Nikodym Theorem

A basic result in the theory of L^p–spaces asserts that, under the appropriate conditions, the dual of L^p is L^q where p and q are conjugate exponents. This theorem is equivalent to the Radon–Nikodym theorem which asserts that if the measure ν is absolutely continuous with respect to the measure μ, then ν can be represented as the integral of a function against the measure μ. There are several ways to prove these results but the classical approach, which we will present here, is via signed measures. It follows Royden [Ro] and Rudin [Ru2].

2.1 Signed Measures

If μ_1 and μ_2 are two measures, a new measure μ can be obtained by setting $\mu = \mu_1 + \mu_2$. Thus adding measures presents no difficulties. What about subtracting measures? This and other problems immediately lead to the notion of *signed measures*.

Definition 2.1. By a signed measure on a measurable space $(X, \mathcal{F})$ we mean an extended real valued function ν defined on $\mathcal{F}$ such that

(i) ν assumes at most one of the values $+\infty$ or $-\infty$,

(ii) $\nu(\emptyset) = 0$,

(iii) $\nu\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \nu(E_j)$,

shenever the E_j are disjoint and in $\mathcal{F}$. By (iii) we mean that the series is absolutely convergent if $\nu(\bigcup_{j=1}^{\infty} E_j)$ is finite and properly divergent otherwise.

Example 2.1. Let $f \in L^1[a,b]$ and set

$$\nu(E) = \int_E f\,dx.$$

ν is a signed measure on $[a,b]$ with $|\nu(E)| < \infty$ for all E.

Definition 2.2. A set $B \in \mathcal{F}$ is said to be

(i) ν*-positive* if $\nu(E) \geq 0$ for every measurable subset $E \subset B$.

(ii) ν*-negative* if $\nu(E) \leq 0$ for every measurable subset $E \subset B$.

(iii) A set which is both positive and negative is a null set.

Thus a measurable set is null if and only if every measurable subset of the set has measure zero.

Remark 2.1. We warn the reader that null sets are not the same as sets of measure zero, as can be easily seen from Example 2.1.

Our first goal in understanding signed measures is to prove that the space X can be written as the disjoint union of a positive set and a negative set. This is called the *Hahn Decomposition*. Before we state this more precisely, we need some lemmas.

Lemma 2.3. *The countable union of positive sets is positive and every measurable subset of a positive set is positive.*

Proof. Suppose that A is positive and that E is a measurable subset of A. If $E_1 \in \mathcal{F}$ and $E_1 \subset E$ then $E_1 \subset A$ and so $\nu(E_1) \geq 0$, since A is positive. So, E is positive.

Next, let $A = \bigcup_{n=1}^{\infty} A_n,\ A_n$ positive. Let $E \subset A$ be measurable. Suppose we can write

$$E = \bigcup_{n=1}^{\infty} E_n, \quad E_n \subset A_n, \quad E_n \cap E_m = \emptyset, \quad n \neq m, \tag{2.1}$$

with each E_n measurable. Since each of the A_n are positive, $\nu(E_n) \geq 0$ for all n and

$$\nu(E) = \sum_{n=1}^{\infty} \nu(E_n) \geq 0.$$

It would then follow that A is positive.

Thus it remains to prove (2.1). Set $E_1 = E \cap A_1$ and for $n \geq 2$, $E_n = E \cap A_n \cap A_{n-1}^c \cap \cdots \cap A_1^c$. Let $x \in E_n$. Then $x \in E$ and $x \in A_n$ but $x \notin A_{n-1}, \ldots, A_1$. Thus if $n > m$, $x \notin E_m$. Therefore $E_n \cap E_m = \emptyset$, if $n \neq m$. If $x \in E$, let n be the first n such that $x \in A_n$. Then $x \in E_n$ and we are done. □

Lemma 2.4. *Suppose E is a measurable set with $0 < \nu(E) < \infty$. Then there exists a measurable, positive subset of E of positive measure.*

Proof. Either E is positive, in which case we are done, or it contains a set of negative measure. Let n_1 be the smallest positive integer such that there exists a set $E_1 \subset E$ with

$$\nu(E_1) < -\frac{1}{n_1}. \tag{2.2}$$

Now consider $E \setminus E_1 \subset E$. Again, if $E \setminus E_1$ is positive with $\nu(E \setminus E_1) > 0$ we are done. If not, let n_2 be the smallest positive integer larger than or equal to n_1 such that there exists a set $E_2 \subset E \setminus E_1$ with

$$\nu(E_2) < -\frac{1}{n_2}. \tag{2.3}$$

Continuing this way we obtain either a positive subset E with positive measure or an infinite sequence of integers $\{n_k\}$ and sets $\{E_k\}$ with the property that $E_k \subset E \setminus \bigcup_{j=1}^{k-1} E_j$, and

$$\nu(E_k) < -\frac{1}{n_k}. \tag{2.4}$$

Set

$$A = E \setminus \bigcup_{k=1}^{\infty} E_k. \tag{2.5}$$

We claim this A has the desired properties.

To show that $\nu(A) > 0$ observe that $E = A \cup \bigcup_{k=1}^{\infty} E_k$, where the sets are disjoint. Thus

$$\nu(E) = \nu(A) + \sum_{k=1}^{\infty} \nu(E_k).$$

Since $0 < \nu(E) < \infty$, $\sum_{k=1}^{\infty} \nu(E_k)$ converges and we have

$$\sum_{k=1}^{\infty} \frac{1}{n_k} < -\sum_{k=1}^{\infty} \nu(E_k) < \infty.$$

But $\nu(E_k) \leq 0$, and hence $\nu(A) > 0$. This shows that $n_k \to \infty$. Now, use this fact to prove the following exercise, which finishes the proof of the lemma. $\square$

Exercise 2.1.1.
Prove that the set A in (2.5) is positive.

Theorem 2.5 (Hahn–Decomposition)**.** *Let ν be a signed measure on the measurable space $(X, \mathcal{F})$. There is a positive set A and a negative set B such that $A \cap B = \emptyset$, $A \cup B = X$.*

Proof. Let us denote the collection of all positive sets by $\mathcal{P}$ and the collection of all negative sets by $\mathcal{N}$. We may assume that ν does not take the value $+\infty$ (otherwise we consider $-\nu$). Let $\lambda = \sup\{\nu(A) \colon A \in \mathcal{P}\}$. Since $\emptyset \in \mathcal{P}$, $\lambda \geq 0$. Chose a sequence of sets $A_n \in \mathcal{P}$ such that $\lambda = \lim_{n\to\infty} \nu(A_n)$ and set $A = \cup A_n$. Then $A \in \mathcal{P}$, by Lemma 2.3 and $\nu(A) \leq \lambda$. Since $A \setminus A_k \subset A$, we have that $\nu(A \setminus A_k) \geq 0$ and $\nu(A) = \nu(A_k) + \nu(A \setminus A_k) \geq \nu(A_k)$ for all k. Thus, $\nu(A) \geq \lambda$ from which we conclude that $0 \leq \nu(A) = \lambda < \infty$.

Let $B = A^c$. We will show that B is negative. To do this, let $E \subset B$ and $\nu(E) \geq 0$. We will show that in fact $\nu(E) = 0$. For suppose $0 < \nu(E)$. Since ν does not take the value $+\infty$ we also have $\nu(E) < \infty$. By Lemma 2.4 there is a positive set $\widetilde{E} \subset E$ with $0 < \nu(\widetilde{E}) < \infty$. But then $\widetilde{E} \cup A \in \mathcal{P}$ and $\widetilde{E} \cap A = \emptyset$ and it follows that

$$\lambda \geq \nu(\widetilde{E} \cup A) = \nu(\widetilde{E}) + \nu(A) = \nu(\widetilde{E}) + \lambda.$$

This implies that $\nu(E) = 0$, giving a contradiction. $\square$

Remark 2.2. The Hahn decomposition gives rise to the Jordan decomposition of signed measures in the following way. For $E \in \mathcal{F}$, we set

$$\nu^+(E) = \nu(E \cap A)$$

and

$$\nu^-(E) = -\nu(E \cap B).$$

Then ν^+ and ν^- are both (positive) measures and $\nu^+(B) = 0$ and $\nu^-(A) = 0$. The measure $|\nu| = \nu^+ + \nu^-$ is called the total variation of ν. If we define two measures ν_1 and ν_2 to be mutually singular, written as $\nu_1 \perp \nu_2$, if there exist two measurable sets A and B such that $A \cup B = X$, $A \cap B = \emptyset$ and $\nu_1(A) = \nu_2(B) = 0$, we see that ν^+ and ν^- are mutually singular.

This remark immediately leads to

Theorem 2.6 (Jordan Decomposition for signed measures)**.** *Let ν be a signed measure. There are two mutually singular measures ν^+, ν^- such that $\nu = \nu^+ - \nu^-$. This decomposition is unique.*

Example 2.2. $f \in L^1[a,b]$ and set

$$\nu(E) = \int_E f\,dx.$$

Then ν is a signed measure and the Jordan Decomposition is given by

$$\nu^+(E) = \int_E f^+\,dx$$

and

$$\nu^-(E) = \int_E f^-\,dx,$$

where $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$

2.2 The Radon–Nikodym Theorem

Definition 2.7. Let ν and μ be two measures on $(X, \mathcal{F})$. We will say that ν is absolutely continuous with respect to μ, written as $\nu \ll \mu$, if $\nu(E) = 0$ for every $E \in \mathcal{F}$ for which $\mu(E) = 0$. If ν is a signed measure and μ is a measure we will say that ν is absolutely continuous with respect to μ, again denoted by $\nu \ll \mu$, if $|\nu| \ll \mu$. That is, if both ν^+ and ν^- are absolutely continuous with respect to μ.

Example 2.3. Let $(X, \mathcal{F}, \mu)$ be a measure space and suppose that f is a nonnegative measurable function on X. Then the measure ν defined on $\mathcal{F}$ by

$$\nu(E) = \int_E f\,d\mu \tag{2.6}$$

is absolutely continuous with respect to μ.

The next theorem says that the measures ν of (2.1) are more the rule than the exception. First, recall that a space $(X, \mathcal{F}, \mu)$ is σ–finite if X can be written as the countable union of sets of finite μ–measure.

Theorem 2.8 (Radon–Nikodym)**.** *Let $(X, \mathcal{F}, \mu)$ be a σ–finite measure space. Assume that ν is a measure on $(X, \mathcal{F})$ with $\nu \ll \mu$. Then there exists a nonnegative measurable function f such that*

$$\nu(E) = \int_E f d\mu,$$

for all $E \in \mathcal{F}$. The function f is unique a.e. $[\mu]$. We call f the Radon–Nikodym derivative of ν with respect to μ. This is often written as $f = d\nu/d\mu$.

Remark 2.3. Let $X = [0, 1]$ and let $\mathcal{F}$ be the Borel σ–algebra on X. Let μ be be the counting measure (which is not σ–finite) and m the Lebesgue measure on $(X, \mathcal{F})$, respectively. Clearly $m \ll \mu$. However, if it were the case that

$$m(E) = \int_E f \, d\mu,$$

for all $E \in \mathcal{F}$, then $f(x) = 0$ for every $x \in [0, 1]$, since then

$$f(x) = \int_{\{x\}} f \, d\mu.$$

Thus the hypothesis of σ–finiteness cannot be removed.

Lemma 2.9. *Let $(X, \mathcal{F})$ be a measure space and suppose $\{B_\alpha\}_{\alpha \in D}$ is a collection of measurable sets indexed by a countable set of real numbers D. Suppose that $B_\alpha \subset B_\beta$ whenever $\alpha < \beta$. Then there exists a measurable function f on X such that $f(x) \le \alpha$ on B_α and $f(x) \ge \alpha$ on B_α^c for each $\alpha \in D$.*

Proof. For $x \in X$, set $f(x)$ to be the first α such that $x \in B_\alpha$. That is, $f(x) = \inf\{\alpha\colon x \in B_\alpha\}$ where $\inf\{\emptyset\} = \infty$. If $x \notin B_\alpha$ then $x \notin B_\beta$ for any $\beta < \alpha$ and so $f(x) \ge \alpha$. If $x \in B_\alpha$, then $f(x) \le \alpha$. It remains to show that such an f is measurable. To see this, let λ be any real number. Clearly,

$$\{x\colon f(x) < \lambda\} = \bigcup_{\beta < \lambda, \beta \in D} B_\beta.$$

Since the right hand side is a countable union of measurable sets, the left hand side is also measurable and the lemma is proved. □

We need the measure theoretic versions of the above result. The reader can consult [Ro, p. xx] which proof we leave as an exercise.

Lemma 2.10. *Suppose $\{B_\alpha\}_{\alpha\in D}$ is as in Lemma 2.9 but this time $\alpha < \beta$ only implies $\mu\left(B_\alpha \setminus B_\beta\right) = 0$. There exists a measurable function f on X such that $f(x) \leq \alpha$ a.e. on B_α and $f(x) \geq \alpha$ a.e. on B_α^c, for each $\alpha \in D$. If, in addition, the index set D is dense in $\mathbb{R}$, the function f is unique μ–almost everywhere.*

Proof of Radon–Nikodym theorem. Assume first $\mu(X) = 1$. For any rational α set

$$\nu_\alpha = \nu - \alpha\mu. \tag{2.7}$$

Then ν_α is a signed measure. Let $\{A_\alpha, B_\alpha\}$ be a Hahn–decomposition for ν_α where we can set $A_\alpha = X$, $B_\alpha = \emptyset$ for all $\alpha \leq 0$. Notice that for arbitrary $\alpha, \beta \geq 0$,

$$B_\alpha \setminus B_\beta = B_\alpha \cap (X \setminus B_\beta) = B_\alpha \cap A_\beta. \tag{2.8}$$

Thus the set $B_\alpha \setminus B_\beta$ is negative with respect to ν_α and positive with respect to ν_β. Hence

$$\nu_\alpha(B_\alpha \setminus B_\beta) \leq 0 \text{ and } \nu_\beta(B_\alpha \setminus B_\beta) \geq 0. \tag{2.9}$$

The inequalities in (2.9) are equivalent to the following two inequalities:

$$\nu(B_\alpha \setminus B_\beta) - \alpha\mu(B_\alpha \setminus B_\beta) \leq 0$$

and

$$\nu(B_\alpha \setminus B_\beta) - \beta\mu(B_\alpha \setminus B_\beta) \geq 0,$$

which give

$$\beta\mu(B_\alpha \setminus B_\beta) \leq \nu(B_\alpha \setminus B_\beta) \leq \alpha\mu(B_\alpha \setminus B_\beta). \tag{2.10}$$

We can conclude from (2.10) that

$$\mu(B_\alpha \setminus B_\beta) = 0,$$

whenever $\alpha < \beta$. By Lemma 2.10 there is a measurable function f such that for each $\alpha \in \mathbb{Q}$, $f \geq \alpha$ a.e. on A_α and $f \leq \alpha$ a.e. on B_α. Since $A_\alpha = X$, $B_\alpha = \emptyset$ for all $\alpha \leq 0$, we see that $f \geq 0$ a.e. Let $E \in \mathcal{F}$ and let N be any positive integer. For $k = 0, 1, 2, \ldots$, set

$$E_k = E \cap \left(B_{(k+1)/N} \setminus B_{k/N}\right)$$

and

$$E_\infty = E \setminus \bigcup_{k=0}^{\infty} B_{k/N}.$$

Then $E_0, \ldots, E_\infty$ are disjoint a.e. and $E = \left(\bigcup_{k=0}^\infty E_k\right) \cup E_\infty$. So,

$$\nu(E) = \nu(E_\infty) + \sum_{k=0}^\infty \nu(E_k).$$

By the properties of the function f, we see that, for almost every $x \in E_k \subset B_{(k+1)/N} \setminus B_{k/N} = B_{(k+1)/N} \cap A_{k/N}$,

$$\frac{k}{N} \leq f(x) \leq \frac{k+1}{N}.$$

Thus, integrating over E_k,

$$\frac{k}{N}\mu(E_k) \leq \int_{E_k} f d\mu \leq \frac{k+1}{N}\mu(E_k). \tag{2.11}$$

Also, since the pair (A_α, B_α) is the Hahn decomposition for ν_α, $E_k \subset A_{k/N}$ implies that

$$\frac{k}{N}\mu(E_k) \leq \nu(E_k) \tag{2.12}$$

and $E_k \subset B_{k+1/N}$ implies that

$$\nu(E_k) \leq \frac{k+1}{N}\mu(E_k). \tag{2.13}$$

(2.11), (2.12) and (2.13) give

$$\begin{aligned}\nu(E_k) - \frac{1}{N}\mu(E_k) \leq \frac{k}{N}\mu(E_k) \leq \int_{E_k} f\, d\mu \\ \leq \quad \frac{k+1}{N}\mu(E_k) \leq \nu(E_k) + \frac{1}{N}\mu(E_k),\end{aligned} \tag{2.14}$$

valid for all k.

On the other hand, on E_∞, $f \equiv \infty$ a.e. If $\mu(E_\infty) > 0$, then $\nu(E_\infty) = \infty$, since $(\nu - \alpha\mu)(E_\infty) \geq 0$ for all α. If $\mu(E_\infty) = 0$ then $\nu(E_\infty) = 0$. (This is the place where we use absolute continuity.) Thus in any case,

$$\nu(E_\infty) = \int_{E_\infty} f\, d\mu. \tag{2.15}$$

Adding (2.14) and (2.15) gives

$$\nu(E) - \frac{1}{N}\mu(E) \le \int_E f\,d\mu \le \nu(E) + \frac{1}{N}\mu(E).$$

Since $\mu(X) = 1$ and N is arbitrary, we conclude that

$$\nu(E) = \int_E f\,d\mu.$$

It remains to prove the uniqueness and to remove the assumption that $\mu(X) < \infty$. Suppose that μ is σ–finite and that $\nu \ll \mu$. Let $\{X_i, i = 1, 2, \dots\}$ be such that $\mu(X_i) < \infty$ for all i, $X_i \cap X_j = \emptyset$ whenever $i \neq j$ and $\bigcup_{j=1}^{\infty} X_j = X$. Put

$$\mu_j(E) = \mu(E \cap X_j),\ \nu_j(E) = \nu(E \cap X_j).$$

Let us denote here by 1_E the characteristic function of E, more frequently written in these book as χ_E. Clearly $\nu_j \ll \mu_j$ and by what we have already shown there is an $f_j \ge 0$ such that

$$\nu_j(E) = \int_E f_j\,d\mu_j$$

or

$$\nu(E \cap X_j) = \int_{E \cap X_j} f_j\,d\mu = \int_E f_j 1_{X_j}\,d\mu.$$

Put

$$f = \sum_{j=1}^{\infty} f_j 1_{X_j}.$$

This and exercise 2.2.3 for dealing with the uniqueness finishes the proof. □

Remark 2.4. Suppose ν is a signed measure and μ a σ–finite measure such that $\nu \ll \mu$. That is, $|\nu| \ll \mu$. Since $|\nu| = \nu^+ + \nu^-$, $\nu^+ \ll \mu$ and $\nu^- \ll \mu$. Let f^+ and f^- be such that

$$\nu^+(E) = \int_E f^+\,d\mu \text{ and } \nu^-(E) = \int_E f^-\,d\mu.$$

Since either ν^+ or ν^- is finite, we have

$$\nu(E) = \nu^+(E) - \nu^-(E) = \int_E (f^+ - f^-)\,d\mu = \int_E f\,d\mu.$$

Thus the Radon–Nikodym theorem also holds in this case.

Theorem 2.11 (Lebesgue Decomposition for Measures)**.** *Let μ and ν be two σ–finite measures on $(X, \mathcal{F})$. There exist two σ–finite measures ν_s and ν_a such that*

(i) $\nu_a \ll \mu,\ \nu_s \perp \mu,$

(ii) $\nu = \nu_a + \nu_s.$

The measures ν_a and ν_s are unique.

Proof. Consider the σ–finite measure $\lambda = \mu + \nu$. Clearly $\lambda(E) = 0$ implies $\mu(E) = \nu(E) = 0$. Thus the Radon–Nikodym theorem gives two nonnegative functions f and g such that for every $E \in \mathcal{F}$,

$$\mu(E) = \int_E f\, d\lambda \text{ and } \nu(E) = \int_E g\, d\lambda.$$

Let

$$A = \{f > 0\}, \quad B = \{f = 0\}.$$

Then

$$X = A \cup B,\ A \cap B = \emptyset,$$

and $\mu(B) = 0$. Set

$$\nu_s(E) = \nu(E \cap B), \quad \text{for} \quad E \in \mathcal{F},$$

and

$$\nu_a(E) = \nu(E \cap A) = \int_{E \cap A} g\, d\lambda.$$

Then $\nu_s(A) = 0$ so that $\nu_s \perp \mu$. Clearly $\nu = \nu_a + \nu_s$ and it only remains to show that $\nu_a \ll \mu$. Assume $\mu(E) = 0$. Then

$$\int_E f\, d\lambda = 0$$

and since f is nonnegative, $f \equiv 0$ a.e. $[\lambda]$ on E. Since $f > 0$ on $E \cap A$, this forces $\lambda(E \cap A) = 0$. Thus,

$$\nu_a(E) = \int_{E \cap A} g\, d\lambda = 0.$$

□

Exercise 2.2.1. (i) Suppose μ, ν and λ are measures satisfying $\nu \ll \mu$ and $\mu \perp \lambda$. Then $\nu \perp \lambda$.

(ii) Give an example of two measures ν and μ with $\nu \ll \mu$ but μ not absolutely continuous with respect to ν.

Exercise 2.2.2.
Prove that the Hahn decomposition need not be unique.

Exercise 2.2.3.
Prove the a.e. uniqueness of the function in Theorem 2.8.

Exercise 2.2.4.
Prove the uniqueness of the measures in the Lebesgue decomposition.

Exercise 2.2.5.
Let ν_1 and ν_2 be two signed measures and let μ be a σ–finite measure such that $\nu_1, \nu_2 \ll \mu$. Then $\nu_1 + \nu_2 \ll \mu$ and

$$d(\nu_1 + \nu_2)/d\mu = d\nu_1/d\mu + d\nu_2/d\mu$$

Exercise 2.2.6.
Let ν and μ be two σ–finite measures on $(X, \mathcal{F})$ with $\nu \ll \mu$. Prove

$$\int_X f\,d\nu = \int_X f\,(d\nu/d\mu)\,d\mu,$$

for all $f \in L^1(\nu)$.

Exercise 2.2.7.
The Conditional Expectation. Let $(X, \mathcal{F}, \mu)$ be a σ–finite measure space. Let $\mathcal{F}_1 \subset \mathcal{F}$ be another σ–algebra of subsets of X. Let $f \in L^1(\mu)$ and define a new measure on $\mathcal{F}_1$ by

$$\nu(E) = \int_E f\,d\mu, \quad E \in \mathcal{F}_1$$

Clearly $\nu \ll \mu$. In this case $d\nu/d\mu$ is called the *conditional expectation* of f relative to $\mathcal{F}_1$ and it is denoted by $E(f|\mathcal{F}_1)$. Prove the following elementary properties of the conditional expectation.

(i) For all $A \in \mathcal{F}_1$, $\int_A E(f|\mathcal{F}_1)\,d\mu = \int_A f\,d\mu$.

(ii) Let $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}$. Then

$$E(E(f|\mathcal{F}_1)|\mathcal{F}_2) = E(E(f|\mathcal{F}_2)|\mathcal{F}_1) = E(f|\mathcal{F}_1).$$

(iii) If $f, g \in L^1(\mu)$ and $fg \in L^1(\mu)$ with g measurable relative to $\mathcal{F}_1$, then $E(fg|\mathcal{F}_1) = gE(f|\mathcal{F}_1)$.

Exercise 2.2.8.
Let ν be a signed measure on $(X, \mathcal{F})$ such that $|\nu|(X) < \infty$. Prove that there is a measurable function h with $|h(x)| = 1$ for all $x \in X$ such that for all $E \in \mathcal{F}$,

$$\nu(E) = \int_E h(x)\, d|\nu|(x)$$

2.3 The Riesz Representation Theorem for L^p

We begin this section by reviewing some basic notions of normed linear spaces.

Definition 2.12. Let X be a vector space over the real or complex numbers, the scalar field. A function $\|\cdot\|\colon X \to \mathbb{R}^+$ is called a norm if

(i) $\|x + y\| \leq \|x\| + \|y\|$, $x,\ y \in Y$.

(ii) $\|\lambda x\| = |\lambda|\|x\|$, $x \in Y,\ \lambda$ is a scalar.

(iii) $\|x\| = 0$ if and only if $x = 0$.

The pair $(X, \|\cdot\|)$ is called a normed linear space.

Let us recall that for any measure space $(\mathcal{M}, \mathcal{F}, \mu)$, the space $L^p(\mu)$, $1 \leq p < \infty$ is the collection of measurable functions f for which

$$\|f\|_p = \left(\int_X |f|^p \, d\mu\right)^{\frac{1}{p}} < \infty.$$

For a measurable function f the essential supremum is defined by

$$\|f\|_\infty = \inf\{\lambda > 0 : \mu(\{|f| > \lambda\})\}$$

and we say that the function belongs to $L^\infty(\mu)$ if this quantity is finite.

Example 2.4. The following are the classical examples of normed linear spaces.

(i) $X = \mathbb{R}^n$, $\|x\| = \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}$,

(ii) $X = \mathbb{C}^n$, $\|z\| = \left(\sum_{i=1}^n |z_i|^2\right)^{1/2}$,

(iii) $X = L^p(\mu)$, $1 \le p \le \infty$, $\|f\| = \|f\|_p$

(iv) $X = BV[a,b]$, $\|f\| = |f(a)| + V(f;a,b)$.

If $(X, \|\cdot\|)$ is a normed linear space the function defined on $X \times X$ by $d(x,y) = \|x-y\|$ is a metric. That is, it is (*i*) symmetric, (*ii*) $d(x,y) = 0$ if and only if $x = y$ and (*iii*) it satisfies the triangle inequality, $d(x,y) \le d(x,z) + d(z,y)$ for all x, y, $z \in X$. If X is complete in this metric, that is, if every Cauchy sequence in X is convergent in X, then X is called a Banach space. The spaces in Example 2.4 are all Banach spaces.

An important class of normed linear spaces which are also Banach spaces are the Hilbert spaces, often denoted by H. More precisely, suppose that H is a vector space over the complex numbers which is an inner product space. That is, for each pair x, $y \in H$ there is a complex number $\langle x, y\rangle$, called the inner product of x and y, such that

(i) $\langle x, y\rangle = \overline{\langle y, x\rangle}$,

(ii) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$,

(iii) $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$,

(iv) $\langle x, x\rangle$ is nonnegative and is 0 if and only if $x = 0$.

We define the map $\|\cdot\| : H \to \mathbb{R}^+$ by $\|x\| = \sqrt{\langle x, x\rangle}$. By the Cauchy-Schwartz inequality this is a norm. If the metric space (H, d) with $d(x,y) = \|x - y\|$ is complete, then the space H is called a Hilbert space. The spaces in (i) and (ii) of Example 2.4 are Hilbert spaces. The space $H = L^2(\mu)$ is a Hilbert space with inner product $\langle f, g\rangle = \int f\overline{g}\, d\mu$.

We recall that the series $\sum_1^\infty x_n$ of elements of X converges if the sequence $\{S_n = \sum_{k=1}^n x_k\}$ of partial sums converges. The sequence converges absolutely if $\sum_1^\infty \|x_n\|$ converges.

Definition 2.13. A function L on a normed linear space X to its scalar field is called a linear functional if

$$L(\alpha x + y) = \alpha L(x) + L(y)$$

for all $x, y \in X$ and scalars α. L is continuous if whenever $\|x_n - x\| \to 0$, $|L(x_n) - L(x)| \to 0$. L is bounded if there is a constant C such that $|L(x)| \le C\|x\|$, $\forall x \in X$. The norm of L, $\|L\|$ is the smallest such C. That is

$$\|L\| = \sup_{\|x\| \ne 0} \frac{|L(x)|}{\|x\|} = \sup_{\|x\|=1} |L(x)|.$$

Proposition 2.14. *A linear functional L on a normed linear space is continuous if and only if it is bounded.*

Proof. Suppose that L is bounded and that $\|x_n - x\| \to 0$. Linearity gives

$$|L(x_n) - L(x)| = |L(x_n - x)| \leq \|L\|\|x_n - x\|$$

and hence L is continuous. In fact, L is Lipschitz continuous. Conversely, suppose that L is continuous but L is not bounded. Then there is a sequence $\{x_n\}$ of points in X such that

$$|L(x_n)| \geq n\|x_n\|$$

Set

$$y_n = \frac{x_n}{n\|x_n\|}.$$

By linearity, this sequence satisfies

$$|L(y_n)| \geq 1.$$

But

$$\|y_n\| = \frac{1}{n} \to 0$$

and this gives a contradiction to the continuity of L. □

Definition 2.15. If X is a normed linear space, X^* denotes the collection of all continuous linear functionals on X.

Exercise 2.3.1.
Let X be a normed linear space. Prove that X is a Banach space if and only if every absolutely convergent series of elements in X converges.

Exercise 2.3.2.
Let $X = BV[a,b]$ with the norm $\|f\| = |f(a)| + V(f;a,b)$. Prove that X is a Banach space.

Exercise 2.3.3.
Prove that $(X^*, \|\cdot\|)$ with the norm

$$\|L\| = \sup_{|x|\neq 0} \frac{|L(x)|}{\|x\|}$$

is a normed linear space.

Proposition 2.16. *Let* $(X, \|\cdot\|)$ *be any normed linear space. Then* X^* *is a Banach space under the norm* $\|L\|$ *of Exercise 2.3.3. This space is refer to as the dual space of* X.

Proof. By Exercises 2.3.1 and 2.3.3 it is enough to prove that if $\sum_{k=1}^{\infty} \|L_k\|$ converges then $\sum_{k=1}^{\infty} L_k$ converges in X^*. Towards this end, let $x \in X$. Then since $|L_k(x)| \leq \|L_k\|\|x\|$ we have

$$\sum_{k=1}^{\infty} |L_k(x)| \leq \|x\| \sum_{k=1}^{\infty} \|L_k\|.$$

Thus for each $x \in X$, $\sum_{k=1}^{\infty} L_k(x)$ is absolutely convergent. Set

$$L(x) = \sum_{k=1}^{\infty} L_k(x).$$

We shall now verify that L is a bounded linear functional and that the partial sums of the linear functionals L_k converges to L in the norm of X^*. The linearity of L follows trivially from the linearity of the L_k's and since $|Lx| \leq \|x\| \sum_{k=1}^{\infty} \|L_k\|$ we see that L is bounded and its norm does not exceed $\sum_{k=1}^{\infty} \|L_k\|$. It remains to prove the convergence in X^*. For this, observe that

$$\left\| L - \sum_{k=1}^{m} L_k \right\| \leq \sum_{k=m+1}^{\infty} \|L_k\|.$$

The right hand side of this inequality goes to 0 as $m \to \infty$ by the convergence of the series of real numbers $\sum_{k=1}^{\infty} \|L_k\|$. This completes the proof. $\square$

Example 2.5. Let $1 \leq q \leq \infty$. Let $g \in L^q(\mu)$ and define

$$L_g(f) = \int_X fg\, d\mu, \text{ for } f \in L^p(\mu), \quad 1/p + 1/q = 1.$$

By Hölder's inequality $\|L_g\| \leq \|g\|_q$ and hence L_g is a continuous linear functional on $L^p(\mu)$. Thus, if $g \in L^q$, $L_g \in (L^p)^*$. We write this as

$$L^q \subset (L^p)^*.$$

The other inclusion is given by

Theorem 2.17 (F. Riesz Representation Theorem)**.** *Let $(X, \mathcal{F}, \mu)$ be a σ–finite measure space. Let $1 \leq p < \infty$ and q be its conjugate exponent. Then for every continuous linear functional on $L^p(\mu)$ there is a unique $g \in L^q(\mu)$ such that*

$$L(f) = \int_X fg\, d\mu.$$

Furthermore, $\|L\| = \|g\|_q$.

Remark 2.5. If $1 < p < \infty$, we do not need σ–finiteness. This will be done later. If $p = 1$ and μ is not σ–finite, then the result is not true, (see [To2] p. 227). Also, if $p = \infty$, the result is not true. To produce such an example one needs the Hahn–Banach Theorem.

The theorem will follow easily from the Radon–Nikodym theorem once we prove a few lemmas. The first lemma is proved in introductory measure theory or functional analysis courses; see [**?**, p. xx].

Lemma 2.18. *Let $f \in L^p(\mu)$, $1 \leq p < \infty$. Given $\varepsilon > 0$ there is a simple function s such that*

$$\|f - s\|_p < \varepsilon.$$

Lemma 2.19. *Suppose that $(X, \mathcal{F}, \mu)$ is a finite measure space. Suppose that $g \in L^1(\mu)$, that $1 \leq p < \infty$, and that*

$$\left| \int_X gs\, d\mu \right| \leq M\|s\|_p,$$

for all simple functions s. Then $g \in L^q(\mu)$ and $\|g\|_q \leq M$.

Proof. Assume $1 < p < \infty$. Let s_n be a sequence of nonnegative simple functions converging pointwise up to $|g|^q$. Set $\widetilde{s}_n = s_n^{1/p}\operatorname{sign}(g)$. Then $\widetilde{s}_n$ is also simple and

$$\|\widetilde{s}_n\|_p = \left(\int_X |s_n|\, d\mu \right)^{1/p} < \infty.$$

Also, $s_n = |s_n|^{1/p+1/q} \leq |s_n^{1/p}|\, |g| = \widetilde{s}_n g$ and so

$$\int_X s_n\, d\mu \leq \int_X g\widetilde{s}_n\, d\mu \leq M \left(\int_X (\widetilde{s}_n)^p\, d\mu \right)^{1/p} = M \left(\int_X s_n\, d\mu \right)^{1/p}.$$

Thus

$$\left(\int_X s_n\,d\mu\right)^{1/q} \le M.$$

Let $n \uparrow \infty$. It follows from the Monotone Convergence Theorem that

$$\int_X |g|^q\,d\mu \le M^q.$$

This proves the lemma for the case $1 < p < \infty$ which together with the exercise 2.3.4 completes the proof. $\square$

Proof of Riesz Representation Theorem. Assume $\mu(X) < \infty$. For any $E \in \mathcal{F}$, define

$$\nu(E) = L(1_E).$$

We claim that ν is a signed measure and that $\nu \ll \mu$. Clearly, $\nu(\emptyset) = 0$. Also,

$$|\nu(E)| \le \|L\|\big(\mu(E)\big)^{1/p}$$

for any set E and since μ is finite, so is ν. it also follows that if $\mu(E) = 0$, then $\nu(E) \equiv 0$ and so $\nu \ll \mu$. Let $\{E_n\}$ be disjoint and set $E = \bigcup_{n=1}^\infty E_n$. Since

$$E \setminus \bigcup_{n=1}^m E_n = \bigcup_{n=m+1}^\infty E_n,$$

we have

$$\begin{aligned}\Big|\nu(E) - \sum_{n=1}^m \nu(E_n)\Big| &= \Big|\nu\big(E \setminus \bigcup_{n=1}^m E_n\big)\Big| = \Big|L\big(\chi_{\bigcup_{n=m+1}^\infty E_n}\big)\Big| \\ &\le \ \|L\|\Big[\mu\big(\bigcup_{n=m+1}^\infty E_n\big)\Big]^{1/p} = \|L\|\big(\sum_{n=m+1}^\infty \mu(E_n)\big)^{1/p},\end{aligned}$$

which converges to 0 as $n \to \infty$, since μ is finite. Thus, $\nu(E) = \sum_{n=1}^\infty \nu(E_n)$ proving that ν is a signed measure. By the Radon–Nikodym Theorem, there is $g \in L^1(\mu)$ such that

$$\nu(E) = L(\chi_E) = \int_E g\,d\mu$$

for all $E \in \mathcal{F}$. By the linearity of L,

$$L(s) = \int_X sg\,d\mu$$

for all simple functions s. Since

$$|L(s)| \leq \|L\| \, \|s\|_p,$$

Lemma 2.19 implies that

$$g \in L^q(\mu) \text{ and } \|g\|_q \leq \|L\|.$$

Define another linear functional Λ_g in $L^p(\mu)$ by

$$\Lambda_g(f) = \int_X fg \, d\mu.$$

Then Λ_g is continuous and $L(s) = \Lambda_g(s)$ for all simple functions s. By the density of the simple functions in L^p and Lemma 2.18,

$$L = \Lambda_g.$$

Since also,

$$\|\Lambda_g\| \leq \|g\|_q$$

we have

$$\|L\| = \|g\|_q.$$

This gives the existence of the function g with the desired properties when the measure of the space is finite. Now suppose $\mu(X)$ is σ–finite. We decompose

$$X = \bigcup_{n=1}^{\infty} X_n, \quad X_n \cap X_m = \emptyset, \quad n \neq m$$

and $0 < \mu(X_n) < \infty$. Let $h \colon X \to (0, \infty)$ be defined by

$$h(x) = \frac{1}{n^2 \mu(X_n)}, \text{ if } x \in X_n.$$

Then

$$\int_X h(x) \, d\mu = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

Consider the finite measure $\widetilde{\mu}$ defined by

$$\widetilde{\mu}(E) = \int_E h \, d\mu.$$

Then if $f \in L^p(\widetilde{\mu})$, $h^{1/p} f \in L^p(\mu)$ and

$$\int_X h f^p \, d\mu = \int_X f^p \, d\widetilde{\mu}.$$

If we define the linear functional Λ_h on $L^p(\widetilde{\mu})$ by

$$\Lambda_h(F) = L(h^{1/p} F),$$

we have

$$\begin{aligned} \|\Lambda_h\| &= \sup_{F \in L^p(\widetilde{\mu})} \frac{|\Lambda_h(F)|}{\|F\|_{L^p(\widetilde{\mu})}} \\ &= \sup_{F \in L^p(\widetilde{\mu})} \frac{|L(h^{1/p}F)|}{\|h^{1/p}F\|_{L^p(\mu)}} \\ &= \sup_{\{F:\, Fh^{1/p} \in L^p(\mu)\}} \frac{|L(h^{1/p}F)|}{\|h^{1/p}F\|_{L^p(\mu)}} \\ &= \sup_{f \in L^p(\mu)} \frac{|L(f)|}{\|f\|_p} = \|L\|. \end{aligned}$$

By what we have already proved, there is a $G \in L^q(\widetilde{\mu})$ such that

$$\Lambda_h(F) = \int_X FG \, d\widetilde{\mu}$$

for all $F \in L^p(\widetilde{\mu})$. Setting $g = h^{1/q} G$, (if $p = 1$, take $g = G$) we see that $g \in L^q(\mu)$ and that

$$\left(\int_X |g|^q \, d\mu \right)^{1/q} = \left(\int_X |G|^q \, d\widetilde{\mu} \right)^{1/q} = \|\Lambda_h\| = \|L\|.$$

Also, if $f \in L^p(\mu)$, then $h^{-1/p} f \in L^p(\widetilde{\mu})$ and

$$\begin{aligned} L(f) &= L(h^{1/p} h^{-1/p} f) \\ &= \Lambda_h(h^{-1/p} f) = \int_X h^{-1/p} f G \, d\widetilde{\mu} \\ &= \int_X f h^{1-1/p} G \, d\mu = \int_X f g \, d\mu. \end{aligned}$$

The uniqueness of the function follows trivially from the fact that if g_1 and g_2 both satisfy the conclusion then the function $g_1 - g_2$ gives the zero functional in L^p and $\|g_1 - g_2\|_q = 0$. This implies that $g_1 = g_2$, as functions in L^q, and completes the proof of the Theorem. $\square$

Exercise 2.3.4.
Prove Lemma 2.19 for $p = 1$.

Exercise 2.3.5.
Let X be the two point set $\{0, 1\}$, $\mathcal{F} = \{\{0\}, \{1\}, \emptyset, X\}$. Define $\mu\{0\} = 1$, $\mu\{1\} = \mu\{X\} = \infty$, $\mu\{\emptyset\} = 0$. Is it true that $L^\infty(\mu)$ is the dual of $L^1(\mu)$?

Exercise 2.3.6.
Let $1 < p < \infty$. Prove that $L^p(\mu)$ is the dual of $L^q(\mu)$, $1/p + 1/q = 1$, even if μ is not σ–finite.

Exercise 2.3.7.
Let $(X, \mathcal{F}, \mu)$ be a measure space with $\mu(X) = 1$. Let $\{f_k\}$ be a sequence of functions in $L^1(\mu)$ and suppose $f_k \to f$ a.e. as $k \to \infty$ and that there is a $p > 1$ and a constant C such that $\int_X |f_k(x)|^p \, d\mu(x) < C$ for all k. Prove that $f_k \to f$ in $L^1(\mu)$.

Chapter 3

Product Measures

Our goal in this chapter is to present the essentials of integration in product space. We begin by defining the product measure. Many of the definitions and properties of product measures are, in some sense, obvious. However, we need to state them properly and prove them carefully so that they may be freely used in subsequent chapters. Further results and applications can be found in [Ru2].

3.1 Preliminaries

Recall that if X and Y are any two sets, their Cartesian product $X \times Y$ is the set of all ordered pairs $\{(x,y)\colon x \in X, y \in Y\}$. If $A \subset X,\ B \subset Y$, the subset $A \times B$ of $X \times Y$ is called a *rectangle*. Suppose that $(X, \mathcal{A})$ and that $(Y, \mathcal{B})$ are measurable spaces. A *measurable rectangle* is a set of the form $A \times B$ with $A \in \mathcal{A}$, $B \in \mathcal{B}$. A set of the form

$$Q = R_1 \cup \cdots \cup R_n,$$

where the R_i are disjoint measurable rectangles, is called an *elementary* set. We will denote the collection of elementary sets by $\mathcal{E}$.

We shall denote by $\mathcal{A} \times \mathcal{B}$ the σ–algebra generated by the measurable rectangles which is the same as the σ–algebra generated by the elementary sets.

Theorem 3.1. *Let $E \subset X \times Y$ and define the projections*

$$E_x = \{y \in Y\colon (x,y) \in E\}, \text{ and } E^y = \{x \in X\colon (x,y) \in E\}.$$

If $E \in \mathcal{A} \times \mathcal{B}$, then $E_x \in \mathcal{B}$ and $E^y \in \mathcal{A}$ for all $x \in X$ and $y \in Y$.

Proof. We shall only prove that if $E \in \mathcal{A} \times \mathcal{B}$ then $E_x \in \mathcal{B}$, the case of E^y being practically the same. For this, let Ω be the collection of all sets $E \in \mathcal{A} \times \mathcal{B}$ for which $E_x \in \mathcal{B}$ for every $x \in X$. We show Ω is a σ–algebra containing all measurable rectangles. To see this, note that if E is a measurable rectangle, say

$$E = A \times B,$$

then

$$E_x = \begin{cases} B & \text{if } x \in A \\ \emptyset & \text{if } x \notin A \end{cases}$$

Thus, $E \in \Omega$. The collection Ω also has the following properties:

(i) $X \times Y \in \Omega$.

(ii) If $E \in \Omega$ then $E^c \in \Omega$.
This follows from the facts that $(E^c)_x = (E_x)^c$ and that $\mathcal{B}$ is a σ–algebra.

(iii) If $E_i \in \Omega$ then $E = \bigcup_{i=1}^{\infty} E_i \in \Omega$.

For (iii), observe that $E_x = \bigcup_{i=1}^{\infty}(E_i)_x$ where $(E_i)_x \in \mathcal{B}$. Once again, the fact that $\mathcal{B}$ is a σ–algebra shows that $E \in \Omega$. (i)–(iii) show that Ω is a σ–algebra and the theorem follows. $\square$

We next show that the projections of measurable functions are measurable. Let $f \colon X \times Y \to \mathbb{R}$. For a fixed $x \in X$, define $f_x \colon Y \to \mathbb{R}$ by $f_x(y) = f(x, y)$ with a similar definition for f^y.

In situations where we have several σ–algebras it will be important to clearly distinguish between measurability relative to each one of these σ–algebras. We shall use the notation $f \in \sigma(\mathcal{F})$ to mean that the function f is measurable relative to the σ–algebra $\mathcal{F}$.

Theorem 3.2. *Suppose $f \in \sigma(\mathcal{A} \times \mathcal{B})$. Then*

(i) For each $x \in X$, $f_x \in \sigma(\mathcal{B})$

(ii) For each $y \in Y$, $f^y \in \sigma(\mathcal{A})$

Proof. Fix $x \in X$ and let V be an open set in $\mathbb{R}$. We need to show that $f_x^{-1}(V) \in \mathcal{B}$. Put

$$Q = f^{-1}(V) = \{(\widetilde{x}, y) \colon f(\widetilde{x}, y) \in V\}.$$

Since $f \in \sigma(\mathcal{A} \times \mathcal{B})$, $Q \in \mathcal{A} \times \mathcal{B}$. However,

$$Q_x = f_x^{-1}(V) = \{y\colon f_x(y) \in V\},$$

and it follows by Theorem 3.1 that $Q_x \in \mathcal{B}$ and hence $f_x \in \sigma(\mathcal{B})$. The same argument proves (ii). □

Definition 3.3. A *monotone class* $\mathcal{M}$ is a collection of sets which is closed under countable increasing unions and countable decreasing intersections. That is:

(i) If $A_1 \subset A_2 \subset \cdots$ and $A_i \in \mathcal{M}$, then $\cup A_i \in \mathcal{M}$

(ii) If $B_1 \supset B_2 \supset \cdots$ and $B_i \in \mathcal{M}$, then $\cap B_i \in \mathcal{M}$.

Lemma 3.4 (Monotone Class Theorem)**.** *Let $\mathcal{F}_0$ be an algebra of subsets of X and let $\mathcal{M}$ be a monotone class containing $\mathcal{F}_0$. Then $\mathcal{M}$ contains the σ–algebra generated by $\mathcal{F}_0$.*

Proof. Let $\mathcal{M}_0$ be the smallest monotone class containing $\mathcal{F}_0$. That is, $\mathcal{M}_0$ is the intersection of all the monotone classes which contain $\mathcal{F}_0$. It is enough to show that the σ–algebra $\mathcal{F}$ generated by $\mathcal{F}_0$ is contained in $\mathcal{M}_0$. By Exercise 3.1.2, we only need to prove that $\mathcal{M}_0$ is an algebra. First we prove that $\mathcal{M}_0$ is closed under complementation. For this let $\Omega = \{E : E^c \in \mathcal{M}_0\}$. It follows from the fact that $\mathcal{M}_0$ is a monotone class that Ω is also a monotone class and since $\mathcal{F}_0$ is an algebra, if $E \in \mathcal{F}_0$ then $E \in \Omega$. Thus, $\mathcal{M}_0 \subset \Omega$ and so closed under complementation.

Next, let $\Omega_1 = \{E : E \cup F \in \mathcal{M}_0$ for all $F \in \mathcal{F}_0\}$. Again the fact that $\mathcal{M}_0$ is a monotone class implies that Ω_1 is also a monotone class and since clearly $\mathcal{F}_0 \subset \Omega_1$, we have $\mathcal{M}_0 \subset \Omega_1$. Define $\Omega_2 = \{F : F \cup E \in \mathcal{M}_0$ for all $E \in \mathcal{M}_0\}$. Again Ω_2 is a monotone class. Let $F \in \mathcal{F}_0$. Since $\mathcal{M}_0 \subset \Omega_1$, if $E \in \mathcal{M}_0$, then $E \cup F \in \mathcal{M}_0$. Thus $\mathcal{F}_0 \subset \Omega_2$ and hence $\mathcal{M}_0 \subset \Omega_2$. Thus, if $E, F \in \mathcal{M}_0$ then $E \cup F \in \mathcal{M}_0$. This shows that $\mathcal{M}_0$ is an algebra and completes the proof. □

Exercise 3.1.1.
Prove that the elementary sets form an algebra. That is, $\mathcal{E}$ is closed under complementation and finite unions.

Exercise 3.1.2.
Suppose that $\mathcal{M}$ is a monotone class which is also an algebra, then $\mathcal{M}$ is a σ–algebra.

Exercise 3.1.3.
Let $\mathcal{F}_0$ be an algebra and suppose the two measures μ_1 and μ_2 on $\sigma(\mathcal{F}_0)$ agree on $\mathcal{F}_0$. Prove that they agree on the σ–algebra $\mathcal{F}$ generated by $\mathcal{F}_0$.

3.2 Fubini's Theorem

We begin this section with a lemma that will allow us to define the product of two measures.

Lemma 3.5. *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be two σ–finite measure spaces. Suppose that $Q \in \mathcal{A} \times \mathcal{B}$. If we set*

$$\varphi(x) = \nu(Q_x) \text{ and } \psi(y) = \mu(Q^y),$$

then

$$\varphi \in \sigma(\mathcal{A}) \text{ and } \psi \in \sigma(\mathcal{B})$$

and

$$\int_X \varphi(x)\, d\mu(x) = \int_Y \psi(y)\, d\nu(y). \tag{3.1}$$

Remark 3.1. With the above notation we can also write

$$\nu(Q_x) = \int_Y \chi_Q(x,y)\, d\nu(y) \tag{3.2}$$

and

$$\mu(Q^y) = \int_X \chi_Q(x,y)\, d\mu(x). \tag{3.3}$$

Thus (3.1) is equivalent to

$$\int_X \left(\int_Y \chi_Q(x,y)\, d\nu(y) \right) d\mu(x) = \int_Y \left(\int_X \chi_Q(x,y)\, d\mu(x) \right) d\nu(y).$$

Remark 3.2. Lemma 3.5 allows us to define a new measure $\mu \times \nu$ on $\mathcal{A} \times \mathcal{B}$ by

$$(\mu \times \nu)(Q) = \int_X \nu(Q_x)\, d\mu(x) = \int_Y \mu(Q^y)\, d\nu(y). \tag{3.4}$$

In particular, the $\mu \times \nu$ measure of a measurable rectangle $A \times B$ is $\mu(A)\nu(B)$.

To see that $\mu \times \nu$ is indeed a measure let $\{Q_j\}$ be a disjoint sequence of sets in $\mathcal{A} \times \mathcal{B}$. Recalling that $(\cup Q_j)_x = \cup (Q_j)_x$ and using the fact that ν is a measure we have

$$(\mu \times \nu)\Big(\bigcup_{j=1}^{\infty} Q_j\Big) = \int_X \nu\Big(\Big(\bigcup_{j=1}^{\infty} Q_j\Big)_x\Big)\, d\mu(x)$$

$$
\begin{aligned}
&= \int_X \nu\Big(\bigcup_{j=1}^{\infty} (Q_j)_x\Big)\, d\mu(x) = \int_X \sum_{j=1}^{\infty} \nu\big((Q_j)_x\big)\, d\mu(x) \\
&= \sum_{j=1}^{\infty} \int_X \nu\big((Q_j)_x\big)\, d\mu(x) = \sum_{j=1}^{\infty} (\mu \times \nu)(Q_j),
\end{aligned}
$$

where the second to the last equality follows from the Monotone Convergence Theorem.

Proof of Lemma 3.5. We assume $\mu(X) < \infty$ and $\nu(Y) < \infty$. Let $\mathcal{M}$ be the collection of all $Q \in \mathcal{A} \times \mathcal{B}$ for which the conclusion of the Lemma is true. We will prove that $\mathcal{M}$ is a monotone class which contains the elementary sets; $\mathcal{E} \subset \mathcal{M}$. By Exercise 3.1.1 and the Monotone class Theorem, this will show that $\mathcal{M} = \mathcal{A} \times \mathcal{B}$. This will be done in several stages. First we prove that rectangles are in $\mathcal{M}$. That is,

(i) Let $Q = A \times B$, $A \in \mathcal{A}$, $B \in \mathcal{B}$. Then $Q \in \mathcal{M}$. To prove (i) observe that

$$
Q_x = \begin{cases} B & \text{if } x \in A \\ \emptyset & \text{if } x \notin A. \end{cases}
$$

Thus

$$
\varphi(x) = \chi_A(x)\nu(B) = \begin{cases} \nu(B) & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}
$$

and clearly $\varphi \in \sigma(\mathcal{A})$. Similarly,

$$
\psi(y) = \chi_B(y)\mu(A) \in \mathcal{B}.
$$

Integrating we obtain that

$$
\begin{aligned}
\int_X \varphi(x)\, d\mu(x) &= \mu(A)\nu(B) \\
\int_Y \psi(y)\, d\nu(y) &= \mu(A)\nu(B),
\end{aligned}
$$

proving (i).

(ii) Let $Q_1 \subset Q_2 \subset \cdots$, $Q_j \in \mathcal{M}$. Then $Q = \bigcup_{j=1}^{\infty} Q_j \in \mathcal{M}$. To prove this, let

$$
\varphi_n(x) = \nu\big((Q_n)_x\big) = \nu\Big(\big(\textstyle\bigcup_{j=1}^{n} Q_j\big)_x\Big)
$$

and
$$\psi_n(y) = \mu\big((Q_n)^y\big) = \mu\Big(\big(\bigcup_{j=1}^n Q_j\big)^y\Big).$$
Then
$$\varphi_n(x) \uparrow \varphi(x) = \nu(Q_x)$$
and
$$\psi_n(y) \uparrow \psi(y) = \mu(Q^y).$$
Since $\varphi_n \in \sigma(\mathcal{A})$ and $\psi_n \in \sigma(\mathcal{B})$, we have $\varphi \in \sigma(\mathcal{A})$ and $\psi \in \sigma(\mathcal{B})$. Also by assumption
$$\int_X \varphi_n(x)\, d\mu(x) = \int_Y \psi_n(y)\, d\nu(y),$$
for all n. By the monotone convergence theorem,
$$\int_X \varphi(x)\, d\mu(x) = \int_Y \psi(y)\, d\nu(y)$$
and we have proved (ii).

(iii) Let $Q_1 \supset Q_2 \supset \cdots$, $Q_j \in \mathcal{M}$. Then $Q = \bigcap_{j=1}^\infty Q_j \in \mathcal{M}$. The proof of this is the same as (ii) except this time we use the Dominated Convergence Theorem. That is, this time the sequences $\varphi_n(x) = \nu\big((Q_n)_x\big)$, $\psi_n(y) = \mu\big((Q_n)^y\big)$ are both decreasing to $\varphi(x) = \nu(Q_x)$ and $\psi(y) = \mu(Q^y)$, respectively, and since both measures are finite, both sequences of functions are uniformly bounded.

(iv) Let $\{Q_i\} \in \mathcal{M}$ with $Q_i \cap Q_j = \emptyset$, $i \neq j$. Then $\bigcup_{j=1}^\infty Q_i \in \mathcal{M}$. For the proof of this, let $\widetilde{Q}_n = \bigcup_{i=1}^n Q_i$. Then $\widetilde{Q}_n \in \mathcal{M}$, since the sets are disjoint. However, the sets $\widetilde{Q}_n$ are increasing and it follows from (ii) that their union is in $\mathcal{M}$, proving (iv).

It follows from (i)–(iv) that $\mathcal{M}$ is a monotone class containing the elementary sets $\mathcal{E}$. By the Monotone Class Theorem and Exercise 3.1.1, $\mathcal{A} \times \mathcal{B} = \sigma(\mathcal{E}) = \mathcal{M}$. This proves the lemma for finite measures and Exercise 3.2.1 does the rest. □

Theorem 3.6 (Fubini's Theorem)**.** *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ–finite measure spaces. Let $f \in \sigma(\mathcal{A} \times \mathcal{B})$.*

(i) **(Tonelli)** *If f is nonnegative and if we set*

$$\varphi(x) = \int_Y f_x(y)\, d\nu(y), \quad \psi(y) = \int_X f^y(x)\, d\mu(x), \tag{3.5}$$

then

$$\varphi \in \sigma(\mathcal{A}), \quad \psi \in \sigma(\mathcal{B})$$

and

$$\int_X \varphi(x)\, d\mu(x) = \int_{X\times Y} f(x,y)\, d(\mu \times \nu) = \int_Y \psi(y)\, d\nu(y). \tag{3.6}$$

(ii) *If f is complex valued such that*

$$\varphi^*(x) = \int_Y |f_x(y)|\, d\nu(y) = \int_Y |f(x,y)|\, d\nu(y) < \infty,$$

for each x and

$$\int_X \varphi^*(x)\, d\mu(x) < \infty$$

then

$$f \in L^1(\mu \times \nu).$$

A similar statement holds with y in place of x.

(iii) *If $f \in L^1(\mu \times \nu)$, then $f_x \in L^1(\nu)$ for a.e. $x \in X$, $f^y \in L^1(\mu)$ for a.e. $y \in Y$, the functions defined in (3.5) are measurable and (3.6) holds.*

Proof. If $f = \chi_Q$, $Q \in \mathcal{A} \times \mathcal{B}$, the result follows from Lemma 3.5. By linearity we also have the result for simple functions. Let $0 \le s_1 \le \cdots$ be nonnegative simple functions such that $s_n(x,y) \uparrow f(x,y)$ for every $(x,y) \in X \times Y$. Let

$$\varphi_n(x) = \int_Y (s_n)_x(y)\, d\nu(y)$$

and

$$\psi_n(y) = \int_X (s_n)^y(x)\, d\mu(x).$$

Then

$$\int_X \varphi_n(x)\, d\mu(x) = \int_{X\times Y} s_n(x,y) d(\mu \times \nu) = \int_Y \psi_n(y)\, d\nu(y).$$

Since $s_n(x,y) \uparrow f(x,y)$ for every $(x,y) \in X \times Y$, by monotone convergence, $\varphi_n(x) \uparrow \varphi(x)$ and $\psi_n(y) \uparrow \psi(y)$. One more application of the Monotone Convergence Theorem once again implies (i).

Parts (ii) and (iii) follow directly from (i) and we leave these as exercises. $\square$

The assumption of σ–finiteness is needed as the following example shows.

Example 3.1. $X = Y = [0,1]$ with μ being the Lebesgue measure and ν being the counting measure. Let $f(x,y) = 1$ if $x = y$, $f(x,y) = 0$ if $x \neq y$. That is, the function f is the characteristic function of the diagonal of the square. Then $\int_X f(x,y)\,d\mu(x) = 0$, for every y, and $\int_Y f(x,y)\,d\nu(y) = 1$, for every x.

Remark 3.3. Before we can integrate the function f in this example, however, we need to verify that it (and hence its projections) is (are) measurable. This can be seen as follows: Set

$$I_j = \left[\tfrac{j-1}{n}, \tfrac{j}{n}\right]$$

and

$$Q_n = (I_1 \times I_1) \cup (I_2 \times I_2) \cup \cdots \cup (I_n \times I_n).$$

Then each Q_n is measurable and so is the diagonal of the square, which is their intersection. Hence also f is a measurable function.

Example 3.2. Consider the function

$$f(x,y) = \frac{x^2 - y^2}{(x^2+y^2)^2} \text{ on } (0,1) \times (0,1).$$

with the μ and ν the Lebesgue measure. Then

$$\int_0^1 \int_0^1 f(x,y)\,dy\,dx = \frac{\pi}{2}$$

but

$$\int_0^1 \int_0^1 f(x,y)\,dx\,dy = -\frac{\pi}{2}.$$

The problem here is that $f \notin L^1\{(0,1) \times (0,1)\}$ since

$$\int_0^1 |f(x,y)|\,dy \geq \frac{1}{2x}.$$

Let m_k be Lebesgue measure in $\mathbb{R}^k$ and recall that m_k is complete. That is, if $m_k(E) = 0$ then any subset of E is also Lebesgue measurable. However, the product measure $m_1 \times m_1$ as not constructed above is complete since $\{x\} \times B$, for any set $B \subset \mathbb{R}$ (including non-measurable sets) has $m_1 \times m_1$– measure zero. If this set were measurable in the product then its projection B would also be measurable. Thus $m_2 \neq m_1 \times m_1$. What is needed here is the notion of the completion of a measure. We leave the proof of the first two theorems as exercises.

Theorem 3.7. *If $(X, \mathcal{F}, \mu)$ is a measure space we let*

$$\tilde{\mathcal{F}} = \{E \subset X \colon \exists A \textit{ and } B \in \mathcal{F}, A \subset E \subset B \textit{ and } \mu(B \backslash A) = 0\}.$$

Then $\tilde{\mathcal{F}}$ is a σ–algebra and the function $\tilde{\mu}$ defined on $\tilde{\mathcal{F}}$ by

$$\tilde{\mu}(E) = \mu(A)$$

is a measure. The measure space $(X, \tilde{\mathcal{F}}, \tilde{\mu})$ is complete. This new space is called the completion of $(X, \mathcal{F}, \mu)$.

Theorem 3.8. *Let m_n be the Lebesgue measure on $\mathbb{R}^n$, $n = r + s$. Then $m_n = \widetilde{(m_r \times m_s)}$, the completion of the product of the Lebesgue measures.*

The next Theorem says that as far as Fubini's theorem is concerned, we need not worry about incomplete measure spaces.

Theorem 3.9. *Let $(X, \mathcal{A}, \mu)$, $(Y, \mathcal{B}, \nu)$ be two complete σ–finite measure spaces. Theorem 3.6 remains valid if $\mu \times \nu$ is replaced by $\widetilde{(\mu \times \nu)}$ except that the functions φ and ψ are defined only almost everywhere relative to the measures μ and ν, respectively.*

Proof. The proof of this theorem follows from the following two facts.

(i) Let $(X, \mathcal{F}, \mu)$ be a measure space. Suppose that $f \in \sigma(\tilde{\mathcal{F}})$. Then there is a $g \in \sigma(\mathcal{F})$ such that $f = g$ a.e. with respect to μ.

(ii) Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be two complete and σ–finite measure spaces. Suppose $f \in \sigma(\widetilde{(\mathcal{A} \times \mathcal{B})})$ is such that $f = 0$ almost everywhere with respect to $\mu \times \nu$. Then for almost every $x \in X$ with respect to μ, $f_x = 0$ a.e. with respect to ν. In particular, $f_x \in \sigma(\mathcal{B})$ for almost every $x \in X$. A similar statement holds with y replacing x.

Let us assume (i) and (ii) for the moment. Then if $f \in \sigma(\widetilde{(\mathcal{A} \times \mathcal{B})})$ is nonnegative there is a $g \in \sigma(\mathcal{A} \times \mathcal{B})$ such that $f = g$ a.e. with respect to $\mu \times \nu$. Now, apply Theorem 3.6 to g and the rest follows from (ii).

It remains to prove (i) and (ii). For (i), suppose that $f = \chi_E$ where $E \in \tilde{\mathcal{A}}$. By definition, $A \subset E \subset B$ with $\mu(A \setminus B) = 0$ and A and $B \in \mathcal{A}$. If we set $g = \chi_A$ we have $f = g$ a.e. with respect to μ and we have proved (i) for characteristic function. We now extend this to simple functions and to nonnegative functions in the usual way; the details are left to the reader.

For (ii) let $\Omega = \{(x, y) : f(x, y) \neq 0\}$. Then $\Omega \in (\widetilde{\mathcal{A} \times \mathcal{B}})$ and $(\widetilde{\mu \times \nu})(\Omega) = 0$. By definition there is an $\Omega_0 \in \mathcal{A} \times \mathcal{B}$ such that $\Omega \subset \Omega_0$ and $(\mu \times \nu)(\Omega_0) = 0$. By Theorem 3.6,

$$\int_X \nu((\Omega_0)_x)\, d\mu(x) = 0$$

and so $\nu((\Omega_0)_x) = 0$ for almost every x with respect to μ. Since $\Omega_x \subset (\Omega_0)_x$ and the space $(Y, \mathcal{B}, \nu)$ is complete, we see that $\Omega_x \in \mathcal{B}$ for almost every $x \in X$ with respect to the measure μ. Thus for almost every $x \in X$ the projection function $f_x \in \mathcal{B}$ has $f_x(y) = 0$ almost everywhere with respect to μ. This completes the proof of (ii) and hence the theorem. □

Exercise 3.2.1.
Extend the proof of Lemma 3.5 to the case of σ–finite measures.

Exercise 3.2.2.
Let f be a nonnegative Lebesgue measurable function in $\mathbb{R}^n$. Prove that

$$\int_{\mathbb{R}^n} f(x)\, dx = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, x_2, \ldots, x_n)\, dx_1 \cdots dx_n$$

where we use dx to denote integration with respect to the Lebesgue measure in $\mathbb{R}^n$ and dx_i to denote integration in $\mathbb{R}$.

Exercise 3.2.3.
Let f be a nonnegative measurable function on $(X, \mathcal{F}, \mu)$. Prove that for any $0 < p < \infty$,

$$\int_X f(x)^p\, d\mu(x) = p \int_0^\infty \lambda^{p-1} \mu\{x \in X : f(x) > \lambda\}\, d\lambda.$$

Exercise 3.2.4.
Let $(X, \mathcal{F}, \mu)$ be a measure space. Suppose that f and g are two nonnegative

functions satisfying the following inequality: There exists a constant C such that for all $\varepsilon > 0$ and $\lambda > 0$,

$$\mu\{x \in X\colon f(x) > 2\lambda,\ g(x) \le \varepsilon\lambda\} \le C\varepsilon^2 \mu\{x \in X\colon f(x) > \lambda\}.$$

Prove that

$$\int_X f(x)^p \, d\mu \le C_p \int_X g(x)^p \, d\mu$$

for any $0 < p < \infty$ for which both integrals are finite where C_p is a constant depending on C and p.

Exercise 3.2.5.
For any $\alpha \in \mathbb{R}$ define

$$\operatorname{sign}(\alpha) = \begin{cases} 1 & \text{if } \alpha > 0 \\ 0 & \text{if } \alpha = 0 \\ -1 & \text{if } \alpha < 0 \end{cases}$$

Prove that

$$0 \le \operatorname{sign}(\alpha) \int_0^y \frac{\sin(\alpha x)}{x} \, dx \le \int_0^\pi \frac{\sin(x)}{x} \, dx \tag{3.7}$$

for all $y > 0$, that

$$\int_0^\infty \frac{\sin(\alpha x)}{x} \, dx = \frac{\pi}{2} \operatorname{sign}(\alpha) \tag{3.8}$$

and that

$$\int_0^\infty \frac{1 - \cos(\alpha x)}{x^2} \, dx = \frac{\pi}{2} |\alpha|. \tag{3.9}$$

Exercise 3.2.6.
Prove that

$$e^{-\alpha} = \frac{2}{\pi} \int_0^\infty \frac{\cos(\alpha s)}{1 + s^2} \, ds \tag{3.10}$$

for all $\alpha > 0$. Use (3.10), the fact that

$$\frac{1}{1+s^2} = \int_0^\infty e^{-(1+s^2)t} \, dt,$$

and Fubini's theorem, to prove that

$$e^{-\alpha} = \frac{1}{\sqrt{\pi}} \int_0^\infty \frac{e^{-t}}{\sqrt{t}} e^{-\alpha^2/(4t)} \, dt. \tag{3.11}$$

Exercise 3.2.7.
Let $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$ and for any Borel set $E \in S^{n-1}$ set $\widetilde{E} = \{r\theta : 0 < r < 1, \theta \in E\}$. Define the measure σ on S^{n-1} by $\sigma(E) = n|\widetilde{E}|$. Notice that with this definition the surface area ω_{n-1} of the sphere in $\mathbb{R}^n$ satisfies $\omega_{n-1} = n\gamma_n = 2\pi^{n/2}/\Gamma(n/2)$ where γ_n is the volume of the unit ball in $\mathbb{R}^n$. Prove (integration in polar coordinates) that for all nonnegative Borel functions f on $\mathbb{R}^n$,

$$\int_{\mathbb{R}^n} f(x)\,dx = \int_0^\infty r^{n-1}\left(\int_{S^{n-1}} f(r\theta)\,d\sigma(\theta)\right)\,dr.$$

In particular, if f is a radial function, that is, $f(x) = f(|x|)$, then

$$\int_{\mathbb{R}^n} f(x)\,dx = \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}\int_0^\infty r^{n-1}f(r)\,dr = n\gamma_n\int_0^\infty r^{n-1}f(r)dr.$$

Exercise 3.2.8.
Prove that for any $x \in \mathbb{R}^n$ and any $0 < p < \infty$

$$\int_{S^{n-1}} |\xi \cdot x|^p\,d\sigma(\xi) = |x|^p \int_{S^{n-1}} |\xi_1|^p\,d\sigma(\xi)$$

where $\xi \cdot x = \xi_1 x_1 + \cdots + \xi_n x_n$ is the inner product in $\mathbb{R}^n$.

Exercise 3.2.9.
Let $e_1 = (1, 0, \ldots, 0)$ and for any $\xi \in S^{n-1}$ define θ, $0 \leq \theta \leq \pi$ such that $e_1 \cdot \xi = \cos\theta$. Prove, by first integrating over $L_\theta = \{\xi \in S^{n-1} : e_1 \cdot \xi = \cos\theta\}$, that for any $1 \leq p < \infty$,

$$\int_{S^{n-1}} |\xi_1|^p\,d\sigma(\xi) = \omega_{n-2}\int_0^\pi |\cos\theta|^p(\sin\theta)^{n-2}\,d\theta. \tag{3.12}$$

Use (3.12) and the fact that for any $r > 0$ and $s > 0$,

$$2\int_0^{\pi/2}(\cos\theta)^{2r-1}(\sin\theta)^{2s-1}\,d\theta = \frac{\Gamma(s)\Gamma(r)}{\Gamma(r+s)}$$

([Ru1], p. 194) to prove that for any $1 \leq p < \infty$

$$\int_{S^{n-1}} |\xi_1|^p\,d\sigma(\xi) = \frac{2\pi^{(n-1)/2}\Gamma(\frac{p+1}{2})}{\Gamma(\frac{n+p}{2})} \tag{3.13}$$

Chapter 4

Convolutions and Approximations to the Identity

In this Chapter we introduce an important operation on functions called convolution. This operation and its properties play an important role in proving, among other things, that smooth functions of compact support are dense in $L^p(\mathbb{R}^n)$ for $1 \le p < \infty$. This is one of the main results of this chapter. Here we shall also study approximations to the identity. Of particular interest is the case of the Hardy-Littlewood maximal function and its applications, especially the Calderón–Zygmund decomposition theorem. This will be discussed in Chapter 5.

4.1 Minkowski's Integral Inequality

We begin with a theorem the usefulness of which will become clear shortly.

Theorem 4.1 (Minkowski's Integral Inequality)**.** *Suppose that $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are two σ–finite measure spaces. Let $1 \le p < \infty$ and suppose F is measurable with respect to $\mathcal{A} \times \mathcal{B}$. Then*

$$\left\| \int_X F(x, \cdot)\, d\mu(x) \right\|_{L^p(d\nu)} \le \int_X \|F(x, \cdot)\|_{L^p(d\nu)}\, d\mu(x). \tag{4.1}$$

Remark 4.1. Recall that if f_1, f_2 are two functions then the classical Minkowski inequality gives

$$\|f_1 + f_2\|_p \le \|f_1\|_p + \|f_2\|_p \tag{4.2}$$

and in general

$$\left\|\sum_{j=1}^{\infty} f_j\right\|_p \leq \sum_{j=1}^{\infty} \|f_j\|_p. \tag{4.3}$$

If we let μ be the counting measure on the positive integers $\mathbb{N}$ we can write (4.3) as

$$\sum_{j=1}^{\infty} f_j = \int_{\mathbb{N}} f_j d\mu(j)$$

Thus (4.3) is a special case of (4.1). However, as we will see in a moment, the proof of (4.1) is exactly the same as the proof of (4.2). An alternative proof of this, which we leave to the reader as an exercise, is also possible based on Theorem 2.17 of Chapter 2.

Proof. Assume, as we may, that $F(x, y) \geq 0$ and, by approximating with simple functions if necessary, that both sides of the inequality are finite. The case $p = 1$ follows directly from Fubini's theorem and so we take $1 < p < \infty$ and let $q = p/(p-1)$ be its conjugate exponent. Set

$$G(y) = \left(\int_X F(x, y)\, d\mu(x)\right)^{p-1}.$$

Then

$$\begin{aligned}
\|G\|_{L^q(d\nu)} &= \left\| \left(\int_X F(x, \cdot)\, d\mu(x)\right)^{p-1} \right\|_{L^q(d\nu)} \\
&= \left(\int_Y \left(\int_X F(x, \cdot)\, d\mu(x)\right)^p d\nu(y)\right)^{1/q} \\
&= \left\|\int_X F(x, \cdot)\, d\mu(x)\right\|_{L^p(d\nu)}^{p-1}.
\end{aligned}$$

By Fubini's theorem and Hölder's inequality,

$$\begin{aligned}
&\left\|\int_X F(x, \cdot)\, d\mu(x)\right\|_{L^p(d\nu)}^{p} \\
&= \int_Y G(y) \int_X F(x, y)\, d\mu(x)\, d\nu(y) \\
&= \int_X \int_Y G(y) F(x, y)\, d\nu(y)\, d\mu(x)
\end{aligned}$$

$$\begin{aligned} &\leq \int_X \|G\|_{L^q(d\nu)}\|F(x,\cdot)\|_{L^p(d\nu)}\,d\mu(x) \\ &= \|G\|_{L^q(d\nu)}\int_X \|F(x,\cdot)\|_{L^p(d\nu)}\,d\mu(x). \end{aligned}$$

Dividing by $\|G\|_{L^q(d\nu)}$ proves (4.1). □

Exercise 4.1.1.
Let $1 < p < \infty$ and $q = p/(p-1)$ be its conjugate exponent. Let $K(x,y)$, $x, y \in (0,\infty)$ be nonnegative and homogeneous of degree -1. That is, K has the property that for all $\lambda > 0$, $K(\lambda x, \lambda y) = \lambda^{-1}K(x,y)$. Suppose that

$$\int_0^\infty K(x,1)x^{-1/p}\,dx = \int_0^\infty K(1,y)y^{-1/q}\,dy = a$$

and denote the $L^p(0,\infty)$ norm of a function f by $\|f\|_p$. Prove that

$$\left\|\int_0^\infty K(x,y)f(x)\,dx\right\|_{L^p(dy)} \leq a\|f\|_p$$

and that

$$\left|\int_0^\infty\int_0^\infty K(x,y)f(x)g(y)\,dx\,dy\right| \leq a\|f\|_p\|g\|_q.$$

Exercise 4.1.2.
Let a_j and b_k be two sequences of nonnegative numbers. Prove that

$$\sum_{j=1}^\infty\sum_{k=1}^\infty \frac{a_j b_k}{j+k} \leq \frac{\pi}{\sin(\pi/p)}\Big(\sum_{j=1}^\infty a_j^p\Big)^{1/p}\Big(\sum_{k=1}^\infty b_k^q\Big)^{1/q}.$$

Exercise 4.1.3.
Hardy's Inequality. Let f be a nonnegative function on $[0,\infty)$, $1 < p < \infty$ and $0 < r < \infty$. Prove that

$$\left(\int_0^\infty\left(\int_0^x f(y)\,dy\right)^p\frac{dx}{x^{r+1}}\right)^{1/p} \leq \frac{p}{r}\left(\int_0^\infty (yf(y))^p\frac{dy}{y^{r+1}}\right)^{1/p}. \tag{4.4}$$

Exercise 4.1.4.
Suppose $r = p-1$ in Exercise 4.1.3. In this case we can write the inequality (4.4) as

$$\left(\int_0^\infty\Big(\frac{1}{x}\int_0^x f(y)\,dy\Big)^p\,dx\right)^{1/p} \leq \frac{p}{p-1}\left(\int_0^\infty\big(f(y)\big)^p\,dy\right)^{1/p}.$$

Prove that the constant $p/(p-1)$ cannot be improved.

Exercise 4.1.5.
Let $\{a_k\}$ be a sequence of nonnegative numbers. Prove that

$$\left(\sum_{N=1}^{\infty}\Big(\frac{1}{N}\sum_{j=1}^{N}a_j\Big)^p\right)^{1/p} \le \frac{p}{p-1}\Big(\sum_{j=1}^{\infty}a_j^p\Big)^{1/p}.$$

Exercise 4.1.6.
Let f be a differentiable function of compact support in $(0,\infty]$. Prove that

$$\int_0^\infty \frac{|f(y)|^2}{y^2}\,dy \le 4\int_0^\infty |f'(y)|^2\,dy.$$

Exercise 4.1.7.
Let $u(x,y)$ be a differentiable function in the upper half space $\mathbb{R}^2_+ = \{(x,y) : x \in \mathbb{R}, y > 0\}$ of compact support. Prove that

$$\int_{\mathbb{R}^2_+} \frac{|u(x,y)|^2}{y^2}\,dx\,dy \le 4\int_{\mathbb{R}^2_+} |\nabla u(x,y)|^2\,dx\,dy,$$

where $\nabla u(x,y)$ denotes the gradient of the function u at the point (x,y).

Exercise 4.1.8. (i) Let D be a simply connected domain in the plane. Use Exercise 4.1.7 and the Koebe 1/4-theorem applied to the conformal mapping $\varphi : \mathbb{R}^2_+ \to D$ (see Duran [**?**]) to conclude that for any smooth function $u(x,y)$ with compact support in D,

$$\int_D \frac{|u(x,y)|^2}{d_D^2(x,y)}\,dx\,dy \le 16\int_D |\nabla u(x,y)|^2\,dx\,dy,$$

where $d_D(x,y)$ represents the distance from the point $(x,y) \in D$ to D^c. That is, the distance from the point (x,y) to the boundary of D.

(ii) Use this inequality to conclude that if we define the inner radius of the domain D to be $R_D = \sup\{d_D(x,y) : (x,y) \in D\}$ and the quantity

$$\lambda_D = \inf\left(\frac{\int_D |\nabla u(x,y)|^2\,dx\,dy}{\int_D |u(x,y)|^2\,dx\,dy}\right),$$

where the infimum is taken over all smooth functions u of compact support in D, then

$$\lambda_D \ge \frac{1}{16R_D^2}.$$

Remark 4.2. The quantity λ_D defined above is the lowest eigenvalue of the Laplacian with Dirichlet boundary conditions in D. The inequality proves that if the "drum" D produces a low tone then it must contain a large disk. The converse is also true and very easy to prove. That is, if it contains a large disk then it produces a low tone. The reader interested in this connection can see [BC].

4.2 The Convolution Operator

Theorem 4.2. *Let* $g \in L^1(\mathbb{R}^n)$*,* $f \in L^p(\mathbb{R}^n)$*,* $1 \leq p \leq \infty$*. Define*

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

Then

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

Remark 4.3. $f * g$ is called the convolution of f and g. Notice that by a simple change of variables $f * g = g * f$. In addition, the convolution operation satisfies the associative law, as the reader can easily verify. That is, $(f*g)*h = f*(g*h)$.

Proof. The case $p = \infty$ is trivial and the case $p = 1$ follows directly from Fubini's theorem. However, before we apply Fubini's theorem we must verify that the function $F(x,y) = f(x-y)g(y)$ is measurable in $\mathbb{R}^n \times \mathbb{R}^n$. We leave this as an exercise to the reader. Assume now that $1 < p < \infty$. Since $g \in L^1(\mathbb{R}^n)$, the measure $d\mu(y) = |g(y)|\,dy$ is finite and applying the Minkowski integral inequality we obtain

$$\begin{aligned} \|f * g\|_p &\leq \left\| \int_{\mathbb{R}^n} |f(x-y)|\,|g(y)|\,dy \right\|_{L^p(dx)} \\ &\leq \int_{\mathbb{R}^n} \|f(x-y)\|_{L^p(dx)} |g(y)|\,dy \\ &= \|f\|_{L^p} \|g\|_{L^1}, \end{aligned}$$

which proves the Theorem for all p. □

The above theorem shows that the convolution of two functions inherits the integrability properties of the better function. The next main result shows that this is also the case as far as smoothness is concerned. This will permit us to show that the space of infinitely differentiable functions with compact support is dense in $L^p(\mathbb{R}^n)$, $1 \leq p < \infty$. We shall use the following notation for the rest of these notes. We denote by $C_0(\mathbb{R}^n)$ the space of continuous functions of compact support in $\mathbb{R}^n$ and by $C_0^m(\mathbb{R}^n)$, $m = 1, 2 \ldots$, the space of functions of compact support which also have m continuous derivatives each with compact support. We will denote functions which are m times differentiable (not necessarily of compact support), by $C^m(\mathbb{R}^n)$. Any polynomial is clearly in $C^\infty(\mathbb{R}^n)$ but not in any $L^p(\mathbb{R}^n)$. Examples of functions in $C^\infty(\mathbb{R}^n)$ which also have good integrability properties are: $e^{-|x|}$, $e^{-|x|^2}$ and

$$\frac{1}{(1 + |x|^2)^{(n+1)/2}} \tag{4.5}$$

These functions will play an important role later in these notes. To construct C^∞–functions of compact support, let $\psi(t) = e^{1/t}$ for $t < 0$ and 0 for $t \geq 0$. It follows from L'Hospital's rule that the function $\varphi(x) = \psi(|x|^2 - 1)$ is in $C_0^\infty(\mathbb{R}^n)$.

Proposition 4.3. *The space $C_0(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$, $1 \leq p < \infty$. That is, given $f \in L^p(\mathbb{R}^n)$ and $\varepsilon > 0$ there is a $g \in C_0(\mathbb{R}^n)$ such that*

$$\|f - g\|_p < \varepsilon$$

Remark 4.4. This proposition is proved in an elementary measure theory course; see for example [To2, p. 133]. We observe that because of the density of the simple functions in $L^p(\mathbb{R}^n)$, it reduces to proving the result when f is a simple function and this can be further reduced to the case when f is the characteristic functions of sets of finite measure. By the regularity of the Lebesgue measure this in turn reduces to the case of characteristic functions of open sets.

Lemma 4.4. *The translation operator $y \to f(x + y)$ is continuous from $\mathbb{R}^n$ to $L^p(\mathbb{R}^n)$ for $1 \leq p < \infty$. That is, if $f \in L^p(\mathbb{R}^n)$, $1 \leq p < \infty$, then*

$$\lim_{|h| \to 0} \|f(x + h) - f(x)\|_p = 0.$$

Proof. Let $\varepsilon > 0$ be given. By Proposition 4.3 there is a function $g \in C_0(\mathbb{R}^n)$ such that

$$\|g(x) - f(x)\|_p < \varepsilon/3. \tag{4.6}$$

Thus

$$\|f(x + h) - f(x)\|_p$$

$$
\begin{aligned}
&\leq \|f(x+h)-g(x+h)\|_p \\
&\quad +\|g(x+h)-g(x)\|_p \\
&\quad +\|g(x)-f(x)\|_p \\
&< \tfrac{2}{3}\varepsilon+\|g(x+h)-g(x)\|_p. \qquad (4.7)
\end{aligned}
$$

Let $B=B(0,r)$ be the ball centered at the origin and of radius r with r such that the support of the function g is contained in B. Since g has compact support, g is in fact uniformly continuous. Thus given $\eta>0$ there is a $\delta>0$ such that

$$|g(x+h)-g(x)|\leq \eta, \qquad (4.8)$$

for all $|h|<\delta$. We may suppose that δ is small enough so that for $|h|<\delta$, the function $g(x+h)-g(x)$ vanishes off the set $\widetilde{B}=B(0,2r)$. Then

$$
\begin{aligned}
&\int_{\mathbb{R}^n}|g(x+h)-g(x)|^p\,dx \\
&\quad\leq \int_{\widetilde{B}}|g(x+h)-g(x)|^p\,dx \\
&\quad\leq \eta^p m(\widetilde{B}). \qquad (4.9)
\end{aligned}
$$

If we take η such that $\eta^p m(\widetilde{B})<(\varepsilon/3)^p$ and substitute the estimate from (4.9) into (4.7) we obtain the desired result. □

Theorem 4.5. *Let $f\in L^p(\mathbb{R}^n)$, $1\leq p\leq\infty$ and $\varphi\in C_0^m(\mathbb{R}^n)$, $m\geq 1$. Then $f*\varphi\in L^p(\mathbb{R}^n)\cap C^m(\mathbb{R}^n)$.*

Proof. By Theorem 4.1, $f*\varphi\in L^p(\mathbb{R}^n)$ and all we need to do is show that $f*\varphi\in C^m$. We begin by proving that

(i) $f*\varphi$ is continuous. For this, apply Hölder's inequality to obtain,

$$
\begin{aligned}
&|(f*\varphi)(x+h)-(f*\varphi)(x)| \\
&\quad= |(\varphi*f)(x+h)-(\varphi*f)(x)| \\
&\quad\leq \int_{\mathbb{R}^n}|\varphi(x+h-y)f(y)-\varphi(x-y)f(y)|\,dy \\
&\quad\leq \int_{\mathbb{R}^n}|f(y)\,[\varphi(x+h-y)-\varphi(x-y)]|\,dy \\
&\quad\leq \|f\|_p\|\varphi(x+h-\cdot)-\varphi(x-\cdot)\|_{L^q(dy)}. \qquad (4.10)
\end{aligned}
$$

If $1 < p$ then $q < \infty$ and the continuity of $f * \varphi$ follows from Theorem 4.5. If $p = 1$ then $q = \infty$ and we get as in (4.10),

$$|(f * \varphi)(x + h) - (f * \varphi)(x)| \leq \|f\|_1 \sup_y |\varphi(y - h) - \varphi(y)| . \tag{4.11}$$

By the uniform continuity of φ the right hand side of (4.11) goes to zero as $|h| \to 0$. This proves (i). Notice that in proving (i) we only used the fact that $\varphi \in C_0^1(\mathbb{R}^n)$.

(ii) $f * \varphi$ is differentiable and

$$\frac{\partial}{\partial x_j}(f * \varphi) = f * \frac{\partial \varphi}{\partial x_j}, \text{ for } j = 1, \dots n. \tag{4.12}$$

To prove (ii), let e_j be the jth coordinate vector in $\mathbb{R}^n$. Fix the point $x \in \mathbb{R}^n$ and consider the function

$$F_{x,t,j}(y) = \frac{\varphi(x + te_j - y) - \varphi(x - y)}{t} - \frac{\partial \varphi}{\partial x_j}(x - y)$$

with $t > 0$. Then

$$\frac{f * \varphi(x + te_j) - f * \varphi(x)}{t} - \big(f * \tfrac{\partial \varphi}{\partial x_j}\big)(x) = \int_{\mathbb{R}^n} f(y) F_{x,t,j}(y)\, dy. \tag{4.13}$$

Since the function φ has compact support, is continuous and differentiable, we have that as $t \to 0$, $F_{x,t,j}(y) \to 0$, uniformly in y. Furthermore, for each fixed x, the function $F_{x,t,j}(y)$ is uniformly bounded in y and of compact support. By the dominated convergence theorem the right side of (4.13) goes to zero as t goes to zero and we have proved (ii).

Now repeat the proofs of (i) and (ii) with $\frac{\partial \varphi}{\partial x_j}$ in place of φ and continue to iterate the argument this way to complete the proof. □

Exercise 4.2.1.
Prove that if f and g both have compact support so does $f * g$ and in fact the support of $f * g$ is contained in the union of the supports of f and g.

We end this section by stating another important inequality for convolutions. The proof follows directly from the Riesz–Thorin interpolation theorem given in Section 5.4 of Chapter 5 below.

Theorem 4.6 (Young's Inequality)**.** *Let $r \geq 1$ satisfy*

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

for $1 \leq p \leq q \leq \infty$ and suppose $f \in L^p(\mathbb{R}^n)$ and $g \in L^r(\mathbb{R}^n)$. Then $f * g \in L^q(\mathbb{R}^n)$ and

$$\|f * g\|_q \leq \|g\|_r \|f\|_p.$$

4.3 Approximations to the Identity

For any function φ defined on $\mathbb{R}^n$ and $\varepsilon > 0$, we define the dilation operator $\tau_\varepsilon \varphi(x) = \varphi(\varepsilon x)$ and the function $\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(x/\varepsilon)$.

Lemma 4.7. *Let $\varphi \geq 0$, and $\varphi \in L^1(\mathbb{R}^n)$. Then for any $\varepsilon > 0$,*

$$\int_{\mathbb{R}^n} \varphi_\varepsilon(x)\, dx = \int_{\mathbb{R}^n} \varphi(x)\, dx \tag{4.14}$$

and

$$\lim_{\varepsilon \to 0} \int\limits_{\{|x|>\delta\}} \varphi_\varepsilon(x)\, dx = 0, \text{ for any } \delta > 0. \tag{4.15}$$

Furthermore, if we set $Tf(x) = \varphi * f(x)$ we have

$$(\tau_{\varepsilon^{-1}} T \tau_\varepsilon)\, f(x) = \varphi_\varepsilon * f(x). \tag{4.16}$$

Proof. (4.14) follows from the obvious change of variables. For (4.15) observe that by the dominated convergence theorem

$$\lim_{k \to \infty} \int\limits_{\{|x| \leq k\}} \varphi(x)\, dx = \int_{\mathbb{R}^n} \varphi(x) dx$$

and so

$$\lim_{k \to \infty} \int\limits_{\{|x|>k\}} \varphi(x)\, dx = 0.$$

Thus,

$$\int\limits_{\{|x|>\delta\}} \varphi_\varepsilon(x)\, dx = \frac{1}{\varepsilon^n} \int\limits_{\{|x|>\delta\}} \varphi(x/\varepsilon)\, dx = \int\limits_{\{|x|>\delta/\varepsilon\}} \varphi(x)\, dx \to 0,$$

as $\varepsilon \to 0$.

For (4.16) simply observe that the left hand side is

$$\int_{\mathbb{R}^n} \varphi(\varepsilon^{-1}x - y) f(\varepsilon y)\, dy$$

and change variables. $\square$

Theorem 4.8. *Let $\varphi \geq 0$ with $\int_{\mathbb{R}^n} \varphi(y)\, dy = 1$.*

(i) Suppose $f \in L^p(\mathbb{R}^n)$, $1 \leq p < \infty$. Then

$$\lim_{\varepsilon \to 0} \|f * \varphi_\varepsilon - f\|_p = 0. \tag{4.17}$$

(ii) Suppose $f \in L^\infty(\mathbb{R}^n)$. Then for every x which is a point of continuity of f,

$$\lim_{\varepsilon \to 0} f * \varphi_\varepsilon(x) = f(x). \tag{4.18}$$

Proof. By our assumption on φ and (4.14), $\varphi_\varepsilon\, dy$ is a probability measure. Writing

$$|f * \varphi_\varepsilon(x) - f(x)| = \left| \int_{\mathbb{R}^n} \left(f(x-y) - f(x)\right) \varphi_\varepsilon(y)\, dy \right|$$

we obtain by the Minkowski integral inequality that

$$\begin{aligned} \|f * \varphi_\varepsilon - f\|_p &\leq \left\| \int_{\mathbb{R}^n} |f(x-y) - f(x)| \varphi_\varepsilon(y)\, dy \right\|_p \\ &\leq \int_{\mathbb{R}^n} \|f(x-y) - f(x)\|_{L^p(dx)} \varphi_\varepsilon(y)\, dy. \end{aligned} \tag{4.19}$$

Set

$$I = \int_{\{|y| \leq \delta\}} \|f(x-y) - f(x)\|_{L^p(dx)} \varphi_\varepsilon(y)\, dy$$

and

$$II = \int_{\{|y| > \delta\}} \|f(x-y) - f(x)\|_{L^p(dx)} \varphi_\varepsilon(y)\, dy$$

By the continuity of the translations in $L^p(\mathbb{R}^n)$, Lemma 4.4, given $\eta > 0$ there is a δ such that

$$\|f(\cdot - y) - f(\cdot)\|_{L^p} < \eta$$

for all $|y| < \delta$. Thus with such a δ,

$$I < \eta \int\limits_{\{|y|<\delta\}} \varphi_\varepsilon(y)\, dy \leq \eta.$$

From the fact that

$$\|f(\cdot - y) - f(\cdot)\|_{L^p} \leq 2\|f\|_p$$

it follows that

$$II \leq 2\|f\|_p \int\limits_{\{|y|>\delta\}} \varphi_\varepsilon(y)\, dy,$$

which goes to zero as ε goes to zero by (4.15). This and (4.19) completes the proof of (4.17).

For (4.18) we begin exactly in the same way and obtain

$$|f * \varphi_\varepsilon(x) - f(x)| \leq \int\limits_{\{|y|<\delta\}} |f(x-y) - f(x)|\varphi_\varepsilon(y)\, dy$$
$$+ \int\limits_{\{|y|\geq\delta\}} |f(x-y) - f(x)|\varphi_\varepsilon(y)\, dy = I + II.$$

II is handled exactly as in (i). For I, suppose f is continuous at x. There is a $\delta > 0$ (depending on x) such that

$$|f(x-y) - f(x)| \leq \eta, \text{ if } |y| < \delta.$$

Thus

$$I < \eta \int\limits_{\{|y|<\delta\}} \varphi_\varepsilon(y)\, dy \leq \eta,$$

which proves (4.18) and hence the Theorem. □

In subsequent chapters we will make repeated applications of the following corollary.

Corollary 4.9. *The space of infinitely differentiable functions with compact support, $C_0^\infty(\mathbb{R}^n)$, is dense in $L^p(\mathbb{R}^n)$, $1 \leq p < \infty$. That is, given an $f \in L^p(\mathbb{R}^n)$ and and $\eta > 0$ there is a $g \in C_0^\infty(\mathbb{R}^n)$ such that*

$$\|f - g\|_p < \eta.$$

Proof. Fix a nonnegative $\varphi \in C_0^\infty(\mathbb{R}^n)$, $\varphi \geq 0$, with integral 1. Let $f \in L^p(\mathbb{R}^n)$ and $\eta > 0$. Let r be large enough so that

$$\int\limits_{\{|y|>r\}} |f(y)|^p \, dy < \eta^p \tag{4.20}$$

and set

$$f^\varepsilon(x) = (f\chi_{B(0,r)}) * \varphi_\varepsilon(x),$$

where $B(0,r)$ is the ball centered at 0 and radius r. By Theorem 4.5 and Exercise 4.2.1, $f^\varepsilon \in C_0^\infty(\mathbb{R}^n)$. By the Minkowski inequality and (4.20),

$$\begin{aligned} \|f - f^\varepsilon\|_p &\leq \left\|f\chi_{B(0,r)} - f^\varepsilon\right\|_p + \left\|f(1 - \chi_{B(0,r)})\right\|_p \\ &\leq \ \left\|f\chi_{B(0,r)} - f^\varepsilon\right\|_p + \eta = \left\|f\chi_{B(0,r)} - \varphi_\varepsilon * (f\chi_{B(0,r)})\right\|_p + \eta. \end{aligned}$$

By Theorem 4.8, the first term on the right hand side of (4.21) goes to zero as $\varepsilon \to 0$. Since $\eta > 0$ was arbitrary, this proves the corollary. $\square$

It follows from Theorem 4.8 that given $f \in L^p(\mathbb{R}^n)$ there is a sequence ε_j converging to zero such that

$$\lim_{\varepsilon_j \to 0} f * \varphi_{\varepsilon_j}(x) = f(x) \text{ a.e.}$$

The following stronger result will be proved in the next chapter where we will study the *Hardy–Littlewood Maximal Function*, one of the most fundamental operators in harmonic analysis.

Theorem 4.10. *With φ as in Theorem 4.8 and $f \in L^p(\mathbb{R}^n)$, $1 \leq p \leq \infty$, we have*

$$\lim_{\varepsilon \to 0} f * \varphi_\varepsilon(x) = f(x) \textit{ a.e.}$$

Chapter 5

The Hardy–Littlewood Maximal Function

In this chapter we study the Hardy-Littlewood maximal function. This is one of the most fundamental operators in real analysis with many applications in different areas. The Hardy-Littlewood maximal function will play an important role in the rest of these notes. We will first prove its L^p–boundedness and then apply this to prove some classical differentiation theorems. The Hardy-Littlewood maximal function is a maximal operator obtained by convolution, as in Chapter 4, with the characteristic function of the unit ball normalized to have integral one. It is remarkable that all other such convolutions can be controlled by this one. This is presented in Theorem 5.4 below; Theorem 4.10 in Chapter 4 immediately follows from this. The Calderón–Zygmund decomposition, another result of fundamental importance in real analysis, will be proved in Section 5.2. As an illustration of the usefulness of the Calderón–Zygmund decomposition, we will discuss in Section 5.3 some of its applications to BMO. This is the John–Nirenberg space of functions with *bounded mean oscillation*. The Marcinkiewicz and Riesz–Thorin interpolation theorems are presented in Section 5.4. These interpolation theorems will be important in proving the L^p–boundedness of singular integrals in Chapter 7.

5.1 The L^p–inequalities

Let us denote by $L^1_{\text{loc}}(\mathbb{R}^n)$ the space of measurable functions on $\mathbb{R}^n$ which are integrable on any bounded subset of $\mathbb{R}^n$. Of fundamental importance for the rest of these notes is the Hardy-Littlewood maximal operator (also often referred to in this book as the Hardy-Littlewood maximal function) defined for any $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ by

$$Mf(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int\limits_{B(x,r)} |f(y)|\,dy, \tag{5.1}$$

where as before $|B(x,r)|$ denotes the volume of the ball $B(x,r)$. Often we will write this simply as $\gamma_n r^n$ where γ_n is the volume of the unit ball. This operator arises naturally in studying the following question on differentiation: *If $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, for what x's is it true that*

$$\lim_{r\to 0} \frac{1}{|B(x,r)|} \int\limits_{B(x,r)} f(y)\,dy = f(x)? \tag{5.2}$$

The Lebesgue Differentiation Theorem asserts that (5.2) holds for almost every x. To write M as a a convolution recall that $\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(x/\varepsilon)$, for any $\varepsilon > 0$. In this chapter we will often use r in place of ε to conform to standard notation. Notice that if we take

$$\varphi(x) = \frac{1}{|B(0,1)|}\chi_{B(0,1)}(x) \tag{5.3}$$

then a simple change of variables gives

$$\varphi_r * f(x) = \frac{1}{|B(x,r)|} \int\limits_{B(x,r)} f(y)\,dy$$

and

$$Mf(x) = \sup_{r>0} \big(\varphi_r * |f|\big)(x);$$

hence the connection to approximations to the identity and to Theorem 4.10 of Chapter 4. However, as mentioned above, it is a remarkable fact that other approximations to the identity can be controlled by this one, as we shall see in Theorem 5.4 below.

We begin with the basic boundedness properties of the Hardy-Littlewood maximal function. First, let us observe that if $f \not\equiv 0$, then $Mf \notin L^1(\mathbb{R}^n)$. We leave it as an exercise to the reader to check that if $f \not\equiv 0$ then $Mf(x) > C^1_f/|x|^n$ for all $|x| > C^2_f$, where the constant C^1_f and C^2_f depend on f. Thus by integrating in polar coordinates (Exercise 3.2.7 in page 50) we have

$$\int_{\mathbb{R}^n} Mf(x)\,dx \ge C^1_f \int\limits_{\mathbb{R}^n\setminus B(0,C^2_f)} \frac{dx}{|x|^n} = C^1_f \sigma(S^{n-1}) \int_{C^2_f}^{\infty} \frac{r^{n-1}}{r^n}\,dr = \infty. \tag{5.4}$$

We next introduce some standard terminology which is used throughout this book. Let $(X, \mathcal{F}, \mu)$ and $(Y, \mathcal{G}, \nu)$ be two measure spaces. For $1 \leq p \leq \infty$ denote their corresponding L^p–spaces by $L^p(\mu)$ and $L^p(\nu)$, respectively, with corresponding norms $\|f\|_{L^p(\mu)}$ and $\|f\|_{L^p(\nu)}$. For any $0 < p < \infty$ and any measure space (X, μ), the weak–$L^p(\mu)$ space is the collection of all measurable functions with the property that

$$\sup_{\lambda>0} \left\{ \lambda^p \mu\{x \in X : |f(x)| > \lambda\} \right\} < \infty.$$

For $p = \infty$, weak–L^∞ is defined to be the same as L^∞. A mapping T taking measurable functions from X to Y is said to be of strong–type (p, q) if there is a constant $A = A_{p,q}$ independent of f such that

$$\|Tf\|_{L^q(\nu)} \leq A\|f\|_{L^p(\mu)}.$$

The mapping T is said to be weak–type (p, q) if there is a constant A_1 independent of f such that

$$\nu\{y \in Y : |Tf(y)| > \alpha\} \leq \left(\frac{A_1}{\alpha} \|f\|_{L^p(\mu)} \right)^q,$$

for $q < \infty$. If $q = \infty$, weak–type (p, q) is defined to be the same as strong–type (p, q)

It follows trivially from Chebychev's inequality that if an operator T is strong-type (p, q) then it is weak-type (p, q). But the opposite is not true as the Hardy-Littlewood operator shows.

Theorem 5.1. *The Hardy–Littlewood maximal operator M is weak–type $(1, 1)$ and strong–type (p, p) for $1 < p \leq \infty$. In fact,*

(i)

$$m\{x \in \mathbb{R}^n : Mf(x) > \alpha\} \leq \frac{C_n}{\alpha} \int_{\mathbb{R}^n} |f(y)| \, dy$$

for $f \in L^1(\mathbb{R}^n)$ and

(ii)

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

$1 < p \leq \infty$, $f \in L^p(\mathbb{R}^n)$. *Furthermore, we can take*

$$C_n = 5^n \tag{5.5}$$

and

$$C_{n,p} = 2\left(\frac{5^n p}{p-1}\right)^{1/p}. \tag{5.6}$$

Proof. Let $E_\alpha = \{x \in \mathbb{R}^n \colon Mf(x) > \alpha\}$. By Exercise 5.1.2, this set is open hence measurable. For each $x \in E_\alpha$ there exists a ball centered at x and radius r_x, which we denote by $B_{r_x}(x)$, such that

$$\int_{B_{r_x}(x)} |f(y)|\, dy > \alpha |B_{r_x}(x)|.$$

It follows that $|B_{r_x}(x)| < \frac{1}{\alpha}\|f\|_1$ and hence these balls have bounded diameter. Clearly we have

$$E_\alpha \subset \bigcup_{x \in E_\alpha} B_{r_x}(x).$$

We may therefore, by Theorem 1.4 in Chapter 1, pick a disjoint sequence $B_1, B_2, \ldots$ of these balls such that $|E_\alpha| \le 5^n \sum_j |B_j|$. By the way the balls were chosen, we know that

$$\sum_j |B_j| \le \frac{1}{\alpha} \sum_j \int_{B_j} |f(y)|\, dy = \frac{1}{\alpha} \int_{\cup B_j} |f(y)|\, dy \le \frac{1}{\alpha} \int_{\mathbb{R}^n} |f(y)|\, dy,$$

proving (i) with $C_n = 5^n$.

For (ii), observe first that since $\|Mf\|_\infty \le \|f\|_\infty$, the case $p = \infty$ is trivial With the case $p = \infty$ and (i) proved, we now follow the argument for the the Marcinkiewicz Interpolation Theorem given in 5.17 below. Assume that $1 < p < \infty$. Let $\alpha > 0$ and define

$$f_1(x) = \begin{cases} f(x), & \text{if } |f(x)| \ge \alpha/2 \\ 0, & \text{otherwise.} \end{cases}$$

Since

$$\int_{\mathbb{R}^n} |f_1(y)|\, dy = \int\limits_{\{|f(x)| \ge \alpha/2\}} |f(x)|\, dx,$$

we see that $f_1 \in L^1(\mathbb{R}^n)$. It is also clear that

$$|f(x)| \le |f_1(x)| + \alpha/2$$

from which we conclude that

$$Mf(x) \leq Mf_1(x) + \alpha/2.$$

Thus by (i) we have that

$$m\{x \in \mathbb{R}^n \colon Mf(x) > \alpha\} \leq m\{Mf_1(x) > \alpha/2\} \leq \tfrac{2\cdot 5^n}{\alpha} \int_{\mathbb{R}^n} |f_1(x)|\, dx.$$

This together with Exercise 3.2.3 on page 48, and Fubini's theorem imply

$$\begin{aligned}
\int_{\mathbb{R}^n} Mf(x)^p\, dx &= p \int_0^\infty \alpha^{p-1} m\{Mf > \alpha\}\, d\alpha \\
&\leq p \int_0^\infty \alpha^{p-1} \frac{2\cdot 5^n}{\alpha} \int\limits_{\{|f|\geq \alpha/2\}} |f(x)|\, dx\, d\alpha \\
&= 2\cdot 5^n p \int_{\mathbb{R}^n} |f(x)| \int_0^{2|f(x)|} \alpha^{p-2}\, d\alpha\, dx \\
&= 2^p \cdot \frac{5^n p}{p-1} \int_{\mathbb{R}^n} |f(x)|^p\, dx
\end{aligned}$$

and we obtains the bound

$$\|Mf\|_p \leq 2 \left(\frac{5^n p}{p-1} \right)^{1/p} \|f\|_p, \quad 1 < p < \infty,$$

which completes the proof of the theorem. □

Remark 5.1. Note that for fixed n, $C_{n,p} \to \infty$, as $p \to 1$ and $C_{n,p} \to 1$, as $p \to \infty$, reflecting the fact that the operator is bounded on $L^\infty(\mathbb{R}^n)$ but not on $L^1(\mathbb{R}^n)$. Also, as far as the asymptotic behavior in p is concerned, the constants in (5.5) and (5.6) are already best possible. However, this is not the case with respect to the dimension n. It has been proved by Stein and Strönberg [St-St] that the constant in (ii) can be taken to be independent of the dimension, and that the constant in (i) can be taken to be $Cn \log n$ with C independent of n. Whether the constant in (i) can be taken independent of the dimension remains an interesting and challenging open problem. For other interesting problems concerning the dependence on n in other clasical operators in analysis, see Stein [St2].

Corollary 5.2 (Lebesgue Differentiation Theorem)**.** *Let* $f \in L^p(\mathbb{R}^n)$, $1 \leq p \leq \infty$. *Then*

$$\lim_{r\to 0} \frac{1}{|B(x,r)|} \int\limits_{B(x,r)} f(y)\, dy = f(x) \text{ a.e.}$$

Proof. Let us first deal with the case $p = 1$. Let

$$\omega(f)(x) = \left| \limsup_{r\to 0} f * \varphi_r(x) - \liminf_{r\to 0} f * \varphi_r(x) \right|,$$

where

$$\varphi = \frac{1}{|B(0,1)|}\chi_{B(0,1)}$$

By Theorem 4.8, Chapter 4, $\|f * \varphi_r - f\|_1 \to 0$ as $r \to 0$ and if $h \in C_0^\infty(\mathbb{R}^n)$, $h * \varphi_r(x) \to h(x)$ for every x. Thus for such an h we have $\omega(h)(x) = 0$ for all x. Given $f \in L^1(\mathbb{R}^n)$ and $\eta > 0$, choose an $h \in C_0^\infty(\mathbb{R}^n)$ such that

$$\|f - h\|_1 < \eta.$$

With this we arrive at

$$\omega(f)(x) \le \omega(f-h)(x) \le 2M(f-h)(x).$$

Fix $\varepsilon > 0$. By (i) of Theorem 5.1 we have that

$$\begin{aligned} |\{\omega(f)(x) > \varepsilon\}| &\le |\{\omega(f-h) > \varepsilon\}| \\ &\le |\{M(f-h) > \varepsilon/2\}| \\ &\le \frac{C_n}{\varepsilon}\int_{\mathbb{R}^n} |f-h|\,dx \le \frac{C_n\eta}{\varepsilon}. \end{aligned}$$

Since this is true for any $\eta > 0$ we obtain that

$$|\{\omega(f)(x) > \varepsilon\}| = 0.$$

Since $\varepsilon > 0$ was arbitrary, this can only happen if $\omega(f)(x) = 0$ for almost every x, which proves that $\lim_{r\to 0} \varphi_r * f$ exists almost everywhere. By the convergence of $\varphi_r * f$ to f in L^1, there exists a sequence $r_k \downarrow 0$ such that $\lim_{r_k\to 0} \varphi_{r_k} * f \to f$ almost everywhere. Thus

$$\lim_{r\to 0} \varphi_r * f = f \text{ a.e.}$$

and we have proved the case $p = 1$. The case $1 < p < \infty$ is exactly the same except that we use Chebychev's inequality and (ii) of Theorem 5.1.

We now present the case $p = \infty$ which is a little different due to the fact that we cannot approximate functions in L^∞ by functions in $C_0^\infty(\mathbb{R}^n)$. Suppose $f \in L^\infty$. We shall show that $f * \varphi_r(x) \to f(x)$ for almost every $x \in B$ for any fixed ball B. Let B_1 be another ball satisfying $B \subset B_1$ and with $\delta =$ distance$(B, B_1^c) > 0$. Set

$$f_1(x) = \begin{cases} f(x) & \text{if } x \in B_1 \\ 0 & \text{if } x \notin B_1 \end{cases}$$

Then $f_1 \in L^1$ and by what we have already proved, $f_1 * \varphi_r(x) \to f_1(x)$, as $r \to 0$, for almost every x. In particular,

$$f_1 * \varphi_r(x) \to f_1(x) = f(x), \text{ a.e. in } B,$$

as $r \to 0$. Set $f_2(x) = f(x) - f_1(x)$. We claim that

$$f_2 * \varphi_r(x) \to 0, \quad x \in B,$$

as $r \to 0$. To see this observe that for $x \in B$,

$$\begin{aligned} |f_2 * \varphi_r(x)| &= \left| \int_{\mathbb{R}^n} f_2(x-y)\varphi_r(y)\, dy \right| \\ &\leq \int\limits_{\{|y| \geq \delta > 0\}} |f_2(x-y)||\varphi_r(y)|\, dy \\ &\leq \|f\|_{L^\infty} \int\limits_{\{|y| \geq \delta\}} |\varphi_r(y)|\, dy \\ &= \|f\|_{L^\infty} \int\limits_{\{|y| \geq \delta/r\}} |\varphi(y)|\, dy \end{aligned}$$

and this last quantity goes to zero as $r \to 0$. $\square$

Exercise 5.1.1.
Adapt the argument used in the proof of (i) in Theorem 5.1 to prove that there are constants C_1 band C_2 depending only on n such that for all $\alpha > 0$ the following inequality holds:

$$m\{x \in \mathbb{R}^n : Mf(x) > \alpha\} \leq \frac{C_1}{\alpha} \int\limits_{\{x \in \mathbb{R}^n : |f(x)| > C_2\alpha\}} |f(x)|\, dx.$$

As we mentioned earlier, the above result for the characteristic (indicator) function of the ball implies the more general result stated in Theorem 4.10 of Chapter IV. Before we state this result we recall a definition.

Definition 5.3. For any $\varphi \in L^1(\mathbb{R}^n)$ we define its least decreasing radial majorant of φ by

$$\psi(x) = \sup_{|y|\geq|x|} |\varphi(y)|.$$

Theorem 5.4. *Let $\varphi \in L^1(\mathbb{R}^n)$ and suppose*

$$\int_{\mathbb{R}^n} \psi(x)\, dx = A < \infty.$$

The for any $f \in L^p(\mathbb{R}^n)$, $1 \leq p \leq \infty$, we have

$$\sup_{\varepsilon>0} |f * \varphi_\varepsilon(x)| \leq AMf(x).$$

Proof. By bringing absolute values inside the integral and majorizing φ by ψ, it is enough to prove

$$|f * \psi_\varepsilon(x)| \leq AMf(x) \tag{5.7}$$

for $f \geq 0$. Since (5.7) is translation invariant with respect to f and dilation invariant with respect to ψ, we only need to show that

$$|f * \psi_1(0)| \leq AMf(0). \tag{5.8}$$

That is, once (5.8) is proved we apply it with f replaced by $f_x(y) = f(x-y)$ and ψ replaced by ψ_ε which has the same A. To prove (5.8), assume our function ψ is a simple function of the form

$$\psi(x) = \sum_{k=1}^{m} c_k \chi_{B(0,r_k)}(x) \tag{5.9}$$

where the numbers c_k are positive. With this we see that

$$A = \sum_{k=1}^{m} c_k |B(0, r_k)|.$$

Then as $\psi(x) = \psi(-x)$,

$$|f * \psi_1(0)| = \left| \int_{\mathbb{R}^n} f(x)\psi(x)\, dx \right| \leq \sum_{k=1}^{m} c_k \int_{B(0,r_k)} |f(x)|\, dx$$

$$= \sum_{k=1}^{m} c_k |B(0,r_k)| \frac{1}{|B(0,r_k)|} \int_{B(0,r_k)} |f(x)|\,dx \le AMf(0). \quad (5.10)$$

The case of general ψ is accomplished by Exercise 5.1.10 below. Indeed, given the exercise we have by the monotone convergence theorem that

$$\begin{aligned} \left| \int_{\mathbb{R}^n} f(x)\psi(x)\,dx \right| &\le \int_{\mathbb{R}^n} |f(x)|\psi(x)\,dx \\ &= \lim_{j\to\infty} \int_{\mathbb{R}^n} |f(x)|\psi_j(x)\,dx \le AMf(0), \end{aligned}$$

and this proves the theorem. □

An alternative proof of Theorem 5.4 goes as follows. Set $E = \{(y,t) : \psi(y) > t > 0\}$. Note that

$$\psi(y) = \int_0^\infty \chi_E(y,t)\,dt$$

By Fubini's Theorem and the fact that for each $t > 0$ the set $\{y \in \mathbb{R}^n : \psi(y) > t\}$ is a ball centered at the origin of radius r_t, we have

$$\begin{aligned} |f * \psi(x)| &\le \int_{\mathbb{R}^n} |f(x-y)| \int_0^\infty \chi_E(y,t)\,dt\,dy \\ &= \int_0^\infty \int_{\{y\in\mathbb{R}^n:\psi(y)>t\}} |f(x-y)|\,dy\,dt \\ &= \int_0^\infty \left(\frac{1}{|B_{r_t}|} \int_{B_{r_t}} |f(x-y)|\,dy \right) |B_{r_t}|\,dt \\ &\le \|\psi\|_1 Mf(x) = AMf(x), \end{aligned}$$

which proves the theorem.

The following two corollaries are direct from Theorems 5.1, 5.4 and the proof of Corollary 5.2. Observe that Corollary 5.5 below is in fact more general than Theorem 4.10 of Chapter 4.

Corollary 5.5. *Let φ and A be as in Theorem 5.4. Set*

$$f^*(x) = \sup_{\varepsilon>0} |\varphi_\varepsilon * f(x)|\,.$$

Then

$$m\{x \in \mathbb{R}^n : f^*(x) \ge \alpha\} \le \frac{AC_n}{\alpha}\|f\|_1$$

and

$$\|f^*\|_p \leq AC_{p,n}\|f\|_p,$$

for $1 < p \leq \infty$*. The constants* C_n *and* $C_{p,n}$ *are the same as those in Theorem 1.1.*

Corollary 5.6. *Let* φ *be as in Theorem 5.4 with* $\int_{\mathbb{R}^n} \varphi(x)\,dx = 1$*.*

(i) If $f \in L^p(\mathbb{R}^n)$, $1 \leq p \leq \infty$*, then*

$$\lim_{\varepsilon\to 0} f * \varphi_\varepsilon(x) = f(x) \text{ a.e.}$$

(ii) If $f \in L^p(\mathbb{R}^n)$, $1 \leq p < \infty$*, then*

$$\lim_{\varepsilon\to 0} \|f * \varphi_\varepsilon - f\|_p = 0.$$

Exercise 5.1.2.
Prove that the maximal function is lower semi-continuous. That is, for all $\lambda > 0$, $\{x \in \mathbb{R}^n : Mf(x) > \lambda\}$ is open.

Exercise 5.1.3.
Prove that

$$\lim_{r\to 0} \frac{1}{|B(x,r)|}\int_{B(x,r)} f(y)\,dy = f(x) \text{ a.e.}$$

for $f \in L^1_{loc}(\mathbb{R}^n)$.

Exercise 5.1.4.
Suppose the function f is supported in the ball B with $f \in L^1(B)$ and $Mf \in L^1(B)$. Prove that

$$\int_{\{x\in\mathbb{R}^n: Mf(x)>\alpha\}} Mf(x)\,dx < \infty,$$

for any $\alpha > 0$.

Exercise 5.1.5.
Prove that in the definition of the maximal function the ball $B(x,r)$ can be replaced by a cube. That is if instead we define

$$Mf(x) = \sup_Q \frac{1}{|Q|}\int_Q |f(y)|\,dy,$$

where the sup is taken over all cubes Q centered at x, then M satisfies the conclusions of Theorem 5.1.

Exercise 5.1.6.
For $f \in L^1_{loc}(\mathbb{R}^n)$. For $1 \le p < \infty$, define

$$M_p(f)(x) = \sup_{r>0} \left(\frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|^p \, dy \right)^{1/p}.$$

Prove that M_p is weak–type (p,p) and strong–type (q,q) for $p < q \le \infty$.

Exercise 5.1.7.
Let μ be a Borel measure on $\mathbb{R}^n$ satisfying the doubling property. That is, there is a constant c such that $\mu(B(x,2r)) \le c\mu(B(x,r))$ for all r. Define

$$M_\mu f(x) = \sup_{r>0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f(y)| \, d\mu(y).$$

Prove that

(i) $\mu\{x \in \mathbb{R}^n \colon M_\mu f(x) > \alpha\} \le \frac{C_n}{\alpha} \int_{\mathbb{R}^n} |f(y)| \, d\mu(y)$, for $f \in L^1(\mu)$

(ii) $\|M_\mu f\|_{L^p(\mu)} \le C_{n,p} \|f\|_{L^p(\mu)}$, $1 < p \le \infty$, $f \in L^p(\mathbb{R}^n)$.

Exercise 5.1.8.
Kolmogorov's inequality. Let $f \in L^1(\mathbb{R}^n)$ and $0 < p < 1$. Prove that

$$\int_E (Mf(x))^p \, dx \le C_p |E|^{1-p} \|f\|_1^p,$$

for all measurable $E \in \mathbb{R}^n$.

Exercise 5.1.9.
Let μ be a finite Borel measure on $\mathbb{R}^n$. We define its maximal function by

$$M\mu(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \mu(B(x,r)).$$

Prove that

$$m\{x \in \mathbb{R}^n : M\mu(x) > \alpha\} \le \frac{C_n}{\alpha} \mu(\mathbb{R}^n).$$

Exercise 5.1.10.
Let $\psi \in L^1(\mathbb{R}^n)$ be nonnegative, radial, and decreasing. Prove that there is a sequence of simple functions ψ_j of the form given in (5.9) such that $\psi_j(x) \uparrow \psi(x)$ for each x.

Exercise 5.1.11.

Consider the function P defined on $\mathbb{R}^n$ by

$$P(x) = \frac{C_n}{(1+|x|^2)^{(n+1)/2}}$$

where the constant $C_n = \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}}$ is chosen so that

$$\int_{\mathbb{R}^n} P(x)\,dx = 1.$$

Set

$$P_y(x) = \frac{1}{y^n}P(x/y) = C_n \frac{y}{(y^2+|x|^2)^{(n+1)/2}} \quad y>0. \tag{5.11}$$

The function $P_y(x)$ defined on the upper half space of $\mathbb{R}^n$,

$$\mathbb{R}^{n+1}_+ = \{(x,y) : x \in \mathbb{R}^n, y>0\}$$

is called the Poisson kernel. By Corollaries 5.5 and 5.6 the function u_f^*, where $u_f(x,y) = P_y * f(x)$ and

$$u_f^*(x) = \sup_{y>0} |u_f(x,y)|,$$

satisfies

$$\begin{aligned} m\{x \in \mathbb{R}^n : u_f^*(x) \geq \alpha\} &\leq \frac{C_n}{\alpha}\|f\|_1, & (5.12)\\ (1<p\leq\infty)\ \|u_f^*(x)\|_p &\leq C_{p,n}\|f\|_p, & (5.13)\\ \lim_{y\to 0} u_f(x,y) &= f(x) \text{ a.e.} & (5.14) \end{aligned}$$

Prove that the function

$$u_{f,k}^*(x) = \sup_{y>0}\left| y^k \frac{\partial^k u_f}{\partial y^k}(x,y)\right|, \text{ for any positive integer } k,$$

also satisfies (5.12) and (5.13) with the constants C_n and $C_{p,n}$ replaced by constants depending on k, n, and p.

Exercise 5.1.12.
Consider the function H defined on $\mathbb{R}^n$ by

$$H(x) = \frac{1}{(4\pi)^{n/2}} e^{-|x|^2/4}$$

where once again the constants are chosen so that

$$\int_{\mathbb{R}^n} H(x)\,dx = 1.$$

Set

$$H_t(x) = \frac{1}{t^{n/2}} H(x/\sqrt{t}),$$

for $t > 0$. The function $H_t(x)$ defined on $\mathbb{R}^{n+1}_+$ is called the heat kernel. If $f \in L^p(\mathbb{R}^n)$, $1 \le p \le \infty$, let $u_f(x,t) = H_t * f(x)$ and $u_f^*(x) = \sup_{t>0} |u_f(x,t)|$. Prove that

$$m\{x \in \mathbb{R}^n : u_f^*(x) \ge \alpha\} \le \frac{C_n}{\alpha} \|f\|_1,$$

$$\|u_f^*(x)\|_p \le C_{p,n} \|f\|_p, \quad 1 < p \le \infty$$

and that

$$\lim_{t\to 0} u_f(x,t) = f(x) \quad \text{a.e.}$$

Exercise 5.1.13.
For any $f \in C^2(\mathbb{R}^n)$ define the Laplacian operator Δ by

$$\Delta f(x) = \sum_{k=1}^{n} \frac{\partial^2 f}{\partial x_k^2}(x).$$

If $u(x,y)$, $x \in \mathbb{R}^n$ and $y > 0$, is defined in the upper-half space, we denote by Δ_x the Laplacian of u in the x variable. Let $f \in L^p(\mathbb{R}^n)$, $1 \le p \le \infty$ and define the functions on $\mathbb{R}^{n+1}_+$ by $u_f(x,y) = P_y * f(x)$, $x \in \mathbb{R}^n$, $y > 0$, and $v_f(x,t) = H_t * f(x)$, $x \in \mathbb{R}^n$, $t > 0$. It follows from the previous two exercises that both of these functions have boundary values f in the sense that $\lim_{y\to 0} u_f(x,y) = f(x)$ and that $\lim_{t\to 0} v_f(x,t) = f(x)$, a.e. Prove, by justifying the differentiation under the integral sign, that these functions also satisfy

$$\frac{\partial^2 u_f}{\partial y^2}(x,y) + \Delta_x u_f(x,y) = 0$$

and

$$\frac{\partial v_f}{\partial t}(x,t) = \Delta_x v_f(x,t).$$

That is, u_f is harmonic in $\mathbb{R}^{n+1}_+$ with almost everywhere boundary values f and v_f satisfies the heat equation in $\mathbb{R}^{n+1}_+$ with almost everywhere boundary values f. Equivalently, u_f and v_f solve the Dirichlet problem for the Laplacian and the heat equation, respectively, in the upper-half space.

Exercise 5.1.14.
Define the functions

$$Q^j_y(x) = C_n \frac{x_j}{(|x|^2 + y^2)^{(n+1)/2}}, \quad j = 1, \dots, n. \tag{5.15}$$

where C_n is the constant of (5.11). For $f \in L^2(\mathbb{R}^n)$ define

$$u_j(x,y) = Q^j_y * f(x), \quad j = 1, \dots, n.$$

Prove that $u_j(x,y)$ is also harmonic in $\mathbb{R}^{n+1}_+$. We will return to these functions in Chapter 8 where we will study the Riesz transforms and conjugate harmonic functions in $\mathbb{R}^{n+1}_+$.

Exercise 5.1.15.
For any $\alpha > 0$ and $x \in \mathbb{R}^n$, define the cone in $\mathbb{R}^{n+1}_+$ with vertex at x and aperture $2\arctan(\alpha)$, to be $\Gamma_\alpha(x) = \big\{(x_0,y) \in \mathbb{R}^{n+1}_+ : |x_0 - x| < \alpha y\big\}$. If $f \in L^p(\mathbb{R}^n), 1 \le p \le \infty$, the nontangential maximal function of f is defined to be

$$N_\alpha f(x) = \sup_{(x_0,y)\in\Gamma_\alpha(x)} |P_y * f(x_0)|.$$

Prove that there is a constant $C_{\alpha,n}$ such that

$$N_\alpha f(x) \le C_{\alpha,n} M f(x)$$

and therefore the nontangential maximal function satisfies the boundedness properties (i) and (ii) of Theorem 5.1.

Exercise 5.1.16.
Let $f \in L^p(\mathbb{R}^n)$, $1 \le p \le \infty$. Prove that the Poisson integral of f converges nontangentially to f almost everywhere. That is, prove that

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

5.2 The Calderón–Zygmund Decomposition

In this section we state one of the most important consequences of the boundedness of the maximal operator. Namely, the Calderón–Zygmund decomposition of functions in R^n. We will see that this result, together with the Plancherel's Theorem which we shall prove in the Chapter 6, imply the boundedness of singular integrals, and more. First, by a cube Q in $\mathbb{R}^n$ we will mean a cube with sides parallel to the coordinate axes and denote its interior by Q^o.

Theorem 5.7 (The Calderón–Zygmund Decomposition)**.** *Let f be a nonnegative function in $L^1(\mathbb{R}^n)$ and α a positive number. There exist two sets Ω and F and a countable collection of cubes $\{Q_k\}$ such that*

(i) $\mathbb{R}^n = F \cup \Omega,\ F \cap \Omega = \emptyset$*,*

(ii) $f(x) \leq \alpha$ *a.e. on* F*,*

(iii) $\Omega = \bigcup_k Q_k,\ Q_k^o \cap Q_j^o = \emptyset$*,* $k \neq j$*.*

In addition, for every cube Q_k,

$$\alpha < \frac{1}{|Q_k|} \int_{Q_k} f(x)\, dx \leq 2^n \alpha. \tag{5.16}$$

Remark 5.2. We will often refer to this decomposition as the *Calderón–Zygmund decomposition* of f at level α.

Corollary 5.8. *Let $f \in L^1(\mathbb{R}^n)$ and let $\{Q_k\}$ be the cubes obtained in Theorem 5.7 when applied to $|f|$ and $\alpha > 0$. Define the "good" function g by*

$$g(x) = \begin{cases} f(x) & \text{if } x \in F \\ \frac{1}{|Q_k|} \int_{Q_k} f(x)\, dx & \text{if } x \in Q_k. \end{cases}$$

That is,

$$g(x) = f\chi_F(x) + \sum_{k=1}^{\infty} \left(\frac{1}{|Q_k|} \int_{Q_k} f(x)\, dx \right) \chi_{Q_k}(x). \tag{5.17}$$

Define the "bad" function b by

$$b(x) = f(x) - g(x). \tag{5.18}$$

Then

(i) The "good" function belongs to L^2 and satisfies

$$\|g\|_2^2 \leq C\alpha \|f\|_1, \tag{5.19}$$

(ii) The "bad" function satisfies $b(x) = 0$ on F with

$$\int_{Q_k} b(x)\, dx = 0 \textit{ for all } k \tag{5.20}$$

and

$$\int_{Q_k} |b(y)|\, dy \leq C\alpha |Q_k|. \tag{5.21}$$

Proof. Clearly $b(x) = 0$ on F. Also,

$$\begin{aligned}\int_{Q_k} b(x)\, dx = \int_{Q_k} f(x)\, dx - \int_{Q_k} g(x)\, dx \\ = \int_{Q_k} f(x)\, dx - \int_{Q_k} f(x)\, dx = 0,\end{aligned}$$

which proves (5.20). Also by (5.16) and the definition of g,

$$\begin{aligned}\|g\|_2^2 &= \int_{\mathbb{R}^n} |g(x)|^2\, dx \\ &= \int_F |g(x)|^2\, dx + \int_\Omega |g(x)|^2\, dx \\ &\leq \alpha \int_F |f(x)|\, dx + \sum_k \int_{Q_k} |g(x)|^2\, dx \\ &= \alpha \int_F |f(x)|\, dx + C\alpha^2 \sum_k |Q_k| \\ &= \alpha \int_F |f(x)|\, dx + C\alpha^2 m(\Omega).\end{aligned}$$

But by the left hand side of (5.16),

$$m(\Omega) \leq \frac{C}{\alpha} \int_\Omega |f(x)|\, dx. \tag{5.22}$$

Hence, $\|g\|_2^2 \leq C\alpha \|f\|_1$, as asserted in (5.18). Finally,

$$\int_{Q_k} |b(y)|\, dy \leq \int_{Q_k} |f(y)|\, dy + \int_{Q_k} |g(y)|\, dy$$

$$
\begin{aligned}
&\leq \int_{Q_k} |f(y)|\,dy + C\alpha \int_{Q_k} dy \\
&\leq C\alpha|Q_k| + C\alpha|Q_k| = C\alpha|Q_k|,
\end{aligned}
$$

which proves (5.21) and hence the corollary. □

Proof of Theorem 5.7. The proof is based on a *stopping time* argument. We start by decomposing $\mathbb{R}^n$ into a mesh of equal cubes $\{Q\}$ whose interiors are disjoint and with diameter so large that for each cube in this mesh,

$$\frac{1}{|Q|}\int_Q f(x)\,dx \leq \alpha.$$

Let Q be one such cube in this mesh. We divide Q into 2^n equal cubes. Let Q' be one of the cubes in this subdivision. We have two cases:

Case 1: $\displaystyle\frac{1}{|Q'|}\int_{Q'} f(x)\,dx \leq \alpha.$

Case 2: $\displaystyle\alpha < \frac{1}{|Q'|}\int_{Q'} f(x)\,dx.$

If the cube Q' falls into case 2, select it as one of the cubes Q_k for our collection. In this case note that

$$\alpha < \frac{1}{|Q'|}\int_{Q'} f(x)\,dx \leq \frac{|Q|}{|Q'|}\,\frac{1}{|Q|}\int_Q f(x)\,dx \leq 2^n\alpha.$$

If the cube falls into the first case, we continue with these subdivisions until (if ever), we fall into the second case. Let $\Omega = \bigcup_k Q_k$, where the Q_k's are selected with the above process where this starts with any Q in the original mesh of cubes. Let $F = \Omega^c$. Then, if $x \in F$ we have that

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

for all such cubes containing x. Since we can take a sequence of such cubes with their diameter tending to zero by our construction, Corollary 5.2 gives that $f(x) \leq \alpha$ a.e. This proves the Theorem. □

We give a different proof of Corollary 5.8 (is slightly different form) using (i) of Theorem 5.1 and the following Whitney decomposition valid for any open set in R^n. See Stein [St1] for the proof.

Theorem 5.9 (Whitney Decomposition)**.** *Let Ω be an open set in $\mathbb{R}^n$. There exists a collection of cubes $Q_k, k = 1, 2, \ldots$, such that*

(i) $\Omega = \bigcup_{k=1}^{\infty} Q_k$,

(ii) $Q_j^0 \cap Q_k^0 = \emptyset$, $j \neq k$,

(iii) and there are two constants C_1, C_2, depending only on n, such that

$$C_1 \operatorname{diam}(Q_k) \leq \operatorname{dist}(Q_k, \Omega^c) \leq C_2 \operatorname{diam}(Q_k),$$

where $diam(Q)$ and $\operatorname{dist}(Q, \Omega^c)$ denote the diameter of the cube and the distance from the cube to the set Ω^c, respectively.

Corollary 5.10. *Let $f \in L^1(\mathbb{R}^n)$ and α be a positive number. There is an open set Ω and a closed set F such that*

(i) $\mathbb{R}^n = F \cup \Omega$, $F \cap \Omega = \emptyset$,

(ii) $|f(x)| \leq \alpha$ *a.e. on* F,

(iii) $\Omega = \bigcup_k Q_k$, $Q_k^o \cap Q_j^o = \emptyset$.

Furthermore, there are constants C_1 and C_2, both depending only on n, such that

$$m(\Omega) \leq \frac{C_1}{\alpha} \|f\|_1 \tag{5.23}$$

and

$$\frac{1}{|Q_k|} \int_{Q_k} f \, dx \leq C_2 \alpha \tag{5.24}$$

Proof. Let $\Omega = \{x \in \mathbb{R}^n \colon Mf(x) > \alpha\}$ and $F = \{x \colon Mf(x) \leq \alpha\}$. By Exercise 5.1.2, Ω is open and hence F is closed. Clearly (i) and (ii) hold and the Whitney decomposition gives a collection of cubes for which (iii) also holds. By (i) of Theorem 5.1, Ω satisfies (5.23). To prove (5.24) we use the structure of the cubes. Let Q_k be one of this cubes and let $x_k \in F$ be such that

$$\operatorname{dist}(Q_k, F) = \operatorname{dist}(x_k, Q_k).$$

Let B_k be the ball centered at x_k with $B_k \supset Q_k$ and such that $\operatorname{diam}(B_k) \approx \operatorname{diam}(Q_k)$ where by $A \approx B$ we mean there are absolute constants c_1, c_2 such that $c_1 A \le B \le c_2 A$. The latter is possible because of (iii) in the Whitney decomposition. Then

$$\begin{aligned}\alpha &\ge (Mf)(x_k)\\ &\ge \frac{1}{|B_k|}\int_{B_k} f(x)\,dx\\ &\ge \frac{|Q_k|}{|B_k|}\,\frac{1}{|Q_k|}\int_{Q_k} f(x)\,dx\\ &\ge C\,\frac{1}{|Q_k|}\int_{Q_k} f(x)\,dx.\end{aligned}$$

This completes the proof. □

We showed above that the maximal function Mf is never integrable unless f is identically zero. However, we do have the following result of Stein [St1] concerning local integrability of the maximal function. We present it here as an application of Exercise 5.2.2.

Theorem 5.11. *Fix the ball $B = B(0,1)$ and let $f \in L^1(B)$ be supported in B. Then $Mf \in L^1(B)$ if and only if*

$$\int_B |f(x)| \log^+ |f(x)|\,dx < \infty. \tag{5.25}$$

Proof. Suppose $Mf \in L^1(B)$. By Exercise 5.1.4 and Fubini's theorem,

$$\int_1^\infty m\{x \in \mathbb{R}^n \colon Mf(x) > \alpha\}\,d\alpha \le \int_{\{x\in\mathbb{R}^n \colon Mf(x)>1\}} Mf(x)\,dx < \infty.$$

On the other hand, by Fubini's theorem,

$$\begin{aligned}\int_B |f(x)| \log^+ |f(x)|\,dx &= \int_B |f(x)| \int_1^{|f(x)|} \frac{d\alpha}{\alpha}\,dx\\ &= \int_1^\infty \frac{1}{\alpha}\int_{\{x\in\mathbb{R}^n \colon |f(x)|>\alpha\}} |f(x)|\,dx\,d\alpha\\ &\le 2^n \int_1^\infty m\{x \in \mathbb{R}^n \colon Mf(x) > \alpha\}\,d\alpha,\end{aligned}$$

where the last inequality follows from Exercise 5.2.2. This proves one direction.

Now assume (5.25). Then by Fubini's theorem, Exercise 2.2.4 of Chapter 3, Corollary 5.1.1, and the fact that f is supported in B, we see that

$$\begin{aligned}\int_B Mf(x)\,dx &= 2\int_0^\infty m\{x\in B : Mf(x)>2\alpha\}\,d\alpha\\ &\le \; 2|B| + C_n\int_1^\infty \frac{1}{\alpha}\int_{\{x\in B\colon |f(x)|>\alpha\}} |f(x)|\,dx\,d\alpha\\ &= \; 2|B| + C_n\int_B |f(x)|\int_1^{|f(x)|}\frac{d\alpha}{\alpha}\,dx\\ &= \; 2|B| + C_n\int_B |f(x)|\log^+|f(x)|\,dx.\end{aligned}$$

This completes the proof of the theorem. □

Exercise 5.2.1.
The Calderón–Zygmund Decomposition for cubes. Let Q be a cube in $\mathbb{R}^n$. Suppose f is integrable and nonnegative on Q with

$$\frac{1}{|Q|}\int_Q f(x)\,dx < \alpha.$$

Prove that there is a collection $Q_1, Q_2, \ldots$ of disjoint subcubes of Q such that

(i) $f \le \alpha$ a.e. on $Q \setminus \cup Q_k$,

(ii) $\alpha \le \frac{1}{|Q_k|}\int_{Q_k} f(x)\,dx < 2^n\alpha$ for all k,

(iii) $\sum_k |Q_k| \le \frac{1}{\alpha}\int_Q f(x)\,dx$.

Exercise 5.2.2.
Let $f\in L^1(\mathbb{R}^n)$. Prove that for all $\alpha>0$ the following inequality holds:

$$m\{x\in\mathbb{R}^n\colon Mf(x)>\alpha\} \ge \frac{2^{-n}}{\alpha}\int_{\{x\in\mathbb{R}^n\colon |f(x)|>\alpha\}} |f(x)|dx.$$

The reader should compare this result with that of Corollary 5.1.1.

Exercise 5.2.3.
Let f satisfy

$$\int_{\mathbb{R}^n} |f(x)|\log^+|f(x)|\,dx < \infty. \tag{5.26}$$

Prove that

$$\int_K Mf(x)dx < \infty$$

for any compact set $K \in \mathbb{R}^n$.

Remark 5.3. The collection of measurable functions satisfying (5.26) is denoted by $L\log L(\mathbb{R}^n)$. This space is often a convenient substitute for those results which fail in $L^1(\mathbb{R}^n)$.

In Chapter 4, we discussed approximations to the identity which are obtained by dilation of a single function. We now present, as an application of the Calderón–Zygmund decomposition, a theorem of Zó [] which gives conditions on a family of functions in order to have weak–type results for the maximal function of convolutions with this family.

Theorem 5.12. *Suppose $\{K_\beta\}_{\beta\in A}$ is a collection of functions indexed by the set A and satisfying the following conditions:*

(i) $\sup_{\beta\in A}\int_{\mathbb{R}^n}|K_\beta(x)|\,dx \le B$,

(ii) $\int_{\{|x|\ge 2|y|\}}\sup_{\beta\in A}|K_\beta(x-y)-K_\beta(x)|\,dx \le B$.

Let $f_K^*(x) = \sup_{\beta\in A}|K_\beta * f(x)|$. There are constants A_p and A_1, the first depending only on p and B and the second only on B, such that

$$\|f_K^*(x)\|_p \le A_p\|f\|_p \quad 1 < p \le \infty \tag{5.27}$$

and

$$m\{x\in\mathbb{R}^n\colon f_K^*(x) > \alpha\} \le \frac{A_1}{\alpha}\|f\|_1. \tag{5.28}$$

We begin the proof with a lemma. A slightly different version of this lemma will be used in Chapter 7 in our study of singular integrals satisfying the Hörmander condition.

Lemma 5.13. *Let $K_\beta, \beta\in A$ be a a family of functions satisfying (ii) of Theorem 5.12. Let $x_0\in\mathbb{R}^n$ and Q be a cube centered at x_0. Suppose $a\in L^1(\mathbb{R}^n)$ is supported in Q and has mean value zero over the cube. That is,*

$$\int_Q a(x)\,dx = 0. \tag{5.29}$$

Then

$$\int_{\mathbb{R}^n\setminus Q^*}\sup_{\beta\in A}|K_\beta * a(x)|\,dx \le B\int_Q|a(x)|\,dx,$$

where $Q^* = 2\sqrt{n}Q$.

Proof. We leave it as an exercise (see (i) of Exercise 7.1.1 in Chapter 7) to prove that for every $x \in \mathbb{R}^n \setminus Q^*$ and $y \in Q$, we have $|x - x_0| > 2|y - x_0|$. Using (5.29) we have, for $x \in \mathbb{R}^n \setminus Q^*$,

$$\begin{aligned}
&\sup_{\beta\in A}|K_\beta * a(x)| \\
&\quad= \sup_{\beta\in A}\left|\int_Q (K_\beta(x-y) - K_\beta(x-x_0))\, a(y)\, dy\right| \\
&\quad\leq \int_Q \sup_{\beta\in A}|K_\beta(x-y) - K_\beta(x-x_0)|\, |a(y)|\, dy.
\end{aligned}$$

Integrating both sides of this inequality, using Fubini's theorem and the geometric observation above concerning Q and Q^*, gives

$$\begin{aligned}
&\int_{\mathbb{R}^n\setminus Q^*} \sup_{\beta\in A}|K_\beta * a(x)|\, dx \\
&\leq \int_Q \left(\int_{\mathbb{R}^n\setminus Q^*} \sup_{\beta\in A}|K_\beta(x-y) - K_\beta(x-x_0)|\ dx\right)|a(y)|\, dy \\
&\leq \int_Q \left(\int_{\{|x-x_0|>2|y-x_0|\}} \sup_{\beta\in A}|K_\beta(x-y) - K_\beta(x-x_0)|\ dx\right)|a(y)|\, dy \\
&\leq B\int_Q |a(y)|\, dy,
\end{aligned}$$

where the last inequality follows from our assumption on K_β. □

Proof of Theorem 5.12. We clearly have $\|f_K^*(x)\|_\infty \leq B\|f\|_\infty$, by (i). By the argument of Theorem 5.1, it is enough to prove the $L^1(\mathbb{R}^n)$ case. Fix $\alpha > 0$. Let $f = g + b$ be the Calderón–Zygmund decomposition of f as given in Corollary 5.8 at level α. Since $\|g\|_\infty \leq 2^n\alpha$, we have that $g_K^*(x) \leq B2^n\alpha$. Since $f_K^*(x) \leq g_K^*(x) + b_K^*(x)$ we have

$$m\{x \in \mathbb{R}^n \colon f_K^*(x) \geq \alpha(2^nB+1)\} \leq m\{x \in \mathbb{R}^n \colon b_K^*(x) > \alpha\}.$$

Let Q_k be the cubes given by the Corollary 5.8 and set $\Omega^* = \cup_k Q_k^*$ where the cubes Q_k^* are as in Lemma 5.13. Since $|Q_k^*| = (2\sqrt{n})^n|Q_k|$ we have that $m(\Omega^*) \leq C_n m(\Omega)$ and if we set $F^* = \mathbb{R}^n \setminus \cup Q_k^*$ we see that

$$m\{x \in \mathbb{R}^n \colon b_K^*(x) \geq \alpha\}$$

$$\begin{aligned}
&\leq C_n m(\Omega) + m\{x \in F^*\colon b_K^*(x) > \alpha\} \\
&\leq \frac{C_n}{\alpha}\|f\|_1 + m\{x \in F^*\colon b_K^*(x) > \alpha\}.
\end{aligned}$$

It remains to show that

$$m\{x \in F^*\colon b_K^*(x) > \alpha\} \leq \frac{BC_n}{\alpha}\|f\|_1.$$

By Lemma 5.13, applied with $Q = Q_k$ and $a = b_k = b\chi_{Q_k}$, we have

$$\int_{\mathbb{R}^n \setminus Q_k^*} \sup_{\beta \in A} |K_\beta * b_k(x)|\, dx \leq B \int_{Q_k} |b(x)|\, dx.$$

Since

$$b_K^*(x) \leq \sum_k \sup_{\beta \in A} |K_\beta * b_k(x)|,$$

it follows from Chebychev's inequality that

$$\begin{aligned}
m\{x \in F^*\colon b_K^*(x) > \alpha\} &\leq \frac{1}{\alpha}\sum_k \int_{\mathbb{R}^n \setminus Q_k^*} \sup_{\beta \in A} |K_\beta * b_k(x)|\, dx \\
\leq \frac{BC_n}{\alpha}\sum_k \int_{Q_k} |b_k(x)|\, dx &\leq \frac{BC_n}{\alpha}\int_\Omega |b(x)|\, dx \leq \frac{BC_n}{\alpha}\|f\|_1,
\end{aligned}$$

completing the proof of the theorem. □

5.3 Applications to BMO

In this section we provide a brief introduction to the John–Nirenberg space of functions of bounded mean oscillation in $\mathbb{R}^n$, denoted by $BMO(\mathbb{R}^n)$. This space has played an important role in the development of harmonic analysis and its applications. Our purpose here is to prove some of the most elementary properties of this space in order to illustrate some application of the Calderón–Zygmund decomposition. The reader interested in further applications can consult [To1].

Definition 5.14. Let $f \in L^1_{loc}(\mathbb{R}^n)$. We will say that f is a function of bounded mean oscillation, and denote this space by $BMO(\mathbb{R}^n)$, if there is a constant C such that for any cube Q

$$\frac{1}{|Q|}\int_Q |f(x) - f_Q|\, dx \leq C \tag{5.30}$$

where
$$f_Q = \frac{1}{|Q|}\int_Q f(x)\,dx$$
is the average of the function f on the cube Q. The smallest constant C satisfying (5.30) is the BMO–norm of the function f and we denote it by $\|f\|_*$. In other words,
$$\|f\|_* = \sup_{Q\subset\mathbb{R}^n}\left(\frac{1}{|Q|}\int_Q |f(x)-f_Q|\,dx\right).$$

By identifying functions which differ by a constant we obtain that the normed linear space $(BMO(\mathbb{R}^n), \|\cdot\|_*)$ is a Banach space. If $f\in L^\infty(\mathbb{R}^n)$ then $f\in BMO$ and clearly $\|f\|_*\le 2\|f\|_\infty$. On the other hand, we also have

Exercise 5.3.1.
Prove that the function $f(x)=\log|x|\in BMO(\mathbb{R})\setminus L^\infty(\mathbb{R})$.

Exercise 5.3.2.
Let f and $g\in BMO(\mathbb{R}^n)$. Prove that
$$\max(f,g)=\frac{|f-g|+f+g}{2}\in BMO(\mathbb{R}^n)$$
and
$$\min(f,g)=\frac{f+g-|f-g|}{2}\in BMO(\mathbb{R}^n).$$

Exercise 5.3.3.
Give an example of a function $f\notin BMO(\mathbb{R})$ for which $|f|\in BMO(\mathbb{R})$.

Exercise 5.3.4.
This exercise explains some more the role played by the average of the function f_Q over the cube in the definition of $BMO(\mathbb{R}^n)$.

(i) Prove that $f\in BMO(\mathbb{R}^n)$ if and only if for every cube Q there is a constant C_Q such that
$$\frac{1}{|Q|}\int_Q |f(x)-C_Q|\;dx\le M$$
where M is independent of Q.

(ii) Suppose $f\in BMO(\mathbb{R}^n)$. Let Q_1 and Q_2 be any two cubes in $\mathbb{R}^n$ with $Q_1\subset Q_2$ and $|Q_2|\le C|Q_1|$. Prove that
$$|f_{Q_1}-f_{Q_2}|\le C\|f\|_*.$$

Exercise 5.3.5. (i) Prove that for any $1 \le p < \infty$,

$$\|f\|_* \le \sup_Q \left(\frac{1}{|Q|} \int_Q |f(x) - f_Q|^p \, dx \right)^{1/p} \le A_p \|f\|_*,$$

where $A_p = (Cp\Gamma(p))^{1/p}$ with C independent of p.

(ii) Prove that

$$\sup_Q \left(\frac{1}{|Q|} \int_Q \exp\left(\frac{C\,|f(x) - f_Q|}{\|f\|_*} \right) dx \right) < \infty,$$

where C is a constant independent of f.

(iii) Prove that, as in Exercise 5.3.4, in both (i) and (ii) the mean value of f over the cube, f_Q, can be replaced by any other constant C_Q.

The following is one of the most important results about $BMO(\mathbb{R}^n)$–functions.

Theorem 5.15 (John–Nirenberg)**.** *Let $f \in BMO(\mathbb{R}^n)$. There are constants C_1 and C_2 independent of f such that for all cubes $Q \subset \mathbb{R}^n$ and $\alpha > 0$,*

$$m\{x \in Q \colon |f(x) - f_Q| > \alpha\} \le C_1 |Q| e^{-C_2\alpha/\|f\|_*}. \tag{5.31}$$

Proof. We assume by scaling that $\|f\|_* = 1$. Fix the cube Q and apply the Calderón–Zygmund decomposition for the cube Q, Exercise 5.2.1, to the function $|f - f_Q|$ at level $3/2$ (any level larger than 1 will do here). We obtain a collection of disjoint cubes whose union is Q and which satisfy (i)–(iii) of Exercise 5.2.1. Let us denote this collection by $Q_{1,k}$. For each cube in this collection we have

$$3/2 < \frac{1}{|Q_{1,k}|} \int_{Q_{1,k}} |f(x) - f_Q| \, dx \le 2^n 3/2, \tag{5.32}$$

$$|f(x) - f_Q| \le 3/2 \text{ a.e. } x \in Q \setminus \cup_k Q_{1,k}, \tag{5.33}$$

and

$$\begin{aligned} \sum_k |Q_{1,k}| &\le \frac{2}{3} \sum_k \int_{Q_{1,k}} |f(x) - f_Q| \, dx \\ &\le \frac{2}{3} \int_Q |f(x) - f_Q| \, dx \le \frac{2}{3}|Q|. \end{aligned} \tag{5.34}$$

Also,

$$\begin{aligned}\left|f_{Q_{1,k}} - f_Q\right| &= \left|\frac{1}{|Q_{1,k}|}\int_{Q_{1,k}} f(x)\,dx - f_Q\right| \\ &\le \frac{1}{|Q_{1,k}|}\int_{Q_{1,k}} |f(x) - f_Q|\,dx \le 2^n\frac{3}{2}. \end{aligned} \tag{5.35}$$

Next, we repeat the above argument for each one of the new cubes in the family $\{Q_{1,k}\}$. That is, fix a cube Q_{1,k_0} in the new family and apply the Calderón–Zygmund decomposition for the cube Q_{1,k_0} at the level 3/2 to the function $|f - f_{Q_{1,k_0}}|$. Denote the new resulting family by $\{Q_{2,k}^{k_0}\}$. As before, these new cubes have the property that

$$3/2 < \frac{1}{|Q_{2,k}^{k_0}|}\int_{Q_{2,k}^{k_0}} \left|f(x) - f_{Q_{1,k_0}}\right|\,dx \le 2^n 3/2, \tag{5.36}$$

$$\left|f(x) - f_{Q_{1,k_0}}\right| \le 3/2, \text{ a.e. } x \in Q_{1,k_0} \setminus \cup_k Q_{2,k}^{k_0}, \tag{5.37}$$

and

$$\begin{aligned}\sum_k |Q_{2,k}^{k_0}| &\le \frac{2}{3}\sum_k \int_{Q_{2,k}^{k_0}} \left|f(x) - f_{Q_{1,k_0}}\right|\,dx \\ &\le \frac{2}{3}\int_{Q_{1,k_0}} \left|f(x) - f_{Q_{1,k_0}}\right|\,dx \le \frac{2}{3}|Q_{1,k_0}| \end{aligned} \tag{5.38}$$

Also,

$$\begin{aligned}\left|f_{Q_{2,k}^{k_0}} - f_{Q_{1,k_0}}\right| &= \left|\frac{1}{|Q_{2,k}^{k_0}|}\int_{Q_{2,k}^{k_0}} f(x)\,dx - f_{Q_{1,k_0}}\right| \\ &\le \frac{1}{|Q_{2,k}^{k_0}|}\int_{Q_{2,k}^{k_0}} \left|f(x) - f_{Q_{1,k_0}}\right|\,dx \le 2^n\frac{3}{2}. \end{aligned} \tag{5.39}$$

Thus if we denote by $\{Q_{2,k}\}$ the collection of all the families starting this way from any one of the $Q_{1,k}$ we see, by (5.33)–(5.34) and (5.38), that

$$\sum_k |Q_{2,k}| \le (2/3)^2|Q|$$

and that for $x \notin \cup_k Q_{2,k}$,

$$|f(x) - f_Q| \le \left|f(x) - f_{Q_{1,k}}\right| + \left|f_{Q_{1,k}} - f_Q\right| \le \frac{3}{2} + 2^n \frac{3}{2} \le 2 \cdot 2^n \frac{3}{2}.$$

Continue this way to obtain a collection of disjoint cubes $\{Q_{j,k}\}$ for each $j = 1, 2, \ldots$ such that

$$\sum_k |Q_{j,k}| \le (2/3)^j |Q| = (3/2)^{-j} |Q|$$

and such that for almost every x outside of the union of these cubes we have

$$|f(x) - f_Q| \le (3/2) j \cdot 2^n.$$

Suppose $\alpha \ge 2^n(3/2)$. Pick a $j \ge 1$ such that $(3/2)j \cdot 2^n \le \alpha < (3/2)(j+1) \cdot 2^n$. Then

$$\begin{aligned} m\{x \in Q\colon\ |f(x) - f_Q| > \alpha\} &\le \sum_k |Q_{j,k}| \\ &\le (3/2)^{-j} |Q| = e^{-j \log(3/2)} |Q| \le e^{-C_2 \alpha} |Q|, \end{aligned}$$

with the constant $C_2 = (2/3)\log(3/2)/2^{n+1}$. If $\alpha < 2^n(3/2)$ then $C_2 \alpha < \frac{1}{2}\log(3/2)$ and we use the trivial estimate

$$m\{x \in Q\colon\ |f(x) - f_Q| > \alpha\} \le |Q| \le \sqrt{(3/2)} e^{-C_2 \alpha} |Q|.$$

This proves the theorem with

$$C_2 = \frac{1}{2^{n+1}} (2/3) \log(3/2) \text{ and } C_1 = \sqrt{(3/2)}.$$

$\square$

Theorem 5.16. *Let $f \in L^2(\mathbb{R}^n)$ and fix $1 \le p < \infty$. Then $f \in BMO(\mathbb{R}^n)$ if and only if*

$$\left(\sup_{x_0 \in \mathbb{R}^n, y > 0} \int_{\mathbb{R}^n} |f(x) - P_y * f(x_0)|^p \, P_y(x_0 - x)\, dx \right)^{1/p} = \|f\|_{p,**} < \infty, \quad (5.40)$$

where $P_y(x)$ is the Poisson kernel as defined in Exercise 5.1.11 above.

Proof. Suppose f satisfies (5.40). Fix a cube Q. Let x_Q be the center of Q and $y_Q = \frac{1}{2}\ell_Q$, where ℓ_Q is the length of the cube Q. Since

$$P_y(x) = \frac{C_n y}{(y^2 + |x|^2)^{(n+1)/2}}$$

we easily see that

$$\frac{1}{|Q|}\chi_Q(x) \leq C_n P_{y_Q}(x_Q - x).$$

Thus

$$\begin{aligned} &\frac{1}{|Q|}\int_Q |f(x) - P_y * f(x_0)|^p \, dx \\ &\leq \; C_n \int_{\mathbb{R}^n} |f(x) - P_y * f(x_0)|^p \, P_y(x_0 - x)\, dx, \end{aligned}$$

and this together with Exercise 5.3.5 proves that $f \in BMO(\mathbb{R}^n)$.

Conversely, suppose $f \in BMO(\mathbb{R}^n)$. Fix (x_0, y_0) and let Q_0 be the cube centered at x_0 and with $\ell_{Q_0} = y_0$. Let $Q_1 = 2Q_0$, $Q_2 = 2^2 Q_0, \ldots, Q_k = 2^k Q_0, \ldots$. This gives $\mathbb{R}^n = Q_0 \cup_k Q_k \setminus Q_{k-1}$ and $|Q_k| = 2^{kn} y_0^n$. Once again by our formula for the Poisson kernel $P_y(x)$ we find that

$$P_{y_0}(x_0 - x) \leq \frac{C}{y_0^n}, \text{ for } x \in Q_0 \tag{5.41}$$

and

$$P_{y_0}(x_0 - x) \leq \frac{C}{2^{k(n+1)} y_0^n}, \text{ for } x \in Q_k \setminus Q_{k-1}. \tag{5.42}$$

Since $|Q_k| = 2^n |Q_{k-1}|$, iterating the estimate (ii) of Exercise 5.3.4 we see that

$$|f_{Q_k} - f_{Q_0}| \leq C 2^n k \|f\|_*. \tag{5.43}$$

Using (5.41)-(5.43), exercise 5.3.5 (i), our decomposition of $\mathbb{R}^n$ as $Q_0 \cup_k Q_k \setminus Q_{k-1}$ and the fact that $y_0 = \ell_{Q_0}$ we obtain

$$\begin{aligned} &\int_{\mathbb{R}^n} |f(x) - f_{Q_0}|^p \, P_{y_0}(x_0 - x)\, dx \\ &\leq \; \frac{C}{y_0^n}\int_{Q_0} |f(x) - f_{Q_0}|^p \, dx \end{aligned}$$

$$
\begin{aligned}
&+\sum_{k=1}^{\infty}\frac{C_p}{2^{kn}y_0^n}\int_{Q_k\setminus Q_{k-1}}|f(x)-f_{Q_k}|^p\,dx\\
&+\sum_{k=1}^{\infty}\frac{C_p}{2^{kn}y_0^n}\int_{Q_k\setminus Q_{k-1}}|f_{Q_0}-f_{Q_k}|^p\,dx\\
\leq\quad & C\|f\|_*^p+C_p\|f\|_*^p\sum_{k=1}^{\infty}\frac{1}{2^{kn}}=C_p\|f\|_*^p.
\end{aligned}
$$

We have shown that given (x_0,y_0) there is a constant $C_{(x_0,y_0)}$ such that

$$\int_{\mathbb{R}^n}\left|f(x)-C_{(x_0,y_0)}\right|^p P_{y_0}(x_0-x)\,dx\leq C_p\|f\|_*^p.$$

This and Jensen's inequality immediately show that

$$\left|P_{y_0}*f(x_0)-C_{(x_0,y_0)}\right|^p\leq C_p\|f\|_*^p,$$

which finishes the proof. □

Remark 5.4. The assumption that $f\in L^2(\mathbb{R}^n)$ is just to ensure that the quantity $P_y*f(x)$ is defined for any (x,y), $x\in\mathbb{R}^n$, $y>0$ and can be considerably weakened. The equivalent BMO–norm in (5.31) is closely related to the space of BMO martingales. The reader interested on this fascinating connection can see Durrett [Dur1], Bass [Bas] or Bañuelos and Moore [BaMo].

5.4 Interpolation Theorems

In the course of proving that the maximal function maps $L^p(\mathbb{R}^n)$ into itself for $1<p<\infty$ we proved that if a sublinear operator T maps $L^1(\mathbb{R}^n)$ into weak–$L^1(\mathbb{R}^n)$ and $L^\infty(\mathbb{R}^n)$ into itself, then $T:L^p(\mathbb{R}^n)\to L^p(\mathbb{R}^n)$ for $1<p<\infty$. This is a special case of the following interpolation theorem which we will use in our study of singular integrals in Chapter 6. Recall that a mapping $T:L^p(\mu)\to L^q(\nu)$ is said to be strong–type (p,q), $1\leq p\leq\infty$, $1\leq q\leq\infty$, if there is a constant $A=A_{p,q}$ independent of f such that

$$\|Tf\|_{L^q(\nu)}\leq A\|f\|_{L^p(\mu)}$$

and weak–type (p,q) if there is a constant A_1 independent of f such that

$$\nu\{y\in Y:|Tf(y)|>\alpha\}\leq\left(\frac{A_1}{\alpha}\|f\|_{L^p(\mu)}\right)^q,$$

for $q < \infty$. If $q = \infty$, weak–type (p, q) is the same as strong–type (p, q) since weak–L^∞ is the same as L^∞ by definition. Recall that by $L^1(\mathbb{R}^n) + L^r(\mathbb{R}^n)$ we mean functions f which can be written as $f = h + g$, where $h \in L^1(\mathbb{R}^n)$ and $g \in L^r(\mathbb{R}^n)$.

Theorem 5.17 (The Marcinkiewicz Interpolation Theorem)**.** *Suppose T is a sublinear operator defined on $L^1(\mathbb{R}^n) + L^r(\mathbb{R}^n)$ which is weak–type $(1,1)$ and weak–type (r,r), $1 < r \leq \infty$. Then T is strong–type (p,p), $1 < p < r$. More precisely, suppose that*

(i) $|T(f+g)(x)| \leq |Tf(x)| + |Tg(x)|$,

(ii) $m\{|Tf(x)| > \alpha\} \leq \frac{A_1}{\alpha}\|f\|_1$, $f \in L^1$,

(iii) $m\{|Tf(x)| > \alpha\} \leq \left(\frac{A_r}{\alpha}\|f\|_r\right)^r = \frac{A_r}{\alpha^r}\int_{\mathbb{R}^n} |f|^r\,dx$, $f \in L^r$.

Then for $1 < p < r$,

$$\|Tf\|_p \leq A(p, A_1, A_r)\|f\|_p$$

with

$$A(p, A_1, A_r) = \left(\frac{2A_1 p}{p-1} + \frac{2^r A_r p}{r-p}\right)^{1/p}.$$

Proof. We exclude the case when r is infinity since, as mentioned above, this has already been proved in the course of proving the L^p–boundedness of the maximal function. Let $f \in L^p(\mathbb{R}^n)$, $1 < p < r$, and $\alpha > 0$. Set

$$f_1(x) = \begin{cases} f(x) & \text{if } |f(x)| > \alpha \\ 0 & \text{if } |f(x)| \leq \alpha \end{cases}$$

and

$$f_2(x) = \begin{cases} f(x) & \text{if } |f(x)| \leq \alpha \\ 0 & \text{if } |f(x)| > \alpha. \end{cases}$$

Then $f(x) = f_1(x) + f_2(x)$ with $f_1 \in L^1(\mathbb{R}^n)$ and $f_2 \in L^r(\mathbb{R}^n)$. For the latter, observe that

$$\int_{\mathbb{R}^n} |f_1(x)|\,dx = \int_{\{|f(x)|>\alpha\}} |f_1(x)|^p |f_1(x)|^{1-p}\,dx \leq \alpha^{1-p}\int_{\mathbb{R}^n} |f(x)|^p\,dx$$

and

$$\int_{\mathbb{R}^n} |f_2(x)|^r\,dx = \int_{\{|f(x)|\leq\alpha\}} |f_2(x)|^p |f_2(x)|^{r-p}\,dx \leq \alpha^{r-p}\int_{\mathbb{R}^n} |f(x)|^p\,dx.$$

Since
$$|Tf(x)| \leq |Tf_1(x)| + |Tf_2(x)|$$
we have that
$$\begin{aligned} &m\{x \in \mathbb{R}^n\colon |Tf(x)| > \alpha\} \\ &\quad \leq \quad m\{x \in \mathbb{R}^n\colon |Tf_1(x)| > \alpha/2\} + m\{x \in \mathbb{R}^n\colon |Tf_2(x)| > \alpha/2\} \end{aligned}$$
and integrating we obtain that
$$\int_{\mathbb{R}^n} |Tf(x)|^p\,dx \leq p\int_0^\infty \alpha^{p-1} m\{x \in \mathbb{R}^n\colon |Tf_1(x)| > \alpha/2\}\,d\alpha$$
$$+p\int_0^\infty \alpha^{p-1} m\{x \in \mathbb{R}^n\colon |Tf_2(x)| > \alpha/2\}\,d\alpha = I + II.$$
Let us deal with I first. By the weak–type $(1,1)$ assumption and Fubini's theorem we have
$$\begin{aligned} I &\leq 2A_1 p\int_0^\infty \alpha^{p-2}\int_{\{|f(x)|>\alpha\}} |f(x)|\,dx\,d\alpha \\ &= \quad 2A_1 p\int_{\mathbb{R}^n} |f(x)| \int_0^{|f(x)|} \alpha^{p-2}\,d\alpha\,dx = \frac{2A_1 p}{p-1}\int_{\mathbb{R}^n} |f(x)|^p\,dx. \end{aligned}$$
Similarly
$$\begin{aligned} II &\leq 2^r A_r p\int_0^\infty \alpha^{p-r-1}\int_{\{|f(x)|\leq\alpha\}} |f(x)|^r\,dx\,d\alpha \\ &= \quad 2^r A_r p\int_{\mathbb{R}^n} |f(x)|^r \int_{|f(x)|}^\infty \alpha^{p-r-1}\,d\alpha\,dx \\ &= \quad \frac{2^r A_r p}{r-p}\int_{\mathbb{R}^n} |f(x)|^p\,dx. \end{aligned}$$
Adding I and II we arrive at
$$\|Tf\|_p^p \leq \left(\frac{2A_1 p}{p-1} + \frac{2^r A_r p}{r-p}\right)\|f\|_p^p$$
which proves the Theorem. □

The following more general interpolation theorem has many interesting applications.

Theorem 5.18 (The Riesz–Thorin Interpolation Theorem). *Let T be a linear operator from $L^{p_0}(\mathbb{R}^n) \cap L^{p_1}(\mathbb{R}^n)$ to $L^{q_0}(\mathbb{R}^n) + L^{q_1}(\mathbb{R}^n)$ with $1 \leq p_i,\ q_i \leq \infty$, $i = 0, 1$ and satisfying*

$$\|Tf\|_{q_i} \leq C_i \|f\|_{p_i}.$$

Then T has a bounded extension from $L^{p_t}(\mathbb{R}^n)$ to $L^{q_t}(\mathbb{R}^n)$ satisfying

$$\|Tf\|_{q_t} \leq C_1^t C_0^{1-t} \|f\|_{p_t}$$

for any $0 < t < 1$ where

$$\frac{1}{p_t} = \frac{t}{p_1} + \frac{1-t}{p_0}, \text{ and } \frac{1}{q_t} = \frac{t}{q_1} + \frac{1-t}{q_0}.$$

The proof of this Theorem uses the so called "three lines lemma" from complex analysis. We will use the lemma here without proving it. The interested reader can find its proof in, for example, [SW] p. 180.

Lemma 5.19. *Suppose the function F is an analytic function in the strip $S = \{z \in \mathbb{C} \colon 0 < \Re(z) < 1\}$ which is bounded and continuous on the closure of S satisfying $|F(z)| \leq C_0$ for $\Re(z) = 0$ and $|F(z)| \leq C_1$ for $\Re(z) = 1$. Then $|F(z)| \leq C_0^{1-\Re(z)} C_1^{\Re(z)}$ for any $z \in S$.*

Proof of Theorem 5.18. We may assume $p_t < \infty$ since otherwise the result follows trivially from Hölder's inequality. By density, it is enough to prove the inequality for simple functions s and by duality it is enough to prove that for all simple functions s and $\widetilde{s}$,

$$\left| \int_{\mathbb{R}^n} Ts(x) \widetilde{s}(x)\, dx \right| \leq C_1^t C_0^{1-t} \|s\|_{p_t} \|\widetilde{s}\|_{q_t'} \tag{5.44}$$

where q_t' is the dual exponent of q_t. To prove this we let

$$s(x) = \sum_{j=1}^{k} a_j \chi_{E_j}$$

and

$$\widetilde{s}(x) = \sum_{\ell=1}^{m} b_\ell \chi_{F_\ell}$$

where the $E_j's$ and the $F_\ell's$ are pairwise disjoint. We may also assume that these functions are not zero and in fact that both norms on the right hand side of (5.44) are 1. Now for any $z \in \mathbb{C}$ with $0 \le \Re(z) \le 1$, set

$$\theta_1(z) = \frac{z}{p_1} + \frac{1-z}{p_0} \text{ and } \theta_2(z) = \frac{z}{q_1} + \frac{1-z}{q_0},$$

choose α_j and β_ℓ such that $a_j = e^{i\alpha_j}|a_j|$ and $b_\ell = e^{i\beta_\ell}|b_\ell|$ and define the functions

$$s_z(x) = \sum_{j=1}^{k} |a_j|^{\theta_1(z)/\theta_1(t)} e^{i\alpha_j} \chi_{E_j}(x) \tag{5.45}$$

$$\widetilde{s_z}(x) = \sum_{\ell=1}^{m} |b_\ell|^{\theta_2(z)/\theta_2(t)} e^{i\beta_l} \chi_{F_l}(x). \tag{5.46}$$

If it happens that $q_t = \theta_2(t) = 1$ then we take $\widetilde{s_z} = \widetilde{s}$. By the linearity of T, it follows that the function

$$F(z) = \int_{\mathbb{R}^n} Ts_z(x)\widetilde{s_z}(x)\,dx \tag{5.47}$$

is analytic in $\mathbb{C}$. That is, it is entire. It is also bounded in $0 \le \Re(z) \le 1$. Since $s_t(x) = s(x)$ and $\widetilde{s_t}(x) = \widetilde{s}(x)$, the left hand side of (5.44) is just $F(t)$. The Theorem now follows from Lemma 5.19 and the following exercise. □

Exercise 5.4.1.
Let F be defined by (5.47). Prove that $|F(z)| \le C_0$ for $\Re(z) = 0$ and $|F(z)| \le C_1$ for $\Re(z) = 1$.

Remark 5.5. The Marcinkiewicz and Riesz-Thorin Theorems are, respectively, examples of the real and complex methods of interpolation. This subject is a theory of its own and the interested reader can consult [DC] for much more.

Theorem 5.20. *Let $r \ge 1$ satisfy*

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

for $1 \le p \le q \le \infty$. Define the operator T by

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

where the function $K(x,y)$ satisfies

$$\sup_{x\in\mathbb{R}^n}\left(\int_{\mathbb{R}^n}|K(x,y)|^r\,dy\right)^{1/r}\le M$$

and

$$\sup_{y\in\mathbb{R}^n}\left(\int_{\mathbb{R}^n}|K(x,y)|^r\,dx\right)^{1/r}\le M.$$

Then

$$\|Tf\|_q\le M\|f\|_p.$$

If we take $K(x,y)=g(x-y)$, clearly $M=\|g\|_r$ and we get

Corollary 5.21. *Let $r\ge 1$ satisfy*

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

for $1\le p\le q\le\infty$ and suppose $f\in L^p(\mathbb{R}^n)$ and $g\in L^r(\mathbb{R}^n)$. Then $f*g\in L^q(\mathbb{R}^n)$ and

$$\|f*g\|_q\le\|g\|_r\|f\|_p.$$

Corollary 5.21 is Young's inequality as we stated it in Theorem 4.6 of Chapter 4.

Proof of Theorem 5.20. Let r' be the conjugate exponent of r. By Hölder's inequality,

$$\|Tf\|_\infty\le M\|f\|_{r'}.$$

That is, $T:L^{r'}(\mathbb{R}^n)\to L^\infty(\mathbb{R}^n)$ with operator norm not exceeding M. By duality (integrating against a function in $L^{r'}(\mathbb{R}^n)$) we also have

$$\|Tf\|_r\le M\|f\|_1.$$

Apply Theorem 5.18 with $q_0=\infty$, $q_1=r$, $p_0=r'$ and $p_1=1$, to complete the proof. □

Remark 5.6. The Marcinkiewicz and Riesz–Thorin interpolation Theorems remain true (and with the same proofs) if the space $\mathbb{R}^n$ with the Lebesgue measure is replaced by a more general measure space. For example we have

Theorem 5.22. *Let $(X, \mathcal{F}, \mu)$ and $(Y, \mathcal{G}, \nu)$ be two σ–finite measure spaces and suppose the operator T defined on $L^{r_0}(\mu) + L^{r_1}(\mu)$, $1 \leq r_0 < r_1 \leq \infty$, into ν–measurable functions, is*

(i) sublinear,

(ii) weak–type (r_0, r_0) and

(iii) weak–type (r_1, r_1).

Then T is strong–type (p, p) for any $r_0 < p < r_1$.

An application of this more general interpolation theorem is given in Chapter 7.

Chapter 6

The Fourier Transform

In this chapter we will define the Fourier transform and prove many of its basic properties including the Plancherel theorem. This theorem asserts that the Fourier transform is an isometry in $L^2(\mathbb{R}^n)$. Such a result is essential in the subsequent chapter on singular integrals. Because of the tools we developed in Chapter 4, the Plancherel theorem will follow easily. This material mostly follows [Ru2]. We will also discuss the Fourier transform of regular finite measures in $\mathbb{R}^n$ and its inversion formula. This is a topic not often presented in analysis books but which is very important in applications of the Fourier transform to limit theorems in probability theory. For more on this, see [Dur2].

6.1 The Fourier Transform on $L^1(\mathbb{R}^n)$

Definition 6.1. If $f \in L^1(\mathbb{R}^n)$ we define its Fourier transform by

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{2\pi i x \cdot \xi}\, dx, \tag{6.1}$$

where

$$x \cdot \xi = \sum_{i=1}^{n} x_i \xi_i.$$

Since $f \in L^1(\mathbb{R}^n)$, the function $\hat{f}(\xi)$ is well defined and clearly

$$\|\hat{f}\|_\infty \leq \|f\|_1.$$

Here are some other properties of the Fourier transform which, even though easy, will turn out to be very useful in our applications.

Proposition 6.2. *Let $y \in \mathbb{R}^n$ and $f,\ g \in L^1(\mathbb{R}^n)$.*

(i) If $h(x) = f(x)e^{2\pi i y\cdot x}$, then $\hat{h}(\xi) = \hat{f}(\xi + y)$.

(ii) If $h(x) = f(x-y)$, then $\hat{h}(\xi) = \hat{f}(\xi)e^{2\pi i y\cdot\xi}$.

(iii) If $h(x) = (f * g)(x)$, then $\hat{h}(\xi) = \hat{f}(\xi)\hat{g}(\xi)$.

(iv) If $h_\varepsilon(x) = \frac{1}{\varepsilon^n}h(x/\varepsilon)$, then $\hat{h}_\varepsilon(\xi) = \hat{h}(\varepsilon\xi)$.

Proof. We prove (i) and (iii) leaving (ii) and (iv) as easy exercises. For (i) we have

$$\hat{h}(\xi) = \int_{\mathbb{R}^n} f(x)e^{2\pi i y\cdot x}e^{2\pi i x\cdot\xi}\,dx = \int_{\mathbb{R}^n} f(x)e^{2\pi i x\cdot(\xi+y)}\,dx = \hat{f}(\xi+y).$$

By Theorem 4.2 in Chapter 4, the convolution of two L^1 functions is again in L^1 and hence the Fourier transform of $f * g$ is defined. By Fubini's theorem and the translation invariance of the Lebesgue measure we have

$$\begin{aligned}
\hat{h}(\xi) &= \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi}\left(\int_{\mathbb{R}^n} f(x-y)g(y)\,dy\right)dx \\
&= \int_{\mathbb{R}^n} e^{2\pi i y\cdot\xi}g(y)\left(\int_{\mathbb{R}^n} e^{2\pi i(x-y)\cdot\xi}f(x-y)\,dx\right)dy \\
&= \hat{f}(\xi)\hat{g}(\xi),
\end{aligned}$$

proving (iii). □

Theorem 6.3 (The Riemann–Lebesgue Lemma)**.** *For every function $f \in L^1(\mathbb{R}^n)$, $\hat{f}$ is a continuous function vanishing at ∞.*

Proof. Let $\xi_k \to \xi$. We have

$$\left|\hat{f}(\xi_k) - \hat{f}(\xi)\right| \le \int_{\mathbb{R}^n} |f(x)|\left|e^{2\pi i x\cdot\xi_k} - e^{2\pi i x\cdot\xi}\right|dx.$$

Since the integrand is dominated by $2|f(x)|$, the dominated convergence theorem and the continuity of the exponential function imply the continuity of $\hat{f}$.

Next, since

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

$$\begin{aligned} &= -\int_{\mathbb{R}^n} f(x)e^{2\pi i\xi\cdot(x+\xi/2|\xi|^2)}\,dx \\ &= -\int_{\mathbb{R}^n} f\left(x-\frac{\xi}{2|\xi|^2}\right)e^{2\pi ix\cdot\xi}\,dx, \end{aligned}$$

we have

$$2\hat{f}(\xi) = \int_{\mathbb{R}^n}\left(f(x)-f\left(x-\frac{\xi}{2|\xi|^2}\right)\right)e^{2\pi ix\cdot\xi}\,dx$$

which gives

$$2|\hat{f}(\xi)| \le \left\| f(x)-f\left(x-\frac{\xi}{2|\xi|^2}\right)\right\|_1.$$

By Lemma 4.4 in Chapter 4, this last quantity goes to 0 as $|\xi|\to\infty$. □

Exercise 6.1.1.
Let $f\in L^1(\mathbb{R}^n)$ be nonnegative. Prove that

$$\|\hat{f}\|_\infty = \hat{f}(0) = \|f\|_1.$$

Exercise 6.1.2.
The Fourier Transform of Gaussian is Gaussian.
Let

$$H(x) = \frac{1}{(4\pi)^{n/2}}e^{-|x|^2/4}.$$

Prove that

$$\hat{H}(\xi) = e^{-4\pi^2|\xi|^2}. \tag{6.2}$$

Hint: It is enough to do the case $n=1$. For this, consider the function

$$\phi(\xi) = \frac{1}{(4\pi)^{1/2}}\int_{\mathbb{R}}\cos(2\pi x\xi)e^{-x^2/4}\,dx$$

and show that

$$\phi'(\xi) = -8\pi^2\xi\phi(\xi).$$

Remark 6.1. This exercise also gives that

$$H(x) = \int_{\mathbb{R}^n} e^{-4\pi^2|\xi|^2}e^{-2\pi ix\cdot\xi}\,d\xi = \int_{\mathbb{R}^n}\hat{H}(\xi)e^{-2\pi ix\cdot\xi}\,d\xi, \tag{6.3}$$

which is a special case of the following inversion formula. In fact, this special case will give the general case.

Theorem 6.4 (The Fourier Inversion Formula in L^1)**.** *If $f \in L^1(\mathbb{R}^n)$, $\hat{f} \in L^1(\mathbb{R}^n)$ and*

$$g(x) = \int_{\mathbb{R}^n} \hat{f}(\xi) e^{-2\pi i x\cdot\xi}\, d\xi,$$

then g is continuous, $|g(x)| \to 0$ as $|x| \to \infty$ and $g(x) = f(x)$ a.e.

Proof. We only need to show that $g(x) = f(x)$ almost everywhere. The other parts are proved exactly as Theorem 6.3. For this consider the approximation to the identity given by

$$\varphi_\varepsilon(x) = \frac{1}{(\sqrt{\varepsilon})^n} H(x/\sqrt{\varepsilon}) = \frac{1}{(4\varepsilon\pi)^{n/2}} e^{-|x|^2/4\varepsilon}$$

Notice that the only difference from this and the dilations we presented in Chapter 4 is that here we replaced ε by $\sqrt{\varepsilon}$. By (6.3) we have

$$\varphi_\varepsilon(x) = \int_{\mathbb{R}^n} \hat{H}(\sqrt{\varepsilon}\xi) e^{-2\pi i x\cdot\xi}\, d\xi. \tag{6.4}$$

By Theorem 4.8 in Chapter 4, there exists a sequence $\varepsilon_k \to 0$ so that

$$f * \varphi_{\varepsilon_k}(x) \to f(x) \text{a.e.}$$

On the other hand, by (6.4) and Fubini's theorem we have

$$\begin{aligned} f * \varphi_{\varepsilon_k}(x) &= \int_{\mathbb{R}^n} f(x-y)\varphi_{\varepsilon_k}(y)\, dy \\ &= \int_{\mathbb{R}^n} \hat{H}(\sqrt{\varepsilon_k}\xi) \int_{\mathbb{R}^n} f(x-y) e^{-2\pi i y\cdot\xi}\, dy\, d\xi \\ &= \int_{\mathbb{R}^n} \hat{H}(\sqrt{\varepsilon_k}\xi) \left(\int_{\mathbb{R}^n} f(y) e^{-2\pi i \xi\cdot(x-y)}\, dy \right) d\xi \\ &= \int_{\mathbb{R}^n} \hat{H}(\sqrt{\varepsilon_k}\xi) \hat{f}(\xi) e^{-2\pi i x\cdot\xi}\, d\xi. \end{aligned} \tag{6.5}$$

Since $\hat{H}(\sqrt{\varepsilon_k}\xi) = e^{-4\pi^2\varepsilon_k|\xi|^2}$ and $\hat{f} \in L^1$, we may apply the dominated convergence to finish the proof. $\square$

Corollary 6.5 (Uniqueness)**.** *Suppose $f \in L^1(\mathbb{R}^n)$ and $\hat{f} = 0$. Then $f = 0$ a.e.*

Exercise 6.1.3.
Suppose $f \in L^1(\mathbb{R}^n)$, f is continuous at 0, and $\hat{f}$ is nonnegative. Prove that $\hat{f} \in L^1(\mathbb{R}^n)$ and $\|\hat{f}\|_1 = f(0)$

Remark 6.2. If we define the inverse Fourier transform of $g \in L^1(\mathbb{R}^n)$ by

$$g^\vee(x) = \int_{\mathbb{R}^n} g(\xi) e^{-2\pi i x\cdot\xi}\, d\xi,$$

we can restate Theorem 6.4 as

$$(\hat{f})^\vee = f$$

a.e. for all f with f and $\hat{f}$ both in $L^1(\mathbb{R}^n)$. This important formula will be extended to functions in $L^2(\mathbb{R}^n)$ in the next sections.

We have just seen the important role played above by the Fourier transform of the gaussian. This will be further used later. Another function whose Fourier transform is important to know explicitly is the Poisson kernel already introduced in Exercise 5.1.11, Chapter 5.

Exercise 6.1.4.
Let

$$P(x) = \frac{C_n}{(1+|x|^2)^{n+1/2}}$$

be the Poisson kernel as defined in Exercise 5.1.11 of Chapter 5. Prove that

$$P(x) = \int_{\mathbb{R}^n} e^{-2\pi|\xi|} e^{-2\pi i x\cdot\xi}\, d\xi. \tag{6.6}$$

Thus $\hat{P}(\xi) = e^{-2\pi|\xi|}$.
Hint: It follows from Exercise 3.2.6, Chapter 3, that

$$\int_{\mathbb{R}^n} e^{-2\pi|\xi|} e^{-2\pi i x\cdot\xi}\, d\xi = \int_{\mathbb{R}^n} \left(\frac{1}{\sqrt{\pi}} \int_0^\infty \frac{e^{-t}}{\sqrt{t}} e^{-4\pi^2|\xi|^2/4t}\, dt \right) e^{-2\pi i x\cdot\xi}\, d\xi.$$

Now use Fubini's theorem and the formula for $\hat{H}(\xi)$.

The definition of the Fourier transform and many of its properties can be extended to finite Borel measures on $\mathbb{R}^n$. Let μ be a finite regular Borel measure on $\mathbb{R}^n$. The Fourier Transform of μ is the function

$$\hat{\mu}(\xi) = \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi}\, d\mu(x).$$

If the measure μ is absolutely continuous with respect to the Lebesgue measure and we denote its Radon-Nikodym derivative by f, then $\hat{\mu} = \hat{f}$. Also as in the case of functions,

$$\|\hat{\mu}\|_\infty \le \mu(\mathbb{R}^n)$$

and

$$|\hat{\mu}(\xi_1) - \hat{\mu}(\xi_2)| \leq \int_{\mathbb{R}^n} \left| e^{2\pi i x \xi_1} - e^{2\pi i x \xi_2} \right| \, d\mu(x)$$

which proves that $\hat{\mu}$ is also continuous.

Exercise 6.1.5.
Suppose the measure μ has finite absolute moments of order m. That is, suppose

$$\int_{\mathbb{R}^n} |x|^m \, d\mu(x) < \infty.$$

Prove that $\hat{\mu} \in C^m(\mathbb{R}^n)$ and that

$$\frac{\partial^k \hat{\mu}}{\partial \xi_j^k}(\xi) = \int_{\mathbb{R}^n} (2\pi i x_j)^k e^{2\pi i x \cdot \xi} \, d\mu(x),$$

for any $k \leq m$.

Exercise 6.1.6.
The Schwartz class, denoted by $\mathcal{S}(\mathbb{R}^n)$, is the collection of all $f \in C^\infty(\mathbb{R}^n)$ such that f and all of its derivatives remain bounded when they are multiplied by any polynomial. Prove that if $f \in \mathcal{S}(\mathbb{R}^n)$, then $\hat{f} \in \mathcal{S}(\mathbb{R}^n)$.

We end this section by proving the Fourier inversion formula for measures on $\mathbb{R}$. This result is not really necessary for our subsequent applications but it is very useful in, for example, the study of characteristic functions of random variables. Once again, all the needed tools for this result are in place and thus our proof will follow very easily.

Theorem 6.6 (The Fourier inversion formula for measures on $\mathbb{R}$)**.** *Let μ be a finite regular Borel measure on $\mathbb{R}$ and let $a < b$. Then*

$$\mu(a,b) + \frac{\mu\{a\} + \mu\{b\}}{2} = \lim_{k \to \infty} \frac{1}{2\pi} \int_{-k}^{k} \frac{e^{-2\pi i a \xi} - e^{-2\pi i b \xi}}{i\xi} \hat{\mu}(\xi) \, d\xi \tag{6.7}$$

Proof. Since

$$\left| \frac{e^{2\pi i \xi (x-a)} - e^{2\pi i \xi (x-b)}}{2\pi i \xi} \right| = \frac{1}{2\pi} \left| \int_a^b e^{-2\pi i \xi s} \, ds \right| \leq \frac{1}{2\pi}(b-a) \tag{6.8}$$

we see that

$$\int_{\mathbb{R}} \int_{-k}^{k} \left| \frac{e^{2\pi i \xi (x-a)} - e^{2\pi i \xi (x-b)}}{i\xi} \right| \, d\xi \, d\mu(x) \leq Ck(b-a)$$

and hence by Fubini's theorem

$$\begin{aligned}
&\int_{-k}^{k} \frac{e^{-2\pi a\xi} - e^{-2\pi b\xi}}{i\xi} \hat{\mu}(\xi)\, d\xi \\
&\quad = \int_{\mathbb{R}} \int_{-k}^{k} \frac{e^{-2\pi i\xi a} - e^{-2\pi i\xi b}}{i\xi} e^{2\pi i\xi x}\, d\xi\, d\mu(x) \\
&\quad = \int_{\mathbb{R}} F(k,a,b,x)\, d\mu(x), \qquad (6.9)
\end{aligned}$$

where

$$\begin{aligned}
F(k,a,b,x) &= \int_{-k}^{k} \frac{e^{-2\pi i\xi a} - e^{-2\pi i\xi b}}{i\xi} e^{2\pi i\xi x}\, d\xi \\
&= \frac{1}{i}\int_{-k}^{k} \frac{\cos(2\pi\xi(x-a))}{\xi}\, d\xi + \int_{-k}^{k} \frac{\sin(2\pi\xi(x-a))}{\xi}\, d\xi \\
&\quad -\frac{1}{i}\int_{-k}^{k} \frac{\cos(2\pi\xi(x-b))}{\xi}\, d\xi - \int_{-k}^{k} \frac{\sin(2\pi\xi(x-b))}{\xi}\, d\xi \\
&= 2\int_{0}^{k} \frac{\sin(2\pi\xi(x-a))}{\xi}\, d\xi - 2\int_{0}^{k} \frac{\sin(2\pi\xi(x-b))}{\xi}\, d\xi.
\end{aligned}$$

The last equality follows from the fact that $\sin(t)/t$ is even and $\cos(t)/t$ is odd. By Exercise 3.2.5 ((3.2.8)) in Chapter 3.

$$\lim_{k\to\infty} F(k,a,b,x) = \begin{cases} 0 & \text{if } x < a \\ \pi & \text{if } x = a \\ 2\pi & \text{if } a < x < b \\ \pi & \text{if } x = b \\ 0 & \text{if } x > b \end{cases} \qquad (6.10)$$

Also, by (3.7), Chapter 3,

$$|F(k,a,b,x)| \leq \frac{2}{\pi}\int_0^{\pi} \frac{\sin x}{x}\, dx < \infty.$$

Applying the dominated convergence theorem, (6.8) and (6.10) give the result. □

Exercise 6.1.7.
Let μ be a finite regular Borel measure on $\mathbb{R}^n$. Let $Q = \{x \in \mathbb{R}^n \colon a_j \leq x_j \leq b_j\}$

with $\mu(\partial Q) = 0$. Let

$$h_j(\xi_j) = \frac{e^{-2\pi i a_j \xi_j} - e^{-2\pi i b_j \xi_j}}{i\xi_j}.$$

Then

$$\mu(Q) = \lim_{k\to\infty} \frac{1}{(2\pi)^n} \int_{Q_k} \prod_{j=1}^{n} h_j(\xi_j)\hat{\mu}(\xi)\, d\xi,$$

where $Q_k = \{x \in \mathbb{R}^n \colon -k \le x_j \le k, \text{ for all } j = 1,\dots,n\}$.

We now continue to state several more results for measures on $\mathbb{R}$. Let us assume, without loss of generality, that our measure μ is a probability measure. That is, $\mu(\mathbb{R}) = 1$. The function $F(x) = \mu(-\infty, x]$ is called the distribution function of the measure μ. This function has several interesting properties which are described in the next exercise.

Exercise 6.1.8. (i) F is nondecreasing.

(ii) $\lim_{x\to\infty} F(x) = 1$ and $\lim_{x\to-\infty} F(x) = 0$

(iii) F is right continuous. That is, $\lim_{y\downarrow x} F(y) = F(x)$.

(iv) Set $F(x^-) = \lim_{y\uparrow x} F(y)$. Then

$$\mu(a, b) = F(b^-) - F(a), \text{ and } \mu\{a\} = F(a^-) - F(a).$$

Several interesting properties of the function F are obtained directly from the Fourier transform of μ via the inversion formula. As an illustration we have the following

Corollary 6.7. *Suppose $\hat{\mu} \in L^1(\mathbb{R})$. Then the function F is differentiable and*

$$F'(x) = \int_{\mathbb{R}} \hat{\mu}(\xi)e^{-2\pi i x\xi}\, d\xi. \tag{6.11}$$

Proof. Applying (6.7) and the Fourier inversion formula we see that

$$\mu(a, b) + \frac{\mu\{a\} + \mu\{b\}}{2} \le \frac{1}{2\pi}(b - a) \int_{\mathbb{R}} |\hat{\mu}(\xi)|\, d\xi.$$

Thus the measure μ has no point masses. That is, $\mu\{a\} = 0$ for all a. By Exercise 6.1.8 the function F is continuous and

$$\begin{aligned} \frac{F(x+h)-F(x)}{h} &= \frac{1}{2h\pi}\int_{\mathbb{R}} \frac{e^{-2\pi i x\xi} - e^{-2\pi i(x+h)\xi}}{i\xi}\hat{\mu}(\xi)\,d\xi \\ &= \int_{\mathbb{R}} \left(\frac{1}{h}\int_x^{x+h} e^{-2\pi i y\xi}\,dy\right)\hat{\mu}(\xi)\,d\xi \\ &= \frac{1}{h}\int_x^{x+h}\left(\int_{\mathbb{R}} e^{-2\pi i y\xi}\hat{\mu}(\xi)\,d\xi\right)dy, \end{aligned}$$

where we have used Fubini's Theorem. The identity (6.11) follows from this. □

Remark 6.3. Notice the corollary also gives that f' is continuous. We can rewrite the Corollary 6.7 as the statement that the measure μ is absolutely continuous with respect to the Lebesgue measure with a continuous Radon-Nikodym derivative. The function

$$\varphi(t) = \int_{\mathbb{R}} e^{itx}\,d\mu(x),$$

which in our previous notation is nothing more than $\hat{\mu}(t/(2\pi))$, is called, in probability theory, the *characteristic function* of μ. This function plays a very important role in studying various classical limit theorems (such as the central limit theorem) in probability.

Exercise 6.1.9.
Give an example of a probability measure μ in $\mathbb{R}$ which is absolutely continuous with respect to the Lebesgue measure for which

$$\int_{\mathbb{R}} |\hat{\mu}(\xi)|\,d\xi = \infty$$

Exercise 6.1.10.
Let μ be a probability measure in $\mathbb{R}$. Argue as in the proof of Theorem 6.6 to prove that

$$\lim_{k\to\infty}\int_{-k}^{k} e^{-2\pi i a\xi}\hat{\mu}(\xi)\,d\xi = \mu\{a\}$$

and that

$$\lim_{k\to\infty}\frac{1}{2\pi}\int_{-k}^{k} |\hat{\mu}(\xi)|^2\,d\xi = \sum_a \mu\{a\}^2$$

(Since μ is a probability measure, at most countably many of the $\mu\{a\}$ are not zero).

Exercise 6.1.11.
Let μ be a probability measure in $\mathbb{R}$. Prove that for any $R > 0$,

$$\frac{1}{2R}\int_{-R}^{R}\hat{\mu}(\xi)\,d\xi = \int_{-R}^{R}\frac{\sin(2\pi Rx)}{2\pi Rx}\,d\mu(x) \tag{6.12}$$

Exercise 6.1.12.
Use (6.12) to prove that for all $R > 0$,

$$\left|\frac{1}{R}\int_{-R}^{R}\hat{\mu}(\xi)\,d\xi\right| \le \mu\left[\frac{-1}{\pi R},\frac{1}{\pi R}\right] + 1$$

6.2 The Fourier Transform on $L^2(\mathbb{R}^n)$

Our goal in this section is to understand the Fourier transform for functions on $L^2(\mathbb{R}^n)$. This will take some work since clearly the definition in (6.1) above does not directly apply if the function is not in $L^1(\mathbb{R}^n)$. We will do this by first proving that for functions $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, $\hat{f} \in L^2(\mathbb{R}^n)$ and $\|\hat{f}\|_2 = \|f\|_2$. We will then extend this to any function in $L^2(\mathbb{R}^n)$ using density arguments. This is all contained in the proof of the following fundamental result which also gives the inversion formula for functions in $L^2(\mathbb{R}^n)$.

Theorem 6.8. *For every $f \in L^2(\mathbb{R}^n)$ there exists an $\hat{f} \in L^2(\mathbb{R}^n)$, called the Fourier transform of f, satisfying the following properties:*

(i) ***Plancherel's Theorem:*** $\|\hat{f}\|_2 = \|f\|_2$.

(ii) If we define the functions

$$\mathcal{F}_k(f)(\xi) = \int_{B(0,k)} f(x)e^{2\pi i x\cdot\xi}\,dx = (\widehat{f\chi_{(B(0,k))}})(\xi)$$

and

$$\mathcal{F}_k^{-1}(\hat{f})(x) = \int_{B(0,k)} \hat{f}(\xi)e^{-2\pi i x\cdot\xi}\,d\xi = (\hat{f}\chi_{B(0,k)})^{\vee}(x),$$

then

$$\lim_{k\to\infty}\left\|\mathcal{F}_k(f) - \hat{f}\right\|_2 = 0 \tag{6.13}$$

and

$$\lim_{k\to\infty}\left\|\mathcal{F}_k^{-1}(\hat{f}) - f\right\|_2 = 0. \tag{6.14}$$

(iii) The map $f \to \hat{f}$ preserved the inner product

$$\langle f, \bar{g} \rangle = \int_{\mathbb{R}^n} f \bar{g} \, dx$$

and is an isomorphism of $L^2(\mathbb{R}^n)$ onto $L^2(\mathbb{R}^n)$.

Remark 6.4. By (ii), if $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, $\hat{f}$ agrees with the Fourier transform as defined in (6.1). In addition, (ii) gives the inversion formula for functions in $L^2(\mathbb{R}^n)$.

Proof. We begin by proving that

$$\|f\|_2 = \|\hat{f}\|_2, \text{ if } f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n). \tag{6.15}$$

To prove (6.15), let us define

$$\widetilde{f}(x) = \overline{f}(-x).$$

Then

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

Thus,

$$|g(x)| \leq \|f(x+y)\|_{L^2(dy)} \|\overline{f}(y)\|_{L^2(dy)} = \|f\|_2^2$$

and we conclude that $g \in L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$. In addition,

$$\begin{aligned} |g(x) - g(\overline{x})| &\leq \int_{\mathbb{R}^n} |f(x+y) - f(\overline{x}+y)| \; |\widetilde{f}(y)| \, dy \\ &\leq \; \|f(x+y) - f(\overline{x}+y)\|_{L^2(dy)} \, \|f\|_2. \end{aligned}$$

Hence by Lemma 4.4 in Chapter 4, g is continuous. Thus $g \in L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n) \cap C(\mathbb{R}^n)$. Let H and φ be as in the proof of Theorem 6.4. By (6.5),

$$g * \varphi_\varepsilon(0) = \int_{\mathbb{R}^n} \hat{H}(\sqrt{\varepsilon}\xi)\hat{g}(\xi) \, d\xi. \tag{6.16}$$

However, by (iii) of Proposition 6.2,

$$\hat{g}(\xi) = \hat{f}(\xi)\widehat{\widetilde{f}}(\xi) = \hat{f}(\xi)\overline{\hat{f}}(\xi) = \left|\hat{f}(\xi)\right|^2.$$

The monotone convergence theorem now shows that the right hand side of 6.16 converges to $\int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 \, d\xi$, as $\varepsilon \to 0$. On the other hand, by (ii) of Theorem 4.8 in Chapter 4, $g * \varphi_\varepsilon(0) \to g(0)$ as $\varepsilon \to 0$. Hence

$$g(0) = \|f\|_2^2 = \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 \, d\xi = \|\hat{f}\|_2^2,$$

which proves (6.15).

Now let $f_k = f\chi_{B(0,k)}$. Then $f_k \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and $\|f_k - f\|_{L^2} \to 0$. This implies that $\{f_k\}_{k=1}^\infty$ is a Cauchy sequence in $L^2(\mathbb{R}^n)$. By (6.15) and the linearity of the Fourier transform,

$$\|\hat{f}_k - \hat{f}_{\tilde{k}}\|_2 = \|(f_k - f_{\tilde{k}})^\wedge\|_2 = \|f_k - f_{\tilde{k}}\|_2. \tag{6.17}$$

We see that $\{\hat{f}_k\}_{k=1}^\infty$ is also a Cauchy sequence in $L^2(\mathbb{R}^n)$. Thus $\hat{f}_k$ converges in $L^2(\mathbb{R}^n)$ to a function which we denote by $\hat{f}$. This is the Fourier transform of the function f. Since by (6.15)

$$\|\hat{f}_k\|_2 = \|f_k\|_2,$$

for every k, it follows that

$$\|\hat{f}\|_2 = \|f\|_2. \tag{6.18}$$

This proves (i). Since $\hat{f}_k = \mathcal{F}_k(f)$, we have also proved the first assertion of (ii), (6.13).

We next prove (6.14). Let us set $\mathcal{F}(f) = \hat{f}$. The same argument just given shows that for any $g \in L^2(\mathbb{R}^n)$ there is a function $g^\vee \in L^2(\mathbb{R}^n)$ which satisfies

$$\lim_{k\to\infty} \|\mathcal{F}_k^{-1}(g) - g^\vee\|_2 = 0$$

and with

$$\|g^\vee\|_2 = \|g\|_2.$$

Let us set $\mathcal{F}^{-1}(g) = g^\vee$. We claim that

$$\mathcal{F}^{-1}(\mathcal{F}(f)) = f \text{ and } \mathcal{F}(\mathcal{F}^{-1}(f)) = f, \text{ for all } f \in L^2(\mathbb{R}^n). \tag{6.19}$$

This proves that the map from $L^2(\mathbb{R}^n)$ into $L^2(\mathbb{R}^n)$ given by $f \to \hat{f}$ is actually onto with inverse $f \to f^\vee$. We will prove the first assertion in (6.19), the other being exactly the same. By Theorem 6.4,

$$\mathcal{F}^{-1}(\mathcal{F}(f)) = f \text{ provided that } f \text{ and } \hat{f} \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n). \tag{6.20}$$

We next remove the assumption in (6.20) that $\hat{f} \in L^1(\mathbb{R}^n)$. To do this, let H and φ be as in the proof of Theorem 6.4. If $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, we have $f * \varphi_\varepsilon \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and

$$\widehat{(f * \varphi_\varepsilon)}(\xi) = \hat{f}(\xi) e^{-4\pi^2 \varepsilon |\xi|^2}$$

which is clearly in $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Applying (6.20)) we obtain $\mathcal{F}^{-1}(\mathcal{F}(f * \varphi_\varepsilon)) = f * \varphi_\varepsilon$. Thus

$$\|\mathcal{F}^{-1}(\mathcal{F}(f * \varphi_\varepsilon) - f\|_2 = \|f * \varphi_\varepsilon - f\|_2 \to 0,$$

as $\varepsilon \to 0$. This, together with (i), gives

$$\begin{aligned} &\left\|\mathcal{F}^{-1}(\mathcal{F}(f)) - f\right\|_2 \\ &\quad \leq \left\|\mathcal{F}^{-1}(\mathcal{F}(f - f * \varphi_\varepsilon))\right\| + \left\|\mathcal{F}^{-1}(\mathcal{F}(f * \varphi_\varepsilon)) - f\right\|_2 \\ &\quad = 2\|f - f * \varphi_\varepsilon\|_2 \to 0, \end{aligned}$$

as $\varepsilon \to 0$. We have proved that

$$\mathcal{F}^{-1}(\mathcal{F}(f)) = f \text{ for all } f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n). \tag{6.21}$$

It only remains to remove the assumption that $f \in L^1(\mathbb{R}^n)$. For this, apply (6.21) to the functions f_k defined above. We have

$$\begin{aligned} &\left\|f - \mathcal{F}^{-1}(\mathcal{F}(f))\right\|_2 \\ &\quad \leq \left\|\mathcal{F}^{-1}(\mathcal{F}(f - f_k))\right\|_2 + \left\|f - \mathcal{F}^{-1}(\mathcal{F}(f_k))\right\|_2 \\ &\quad = 2\|f - f_k\|_2 \to 0. \end{aligned}$$

This finally gives

$$\mathcal{F}^{-1}(\mathcal{F}(f)) = f, \text{ for all } f \in L^2(\mathbb{R}^n).$$

The fact that the map $f \to \hat{f}$ preserves the inner product in $L^2(\mathbb{R}^n)$ follows from (i) and the polarization identity

$$4f\overline{g} = |f + g|^2 - |f - g|^2 + i|f + ig|^2 - i|f - ig|^2.$$

□

Remark 6.5. As the reader has observed, the key to extending our definition of the Fourier transform to $L^2(\mathbb{R}^n)$ was to prove that for all $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, $\|\hat{f}\|_2 = \|f\|_2$. The following Theorem provides the necessary inequality to extend the definition of the Fourier transform to any p between 1 and 2.

Theorem 6.9 (Hausdorff-Young)**.** *Suppose* $1 \leq p \leq 2$ *and* $1/p + 1/q = 1$. *Let* $f \in L^1(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$. *Then* $\hat{f} \in L^q(\mathbb{R}^n)$ *and*

$$\|\hat{f}\|_q \leq \|f\|_p, \tag{6.22}$$

Proof. Since we have already proved the case of $p = 1$ and $p = 2$, (6.22) follows directly from the Riesz–Thorin interpolation theorem (Theorem 5.18, Chapter 4). More precisely, define the operator $Tf = \hat{f}$. We know that

$$\|Tf\|_\infty \leq \|f\|_1$$

and

$$\|Tf\|_2 = \|f\|_2.$$

Let $q_0 = \infty$, $p_0 = 1$, and $q_1 = p_1 = 2$ and apply the Riesz–Thorin Theorem to conclude. □

With (6.22) proved, the density argument in the proof of Theorem 6.8 allows us to extend the definition of the Fourier transform to $L^p(\mathbb{R}^n)$ for this range of p.

6.3 Applications

We shall now give, in the form of several exercises, some applications of Plancherel's Theorem.

Exercise 6.3.1.
Prove the following "semigroup" properties of the Poisson and heat kernels:

(i) $P_{y_1} * P_{y_2}(x) = P_{y_1+y_2}(x)$, for all $y_1, y_2 > 0$.

(ii) $H_{t_1} * H_{t_2}(x) = H_{t_1+t_2}(x)$, for all $t_1, t_2 > 0$.

Next, recall that the Laplacian operator Δ in $\mathbb{R}^n$ defined for any $f \in C^2(\mathbb{R}^n)$ is given by

$$\Delta f(x) = \sum_{k=1}^{n} \frac{\partial^2 f}{\partial x_k^2}.$$

Since
$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x)e^{2\pi i x\cdot\xi}\,dx,$$
integration by parts shows that
$$\widehat{\frac{\partial f}{\partial x_k}}(\xi) = -2\pi i \xi_k \hat{f}(\xi)$$
and
$$\widehat{\Delta f}(\xi) = -4\pi^2|\xi|^2\hat{f}(\xi) \tag{6.23}$$
for all $f \in C_0^2(\mathbb{R}^n)$.

Exercise 6.3.2.
For $f \in L^2(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$ define
$$u_f(x,y) = \int_{\mathbb{R}^n} \hat{f}(\xi)e^{-2\pi y|\xi|}e^{-2\pi i x\cdot\xi}\,d\xi$$
and
$$v_f(x,t) = \int_{\mathbb{R}^n} \hat{f}(\xi)e^{-4\pi^2 t|\xi|^2}e^{-2\pi i x\cdot\xi}\,d\xi.$$
Prove that $u_f(x,y) = P_y * f(x)$ and $v_f(x,t) = H_t * f(x)$ and hence by Exercise 5.1.13 in Chapter 5, u_f is harmonic in $\mathbb{R}^{n+1}_+$ and v_f satisfies the heat equation in $\mathbb{R}^{n+1}_+$ both with boundary values f. Notice, however, that these properties can be verified very easily with the above formulas by differentiating under the integral sign.

Exercise 6.3.3.
Let $f \in C_0^1(\mathbb{R}^n)$. Prove

(i) $\int_{\mathbb{R}^n} |\nabla f(x)|^2\,dx = 4\pi^2 \int_{\mathbb{R}^n} |\xi|^2|\hat{f}(\xi)|^2\,d\xi$.

(ii) For all $y \in \mathbb{R}^n$,
$$\int_{\mathbb{R}^n} |f(x+y)+f(x-y)-2f(x)|^2\,dx = 4\int_{\mathbb{R}^n} |\hat{f}(\xi)|^2|1-\cos(2\pi\xi\cdot y)|^2\,d\xi.$$

Exercise 6.3.4.
Let $f \in L^2(\mathbb{R}^n)$ and set $u_f(x,y) = P_y * f(x)$ and $v_f(x,t) = H_t * f(x)$. Set
$$\nabla_x u_f = \left(\frac{\partial u_f}{\partial x_1}, \frac{\partial u_f}{\partial x_2}, \ldots, \frac{\partial u_f}{\partial x_n}\right).$$

Define the Littlewood-Paley functions

$$\begin{aligned}
g_v(f)(x) &= \left(\int_0^\infty y\left|\frac{\partial u_f}{\partial y}(x,y)\right|^2\,dy\right)^{1/2},\\
g_h(f)(x) &= \left(\int_0^\infty y\left|\nabla_x u_f(x,y)\right|^2\,dy\right)^{1/2},\\
g(f)(x) &= \left(\int_0^\infty y\left|\nabla u_f(x,y)\right|^2\,dy\right)^{1/2},\\
G_v(f)(x) &= \left(\int_0^\infty t\left|\frac{\partial v_f}{\partial t}(x,t)\right|^2\,dt\right)^{1/2},\\
G_h(f)(x) &= \left(\int_0^\infty \left|\nabla_x v_f(x,t)\right|^2\,dt\right)^{1/2}.
\end{aligned}$$

Use Plancherel's theorem to prove that

$$\|g_v(f)\|_2 = \|g_h(f)\|_2 = \|G_v(f)\|_2 = \frac{1}{2}\|f\|_2.$$

and that

$$\|G_h(f)\|_2 = \frac{1}{\sqrt{2}}\|f\|_2.$$

Exercise 6.3.5.
Let $f, h \in L^2(\mathbb{R}^n)$. Prove that

$$\int_0^\infty\int_{\mathbb{R}^n} y\frac{\partial u_f}{\partial y}(x,y)\overline{\frac{\partial u_h}{\partial y}}(x,y)\,dx\,dy = \frac{1}{4}\int_{\mathbb{R}^n} f(x)\overline{h(x)}dx$$

and

$$\int_0^\infty\int_{\mathbb{R}^n} y\nabla_x u_f(x,y)\cdot\overline{\nabla_x u_h}(x,y)\,dx\,dy = \frac{1}{4}\int_{\mathbb{R}^n} f(x)\overline{h(x)}\,dx$$

In Chapter 10 we will derive L^P–versions of the above inequalities. For now, let us observe that

$$|\nabla u_f|^2 = \frac{1}{2}\Delta u_f^2,$$

and hence it follows from this exercise that

$$\int_0^\infty\int_{\mathbb{R}^n} y\Delta u_f^2(x,y)\,dx\,dy = \int_{\mathbb{R}^n} f(x)^2\,dx,$$

for all $f \in L^2(\mathbb{R}^n)$. There is a more general version of this result which is proved using Green's second identity. Let us first recall that identity. Let D be a smooth bounded domain in $\mathbb{R}^n$ and let $u, v \in C^2(D) \cap C^1(\overline{D})$. Then

$$\int_D (u\Delta v - v\Delta u)\, dx = \int_{\partial D} \left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n} \right) d\sigma. \tag{6.24}$$

Exercise 6.3.6.
Let $B(0,R)$ be the ball in $\mathbb{R}^{n+1}$ centered at 0 and radius R. Apply (6.24) to $D = B(0,R) \cap \mathbb{R}^{n+1}_+$, where $\mathbb{R}^{n+1}_+$ is the upper–half space of $\mathbb{R}^n$, to conclude that

$$\int_0^\infty \int_{\mathbb{R}^n} y\Delta u_f^p(x,y)\, dx\, dy = \int_{\mathbb{R}^n} f(x)^p\, dx, \tag{6.25}$$

for all nonnegative $f \in C_0^\infty(\mathbb{R}^n)$.

Exercise 6.3.7. (i) Prove that the function

$$F(\xi) = \int_{S^{n-1}} e^{2\pi i r\eta\cdot\xi}\, d\sigma(\eta) \tag{6.26}$$

is radial. That is, $F(\xi_1) = F(\xi_2)$ for all ξ_1, ξ_2 with $|\xi_1| = |\xi_2|$.

(ii) Use (i) to prove that the Fourier transform of a radial function is radial.

Remark 6.6. The function F can in fact be computed explicitly using the argument of Exercise 3.2.9, Chapter 3, and one finds that

$$F(\xi) = 2\pi\, (r|\xi|)^{-(n-2)/2}\, J_{(n-2)/2}(2\pi r|\xi|)$$

where $J_{(n-2)/2}$ is a Bessel function; see Stein and Weiss [SW]. In particular we have that the Fourier transform of a radial function f is given by

$$\hat{f}(\xi) = 2\pi\, (|\xi|)^{-(n-2)/2} \int_0^\infty f(r) r^{n/2} J_{(n-2)/2}(2\pi r|\xi|)\, dr. \tag{6.27}$$

Exercise 6.3.8.
Let φ be a $C_0^\infty(\mathbb{R}^n)$ radial function. By exercise 6.3.7, its Fourier transform is also radial. Suppose

$$\int_0^\infty |\hat{\varphi}'(y)|^2\, y\, dy = A < \infty.$$

Define

$$U(x,y) = (\varphi_y * f)(x)$$

and consider the generalized Littlewood-Paley function

$$g(f)(x) = \left(\int_0^\infty y \left| \frac{\partial U}{\partial y}(x,y) \right|^2 dy \right)^{1/2}.$$

Prove that

$$\|g(f)\|_2 = \sqrt{A}\|f\|_2.$$

Exercise 6.3.9.
Let f and u_f be as in Exercise 6.3.4. Define

$$T(f)(x) = \left(\int_0^\infty \int_0^\infty \frac{|u_f(x,y_1) - u_f(x,y_2)|^2}{|y_1 - y_2|^2} \, dy_1 \, dy_2 \right)^{1/2}.$$

Prove that

$$\|T(f)\|_2 = C\|f\|_2$$

for some constant C.

Exercise 6.3.10.
Let f be a function on the real line $\mathbb{R}$ such that both f and xf are in $L^2(\mathbb{R})$. Prove that $f \in L^1(\mathbb{R})$ and that

$$\left(\int_{\mathbb{R}} |f(x)| \, dx \right)^2 \leq 8 \left(\int_{\mathbb{R}} |f(x)|^2 \, dx \right)^{1/2} \left(\int_{\mathbb{R}} |x|^2 |f(x)|^2 \, dx \right)^{1/2}.$$

Chapter 7

Singular Integrals

Singular integrals arise naturally in many settings and play a fundamental role in the application of harmonic analysis to various areas of mathematics such as boundary value problems and regularity of solutions to PDE's. The basic example which arises from conjugate harmonic functions in the upper half-space $\mathbb{R}^2_+$ is the Hilbert transform defined by

$$Hf(x) = \lim_{\varepsilon \to 0} \frac{1}{\pi} \int_{\{|y| > \varepsilon\}} \frac{f(x-y)}{y}\, dy. \tag{7.1}$$

We will discuss the connection to conjugate harmonic functions in section 8.3 of Chapter 8. In this chapter we will address the questions of existence and properties of such "singular integrals". For example, what do we really mean by (7.1)? If we make sense of (7.1), what are the boundedness properties of the operator $f \to Hf$? Even for relatively nice functions, such as continuous functions of compact support, the existence of these type of principal–value integrals are by no means obvious. This chapter presents some of the most basic results in this large and very sophisticated subject which is often referred to as the Calderón–Zygmund Theory. In section 7.1 we prove the convergence of singular integral in $L^p(\mathbb{R}^n)$ and in section 7.2 we prove their almost everywhere convergence. Many other results, such as computation of their Fourier multipliers, will be obtained along the way. In section 7.4 we derive some vector valued versions of these results similar results for the maximal function.

7.1 $L^p(\mathbb{R}^n)$–Convergence

We begin with a lemma whose importance will become clear shortly.

Lemma 7.1 (The Marcinkiewicz Integral)**.** *Let F be a closed set in $\mathbb{R}^n$ with $m(F^c) < \infty$. Let $\delta_F(x)$ denote the distance from the point x to the set F. That is,*

$$\delta_F(x) = \inf\{|x-y| : y \in F\}.$$

Set

$$I_F(x) = \int_{\mathbb{R}^n} \frac{\delta_F(y)dy}{|x-y|^{n+1}}. \tag{7.2}$$

Then,

$$\int_F I_F(x)\,dx \le Cm(F^c).$$

Proof. By Fubini's Theorem,

$$\begin{aligned}\int_F I_F(x)\,dx &= \int_F \int_{\mathbb{R}^n} \frac{\delta_F(y)}{|x-y|^{n+1}}\,dy\,dx \\ &= \int_F \int_{F^c} \frac{\delta_F(y)}{|x-y|^{n+1}}\,dy\,dx = \int_{F^c} \left(\int_F \frac{dx}{|x-y|^{n+1}} \right) \delta_F(y)\,dy.\end{aligned}$$

Since $x \in F$, implies $|x-y| \ge \delta(y)$, we have by integrating in polar coordinates that

$$\int_F \frac{dx}{|x-y|^{n+1}} \le \int_{\{|x| \ge \delta_F(y)\}} \frac{dx}{|x|^{n+1}} = C(\delta_F(y))^{-1}.$$

Thus

$$\int_{F^c} \int_F \frac{dx}{|x-y|^{n+1}} \delta_F(y)\,dy \le C \int_{F^c} (\delta_F(y))^{-1} \delta_F(y)\,dy = Cm(F^c),$$

which proves the lemma. □

Theorem 7.2. *Let $K \in L^2(\mathbb{R}^n)$. Suppose that*

$$|\hat{K}(\xi)| \le B \text{ for all } \xi \in \mathbb{R}^n \tag{7.3}$$

and that $K \in C^1(\mathbb{R}^n \backslash \{0\})$ with

$$|\nabla K(x)| \le \frac{B}{|x|^{n+1}} \text{ for all } x \in \mathbb{R}^n \backslash \{0\} \tag{7.4}$$

where B is a constant. For $f \in L^1(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$ define

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

There is a constant $A_{p,B,n}$ depending only on p, n, and B, such that

$$\|Tf\|_p \leq A_{p,B,n}\|f\|_p \quad 1 < p < \infty. \tag{7.5}$$

By continuity the operator T has an extension to all of $L^p(\mathbb{R}^n)$ satisfying (7.5) with the same $A_{p,B,n}$. Furthermore,

$$A_{p,B,n} \leq C_{n,B}(p-1), \text{ for } 2 \leq p < \infty,$$

and

$$A_{p,B,n} \leq C_{n,B}\frac{1}{p-1}, \text{ for } 1 < p \leq 2,$$

where the constant $C_{n,B}$ is independent of p.

Remark 7.1. It is very important to observe here that the constant $A_{p,B,n}$ does not depend on the $L^2(\mathbb{R}^n)$–norm of the function K. This assumption is only used to initially define Tf and to apply Plancherel's theorem. The fact that the constant $A_{p,B,n}$ does not depend on the $L^2(\mathbb{R}^n)$–norm of K is what makes this a nontrivial result. We will often refer to the function K as the kernel of the singular integral operator T. This terminology is very common in the literature.

Proof. The basic tools for the proof of this Theorem come from Chapter 5. Key to the proof are the Calderón–Zygmund decomposition and the Marcinkiewicz interpolation theorem. Indeed, by (7.3) and the Plancherel theorem (Chapter 6, (6.13)) we have

$$\int_{\mathbb{R}^n} |Tf(x)|^2\,dx = \int_{\mathbb{R}^n} |\hat{K}(\xi)|^2|\hat{f}(\xi)|^2\,d\xi \leq B^2\|f\|_2^2. \tag{7.6}$$

Thus T is strong–type (2, 2). We shall now prove that T is also weak–type $(1, 1)$. That is, we will prove there is a constant A_1 depending only on n and B such that

$$m\{x \in \mathbb{R}^n \colon |Tf(x)| > \alpha\} \leq \frac{A_1}{\alpha}\|f\|_1, \text{ for all } f \in L^1(\mathbb{R}^n). \tag{7.7}$$

To prove (7.7), apply the Calderón–Zygmund decomposition in the form of Corollary 5.8 of Chapter 5 and write $f = g + b$. Then

$$Tf = Tg + Tb$$

and we have

$$m\{x \in \mathbb{R}^n \colon |Tf(x)| > \alpha\}$$

$$\begin{aligned} &\leq \quad m\{x \in \mathbb{R}^n \colon |Tg(x)| > \alpha/2\} \\ &+ \quad m\{x \in \mathbb{R}^n \colon |Tb(x)| > \alpha/2\}. \end{aligned} \tag{7.8}$$

Now, by Chebychev's inequality, the $L^2(\mathbb{R}^n)$–boundedness of T just proved and (5.19) of Corollary 5.8, Chapter 5, we have

$$\begin{aligned} m\{x \in \mathbb{R}^n \colon |Tg(x)| > \alpha/2\} &\leq \frac{4}{\alpha^2} \int_{\mathbb{R}^n} |Tg(x)|^2 \, dx \\ \leq \quad \frac{4B^2}{\alpha^2} \int_{\mathbb{R}^n} |g(x)|^2 \, dx &\leq \frac{C_n B^2}{\alpha^2} \alpha \|f\|_1 = \frac{C_n B^2}{\alpha} \|f\|_1. \end{aligned}$$

Thus to show weak–type $(1,1)$, it is enough to show that

$$m\{x \in \mathbb{R}^n \colon |Tb(x)| > \alpha/2\} \leq \frac{C}{\alpha} \|f\|_1. \tag{7.9}$$

Since $\mathbb{R}^n = \Omega \cup F$,

$$m\{x \in \mathbb{R}^n \colon |Tb(x)| > \alpha/2\} \leq m\{\Omega\} + m\{x \in F \colon |Tb(x)| > \alpha/2\}$$

and

$$m\{\Omega\} \leq \frac{C}{\alpha} \|f\|_1.$$

Thus (7.9) reduces to proving that

$$m\{x \in F \colon |Tb(x)| > \alpha/2\} \leq \frac{C}{\alpha} \|f\|_1. \tag{7.10}$$

Recall that $b(x) = 0$ on F and write

$$b(x) = \sum_{k=1}^{\infty} \chi_{Q_k}(x) b(x) = \sum_{k=1}^{\infty} b_k(x).$$

Fix $x \in F$. Since

$$\int_{Q_k} b(y) \, dy = 0$$

we have

$$Tb_k(x) = \int_{Q_k} \left[K(x-y) - K(x-y^k) \right] b_k(y) \, dy,$$

where y^k is the center of the cube Q_k. By the mean value property,

$$|K(x-y)-K(x-y^k)| = |y-y^k|\left|\nabla K(x-\overline{y}^k)\right| \le \frac{B\,\mathrm{diam}(Q_k)}{|x-\overline{y}^k|^{n+1}}$$

where $\overline{y}^k$ is a variable point on the line segment from y and y^k. By the proof of Corollary 5.10, Chapter 5, $\mathrm{diam}(Q_k) \approx \mathrm{dist}(Q_k, F)$. Therefore for any $y_1, y_2 \in Q_k$,

$$|x-y_1| = |x-y_2+y_2-y_1| \le |x-y_2| + |y_2-y_1|$$

and

$$|y_2-y_1| \le \mathrm{diam}(Q_k) \le C\,\mathrm{dist}(y_2, F) \le C|x-y_2|.$$

Thus for a fixed point $x \in F$, the numbers $\{|x-y| \colon y \in Q_k\}$ are all comparable. This gives

$$|Tb_k(x)| \le C\frac{\mathrm{diam}(Q_k)}{|x-\overline{y}^k|^{n+1}}\int_{Q_k}|b(y)|\,dy \le C\frac{\mathrm{diam}(Q_k)}{|x-\overline{y}^k|^{n+1}}\alpha|Q_k| \tag{7.11}$$

where the last inequality follows from (5.21) of Chapter 5. If we set $\delta_F(y) = \mathrm{dist}(y, F)$ as above we obtain from (7.11) that

$$|Tb_k(x)| \le C\alpha\frac{\mathrm{diam}(Q_k)|Q_k|}{|x-\overline{y}^k|^{n+1}} \le C\alpha\int_{Q_k}\frac{\delta_F(y)}{|x-y|^{n+1}}\,dy.$$

Since the series $\sum_k b_k$ and $\sum_k Tb_k$ converges in $L^2(\mathbb{R}^n)$ to b and Tb, respectively, we have that $|Tb(x)| \le \sum_k |Tb_k(x)|$ for almost every $x \in \mathbb{R}^n$. Summing over k this gives for $x \in F$,

$$|Tb(x)| \le C\alpha\int_{\Omega}\frac{\delta_F(y)}{|x-y|^{n+1}}\,dy = C\alpha\int_{\mathbb{R}^n}\frac{\delta_F(y)}{|x-y|^{n+1}}\,dy = C\alpha I_F(x).$$

By Chebychev's inequality and lemma 7.9,

$$\begin{aligned} m\{x\in F \colon |Tb(x)| > \alpha/2\} &\le \frac{2}{\alpha}\int_F |Tb(x)|\,dx \\ &\le \; C\int_{\mathbb{R}^n} I_F(x)\,dx \le Cm(\Omega) \le \frac{C}{\alpha}\|f\|_1, \end{aligned}$$

proving (7.9) and hence (7.7). By the interpolation Theorem (Chapter 5, Theorem 5.17)

$$\|Tf\|_p \le A_p\|f\|_p \quad 1<p<2. \tag{7.12}$$

For $2 < p < \infty$, we will use duality. Recall (Chapter 2, Theorem 2.17) that

$$\|g\|_p = \sup_{\substack{\psi \in L^q \\ \|\psi\|_q = 1}} \left| \int_{\mathbb{R}^n} g\psi \, dx \right| \tag{7.13}$$

where $1 < q < 2$ and $1/p + 1/q = 1$. If we set $\widetilde{K}(x) = K(-x)$ we see that $\widetilde{K}$ satisfies the hypothesis of the Theorem. If we denote the T corresponding to this $\widetilde{K}$ by $\widetilde{T}$ we obtain by what we have already proved that

$$\begin{aligned} \left| \int_{\mathbb{R}^n} Tf \cdot \psi \, dx \right| &= \left| \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} K(x-y) f(y) \, dy \psi(x) \, dx \right| \\ &= \left| \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} K(x-y) \psi(x) \, dx f(y) \, dy \right| \\ &= \left| \int_{\mathbb{R}^n} \widetilde{T}\psi(y) f(y) \, dy \right| \le A(q) \|\psi\|_q \|f\|_p, \end{aligned}$$

which proves the case $2 < p < \infty$. The fact that the constant $A_{p,B,n}$ has the behavior stated in the theorem follows from the form of the constant $A(p, A_1, A_r)$ of Theorem 5.17 in Chapter 5 after we observe that our L^2 constant is just B and our weak–type constant depends only on B and and n. □

Remark 7.2. The supremum in (7.13) can be restricted to functions in $C_0^\infty(\mathbb{R}^n)$ by the density of the latter in $L^p(\mathbb{R}^n)$ and this makes, as it is easy to see, the use of Fubini's theorem above valid.

Our goal now is to remove some of the very restrictive assumptions on the kernel K so that we can eventually allow the example of $K(x) = \frac{1}{\pi x}$. The following properties of the above cubes will be used in several places below. Notice that (i) is exactly the geometric fact used in the proof of Lemma 5.13 of Chapter 5.

Exercise 7.1.1.
Fix $\varepsilon > 0$ and for Q_k as in the proof of Theorem 7.2 set $Q_k^* = 2\sqrt{n}\, Q_k$. Prove that

(i) If $x \notin Q_k^*$, then $|x - y^k| \ge 2|y - y^k|$ for any $y \in Q_k$.

(ii) If $x \notin Q_k^*$ and for some $y_0 \in Q_k$, $|x - y_0| = \varepsilon$, then there is a constant C_1 depending only on n such that $Q_k \subset \overline{B(x, C_1\varepsilon)}$.

(iii) If $x \notin Q_k^*$ and for some $y_0 \in Q_k$, $|x - y_0| = \varepsilon$, then there is a constant C_2 depending only on n such that $|x - y| \ge C_2\varepsilon$ for all $y \in Q_k$.

We now begin to remove some of the more restrictive assumptions on the function K.

Definition 7.3. The function K is said to satisfy the Hörmander condition (H) if there is a constant B such that

$$\sup_{|y|>0} \int_{|x|\geq 2|y|} |K(x-y) - K(x)|\, dx \leq B. \tag{7.14}$$

The following proposition is proved exactly as Lemma 5.13 in Chapter 5.

Proposition 7.4. *Let K be a function satisfying the condition H. Let Q be a cube centered at $x_0 \in \mathbb{R}^n$ and let $a \in L^1(\mathbb{R}^n)$ be supported in Q with mean value zero over the cube. That is*

$$\int_Q a(x)\, dx = 0.$$

Then

$$\int_{\mathbb{R}^n \setminus Q^*} |K * a(x)|\, dx \leq B \int_Q |a(x)|\, dx,$$

where as above $Q^* = 2\sqrt{n}Q$.

Proposition 7.5. *Suppose K satisfies*

$$|\nabla K(x)| \leq \frac{B}{|x|^{n+1}}, \quad x \in \mathbb{R}^n \setminus \{0\}$$

for some B. Then K satisfies (H) with a B' depending on B and n.

Proof. By the mean value theorem,

$$|K(x-y) - K(x)| \leq |y|\, |\nabla K(\xi)| \leq \frac{B|y|}{|\xi|^{n+1}}$$

with

$$\xi = tx + (1-t)(x-y) = tx - tx + (t-1)y + x = (t-1)y + x$$

for some $t \in (0,1)$. This gives, for $2|y| \leq |x|$ and $t_0 = 1-t$,

$$|x| \leq |x - t_0 y| + |t_0 y| \leq |\xi| + |y| \leq |\xi| + \frac{1}{2}|x|$$

or $|x| \le 2|\xi|$. Hence using polar coordinates we obtain

$$\begin{aligned}\int_{|x|\ge 2|y|} |K(x-y)-K(x)|\,dx &\le C_n B \int_{|x|\ge 2|y|} \frac{|y|}{|x|^{n+1}}\,dx \\ &= C_n B|y| \int_{2|y|}^{\infty} \frac{r^{n-1}}{r^{n+1}}\,dr = C_n B|y| \int_{2|y|}^{\infty} \frac{dr}{r^2} = C_n B|y|/|y| = C_n B,\end{aligned}$$

which proves the Proposition. □

Theorem 7.6. *Suppose $K \in L^2(\mathbb{R}^n)$, $|\hat{K}(\xi)| \le B$ and K has property (H) with constant B. Then the conclusions of Theorem 7.2 hold.*

Proof. As before, our assumption on the Fourier transform and Plancherel's theorem imply that $T\colon L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ with $\|Tf\|_2 \le B\|f\|_2$. Continuing as in the proof of Theorem 7.2, we find that

$$m\{x \in \mathbb{R}^n \colon |Tf(x)| > \alpha\} \le \frac{C}{\alpha}\|f\|_1 + m\left\{x \in \mathbb{R}^n \colon |Tb(x)| > \frac{\alpha}{2}\right\}. \tag{7.15}$$

Let $Q_k^* = 2\sqrt{n}Q_k$. These new cubes are not disjoint but we do have the equality $|Q_k^*| = (2\sqrt{n})^n|Q|$ and if we set $\Omega^* = \cup Q_j^*$ we have $m(\Omega^*) \le (2\sqrt{n})^n m(\Omega)$. Let $F^* = (\Omega^*)^c$. We obtain

$$\begin{aligned} & m\{x \in \mathbb{R}^n \colon |Tb(x)| > \frac{\alpha}{2}\} \\ \le\ & m\{\Omega^*\} + m\{x \in F^* \colon |Tb(x)| > \frac{\alpha}{2}\} \\ \le\ & (2\sqrt{n})^n m(\Omega) + m\{x \in F^* \colon |Tb(x)| > \frac{\alpha}{2}\} \\ \le\ & \frac{C}{\alpha}\|f\|_1 + m\{x \in F^* \colon |Tb(x)| > \frac{\alpha}{2}\}. \end{aligned} \tag{7.16}$$

Applying Proposition 7.4 with $a = b_k$ we have that

$$\int_{\mathbb{R}^n \setminus Q_k^*} |Tb_k(x)|\,dx \le B \int_{Q_k} |b_k(x)|\,dx.$$

and summing over k gives

$$\int_{F^*} |Tb(x)|\,dx \le B \sum_k \int_{Q_k} |b(y)|\,dy \le c\|f\|_1.$$

Thus,

$$m\{x \in F^* \colon |Tb(x)| > \alpha/2\} \le \frac{C}{\alpha}\|f\|_1.$$

Putting this together with (7.15) and (7.16) proves the Theorem. □

By Proposition 7.5, the kernel $K(x) = \frac{1}{\pi x}$ satisfies the Hörmander condition but it clearly is not in $L^2(\mathbb{R})$ and its Fourier transform is not defined. Thus the above two results do not yet apply to this example. The next theorem corrects this difficulty.

Theorem 7.7. *Suppose the function K satisfies the Hörmander condition (H) with constant B and*

(i) $|K(x)| \leq \dfrac{B}{|x|^n}$,

(ii) $\displaystyle\int_{R_1<|x|<R_2} K(x)\,dx = 0$, *for all* $0 < R_1 < R_2 < \infty$.

Let $\varepsilon > 0$ and define

$$T_\varepsilon f(x) = \int_{|y|\geq\varepsilon} K(y)f(x-y)\,dy, \tag{7.17}$$

for $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$. Then

$$\|T_\varepsilon f\|_p \leq A_{p,B,n}\|f\|_p, \tag{7.18}$$

$A_{p,B,n}$ depends only on p, B, n and in particular is independent of ε. Furthermore,

$$A_{p,B,n} \leq C_{n,B}(p-1), \text{ for } 2 \leq p < \infty,$$

and

$$A_{p,B,n} \leq C_{n,B}\frac{1}{p-1}, \text{ for } 1 < p \leq 2,$$

where the constant $C_{n,B}$ is independent of p.

Corollary 7.8. *Assume K satisfies the hypothesis of Theorem 7.7 and T_ε is defined as in (7.17). Then*

$$\lim_{\varepsilon\to 0} T_\varepsilon f = T(f) \tag{7.19}$$

in $L^p(\mathbb{R}^n)$. The operator defined by (7.19) satisfies

$$\|Tf\|_p \leq A_{p,B,n}\|f\|_p, \quad 1 < p < \infty, \tag{7.20}$$

with the same constant as in (7.18).

Proof. Suppose we can show that $\lim_{\varepsilon\to 0} T_\varepsilon f_1 = Tf_1$ in $L^p(\mathbb{R}^n)$ for $f_1 \in C_0^\infty(\mathbb{R}^n)$. Let $\eta > 0$ and pick an $f_1 \in C_0^\infty(\mathbb{R}^n)$ and an $\varepsilon_0 = \varepsilon(\eta) > 0$ such that

$$\|f - f_1\|_p < \frac{\eta}{4A_{p,B,n}}$$

and

$$\|T_\varepsilon f_1 - Tf_1\|_p < \eta/4,$$

for all $\varepsilon < \varepsilon_0$. Then

$$\begin{aligned}\|T_\varepsilon f - Tf_1\|_p &\le \|T_\varepsilon f - T_\varepsilon f_1\|_p + \|T_\varepsilon f_1 - Tf_1\|_p \\ &\le A_{p,B,n}\|f - f_1\|_p + \eta/2 \le \eta/4 + \eta/4 = \eta/2,\end{aligned}$$

for all $\varepsilon < \varepsilon_0$. This shows that for all $\varepsilon_1, \varepsilon_2$ smaller than ε_0 we have

$$\|T_{\varepsilon_1} f - T_{\varepsilon_2} f\|_p < \eta$$

Since η is arbitrary we have proved the existence of the $\lim_{\varepsilon\to 0} T_\varepsilon f$ in $L^p(\mathbb{R}^n)$ provided such a limit exists for functions in $C_0^\infty(\mathbb{R}^n)$. To prove the existence of the limit under the assumption that $f_1 \in C_0^\infty(\mathbb{R}^n)$, assume $\varepsilon < 1$ and write

$$T_\varepsilon f_1(x) = \int_{\varepsilon\le|y|\le 1} K(y) f_1(x-y)\,dy + \int_{1\le|y|} K(y) f_1(x-y)\,dy. \tag{7.21}$$

For $|y| \ge 1$,

$$|K_\varepsilon(y)| = |K(y)| \le \frac{B}{|y|^n}$$

and we see that

$$\int_{|y|\ge 1} |K(y)|^p\,dy \le B^p \int_1^\infty \frac{r^{n-1}}{r^{np}}\,dr < \infty,$$

since $1 < np - n + 1$. This and the fact that $f \in L^1(\mathbb{R}^n)$ shows that the second integral in (7.21) represents an $L^p(\mathbb{R}^n)$ function.

By the compactness of the support of f_1, the first integral is supported in a ball (independent of ε) and using our assumptions on K and the differentiability of f_1, we have

$$\left|\int_{\varepsilon\le|y|\le 1} K(y) f_1(x-y)\,dy\right|$$

$$
\begin{aligned}
&= \left| \int_{\varepsilon \le |y| \le 1} K(y) \left[f_1(x-y) - f_1(x) \right] \, dy \right| \\
&\le \int_{\varepsilon \le |y| \le 1} |K(y)| \, |f_1(x-y) - f_1(x)| \, dy \\
&\le C \int_{\varepsilon \le |y| \le 1} |K(y)| \, |y| \, dy \le C_n B \int_{\varepsilon}^{1} \frac{r^n dr}{r^n} \le C_n B
\end{aligned}
$$

So, the first integral in (7.21) converges uniformly to a function of compact support and this proves the Corollary. □

The proof of Theorem 7.7 follows immediately from Theorem 7.6 and the following lemma.

Lemma 7.9. *Suppose K satisfies the hypothesis of Theorem 7.7 with constant B. Let*

$$
K_\varepsilon(x) = \begin{cases} K(x) & \text{if } |x| > \varepsilon \\ 0 & \text{if } |x| \le \varepsilon. \end{cases}
$$

Then $K_\varepsilon \in L^2(\mathbb{R}^n)$, $\sup_{\xi \in \mathbb{R}} |\hat{K}_\varepsilon(\xi)| \le C_1 B$, and K_ε satisfies (H) with constant $C_2 B$. The constants C_1 and C_2 are independent of ε.

Proof. By our assumptions on K we trivially have

$$
|K_\varepsilon(x)| \le \frac{B}{|x|^n}.
$$

Using polar coordinates, it follows that $K_\varepsilon \in L^2(\mathbb{R}^n)$. Also we trivially have

$$
\int_{\{R_1 \le |x| \le R_2\}} K_\varepsilon(x) \, dx = 0, \quad 0 < R_1 < R_2 < \infty.
$$

For the condition (H) we have

$$
\begin{aligned}
&\int_{|x| \ge 2|y|} |K_\varepsilon(x-y) - K_\varepsilon(x)| \, dx \\
&= \int_{\{|x| \ge 2|y|, |x-y| < \varepsilon\}} |K_\varepsilon(x-y) - K_\varepsilon(x)| \, dx \\
&+ \int_{\{|x| \ge 2|y|, |x-y| > \varepsilon\}} |K_\varepsilon(x-y) - K_\varepsilon(x)| \, dx \\
&= I + II.
\end{aligned}
\tag{7.22}
$$

By our definition of K_ε,

$$\begin{aligned}
I &= \int_{\{|x|\geq 2|y|, |x-y|<\varepsilon\}} |K_\varepsilon(x)|\,dx \\
&\leq \int_{|x|<2\varepsilon} |K_\varepsilon(x)|\,dx = \int_{\varepsilon<|x|<2\varepsilon} |K(x)|\,dx \\
&\leq B\int_{\varepsilon<|x|<2\varepsilon} \frac{dx}{|x|^n} \leq C_n B.
\end{aligned}$$

In the same way

$$\begin{aligned}
II &= \int_{\{|x|\geq 2|y|, |x-y|>\varepsilon\}} |K(x-y) - K_\varepsilon(x)|\,dx \\
&\leq \int_{\{|x|\geq\varepsilon, |x|>2|y|, |x-y|>\varepsilon\}} |K(x-y) - K(x)|\,dx \\
&\quad + \int_{\{\varepsilon<|x-y|<3/2\varepsilon\}} |K(x-y)|\,dx.
\end{aligned} \tag{7.23}$$

The first integral is dominated by B using the assumption (H). The second quantity in (7.23) is dominated by $C_n B$ by the argument of I. This proves then that K_ε also satisfies the hypothesis (H).

It remains to prove

$$\sup_{\xi\in\mathbb{R}^n} |\hat{K}_\varepsilon(\xi)| \leq C_n B \tag{7.24}$$

with C_n independent of ε. To show (7.24) we first prove the case $\varepsilon = 1$ and then obtain the general case by a scaling argument. That is, we first prove that

$$\sup_{\xi\in\mathbb{R}^n} \left|\hat{K}_1(\xi)\right| \leq C_n B \tag{7.25}$$

For this we start by writing

$$\hat{K}_1(\xi) = \int_{|x|\leq 1/|\xi|} e^{2\pi i x\cdot\xi} K_1(x)\,dx + \int_{1/|\xi|\leq|x|} e^{2\pi i x\cdot\xi} K_1(x)\,dx = I + II.$$

We can think of II as

$$\lim_{R\to\infty} \int_{1/|\xi|\leq|x|\leq R} e^{2\pi i x\cdot\xi} K_1(x)\,dx$$

which exists by the way we constructed the Fourier transform of an $L^2(\mathbb{R}^n)$ function. By our assumption (ii) we have

$$\begin{aligned} I = \int_{\{1\le|x|\le 1/|\xi|\}} & e^{2\pi i x\cdot\xi} K_1(x)\,dx \\ & - \int_{\{1<|x|\le 1/|\xi|\}} K_1(x)\,dx \\ = & \int_{\{|x|\le 1/|\xi|\}} \left[e^{2\pi i x\cdot\xi} - 1\right] K_1(x)\,dx. \end{aligned}$$

Thus

$$|I| \le \int_{\{|x|\le 1/|\xi|\}} \left|e^{2\pi i x\cdot\xi} - 1\right| |K_1(x)|\,dx.$$

Since

$$\begin{aligned} \left|e^{i2\pi x\cdot\xi} - 1\right| &= \sqrt{(\cos 2\pi(x\cdot\xi) - 1)^2 + \sin^2 2\pi(x\cdot\xi)} \\ &= \sqrt{2 - 2\cos(2\pi(x\cdot\xi))} = C|\sin\pi(x\cdot\xi)| \le C|x||\xi|. \end{aligned}$$

From this and our assumption (i) we have

$$\begin{aligned} |I| \le C|\xi| & \int_{\{|x|\le 1/|\xi|\}} |x|\,|K_1(x)|\,dx \\ \le & \;\; BC|\xi| \int_{\{|x|\le 1/|\xi|\}} \frac{|x|}{|x|^n}\,dx \\ = & \;\; C_n B|\xi| \int_0^{1/|\xi|} dr = C_n B. \end{aligned}$$

Next, with $z = \xi/(2|\xi|^2)$, it follows (as in the proof of the Riemann-Lebesgue Lemma, Theorem 6.3, Chapter 6) that

$$\begin{aligned} II = & \int_{\{1/|\xi|\le|x|\}} e^{2\pi i x\cdot\xi} K_1(x)\,dx \\ = & \;\; \frac{1}{2}\int_{\{1/|\xi|\le|x|\}} K_1(x) e^{2\pi i x\cdot\xi}\,dx \qquad (7.26) \\ & - \frac{1}{2}\int_{\{1/|\xi|\le|x-z|\}} K_1(x-z) e^{2\pi i x\cdot\xi}\,dx \\ = & \;\; \frac{1}{2}\int_{\{1/|\xi|\le|x|\}} \left(K_1(x) - K_1(x-z)\right) e^{2\pi i x\cdot\xi}\,dx + III \end{aligned}$$

where
$$|III| \le \int_{A_1 \cup A_2} \left| K_1(x-z) e^{2\pi i x\cdot\xi} \right| \, dx \tag{7.27}$$
with
$$A_1 = \left\{ \frac{1}{|\xi|} \le |x|, |x-z| \le \frac{1}{|\xi|} \right\},$$
and
$$A_2 = \left\{ \frac{1}{|\xi|} \le |x-z|, |x| \le \frac{1}{|\xi|} \right\}.$$
On the set A_1,
$$\frac{1}{2|\xi|} \le |x-z| \le \frac{1}{|\xi|}$$
and on A_2,
$$\frac{1}{|\xi|} \le |x-z| \le \frac{3}{2|\xi|}.$$
Thus the right hand side of (7.27) is dominated by
$$\begin{aligned} &\int\limits_{\{\frac{1}{2|\xi|} \le |x-z| \le \frac{1}{|\xi|}\}} |K_1(x-z)| \, dx + \int\limits_{\{\frac{1}{||\xi|} \le ||x-z| \le \frac{3}{2|\xi|}\}} |K_1(x-z)| \, dx \\ \le \;& C_n B \int\limits_{\{\frac{1}{2|\xi|} \le |x| \le \frac{1}{|\xi|}\}} \frac{dx}{|x|^n} + C_n B \int\limits_{\{\frac{1}{|\xi|} \le |x| \le \frac{3}{2|\xi|}\}} \frac{dx}{|x|^n} = C_n B. \end{aligned}$$

By the property (H) which is satisfied by K_1, the integral on the last line of the right hand side of (7.26) is dominated by
$$\begin{aligned} &\frac{1}{2} \int_{\{|x| \ge 1/|\xi|\}} |K_1(x-z) - K_1(x))| \, dx \\ =\;& \frac{1}{2} \int_{\{|x| \ge 2|z|\}} |K_1(x-z) - K_1(z)| \, dx \le CB. \end{aligned}$$
This proves (7.25).

Let us now consider the function $\widetilde{K}(x) = \varepsilon^n K(\varepsilon x)$. It follows by simple change of variables from what we just proved that $\widetilde{K}$ also satisfies the hypothesis (H) and the assumptions (i) and (ii) of Theorem 7.7. Thus if we define
$$\widetilde{K}_1(x) = \begin{cases} \widetilde{K}(x) & \text{if } |x| > 1 \\ 0 & \text{if } |x| \le 1 \end{cases}$$

then
$$\sup_{\xi\in\mathbb{R}^n}\left|\widehat{\widetilde{K}_1}(\xi)\right| \le C_n B.$$
or
$$\sup_{\varepsilon\xi\in\mathbb{R}^n}\left|\widehat{\widetilde{K}_1}(\varepsilon\xi)\right| \le C_n B. \tag{7.28}$$
However, since $\left(\varepsilon^{-n}\widetilde{K}_1(\varepsilon^{-1}\cdot)\right)^{\wedge}(\xi) = \widehat{\widetilde{K}_1}(\varepsilon\xi)$, (7.28) is equivalent to
$$\sup_{\varepsilon\xi}\left|\left(\varepsilon^{-n}\widetilde{K}_1(\varepsilon^{-1}\cdot)\right)^{\wedge}(\xi)\right| \le C_n B. \tag{7.29}$$
But we have
$$\varepsilon^{-n}\widetilde{K}_1(\varepsilon^{-1}x) = \begin{cases} \varepsilon^{-n}\widetilde{K}(\varepsilon^{-1}x) & \text{if } |\varepsilon^{-1}x| > 1 \\ 0 & \text{if } |\varepsilon^{-1}x| \le 1 \end{cases} = K_\varepsilon(x)$$
which proves the lemma. □

Remark 7.3. If we define
$$K_{\varepsilon,N}(x) = \begin{cases} K(x) & \text{if } \varepsilon < |x| < N \\ 0 & \text{otherwise,} \end{cases}$$
then this integrable function satisfies the Hörmander condition and it has bounded Fourier transform, both independent of ε and N. It follows that the operator $K_{\varepsilon,N}f(x) = K_{\varepsilon,N} * f(x)$ satisfies
$$\|K_{\varepsilon,N} * f\|_p \le A\|f\|_p$$
with A depending only on p, B, and n but not on ε and N.

Exercise 7.1.2.
Let $f(x) = \chi_{[a,b]} \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$. Prove that
$$Hf(x) = \frac{1}{\pi}\log\frac{|x-a|}{|x-b|}$$
and conclude that $Hf \notin L^1(\mathbb{R}) \cup L^\infty(\mathbb{R})$.

Exercise 7.1.3. (i) Let T be as in Theorem 7.7, and suppose f is supported in the ball B with the property that $|f|\log(1+|f|)$ is integrable over B. Prove that Tf is integrable over B.

(ii) Suppose f is bounded and supported on the ball B. Prove, using the bound for the constant A_p of Theorem 7.7, that there is a constant C such that $e^{C|Tf|}$ is integrable over B.

7.2 Almost Everywhere Convergence

In this section we restrict our attention to singular integrals which commute with dilation and obtain several properties for these including their almost everywhere existence. More precisely, define the dilation map by $\tau_\varepsilon f(x) = f(\varepsilon x)$. We seek singular integrals T for which $\tau_{\varepsilon^{-1}} T \tau_\varepsilon = T$. This assumption, by (4.16) of Chapter 4, is equivalent to the requirement that $\varepsilon^{-n} K(\varepsilon^{-1} x) = K(x)$. In other words, the commutativity with dilations is equivalent to the kernel K being of the form

$$K(x) = \frac{\Omega(x)}{|x|^n}. \tag{7.30}$$

where Ω is homogeneous of degree zero. That is, $\Omega(rx) = \Omega(x)$ for all $r > 0$ or equivalently, Ω is constant on rays emanating from the origin.

With this notation, assumptions (i) and (ii) of Theorem 7.7 are equivalent to

(i') Ω is bounded on S^{n-1}, and

(ii') $\int_{S^{n-1}} \Omega(\xi)\, d\sigma(\xi) = 0$.

The condition (H) is not as clear in terms of Ω but we do have a sufficient condition which guarantees that K satisfies (H). First,

Definition 7.10. The function Ω on S^{n-1} is said to be Dini continuous if

$$\omega(\delta) = \sup_{\substack{|x'-x|<\delta \\ |x|=|x'|=1}} |\Omega(x) - \Omega(x')|\,,$$

then

$$\int_0^1 \frac{\omega(\delta)}{\delta}\, d\delta < \infty \tag{7.31}$$

Notice that the kernel K defining the Hilbert transform can be written as

$$K(x) = \frac{\Omega(x)}{|x|}$$

where

$$\Omega(x) = \frac{1}{\pi}\frac{x}{|x|} = \frac{1}{\pi}\,\mathrm{sign}(x), \tag{7.32}$$

which clearly satisfies the Dini condition. Other examples of functions satisfying (7.31) are given by functions which are sufficiently smooth. Also, any function which is Hölder continuous of exponent α, for any $0 < \alpha \le 1$, satisfies the Dini condition.

Theorem 7.11. *Suppose Ω is homogeneous of degree zero, satisfies (i'), (ii'), and is Dini continuous. Define*

$$T_\varepsilon f(x) = \int_{|y|\geq\varepsilon} \frac{\Omega(y)}{|y|^n} f(x-y)\,dy.$$

Then for any $1<p<\infty$, there is a constant A_p independent of ε such that

$$\|T_\varepsilon f(x)\|_p \leq A_p \|f\|_p. \tag{7.33}$$

Furthermore,

$$Tf = \lim_{\varepsilon\to 0} T_\varepsilon f$$

exists in L^p and satisfies

$$\|Tf(x)\|_p \leq A_p\|f\|_p \tag{7.34}$$

with the same constant A_p as in (7.33). As before,

$$A_p \leq C(p-1), \text{ for } 2\leq p<\infty,$$

and

$$A_p \leq C\frac{1}{p-1}, \text{ for } 1<p\leq 2,$$

where the constant C is independent of p.

Proof. By Theorem 7.7 and Corollary 7.8, it is enough to show that

$$K(x) = \frac{\Omega(x)}{|x|^n}$$

satisfies (i), (ii) of Theorem 7.7 and the Hörmander condition (H). Since Ω satisfies (i') and (ii'), it only remains to prove (H). To see this write

$$\begin{aligned} K(x-y)-K(x) &= \frac{\Omega(x-y)}{|x-y|^n} - \frac{\Omega(x)}{|x|^n} \\ &= \frac{\Omega(x-y)-\Omega(x)}{|x-y|^n} + \Omega(x)\left[\frac{1}{|x-y|^n} - \frac{1}{|x|^n}\right]. \end{aligned}$$

Using polar coordinates (or repeating the argument of Proportion 7.5), we find that

$$\int_{|x|\geq 2|y|} \left|\frac{1}{|x-y|^n} - \frac{1}{|x|^n}\right| dx \leq C_n. \tag{7.35}$$

Since Ω is bounded, we have reduced to proving that

$$\int_{|x|\geq 2|y|} \frac{|\Omega(x-y)-\Omega(x)|}{|x-y|^n}\,dx \leq C_n^1. \tag{7.36}$$

By a simple geometric argument, we have

$$\left|\frac{x-y}{|x-y|}-\frac{x}{|x|}\right| \leq C\left|\frac{y}{x}\right|, \tag{7.37}$$

for $|x| \geq 2|y|$. Thus

$$|\Omega(x-y)-\Omega(x)| = \left|\Omega\left(\frac{x-y}{|x-y|}\right)-\Omega\left(\frac{x}{|x|}\right)\right| \leq \omega\left(C_n^2\frac{|y|}{|x|}\right)$$

and we have

$$\int_{|x|\geq 2|y|} \frac{|\Omega(x-y)-\Omega(x)|}{|x-y|^n}\,dx \leq C_n \int_{|x|\geq 2|y|} \frac{\omega\left(C_n^2|y|/|x|\right)}{|x-y|^n}\,dx. \tag{7.38}$$

However, $|x| \leq |x-y|+|y| \leq |x-y|+|x|/2$ and hence

$$\frac{1}{|x-y|^n} \leq \frac{C_n}{|x|^n}.$$

From this we see that the right hand side of (7.38) is dominated by

$$\begin{aligned}
&\leq C_n \int_{|x|\geq 2|y|} \omega\left(C_n^2\frac{|y|}{|x|}\right)\frac{dx}{|x|^n}\\
&\leq C_n \int_{2|y|}^{\infty}\int_{S^{n-1}} \frac{1}{r}\omega\left(C_n^2|y|/r\right)\,d\sigma\,dr\\
&= C_n \int_0^{C_n^2} \frac{\omega(\delta)}{\delta}\,d\delta < \infty,
\end{aligned}$$

which proves that K satisfies the condition (H) and hence the theorem. $\square$

Exercise 7.2.1.
Let $\varphi \in C_0^\infty(\mathbb{R}^n)$ be nonnegative, radial, decreasing and with $\int_{\mathbb{R}^n}\varphi(x)\,dx = 1$. If we set as before

$$K_\varepsilon(x) = \begin{cases} \Omega(x)/|x|^n & \text{if } |x| \geq \varepsilon \\ 0 & \text{if } |x| < \varepsilon, \end{cases}$$

it follows from the proof of Corollary 7.8 that $\varphi * K(x) = \lim_{\varepsilon\to 0} \varphi * K_\varepsilon(x)$ exists for every $x \in \mathbb{R}^n$. Define

$$\Phi(x) = \varphi * K(x) - K_1(x).$$

Prove that the least decreasing radial majorant of Φ is integrable.

Theorem 7.12. *Let Ω be as in Theorem 7.11. For $f \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$ define*

$$T_\varepsilon f(x) = \int_{|y|\ge\varepsilon} \frac{\Omega(y)}{|y|^n} f(x-y)\, dy$$

and the maximal operator

$$T^* f(x) = \sup_{\varepsilon>0} |T_\varepsilon f(x)|.$$

Then

(i) For any $f \in L^1(\mathbb{R}^n)$ and all $\alpha > 0$,

$$m\{x \in \mathbb{R}^n : T^* f(x) > \alpha\} \le \frac{A_1}{\alpha}\|f\|_1. \tag{7.39}$$

(ii) For any $1 < p < \infty$,

$$\|T^* f\|_p \le A_p \|f\|_p. \tag{7.40}$$

(iii) For any $f \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$,

$$\lim_{\varepsilon\to 0} T_\varepsilon f(x) = Tf(x) \text{ a.e.} \tag{7.41}$$

Proof. We will first prove (7.40). Let Tf be the L^p–limit of $T_\varepsilon f$ guarantees by Theorem 7.11 and let $M(Tf)(x)$ be its maximal function. We claim that

$$T^* f(x) \le M(Tf)(x) + CMf(x) \tag{7.42}$$

It follows from this that

$$\begin{aligned} \|T^* f\|_p &\le \|M(Tf)\|_p + C\|M_f\|_p \\ &\le A_p\|Tf\|_p + CA_p\|f\|_p \\ &\le A_pA'_p\|f\|_p + CA_p\|f\|_p = A_p\|f\|_p. \end{aligned}$$

To prove (7.42) let Φ be as in Exercise 7.2.1 and notice that

$$\Phi_\varepsilon(x) = \frac{1}{\varepsilon^n}\Phi(x/\varepsilon) = \varphi_\varepsilon * K - K_\varepsilon(x).$$

Thus,

$$T_\varepsilon f(x) = -\Phi_\varepsilon * f(x) + (\varphi_\varepsilon * K) * f(x).$$

Now, by Theorem 5.4 of Chapter 5 and exercise 7.2.1,

$$|\Phi_\varepsilon * f(x)| \leq AMf(x).$$

Suppose we can show that

$$(\varphi_\varepsilon * K) * f(x) = T(f) * \varphi_\varepsilon(x). \tag{7.43}$$

Then we will also have

$$|(\varphi_\varepsilon * K) * f(x)| \leq M(Tf)(x)$$

which proves (7.42). (7.43) follows from

$$(\varphi_\varepsilon * K_\delta) * f(x) = T_\delta(f) * \varphi_\varepsilon(x), \tag{7.44}$$

after taking the limit as $\delta \to 0$. But

$$\varphi_\varepsilon * K_\delta(x) = \int_{\mathbb{R}^n} K_\delta(x-y)\varphi_\varepsilon(y)\,dy$$

and an application of Fubini's theorem gives

$$\begin{aligned}
&(\varphi_\varepsilon * K_\delta(x)) * f(x)\\
&\quad = \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} K_\delta(\zeta - y)\varphi_\varepsilon(y)\,dy\right) f(x-\zeta)\,d\zeta\\
&\quad = \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} K_\delta(\zeta - y) f(x-\zeta)\,d\zeta\right)\varphi_\varepsilon(y)\,dy\\
&\quad = \int_{\mathbb{R}^n} (T_\delta f)(x-y)\varphi_\varepsilon(y)\,dy,
\end{aligned}$$

which proves (7.44) and hence (7.43).

We will now use (7.40) and the Calderón–Zygmund decomposition to prove (7.39). Let $f \in L^1(\mathbb{R}^n)$, let $f = g+b$ be the Calderón–Zygmund decomposition as given in Corollary 5.8 of Chapter 5. Then $|T_\varepsilon f(x)| \le |T_\varepsilon g| + |T_\varepsilon b|$ and $T^*f(x) \le T^*g + T^*b$. Proceeding as in the proof of Theorem 7.6 and using this time the L^2–boundedness of T^* we obtain,

$$\begin{aligned} & m\{x \in \mathbb{R}^n : T^*f(x) > \alpha\} \\ & \quad \le \; m\{x \in \mathbb{R}^n : T^*g(x) > \alpha/2\}| + m\{x \in \mathbb{R}^n : T^*b(x) > \alpha/2\} \\ & \quad \le \; \frac{C}{\alpha}\|f\|_1 + m\{x \in \mathbb{R}^n : T^*b(x) > \alpha/2\}. \end{aligned}$$

With Q_k^*, Ω^* and F^* as in the proof of Theorem 7.6, we have exactly as before that

$$\begin{aligned} & m\{x \in \mathbb{R}^n \colon T^*b > \alpha/2\} \\ & \quad \le \; m\{\Omega^*\} + m\{x \in F^* \colon T^*b > \alpha/2\} \\ & \quad \le \; \frac{c}{\alpha}\|f\|_1 + m\{x \in F^* \colon T^*b > \alpha/2\}. \end{aligned}$$

Thus (7.39) follows immediately if we can prove that for all $x \in F^*$,

$$\begin{aligned} \sup_{\varepsilon>0} |T_\varepsilon b(x)| \le \sum_{Q_j} \int_{Q_j} |K(x-y) - K(x-y^j)||b(y)|\,dy \\ + C \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |b(y)|\,dy = Sb(x) + CMb(x) \end{aligned} \tag{7.45}$$

where $Mb(x)$ is the maximal function of b and $Sb(x)$ is defined by the sum. Indeed, as in the proof of Lemma 5.13, that

$$\int_{F^*} Sb(x)\,dx \le C \int_{\mathbb{R}^n} |f(x)|\,dx.$$

This together with the weak $(1,1)$ property of the maximal function proves the result.

To prove (7.45), fix $x \in F^*$ and $\varepsilon > 0$. Let Q_j be any cube in our family. We have the following three cases:

Case (a) For all $y \in Q_k$, $|x-y| < \varepsilon$.

Case (b) For all $y \in Q_k$, $|x-y| > \varepsilon$.

Case (c) There exists a $y \in Q_k$ such that $|x-y| = \varepsilon$.

With this we have

$$\begin{aligned}
T_\varepsilon b(x) &= \sum_{Q_k} \int_{Q_k} K_\varepsilon(x-y) b(y)\, dy \\
&= \sum_{Q_k \in \text{ case(a)}} \int_{Q_k} K_\varepsilon(x-y) b(y)\, dy \\
&\quad + \sum_{Q_k \in \text{ case(b)}} \int_{Q_k} K_\varepsilon(x-y) b(y)\, dy \\
&\quad + \sum_{Q_k \in \text{ case(c)}} \int_{Q_k} K_\varepsilon(x-y) b(y)\, dy \\
&= I(x) + II(x) + III(x).
\end{aligned}$$

Since $K_\varepsilon(x-y) = 0$ if $|x-y| < \varepsilon$, $I(x) = 0$. For case (b), we have $K_\varepsilon(x-y) = K(x-y)$ by our definition of K_ε and therefore

$$\begin{aligned}
&\left| \int_{Q_k} K(x-y) b(y)\, dy \right| \\
&\quad = \left| \int_{Q_k} K(x-y) - K(x-y^k) b(y)\, dy \right| \\
&\quad \le \int_{Q_k} |K(x-y) - K(x-y^k)|\, |b(y)|\, dy.
\end{aligned}$$

Thus

$$|II(x)| \le Sb(x). \tag{7.46}$$

For case (c), we have by Exercise 7.1.1 that

$$\begin{aligned}
\left| \int_{Q_k} K_\varepsilon(x-y) b(y)\, dy \right| &\le \int_{Q_k} |K_\varepsilon(x-y)|\, |b(y)|\, dy \\
&= \int_{B(x, C_n\varepsilon) \cap Q_k} |K_\varepsilon(x-y)|\, |b(y)|\, dy \\
&= \int_{B(x, C_n\varepsilon) \cap Q_k} \frac{|\Omega(x-y)|}{|x-y|^n} |b(y)|\, dy \\
&\le C_n B \int_{B(x, C_n\varepsilon) \cap Q_k} \frac{|b(y)|}{|x-y|^n}\, dy
\end{aligned}$$

$$\leq \frac{C_n B}{\varepsilon^n} \int_{B(x, C_n\varepsilon)\cap Q_k} |b(y)|\, dy.$$

Thus

$$|III(x)| \leq C_n BMb(x). \tag{7.47}$$

Adding (7.46) and (7.48) proves (7.45) and hence the weak–type $(1,1)$ estimate for T^*.

It remains to prove (7.41). With (7.39) proved we argue as in the proof of the Lebesgue differentiation theorem, (Corollary 5.2, Chapter 5). Recall that for $f_1 \in C_0^\infty(\mathbb{R}^n)$, $\lim_{\varepsilon\to 0} T_\varepsilon f_1(x)$ exists for every $x \in \mathbb{R}^n$. This was proved in the course of proving the existence of the L^p–limit, Corollary 7.8. Thus, if $f \in L^p(\mathbb{R}^n)$, $1 \leq p < \infty$, take as before a sequence $f_k \in C_0^\infty(\mathbb{R}^n)$ such that $\|f - f_k\|_p \to 0$ as $k \to \infty$ and define

$$\omega(f) = \left|\limsup_{\varepsilon\to 0} T_\varepsilon f - \liminf_{\varepsilon\to 0} T_\varepsilon f\right|.$$

Then

$$\omega(f) \leq \omega(f_k - f) \leq 2T^*(f_k - f)$$

and for any $\eta > 0$,

$$m\{x \in \mathbb{R}^n \colon \omega(f)(x) > \eta\} \leq \frac{C}{\eta^p} \int_{\mathbb{R}^n} |f_k - f|^p\, dx.$$

Letting $k \to \infty$ we find that

$$m\{x \in \mathbb{R}^n \colon \omega(f)(x) > \eta\} = 0.$$

Since $\eta > 0$ is arbitrary, this can only happen if $\omega(f) = 0$ almost everywhere. That is, $\lim_{\varepsilon\to 0} T_\varepsilon f$ exists a.e. This together with the fact that $\lim_{\varepsilon\to 0} T_\varepsilon f = Tf$ in $L^p(\mathbb{R}^n)$, proves (7.41) and hence the theorem. □

The singular integrals discussed above have some additional properties besides commuting with dilations. In particular, they also commute with the translation operator $f(x) \to f(x+y)$. Such operators are characterized in $L^2(\mathbb{R}^n)$ by the following lemma whose proof may be found in [SW].

Lemma 7.13. *Let T be a bounded linear operator on $L^2(\mathbb{R}^n)$. Then T commutes with translations if and only if there is a bounded function m such that $\widehat{Tf}(\xi) = m(\xi)\widehat{f}(\xi)$. That is, T is a Fourier multiplier if and only if it commutes with translations. If we denote the operator norm of T by $\|T\|_{2\to 2}$ we have*

$$\|T\|_{2\to 2} = \|m\|_\infty.$$

Operators $T : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ which have $\widehat{T}f(\xi) = m(\xi)\hat{f}(\xi)$ are called Fourier multipliers with multiplier m. For the singular integral operator of Theorem 7.12 there is an explicit formula for $m(\xi)$ given in terms of Ω. More precisely we have

Theorem 7.14. *Let Ω and T be as in Theorem 7.12. Then for all $f \in L^2(\mathbb{R}^n)$ we have*

$$\widehat{T}f(\xi) = m(\xi)\hat{f}(\xi)$$

where

$$m(\xi) = \int_{S^{n-1}} \left(\frac{\pi i}{2} \operatorname{sign}(\theta \cdot \xi) + \log \frac{1}{|\theta \cdot \xi|} \right) \Omega(\theta)\, d\sigma. \tag{7.48}$$

Let us look at the special case of the Hilbert Transform. Let $f \in L^2(\mathbb{R})$ and recall that

$$Hf(x) = \lim_{\varepsilon \to 0} \frac{1}{\pi} \int_{|y|>\varepsilon} \frac{f(x-y)}{y}\, dy = \lim_{\varepsilon \to 0} \int_{|y|>\varepsilon} \frac{\Omega(y)}{|y|} f(x-y)\, dy,$$

with

$$\Omega(x) = \frac{1}{\pi} sign(x). \tag{7.49}$$

We know now that Hf exists not only in L^p but also almost everywhere and its restriction to L^2 is a Fourier multiplier. Substituting (7.49) in (7.48) we find that

$$\widehat{H}f(\xi) = i \operatorname{sign}(\xi)\hat{f}(\xi). \tag{7.50}$$

It follows from this and Plancherel's theorem that

$$\|Hf\|_2 = \|f\|_2, \text{ for all } f \in L^2(\mathbb{R}^n). \tag{7.51}$$

In addition, (7.50) gives that

$$H^2 = -I, \tag{7.52}$$

where I denotes the identity operator on $L^2(\mathbb{R}^n)$.

Proof of Theorem 7.14. By Lemma 7.13, all we need to do is prove that the multiplier m of the operator T is given by (7.48). Define

$$K_{\varepsilon,N}(x) = \frac{\Omega(x)}{|x|^n} \quad \varepsilon < |x| < N$$

and 0 otherwise. Integration in polar coordinates gives

$$\int_{\mathbb{R}^n} |K_{\varepsilon,N}(x)|\,dx \le C\log(N/\varepsilon) < \infty$$

and $\widehat{K_{\varepsilon,N} * f}(\xi) = \widehat{K_{\varepsilon,N}(\xi)}\hat{f}(\xi)$. We claim that

(i) $\sup_{\xi\in\mathbb{R}^n} |\widehat{K}_{\varepsilon,N}(\xi)| \le C$, with C independent of ε and N, and that

(ii) $\widehat{K}_{\varepsilon,N}(\xi) \to m(\xi)$, as $\varepsilon \to 0$ and $N \to \infty$.

Assuming (i) and (ii) it follows from Plancherel's theorem that $K_{\varepsilon,N} * f$ converges in $L^2(\mathbb{R}^n)$ and the Fourier transform of this limit is $m(\xi)\hat{f}(\xi)$. Now letting $N \to \infty$ and then $\varepsilon \to 0$ proves the theorem. To prove (i) and (7.14), let $\xi = \rho\eta$ and $x = r\theta$, with $0 < \rho < \infty$, $0 < r < \infty$ and $\eta, \theta \in S^{n-1}$. Then using the fact that the integral of Ω is zero on S^{n-1} we obtain

$$\begin{aligned}
\widehat{K}_{\varepsilon,N}(\xi) &= \int_{\varepsilon<|x|<N} \frac{\Omega(x)}{|x|^n} e^{2\pi i x\cdot\xi}\,dx \\
&= \int_{S^{n-1}} \left(\int_\varepsilon^N \Omega(\theta) r^{-1} e^{2\pi i\rho r\eta\cdot\theta}\,dr\right) d\sigma(\theta) \\
&= \int_{S^{n-1}} \Omega(\theta)\left(\int_\varepsilon^N \left(r^{-1}e^{2\pi i\rho r\eta\cdot\theta} - \cos(2\pi r\rho)r^{-1}\right) dr\right) d\sigma(\theta) \\
&= \int_{S^{n-1}} \Omega(\theta) I_{\varepsilon,N}(\eta,\theta)\,d\sigma(\theta), \qquad (7.53)
\end{aligned}$$

with

$$I_{\varepsilon,N}(\eta,\theta) = \int_\varepsilon^N (\cos(2\pi\rho r\eta\cdot\theta) - \cos(2\pi r\rho))\,\frac{dr}{r} + i\int_\varepsilon^N \sin(2\pi\rho r\eta\cdot\theta)\frac{dr}{r}. \qquad (7.54)$$

By Exercise 3.2.5 of Chapter 3, the second integral is uniformly bounded and it converges to $\pi\,\mathrm{sign}(\eta\cdot\theta)/2$, as $\varepsilon \to 0$ and $N \to \infty$. It remains to prove that

$$\int_\varepsilon^N (\cos(2\pi\rho r\eta\cdot\theta) - \cos(2\pi r\rho))\,\frac{dr}{r} \to \log\left(\frac{1}{|\eta\cdot\theta|}\right) \qquad (7.55)$$

as $\varepsilon \to 0$ and $N \to \infty$. This is obtained from the following result. □

Proposition 7.15. *Suppose f is a differentiable function on $(0,\infty)$, continuous at 0 and with*

$$\int_a^b \frac{f(x)}{x}\,dx \to 0, \text{ as } a,b\to\infty. \tag{7.56}$$

Then for all $\lambda,\mu>0$,

$$\int_\varepsilon^N \frac{f(\lambda r)-f(\mu r)}{r}\,dr \to -f(0)\log\frac{\lambda}{\mu},$$

as $\varepsilon\to 0$ and $N\to\infty$.

Notice that the hypothesis of this Proposition are clearly satisfied by the function $f(x)=\cos(x)$.

Proof. Applying Fubini's theorem we see that

$$\begin{aligned}\int_\varepsilon^N \frac{f(\lambda r)-f(\mu r)}{r}\,dr &= \int_\varepsilon^N\int_\mu^\lambda f'(\tau r)\,d\tau\,dr\\ = \int_\mu^\lambda\int_\varepsilon^N f'(\tau r)\,dr\,d\tau &= \int_\mu^\lambda \frac{f(Nr)-f(\varepsilon r)}{r}\,dr.\end{aligned}$$

However, by our assumption,

$$\int_\mu^\lambda \frac{f(Nr)}{r}\,dr = \int_{\mu N}^{\lambda N}\frac{f(r)}{r}\,dr \to 0$$

as $N\to\infty$. Also, by our assumptions and the second (generalized) mean value theorem for integrals,

$$\int_\mu^\lambda \frac{f(\varepsilon r)}{r}\,dr = f(\gamma)\log\frac{\lambda}{\mu}$$

with $\gamma\to 0$ as $\varepsilon\to 0$. This completes the proof. □

7.3 Singular Integrals and BMO

Exercise 7.1.2 in section 7.1 shows that singular integrals do not in general map $L^\infty(\mathbb{R}^n)$ into $L^\infty(\mathbb{R}^n)$. However, as Exercise 7.1.3 shows, the function Hf of Exercise 7.1.2 does belong to $BMO(\mathbb{R})$. This in fact is always the case provided that we make the correct definition for the singular integral on $L^\infty(\mathbb{R}^n)$. We begin with the following

Theorem 7.16. *Let K satisfy the hypotheses of Theorem 7.7 and define for any $f \in L^2(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$,*

$$Tf(x) = \lim_{\varepsilon \to 0} \int_{\{|y|>\varepsilon\}} K(y) f(x-y)\, dy.$$

Then for all such f's we have

$$\|Tf\|_* \leq A \|f\|_\infty,$$

where the constant A depends only on n and the Hörmander constant B of the kernel K.

Proof. Let Q be a cube centered at x_Q and let $Q_2 = 2Q$. For any $f \in L^2(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ write $f = f\chi_{Q_2} + f\chi_{\mathbb{R}^n \setminus Q_2} = f_1 + f_2$ and set

$$C_Q = \int_{\mathbb{R}^n \setminus Q_2} K(x_Q - y) f(y)\, dy, \tag{7.57}$$

which is finite by the fact that $f \in L^2(\mathbb{R}^n)$ and our assumptions on K ((i) in Theorem 7.7). By the boundedness of the operator on L^2 and Hölder's inequality we have

$$\int_Q |Tf_1(x)|\, dx \leq |Q|^{1/2} \|Tf_1\|_2 \leq C_1 |Q| \|f\|_\infty, \tag{7.58}$$

where the constant C_1 depends on the L^2–bound for T and the dimension. Also, for $x \in Q$

$$Tf_2(x) - C_Q = \int_{\mathbb{R}^n \setminus Q_2} \left[K(x-y) - K(x_Q - y)\right] f(y)\, dy. \tag{7.59}$$

Thus the Hörmander condition gives

$$|Tf_2(x) - C_Q| \leq C_2 \|f\|_\infty. \tag{7.60}$$

Putting (7.59) and (7.60) together we see that

$$\frac{1}{|Q|} \int_Q |Tf(x) - C_Q|\ dx \leq (C_1 + C_2) \|f\|_\infty,$$

which together with Exercise 5.3.4 of Chapter 5 proves the theorem. □

Unlike the situation of $L^p(\mathbb{R}^n)$, $1 \leq p < \infty$, we cannot immediately extend the above result to all of $L^\infty(\mathbb{R}^n)$ since $L^2(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ is not dense in $L^\infty(\mathbb{R}^n)$. To extend this to all of $L^\infty(\mathbb{R}^n)$ we modified the definition of the operator T. We begin by replacing the Hörmander condition (H) by a stronger condition.

Definition 7.17. The function K satisfies the *strong Hörmander condition* (SH) if there is a constant B such that

$$|K(x-y) - K(x)| \leq B\frac{|y|}{|x|^{n+1}}, \text{ for all } |x| > 2|y|. \tag{7.61}$$

Remark 7.4. Clearly the functions satisfying the condition (SH) satisfy (H). For many results, including those proved here, the bound on the right hand side of (SH) can be replaced by the quantity $B|y|^\alpha/|x|^{n+\alpha}$ for some $\alpha > 0$.

Theorem 7.18. *Suppose the function K satisfies the condition (SH) and conditions (i) and (ii) of Theorem 7.7. For $f \in L^\infty(\mathbb{R}^n)$ and $k = 1, 2, \ldots$, set $f_k(x) = f\chi_{B(0,k)}$ and with T as in Theorem 7.16, define*

$$T_k f(x) = Tf_k(x) - \int_{\{1<|y|<k\}} K(-y) f(y)\, dy$$

Then the sequence $\{T_k f\}$ converges in $L^1(E)$ for any compact set $E \subset \mathbb{R}^n$ and almost everywhere to a function $\widetilde{T}f$ which belongs to $BMO(\mathbb{R}^n)$, and satisfies

$$\|\widetilde{T}f\|_* \leq C\|f\|_\infty,$$

where the constant C depends on n and B only.

Proof. Given the compact set E, let $A = 2\sup\{|y| : y \in E\}$ and let m be the first integer such that $m \geq A$. Then for all $x \in E$ and $k > m$ we have

$$T_k f(x) - T_m f(x) = \int_{\{m<|y|<k\}} [K(x-y) - K(-y)]\, f(y)\, dy$$

and by the assumption (SH) on K, the limit as $k \to \infty$ on the right hand side exists uniformly and is bounded by $B\|f\|_\infty$. Thus $T_k f$ converges pointwise and in $L^1(E)$. Now, the functions $T_k f$ and Tf_k differ only by a constant. This and Theorem 7.16 give that

$$\|\widetilde{T}f\|_* \leq \limsup_{k\to\infty} \|T_k f\|_* = \limsup_{k\to\infty} \|Tf_k\|_* \leq CB\|f\|_\infty,$$

proving the theorem. $\square$

Notice that if the function f is also in $L^2(\mathbb{R}^n)$ then the $\widetilde{T}f$ and Tf differ only on a constant and hence they are equal as BMO–functions. Finally, it is the case that singular integrals, when properly defined, map BMO into itself.

7.4 Some vector valued inequalities

For some of the applications of singular integrals which we will give below, particularly to the Beurling-Ahlfors operator in Chapter 8 and to square functions in Chapter 10, it is desirable to extend the above inequalities to the case when our functions, and even our kernels, take values in a Hilbert space. We begin with an easy result of Marcinkiewicz and Zygmund which asserts that bounded linear operators on $L^p(\mathbb{R}^n)$ extend to bounded linear operators on $L^p(\mathbb{R}^n, l^2)$ without increasing their norms. First recall that $l^2(\mathbb{R})$ denotes the space of all sequences $A = (a_1, a_2, \ldots)$ of real numbers with

$$|A|_{l^2} = \left(\sum_{k=1}^{\infty} |a_k|^2\right)^{1/2} < \infty$$

and that $L^p(\mathbb{R}^n, l^2)$ denotes the space of all functions $F = (f_1(x), f_2(x), \ldots)$ with

$$\||F|_{l^2}\|_p = \left(\int_{\mathbb{R}^n} \left(\sum_{k=1}^{\infty} |f_k(x)|^2\right)^{p/2} dx\right)^{1/p} < \infty.$$

Theorem 7.19. *Let $T : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ be a linear operator with operator norm $\|T\|$. Then T extends to a bounded linear operator on $L^p(\mathbb{R}^n, l^2)$ with the same norm. That is, set, $F = (f_1(x), f_2(x), \ldots)$ and $\mathcal{T}F = (Tf_1, Tf_2, \ldots)$. Then*

$$\||\mathcal{T}F|_{l^2}\|_p \leq \|T\|\||F|_{l^2}\|_p. \tag{7.62}$$

Proof. It is enough to prove (7.62) for $F = (f_1, f_2, \ldots, f_m, 0, 0, \ldots)$ for a fixed m. Let $\xi = (\xi_1, \ldots, \xi_m) \in S^{m-1}$ and observe that by the linearity of T we have $\xi \cdot \mathcal{T}F = T(\xi \cdot F)$. Thus by our assumption on T we have $\|\xi \cdot \mathcal{T}F\|_p \leq \|T\|\|\xi \cdot F\|_p$. Raising this inequality to the pth power, integrating over the sphere and applying Fubini's theorem leads to

$$\int_{\mathbb{R}^n} \left(\int_{S^{m-1}} |\xi \cdot \mathcal{T}F(x)|^p \, d\sigma(\xi)\right) dx \leq \|T\|^p \int_{\mathbb{R}^n} \left(\int_{S^{m-1}} |\xi \cdot F(x)|^p \, d\sigma(\xi)\right) dx.$$

However, by Exercise 3.2.8 of Chapter 3, the inner integrals on the left hand and the right hand side of the previous inequality are equal to

$$|\mathcal{T}F(x)|_{l^2}^p \int_{S^{m-1}} |\xi_1|^p \, d\sigma(\xi)$$

and

$$|F(x)|_{l^2}^p \int_{S^{m-1}} |\xi_1|^p \, d\sigma(\xi),$$

respectively. This implies the Theorem. □

The following related result for the Hardy–Littlewood maximal function is due to Fefferman and Stein, see [To1] for other similar results and further applications.

Theorem 7.20. *Let $F = (f_1, f_2, \ldots)$ and define $\mathcal{M}F = (Mf_1, Mf_2, \ldots)$ where Mf is the Hardy-Littlewood maximal function of f. Then*

(i) $m\{x \in \mathbb{R}^n : |\mathcal{M}F(x)|_{l^2} > \alpha\} \leq \frac{C}{\alpha} \| |F|_{l^2} \|_1$, *and*

(ii) $\| |\mathcal{M}F|_{l^2} \|_p \leq A_p \| |F|_{l^2} \|_p$, *for* $1 < p < \infty$.

Furthermore, we can take $A_p = C_n \sqrt{p}$ for $p \geq 2$ where C_n depends only on n.

Before we prove this result we need a lemma which generalizes Theorem 5.1 of Chapter 5 and which has some independent interest.

Lemma 7.21. *Let $1 < p < \infty$. There is a constant $A_{p,n}$ depending only on p and n such that*

$$\int_{\mathbb{R}^n} Mf(x)^p \, g(x) \, dx \leq A_{p,n} \int_{\mathbb{R}^n} |f(x)|^p \, Mg(x) \, dx$$

for any measurable functions f and g with $g \geq 0$.

Proof. Let μ and ν be the measures defined by $d\mu = g(x)\,dx$ and $d\nu = Mg(x)\,dx$, respectively. The Lemma asserts that $M : L^p(\mathbb{R}^n, \nu) \to L^p(\mathbb{R}^n, \mu)$. We leave it to the reader to verify the fact that $M : L^\infty(\mathbb{R}^n, \nu) \to L^\infty(\mathbb{R}^n, \mu)$ with norm not exceeding 1. We shall now prove that $M : L^1(\mathbb{R}^n, \nu) \to \mathrm{weak} - L^1(\mathbb{R}^n, \mu)$ and hence the Marcinkiewicz interpolation completes the proof of the lemma. That is, we claim there is a constant C_n depending only on n such that

$$\int_{\{x \in \mathbb{R}^n : \, Mf(x) > \alpha\}} g(x) \, dx \leq \frac{C_n}{\alpha} \int_{\mathbb{R}^n} |f(x)| Mg(x) \, dx. \tag{7.63}$$

To prove (7.63) assume first that $f \in L^1(\mathbb{R}^n)$ is nonnegative. Apply Corollary 5.8 of Chapter 5 to conclude that

$$\int_{\{x \in \mathbb{R}^n : \, Mf(x) > \alpha\}} g(x) \, dx \leq \sum_{k=1}^{\infty} \int_{Q_k} g(x) \, dx$$

$$
\begin{aligned}
&\le C\sum_{k=1}^{\infty}\int_{Q_k} g(x)\,dx\frac{1}{\alpha}\frac{1}{|Q_k|}\int_{Q_k}|f(x)|\,dx\\
&\le \frac{C_n}{\alpha}\sum_{k=1}^{\infty}\int_{Q_k} f(x)Mg(x)\,dx\\
&\le \frac{C_n}{\alpha}\int_{\mathbb{R}^n}|f(x)|Mg(x)\,dx.
\end{aligned}
$$

Our assumption that f is nonnegative is obviously irrelevant. To reduce to the case when $f\in L^1(\mathbb{R}^n, d\nu)$, we may take an increasing sequence of functions $\{f_j\}$ in $L^1(\mathbb{R}^n)$ converging up to f and observe that $\{x\in\mathbb{R}^n\colon Mf(x)>\alpha\}=\bigcup_j\{x\in\mathbb{R}^n\colon Mf_j>\alpha\}$. □

Proof of Theorem 7.20. We shall first prove (i). To do this, however, we need the case (ii) for $p=2$. This is easy. Indeed, by the boundedness of M on $L^2(\mathbb{R}^n)$

$$
\begin{aligned}
\||\mathcal{M}F|_{l^2}\|_2^2 &= \sum_{k=1}^{\infty}\int_{\mathbb{R}^n}(Mf_k(x))^2\,dx\\
&\le A_2\sum_{k=1}^{\infty}\int_{\mathbb{R}^n}|f_k(x)|^2\,dx \qquad (7.64)\\
&= A_2\||F|_{l^2}\|_2^2.
\end{aligned}
$$

We now follow the strategy used in proving the boundedness of singular integrals. Apply the Calderón–Zygmund decomposition to the function $|F|_{l^2}$ at level α to obtain a family of disjoint cubes $\{Q_k\}$ such that $\Omega=\bigcup_k Q_k$ satisfies

$$
\begin{aligned}
m(\Omega) &\le \frac{C}{\alpha}\||F|_{l^2}\|_1,\\
|F(x)|_{l^2} &\le \alpha, \text{a.e. on}\mathbb{R}^n\setminus\Omega, \qquad (7.65)\\
\frac{1}{|Q_k|}\int_{Q_k}|F(x)|_{l^2}\,dx &\le 2^n\alpha.
\end{aligned}
$$

Put

$$F_1=(f_1\chi_{\mathbb{R}^n\setminus\Omega}, f_2\chi_{\mathbb{R}^n\setminus\Omega},\ldots)$$

and

$$F_2=(f_1\chi_{\Omega}, f_2\chi_{\Omega},\ldots).$$

By the triangle inequality

$$|\mathcal{M}F(x)|_{l^2} \leq |\mathcal{M}F_1(x)|_{l^2} + |\mathcal{M}F_2(x)|_{l^2}. \tag{7.66}$$

Now apply (7.66) to conclude that

$$\begin{aligned} &m\{x \in \mathbb{R}^n \colon |\mathcal{M}F(x)|_{l^2} > \alpha\} \\ &\quad\leq\ m\{x \in \mathbb{R}^n \colon |\mathcal{M}F_1(x)|_{l^2} > \alpha/2\} + m\{x \in \mathbb{R}^n \colon |\mathcal{M}F_2(x)|_{l^2} > \alpha/2\} \\ &\quad=\ I + II.s \end{aligned}$$

Applying (7.64) to F_1 and using (7.65) we have

$$I \leq \frac{1}{\alpha^2}\int_{\mathbb{R}^n} |\mathcal{M}F_1(x)|_{l^2}^2\,dx \leq C\frac{1}{\alpha^2}\int_{\mathbb{R}^n} |F_1(x)|_{l^2}^2\,dx \leq C\frac{1}{\alpha}\int_{\mathbb{R}^n} |F(x)|_{l^2}\,dx.$$

To estimate II we consider $\widetilde{\Omega} = \bigcup_j \widetilde{Q_j}$ with $\widetilde{Q_j} = 2nQ_j$. Then $m\{\widetilde{\Omega}\} \leq (2n)^n m\{\Omega\}$. As before

$$II \leq C\frac{1}{\alpha}\||F|_{l^2}\|_1 + m\{x \in \mathbb{R}^n \setminus \widetilde{\Omega} \colon |\mathcal{M}F_2|_{l^2} > \alpha/2\} = C\frac{1}{\alpha}\||F|_{l^2}\|_1 + III.$$

Now consider the function

$$H(x) = (h_1, h_2, \ldots)$$

where, as in the Calderón–Zygmund decomposition, we set

$$h_k(x) = \sum_{j=1}^{\infty}\left(\frac{1}{|Q_j|}\int_{Q_j} |f_k(x)|\,dx\right)\chi_{Q_j}(x).$$

We now claim that

$$|\mathcal{M}F_2(x)|_{l^2}^2 \leq C|\mathcal{M}H(x)|_{l^2}^2, \text{ a.e. } x \in \mathbb{R}^n \setminus \widetilde{\Omega}. \tag{7.67}$$

Suppose we have proved (7.67). Then

$$\begin{aligned} III &\leq \frac{1}{\alpha^2}\int_{\mathbb{R}^n\setminus\widetilde{\Omega}} |\mathcal{M}H(x)|_{l^2}^2\,dx \leq \frac{C}{\alpha^2}\int_{\mathbb{R}^n} |H(x)|_{l^2}^2\,dx \\ &= \frac{C}{\alpha^2}\int_{\Omega} |H(x)|_{l^2}^2\,dx = \frac{C}{\alpha^2}\sum_{j=1}^{\infty}\int_{Q_j} |H(x)|_{l^2}^2\,dx \end{aligned} \tag{7.68}$$

However, for $x \in Q_j$ we have by the Minkowski inequality and the third inequality in (7.65) that

$$|H(x)|_{l^2} \leq \frac{1}{|Q_j|} \int_{Q_j} |F(x)|_{l^2} \, dx \leq 2^n \alpha.$$

Substituting this estimate in (7.68) proves the weak type estimate, provided we prove (7.67) which we shall now do. For this, fix $x \in \mathbb{R}^n \setminus \widetilde{\Omega}$ and let Q be any cube centered at x. Then

$$\frac{1}{|Q|} \int_Q |f_k \chi_\Omega(x)| \, dx = \frac{1}{|Q|} \sum_{j=1}^{\infty} \int_{Q \cap Q_j} |f_k(x)| \, dx.$$

Now, the terms on the sum on the right hand side are not zero only for those cubes $Q_j's$ which intersect the cube Q. Let I be the index set for these $j's$. If $j \in I$, an argument as in Exercise 7.1.1 above shows that $Q_j \subset 2nQ$. Thus

$$\begin{aligned}
\frac{1}{|Q|} \int_Q |f_k \chi_\Omega(x)| \, dx &= \frac{C}{|Q|} \sum_{j \in I} \int_{Q \cap Q_j} |h_k(x)| \, dx \\
&= \frac{C}{|Q|} \sum_{j \in I} \frac{1}{|Q_j|} \int_{Q \cap Q_j} |h_k(x)| \, dx |Q_j| \\
&\leq \frac{C}{|Q|} \sum_{j \in I} \int_{Q_j} |h_k(x)| \, dx \\
&\leq \frac{C}{|Q|} \int_{2nQ} |h_k(x)| \, dx \leq C M h_k(x),
\end{aligned}$$

which gives (7.67) since the cube centered at $x \in \mathbb{R}^n \setminus \widetilde{\Omega}$ was arbitrary. This completes the proof of the weak–type estimate.

We now show (ii). First observe that (i), case $p = 2$ of (ii) which we have already proved above, and the Marcinkiewcz interpolation Theorem proves (ii) for all $1 < p \leq 2$. Assume then that $2 < p < \infty$. Then $p/2 > 1$ and we let $q = \frac{p}{p-2}$ be its conjugate exponent. By duality

$$\| |\mathcal{M} F|_{l^2} \|_p = \| |\mathcal{M} F|_{l^2}^2 \|_{p/2}^{1/2} = \left(\sup_{g \in L^q(\mathbb{R}^n} \left| \int_{\mathbb{R}^n} |\mathcal{M} F(x)|_{l^2}^2 g(x) \, dx \right| \right)^{1/2},$$

where the supremum is taken over all g with norm less than or equal to 1. However, by Lemma 7.19

$$\begin{aligned}\int_{\mathbb{R}^n} |\mathcal{M}F(x)|_{l^2}^2 g(x)\,dx &= \sum_{k=1}^{\infty} \int_{\mathbb{R}^n} (Mf_k(x))^2 g(x)\,dx \\ \leq\ & C\sum_{k=1}^{\infty} \int_{\mathbb{R}^n} |f_k(x)|^2 Mg(x)\,dx = C\int_{\mathbb{R}^n} |F(x)|_{l^2}^2 Mg(x)\,dx \\ \leq\ & C\||F|_{l^2}\|_{p/2}\|Mg(x)\|_q \leq C_n \left(\frac{q}{q-1}\right)^{1/q} \||F|_{l^2}^2\|_{p/2}\|g(x)\|_q \\ =\ & C_n \left(\frac{p}{2}\right)^{(p-2)/p} \||F|_{l^2}^2\|_{p/2}\|g(x)\|_q,\end{aligned}$$

and the result follows. Furthermore, we get

$$\||\mathcal{M}F|_{l^2}\|_p \leq C_n\sqrt{p}\||F|_{l^2}\|_p$$

with the constant C_n depending only on n. $\square$

Exercise 7.4.1.
Prove that the vector valued Hardy–Littlewood maximal function $\mathcal{M}$ is not bounded on $L^\infty(\mathbb{R}^n, l^2)$. (This is already indicated by the behavior of the constant A_p as $p \to \infty$.)
Hint: Consider the function $F = (f_1, f_2, \dots)$ where $f_k = \chi_{(2^{k-1},2^k)}$.

Another interesting result along the lines of Theorems 7.19 and 7.20 is the following Theorem of A. Córdoba and C. Fefferman (see [To1]) which we state here but not prove. The proof uses a variant of Lemma 7.21 which is related to the theory of A_p–weights; an interesting topic which is not at the scope of these notes.

Theorem 7.22. *Let K_j be a sequence of kernels satisfying the hypotheses of Theorem 7.7 with the same constant B and denote by T_j the corresponding singular integral operators. Consider $F(x) = (f_1(x), f_2(x), \dots)$ and define $TF(x) = (T_1f_1(x), T_2f_2, \dots)$. There is a constant C_p independent of F such that*

$$\||TF|_{l^2}\|_p \leq C_p\||F|_{l^2}\|_p,$$

for all $1 < p < \infty$.

We end this chapter by stating some more vector valued inequalities when the Hilbert space l^2 is replace by a more general one. We first recall various facts about integration of Hilbert space valued functions. We will not give the proofs. These are all very simple and follow, once all the relevant definitions have been properly stated, exactly as the scalar cases which have already been discussed. First, we denote by H a separable Hilbert space and by $\langle h_1, h_2\rangle$ the inner product of any two elements $h_1, h_2 \in H$, and by $|h|_H = \sqrt{\langle h, h\rangle}$ the norm of $h \in H$. We recall that a function $f : \mathbb{R}^n \to H$ is Lebesgue measurable if for every $h \in H$ the function $f_h(x) = \langle f(x), h\rangle$ is Lebesgue measurable. We define the space $L^p(\mathbb{R}^n, H)$, for any $1 \leq p < \infty$ to be the collection of H–valued measurable functions for which

$$\|f\|_{L^p(\mathbb{R}^n,H)} = \left(\int_{\mathbb{R}^n} |f(x)|_H^p \, dx\right)^{1/p} < \infty$$

with the usual definition for $L^\infty(\mathbb{R}^n, H)$ as the collection of all measurable functions $f : \mathbb{R}^n \to H$ for which the essential supremum of $|f|_H$ is finite.

Given two separable Hilbert spaces H_1 and H_2 we denote by $B(H_1, H_2)$ the Banach space of bounded linear operators from H_1 into H_2 as discussed in section 2.3 of Chapter 2. Now, given a function $f : \mathbb{R}^n \to B(H_1, H_2)$, we say that f is Lebesgue measurable if for every $h \in H_1$ the function $f^h : \mathbb{R}^n \to H_2$ defined by $f^h(x) = f(x)h$ is measurable as defined above. The space $L^p(\mathbb{R}^n, B(H_1, H_2))$, $1 < p < \infty$ is the collection of measurable functions with values in $B(H_1, H_2)$ such that

$$\|f\|_{L^p(\mathbb{R}^n,B(H_1,H_2))} = \left(\int_{\mathbb{R}^n} |f(x)|_{B(H_1,H_2)}^p \, dx\right)^{1/p} < \infty.$$

The space $L^\infty(\mathbb{R}^n, B(H_1, H_2))$ consists of the measurable functions for which the essential supremum of $|f(x)|_{B(H_1,H_2)}$ is finite. With this definition many of the facts concerning integration of functions continue to hold for Hilbert space value functions and for functions with values in the space $B(H_1, H_2)$. In particular, the convolution of a function $K \in L^1(\mathbb{R}^n, B(H_1, H_2))$ with a function $f \in L^p(\mathbb{R}^n, H_1)$

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

produces a function with values in H_2 satisfying

$$\|K * f\|_{L^p(\mathbb{R}^n,H_2)} \leq \|K\|_{L^1(\mathbb{R}^n,B(H_1,H_2))}\|f\|_{L^p(\mathbb{R}^n,H_1)}.$$

In the same way many of the other results on convolutions stated in Chapter 4 continue to hold in this general setting. In the same way, the Fourier transform of a function $f \in L^1(\mathbb{R}^n, H)$ is defined by

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi} f(x)\, dx$$

and it satisfies the properties of the ordinary Fourier transform for functions with values on $\mathbb{R}$ or $\mathbb{C}$. In particular, it has an extension to $L^2(\mathbb{R}^n, H)$ and the Plancherel theorem holds.

Theorem 7.23. *Let H_1 and H_2 be two separable Hilbert spaces and let $K : \mathbb{R}^n \to B(H_1, H_2)$ satisfy the assumptions of Theorem 7.7 where the absolute of K is replaced here by $|K|_{B(H_1,H_2)}$. (For example, the Hörmander condition now reads as*

$$\int_{\{|x|>2|y|\}} |K(x-y) - K(x)|_{B(H_1,H_2)}\, dx \le B$$

for all $|y| > 0$.) For any $f \in L^p(\mathbb{R}^n, H_1) \cap L^2(\mathbb{R}^n, H_1)$ define

$$T_\varepsilon f(x) = \int_{\{|y|>\varepsilon\}} K(x-y) f(y)\, dy. \tag{7.69}$$

Then

$$\|T_\varepsilon f\|_{L^p(\mathbb{R}^n,H_2)} \le C_p \|f\|_{L^p(\mathbb{R}^n,H_1)}$$

with a constant C_p independent of f and ε and with the same behavior as the constant of Theorem 7.19. Also, the conclusions of Corollary 7.8 hold in this setting.

Theorem 7.24. *If $K : \mathbb{R}^n \to B(H_1, H_2)$ satisfies the conditions of Theorem 7.11 then the conclusions of that theorem continue to hold for the the vector valued operator as defined in* (7.69).

An important application of Theorems 7.23 and 7.24 will be given in Chapter 10 where we study the Littlewood-Paley function of Exercise 6.3.4 in Chapter 6, in $L^p(\mathbb{R}^n)$. This chapter is independent from Chapters 8 and 9 so the interested and impatient reader can read it now if he or she so desires.

Chapter 8

The Riesz and Beurling–Ahlfors Transforms

In this chapter we will first show that the Hilbert transform arises naturally as the only operator mapping $L^2(\mathbb{R})$ into itself which preserves the basic group operations of the real line. We will then look for operators on $L^2(\mathbb{R}^n)$ with similar properties. This will naturally lead to the Riesz transforms. We will then give the connection, mentioned several times before, between the Hilbert transform and conjugate harmonic functions in the upper half space. This is probably the most natural way to discover the Hilbert transform and historically the way it arose. We will then look at the exact analogue of this situation in the upper half space of $\mathbb{R}^n$ and show again that the Riesz transforms are the natural extensions of the Hilbert transform when viewed also from the point of view of conjugate harmonic functions. That is, we will present the generalization of the Cauchy Riemann equations of Stein and Weiss. Section 8.3 is devoted to some standard applications of the Riesz transforms. However, here we will also discuss in detail the Beurling-Ahlfors operator. This singular integral is very important from the point of view of applications and there are several very interesting open problems concerning its operator norm on L^p. This important operator, unlike the Hilbert and Riesz transforms, does not seem to get any special treatment in the standard literature on singular integrals. Our presentation of the Riesz transforms follows [St1].

8.1 Invariant Properties of the Hilbert Transform

As we have just shown in Chapter 7, the Hilbert transform satisfies

$$(H\tau_\delta)(f) = (\tau_\delta H)(f), \qquad \delta > 0 \tag{8.1}$$

and

$$(H\tau_\delta)(f) = -(\tau_\delta H)(f), \qquad \delta < 0. \tag{8.2}$$

where $\tau_\delta f = f(\delta x)$. That is, H commutes with positive dilations and anticommutes with the reflection $f(x) \to f(-x)$. These properties in fact characterize the Hilbert transform. That is we have

Theorem 8.1. *Suppose T is a bounded linear operator on $L^2(\mathbb{R})$ which has the following properties:*

(i) T commutes with translations.

(ii) T commutes with positive dilations.

(iii) T anti–commutes with the reflection $f(x) \to f(-x)$.

Then T is a constant multiple of the Hilbert Transform.

Proof. By (7.50) of Chapter 7 we must show that

$$\widehat{Tf}(\xi) = C\,\mathrm{sign}(\xi)\hat{f}(\xi),$$

for some constant C. By i and Lemma 7.13, Chapter 7, there is an $m \in L^\infty$ with the property that

$$\widehat{Tf}(\xi) = m(\xi)\hat{f}(\xi). \tag{8.3}$$

As before, let us denote the operator : $f \to \hat{f}$ by $\mathcal{F}$. A simple change of variables gives

$$\begin{aligned}
(\mathcal{F}\tau_\delta f)(\xi) &= \int_{-\infty}^{\infty} e^{2\pi i x\cdot\xi} f(\delta x)\,dx \\
&= \frac{1}{|\delta|}\int_{-\infty}^{\infty} e^{2\pi i x\cdot\xi/\delta} f(x)\,dx \\
&= \frac{1}{|\delta|}\hat{f}(\xi/\delta) = \frac{1}{|\delta|}(\tau_{\delta^{-1}}\mathcal{F}f)(\xi).
\end{aligned}$$

Thus as operators on $L^2(\mathbb{R})$, τ_δ and $\mathcal{F}$ satisfy

$$\mathcal{F}\tau_\delta = \frac{1}{|\delta|}\tau_{\delta^{-1}}\mathcal{F} \Leftrightarrow \frac{1}{|\delta|}\mathcal{F}\tau_{\delta^{-1}} = \tau_\delta\mathcal{F} \Leftrightarrow \frac{1}{|\delta|}\tau_{\delta^{-1}}\mathcal{F}^{-1} = \mathcal{F}^{-1}\tau_\delta. \tag{8.4}$$

Continuing with this notation, $\widehat{Tf}(\xi) = m(\xi)\hat{f}(\xi)$ is equivalent to

$$\mathcal{F}T = m\mathcal{F} \text{ or } m = \mathcal{F}T\mathcal{F}^{-1} \tag{8.5}$$

and the assumptions (ii) and (iii) can be written as

$$T\tau_\delta = \operatorname{sign}(\delta)\tau_\delta T. \tag{8.6}$$

From (8.4)–(8.6) we obtain

$$\begin{aligned}\tau_\delta m = \tau_\delta\mathcal{F}T\mathcal{F}^{-1} &= \frac{1}{|\delta|}\mathcal{F}\tau_{\delta^{-1}}T\mathcal{F}^{-1} = \frac{1}{|\delta|}\mathcal{F}(\operatorname{sign}(\frac{1}{\delta})T\tau_{\delta^{-1}}\mathcal{F}^{-1}) \\ &= \frac{\operatorname{sign}(\delta)}{|\delta|}\mathcal{F}T\tau_{\delta^{-1}}\mathcal{F}^{-1} = \operatorname{sign}(\delta)\mathcal{F}T\mathcal{F}^{-1}\tau_\delta = \operatorname{sign}(\delta)m\tau_\delta.\end{aligned}$$

That is, for all $f \in L^2(\mathbb{R})$ and all $\delta \neq 0$, we have

$$\tau_\delta(mf)(x) = \operatorname{sign}(\delta)(m\tau_\delta f)(x)$$

or

$$m(\delta x)f(\delta x) = \operatorname{sign}(\delta)m(x)f(\delta x)$$

which gives $m(\delta) = \operatorname{sign}(\delta)m(1)$. This proves the Theorem. □

8.2 Invariant Properties of the Riesz Transforms

We now look in $\mathbb{R}^n$ for operators with similar properties. First, a change of variables as above leads to

$$(\mathcal{F}\tau_\delta)(f) = \delta^{-n}(\tau_{\delta^{-1}}\mathcal{F})(f)$$

for any $f \in L^2(\mathbb{R}^n)$ and all $\delta > 0$. That is,

$$\mathcal{F}\tau_\delta = \frac{1}{\delta^n}\,\tau_{\delta^{-1}}\mathcal{F} \tag{8.7}$$

as operators on $L^2(\mathbb{R}^n)$. Besides dilations and reflections, another important operation in $\mathbb{R}^n$ is the rotations. If ρ is a rotation in $\mathbb{R}^n$, we define $\rho(f)(x) = f(\rho^{-1}x)$. We have, by change of variables,

$$\begin{aligned}(\mathcal{F}\rho)(f)(\xi) &= \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi} f(\rho^{-1}x)\,dx \\ &= \int_{\mathbb{R}^n} e^{2\pi i \rho x\cdot\xi} f(x)\,dx \\ &= \int_{\mathbb{R}^n} e^{2\pi i x\cdot\rho^{-1}\xi} f(x)\,dx \\ &= (\rho\mathcal{F})(f)(\xi),\end{aligned}$$

which we again write as

$$\mathcal{F}\rho = \rho\mathcal{F}. \tag{8.8}$$

Definition 8.2. Let $m(x) = (m_1(x), \ldots, m_n(x))\colon \mathbb{R}^n \to \mathbb{R}^n$ and let $\rho = (\rho_{jk})$ be a rotation. We say that the function m operates as a vector if

$$m(\rho x) = \rho(m(x)), \tag{8.9}$$

for all rotations ρ.

We can write (8.9) in matrix notation as

$$(m_1(\rho x), \ldots, m_n(\rho x)) = \begin{bmatrix}\rho_{11} & \cdots & \rho_{1n} \\ \vdots & \ddots & \vdots \\ \rho_{n1} & \cdots & \rho_{nn}\end{bmatrix} \begin{bmatrix} m_1(x) \\ \vdots \\ m_n(x)\end{bmatrix}$$

and in coordinates as

$$m_j(\rho x) = \sum_{k=1}^{n} \rho_{jk} m_k(x).$$

Lemma 8.3. *Let $m\colon \mathbb{R}^n \to \mathbb{R}^n$ be homogeneous of degree zero. That is, $m(\delta x) = m(x)$ for all $\delta > 0$. Suppose m operates like a vector. Then $(m_1(x), \ldots, m_n(x)) = m(x) = C\frac{x}{|x|}$ for some constant C.*

Proof. By the homogeneity of m, it is enough to consider x on the unit sphere. Let $C = m_1(e_1)$. We claim $m_j(e_1) = 0$ for all $j \neq 1$. To see this, let ρ be any rotation leaving e_1 fixed. That is, the matrix representing ρ has the form

$$(\rho_{jk}) = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \rho_{n-1} & \\ 0 & & & \end{array}\right]$$

where ρ_{n-1} is a rotation in $\mathbb{R}^{n-1}$. Then

$$m_j(e_1) = \sum_{k=1}^{n} \rho_{jk} m_k(e_1) = \sum_{k=2}^{n} \rho_{jk} m_k(e_1),$$

Thus the vector $\vec{v} = (m_2(e_1), \ldots, m_n(e_1))$ is fixed by all rotations of $\mathbb{R}^{n-1}$. This implies that $\vec{v} = \vec{0}$. (Note that in dimensions 3 and higher we only need the proper rotations; that is, those of determining one). Thus for all rotations ρ in $\mathbb{R}^n$

$$m_j(\rho e_1) = \sum_{k=1}^{n} \rho_{jk} m_k(e_1) = \rho_{j1} m_1(e_1) = C\rho_{j1}.$$

Now pick a rotation with $\rho e_1 = x$. Then $\rho_{j1} = x_j$ and $m_j(x) = Cx_j$. That is,

$$m(x) = C\frac{x}{|x|},$$

as desired. □

We are now ready to define Riesz transforms. First, for any $n \geq 1$ define the functions $K_1, K_2, \ldots K_n$ by

$$K_j(x) = \frac{\Omega_j(x)}{|x|^n}, \quad \Omega_j(x) = \frac{C_n x_j}{|x|},$$

where C_n is a constant depending on n to be chosen in a minute. These $K's$ clearly satisfy the assumptions of Theorem 7.11, Chapter 7. The operators defined by

$$R_j(f)(x) = \lim_{\varepsilon \to 0} \int_{\{|y| > \varepsilon\}} \frac{\Omega_j(y)}{|y|^n} f(x - y)\, dy, \quad j = 1, 2, \ldots, n \tag{8.10}$$

are called the Riesz transforms. Notice if $n = 1$ and $C_1 = 1/\pi$ then

$$K_1(x) = \frac{\Omega_1(x)}{|x|}, \quad \Omega_1(x) = \frac{1}{\pi}\frac{x}{|x|}$$

and in this case the above operators just reduce to the Hilbert transform. By Theorem 7.14, Chapter 7, the multipliers for these operators are given by

$$m_j(\xi) = \int_{S^{n-1}} \left(\frac{\pi i}{2} \operatorname{sign}(x \cdot \xi) + \log \frac{1}{|x \cdot \xi|} \right) \Omega_j(x)\, dx.$$

From

$$\begin{aligned} \rho(m_j)(\xi) &= \int_{S^{n-1}} \left(\frac{\pi i}{2} \operatorname{sign}(x \cdot \rho^{-1}\xi) + \log \frac{1}{|x \cdot \rho^{-1}\xi|} \right) \Omega_j(x)\, d\sigma \\ &= \int_{S^{n-1}} \left(\frac{\pi i}{2} \operatorname{sign}(\rho x \cdot \xi) + \log \frac{1}{|\rho x \cdot \xi|} \right) \Omega_j(x)\, d\sigma \\ &= \int_{S^{n-1}} \left(\frac{\pi i}{2} \operatorname{sign}(x \cdot \xi) + \log \frac{1}{|x \cdot \xi|} \right) \Omega_j(\rho^{-1}x)\, d\sigma \\ &= \int_{S^{n-1}} \left(\frac{\pi i}{2} \operatorname{sign}(x \cdot \xi) + \log \frac{1}{|x \cdot \xi|} \right) \rho(\Omega_j)(x)\, d\sigma \end{aligned}$$

and the fact that

$$\Omega(x) = (\Omega_1(x), \dots \Omega_n(x)) = \left(C_n \frac{x_1}{|x|}, \dots, C_n \frac{x_n}{|x|} \right)$$

operates as a vector, it follows that the function

$$m(\xi) = (m_1(\xi), \dots, m_n(\xi))$$

also operates as a vector. Thus by Lemma 8.3,

$$m_j(\xi) = m_1(e_1) \frac{\xi_j}{|\xi|}, \quad j = 1, \dots, n \tag{8.11}$$

with

$$\begin{aligned} m_1(e_1) &= \int_{S^{n-1}} \left[i\frac{\pi}{2} \operatorname{sign}(\xi_1) + \log \frac{1}{|\xi_1|} \right] C_n \xi_1\, d\sigma(\xi) \\ &= \frac{\pi i}{2} C_n \int_{S^{n-1}} |\xi_1|\, d\sigma(\xi). \end{aligned} \tag{8.12}$$

That is, the Riesz transforms in $L^2(\mathbb{R}^n)$ are Fourier multiplier operators with

$$\widehat{R_j f}(\xi) = m_1(e_1) \frac{\xi_j}{|\xi|} \widehat{f}(\xi).$$

Since

$$(\widehat{R_j(R_j f)})(\xi) = {m_1}^2(e_1)\frac{\xi_j^2}{|\xi|^2}\widehat{f}(\xi)$$

we see that

$$(R_1^2 + \cdots + R_n^2)f = {m_1}^2(e_1)f.$$

We are now in a position to choose the constant C_n. Recall that the Hilbert transform satisfies $H^2 f = -f$. If we choose C_n such that $m_1(e_1) = i$ we also have

$$\widehat{R_j f}(\xi) = i\frac{\xi_j}{|\xi|}\widehat{f}(\xi) \tag{8.13}$$

and

$$R_1^2 + \cdots + R_n^2 f = -f. \tag{8.14}$$

for all $f \in L^2(\mathbb{R}^n)$ (exactly as for H). The requirement on $m_1(e_1)$ and (8.12) forces

$$C_n = \frac{1}{\frac{\pi}{2}\int_{S^{n-1}} |\xi_1|\, d\sigma(\xi)}$$

and if we recall now that by Exercise 3.2.9 at the end of Chapter 3

$$\int_{S^{n-1}} |\xi_1|^p\, d\sigma(\xi) = \frac{2\pi^{\frac{n-1}{2}}\Gamma(\frac{p+1}{2})}{\Gamma(\frac{n+p}{2})}$$

for any $1 \le p < \infty$, we see that

$$C_n = \frac{\Gamma(\frac{n+1}{2})}{\pi^{(n+1)/2}}.$$

Thus the Riesz transforms are the singular integrals

$$R_j f(x) = \lim_{\varepsilon\to 0}\frac{\Gamma(\frac{n+1}{2})}{\pi^{(n+1)/2}}\int_{\{|y|>\varepsilon\}}\frac{y_j}{|y|^{n+1}}f(x-y)\,dy, \quad j = 1, \ldots, n$$

satisfying (8.13) and (8.14).

Exercise 8.2.1.
Let $R_j R_k$ be the composition of two Riesz transforms. Prove that

$$\|R_j R_k f\|_2 \le \frac{1}{2}\|f\|_2, \text{ for any } j \ne k.$$

As in the case of the Hilbert transform, the Riesz transforms also arise naturally as the only operators on $L^2(\mathbb{R}^n)$ which preserve certain properties of $\mathbb{R}^n$. More precisely, with

$$m(\xi) = (m_1(\xi), \ldots, m_n(\xi)) = \left(i\frac{\xi_1}{|\xi|}, \ldots, i\frac{\xi_n}{|\xi|} \right)$$

we have for any rotation ρ of $\mathbb{R}^n$,

$$m_j(\rho\xi) = \sum_{k=1}^{n} \rho_{j_k} m_k(\xi)$$

which gives for any $f \in L^2(\mathbb{R}^n)$

$$m_j(\rho\xi)(\widehat{f})(\rho^{-1}\rho\xi) = \sum_{k=1}^{n} \rho_{j_k} m_k(\xi)(\widehat{f})(\xi) = \sum_{k=1}^{n} \rho_{j_k} \widehat{R_k f}(\xi). \tag{8.15}$$

However, from our definition of ρm and (8.8) we see that the left hand side of (8.15) is

$$\begin{aligned} &\rho^{-1}\left(m_j(\xi)(\widehat{f})(\rho^{-1}\xi) \right) = \rho^{-1}\left(m_j(\xi)\rho(\widehat{f})(\xi) \right) \\ = \;& \rho^{-1}(m_j(\xi)(\widehat{\rho f})(\xi)) = \rho^{-1}(\widehat{R_j \rho f})(\xi) = (\widehat{\rho^{-1} R_j \rho f})(\xi). \end{aligned}$$

Thus for all rotation ρ of $\mathbb{R}^n$,

$$\rho^{-1} R_j \rho = \sum_{k=1}^{n} \rho_{j_k} R_k, \tag{8.16}$$

as operators on $L^2(\mathbb{R}^n)$. The following Theorem is the analogue in $\mathbb{R}^n$ of Theorem 8.1.

Theorem 8.4. *Let $T = (T_1, T_2, \ldots, T_n)$ be bounded linear operators on $L^2(\mathbb{R}^n)$. Suppose*

(i) Each T_j commutes with translations.

(ii) Each T_j commutes with dilations.

(iii) For all rotations $\rho = (\rho_{j_k})$

$$\rho^{-1}T_j\rho = \sum_{k=1}^{n} \rho_{j_k} T_k \tag{8.17}$$

Then $T_j = CR_j,\ C$ *a constant.*

Proof. Proceeding as in the proof of Theorem 8.1, we must verify that for all $f \in L^2(\mathbb{R}^n)$ we have

$$\widehat{T_j f}(\xi) = C\frac{\xi_j}{|\xi|}\widehat{f}(\xi). \tag{8.18}$$

Since T_j commutes with translations, Lemma 7.13, Chapter 6, gives an $m_j \in L^\infty(\mathbb{R}^n)$ such that

$$\widehat{T_j f}(\xi) = m_j(\xi)\widehat{f}(\xi).$$

The fact that T_j commutes with dilations shows that m_j is homogeneous of degree zero. That is, $m_j(\delta\xi) = m_j(\xi)$ for all $\delta > 0$. By (iii)

$$(\widehat{\rho^{-1}T_j\rho(}f)(\xi) = \sum_{k=1}^{n} \rho_{j_k} m_k(\xi)\widehat{f}(\xi)$$

or

$$\begin{aligned} \rho^{-1}\left(m_j(\xi)(\widehat{\rho f})(\xi)\right) &= \rho^{-1}\left(m_j(\xi)\rho(\widehat{f})(\xi)\right) \\ = \quad \rho^{-1}\left(m_j(\xi)\widehat{f}(\rho^{-1}\xi)\right) &= m_j(\rho\xi)\widehat{f}(\xi). \end{aligned}$$

Thus,

$$m_j(\rho\xi) = \sum_{k=1}^{n} \rho_{j_k} m_k(\xi)$$

which means that $m = (m_1, \ldots, m_n)$ operates as a vector. By Lemma 8.3

$$m_j(\xi) = C\frac{\xi_j}{|\xi|},$$

proving (8.18) and hence the Theorem. $\square$

Next we briefly describe the notion of higher Riesz transforms. First, let $P(x)$ be a homogeneous polynomial of degree k on $\mathbb{R}^n$. That is, $p(x)$ has the form

$$P(x) = \sum_{|\alpha|=k} C_\alpha x^\alpha$$

where as usual we take $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$, $\alpha_j \geq 0$, $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$ and $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$. The polynomial is harmonic if $\Delta P = 0$. Examples of homogeneous harmonic polynomial of degree 1 are given by $P(x) = x_j$ and of degree two by $P(x) = x_j x_k$, $P(x) = x_j^2 - x_k^2$, $j \neq k$.

If we set $x = r\xi$, $r > 0$ and $\xi \in S^{n-1}$, we see that $P(x) = r^k P(\xi)$ and thus $\frac{\partial P}{\partial r} = k r^{k-1} P(\xi)$ on the sphere S^{n-1}. It follows from this and the divergence theorem that for any two homogeneous harmonic polynomial P and Q, of degree k and j

$$\begin{aligned}
(k-j) &\int_{S^{n-1}} P(\xi)\overline{Q}(\xi)\, d\sigma(\xi) \\
&= \int_{S^{n-1}} \left(\overline{Q}(\xi) \frac{\partial P}{\partial r}(\xi) - Q(\xi) \frac{\partial \overline{P}}{\partial r}(\xi) \right) d\sigma(\xi) \\
&= \int_B \left(\overline{Q}(x) \Delta P(x) - P(x) \Delta \overline{Q}(x) \right) dx = 0.
\end{aligned}$$

Thus in particular for $k \geq 1$,

$$\int_{S^{n-1}} P(\xi)\, d\xi = 0$$

and we may apply Theorem 7.12, Chapter 7 with

$$\Omega(x) = \frac{P(x)}{|x|^k} \tag{8.19}$$

to conclude that the singular integral defined by

$$T_P(f)(x) = \lim_{\varepsilon \to 0} \int_{\{|y| > \varepsilon\}} \frac{P(y)}{|y|^{n+k}} f(x-y)\, dy \tag{8.20}$$

is weak–type (1,1) and strong–type (p, p) for $1 < p < \infty$. Its Fourier multiplier can be computed (see Stein [St1] or Stein–Weiss [SW]) using the Hecke identities to find that for all $f \in L^2(\mathbb{R}^n)$,

$$\widehat{T_P}(f)(\xi) = i^k C_k \frac{P(\xi)}{|\xi|^k} \widehat{f}(\xi) \tag{8.21}$$

where

$$C_k = \pi^{n/2} \frac{\Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{k+n}{2}\right)}.$$

Since

$$\widehat{R_j f}(\xi) = i\frac{\xi_j}{|\xi|}\widehat{f}(\xi)$$

if we set

$$Rf = (R_1 f, R_2 f, \dots, R_n f)$$

we see that

$$P(R)f(x) = \frac{1}{C_n} T_P f(x)$$

and we call $P(R)$ a higher order Riesz transform. The case $P(x) = x_j$ gives the first order Riesz transforms and $P(x) = x_j x_k$ are called second order Riesz transforms. The polynomial $P(x) = x_1^2 - x_2^2 - 2ix_1x_2$ in $\mathbb{R}^2$ gives the Beurling–Ahlfors transform which will be discussed below.

We end this section with some exercises which show that the operator norm of the Riesz transforms in $L^p(\mathbb{R}^n)$ is the same as the operator norm of the Hilbert transform in $L^p(\mathbb{R})$. This result is based on the method of rotations of Calderón and Zygmund. First we state without proof the following result.

Theorem 8.5 (Pichorides [Pi]). *Let H_p be the operator norm of the Hilbert transform on $L^p(\mathbb{R})$, $1 < p < \infty$. That is,*

$$H_p = \sup_{f \in L^p(\mathbb{R})} \frac{\|Hf\|_p}{\|f\|_p}.$$

Then

$$H_p = \begin{cases} \tan\frac{\pi}{2p}, & 1 < p \le 2 \\ \cot\frac{\pi}{2p}, & 2 \le p < \infty \end{cases}$$

We will also use the following proposition about L^p–multipliers which is of independent interest. First, a function $m \in L^\infty(\mathbb{R}^n)$ is said to be an $L^p(\mathbb{R}^n)$–multiplier, $1 \le p \le \infty$, if the Fourier multiplier operator $T : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ defined by $\widehat{T_m f}(\xi) = m(\xi)\widehat{f}(\xi)$ has a bounded extension to $L^p(\mathbb{R}^n)$. That is, for $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$ define

$$T_m f(x) = \int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi} m(\xi)\widehat{f}(\xi)\, d\xi.$$

The function m is an $L^p(\mathbb{R}^n)$–multiplier if for all such f we have

$$\|T_m f\|_p \le A_p \|f\|_p$$

where the constant A_p is independent of f. We denote the class of $L^p(\mathbb{R}^n)$–multipliers by $\mathcal{M}_p(\mathbb{R}^n)$ and the operator norm of T_m in $L^p(\mathbb{R}^n)$ by $\|T_m\|_p$. We saw in Chapter 6 examples of multipliers in $\mathcal{M}_p(\mathbb{R}^n)$ for $1 < p < \infty$. Also, by Theorem 4.2, Chapter 4, and Proposition 6.2 (iii), Chapter 6, for any $g \in L^1(\mathbb{R}^n)$ the function $m(\xi) = \widehat{g}(\xi) \in \mathcal{M}_p(\mathbb{R}^n)$ for any $1 \leq p \leq \infty$.

Proposition 8.6. *Suppose $m \in \mathcal{M}_p(\mathbb{R}^{n+k})$, $1 < p < \infty$, has the property that $\lim_{r\to 0} m(x_1, rx_2) = m(x_1, 0) = m_0(x_1)$ exists for almost every $x_1 \in \mathbb{R}^n$ where we denote the points in $\mathbb{R}^{n+k}$ by (x_1, x_2), $x_1 \in \mathbb{R}^n$, $x_2 \in \mathbb{R}^k$. Then $m_0 \in \mathcal{M}_p(\mathbb{R}^n)$ and $\|T_{m_0}\|_p \leq \|T_m\|_p$.*

Proof. For any $r > 0$ and any function h on $\mathbb{R}^{n+k}$ define the function $h_r(x_1, x_2) = h(x_1, rx_2)$. Consider the operators T_r defined by

$$T_r f(x_1, x_2) = \int_{\mathbb{R}^{n+k}} e^{-2\pi i (x_1,x_2)\cdot(\xi_1,\xi_2)} m_r(\xi_1, \xi_2) \widehat{f}(\xi_1, \xi_2)\, d\xi_1\, d\xi_2,$$

for $f \in L^2(\mathbb{R}^{n+k}) \cap L^p(\mathbb{R}^{n+k})$. A simple change of variables shows that $T_r f(x_1, x_2) = (T_m f_r)(x_1, x_2/r)$. Thus

$$\begin{aligned}\|T_r f\|_p &= r^{k/p}\|T_m f_r\|_p \leq r^{k/p}\|T_m\|_p\|f_r\|_p \\ &= r^{k/p} r^{-k/p}\|T_m\|_p\|f\|_p = \|T_m\|_p\|f\|_p.\end{aligned}$$

Thus the operators T_r have an extension to $L^p(\mathbb{R}^{n+k})$ satisfying $\|T_r\|_p \leq \|T_m\|_p$. By Plancherel's Theorem and the fact that $m_r(x_1, x_2) \to m(x_1, 0)$, as $r \to 0$, we see that $T_r f \to T_0 f$ in $L^2(\mathbb{R}^{n+k})$ for any $f \in L^2(\mathbb{R}^{n+k}) \cap L^p(\mathbb{R}^{n+k})$, where

$$T_0 f(x_1, x_2) = \int_{\mathbb{R}^{n+k}} e^{-2\pi i (x_1,x_2)\cdot(\xi_1,\xi_2)} m(\xi_1, 0) \widehat{f}(\xi_1, \xi_2)\, d\xi_1\, d\xi_2.$$

Thus we also have

$$\|T_0\|_p \leq \|T_m\|_p.$$

Now let $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$ and consider the function

$$h(x_1, x_2) = f(x_1)g(x_2)$$

with $g \in L^2(\mathbb{R}^k) \cap L^p(\mathbb{R}^k)$. Then

$$T_0 h(x_1, x_2) = g(x_2) T_{m_0} f(x_1)$$

and thus

$$\|g\|_p\|T_{m_0} f\|_p = \|T_0 h\|_p \leq \|T_m\|_p\|h\|_p = \|T_m\|_p\|g\|_p\|f\|_p,$$

which proves the assertion of the proposition. $\square$

Exercise 8.2.2.
For a unit vector $e \in \mathbb{R}^n$, the Hilbert transform of the function $f \in L^p(\mathbb{R}^n)$ in the direction of e is the operator defined by $H_e f(x) = \lim_{\varepsilon\to 0} H_\varepsilon^e f(x)$ where

$$H_\varepsilon^e f(x) = \frac{1}{\pi}\int_{\{|t|>\varepsilon\}} \frac{f(x-te)}{t}\,dt.$$

Prove that $\|H_e f\|_p \leq H_p\|f\|_p$, $1 < p < \infty$, where H_p is the L^p–norm of the Hilbert transform in one dimension which was given by Theorem 8.5.

Proposition 8.7 (Calderón–Zygmund Method of Rotations)**.** *Let Ω satisfy the hypothesis of Theorem 6.9, Chapter 6, and suppose Ω is odd. That is, $\Omega(x) = -\Omega(-x)$. Then*

$$T_\varepsilon f(x) = \frac{\pi}{2}\int_{S^{n-1}} \Omega(e) H_\varepsilon^e f(x)\,d\sigma(e).$$

Proof. By integrating in polar coordinates and using the fact that Ω is odd we obtain

$$\begin{aligned} T_\varepsilon f(x) &= \int_{S^{n-1}}\int_\varepsilon^\infty f(x-re)\frac{\Omega(e)}{r^n}r^{n-1}\,dr\,d\sigma(e) \\ &= \int_{S^{n-1}}\Omega(e)\int_\varepsilon^\infty \frac{f(x-re)}{r}\,dr\,d\sigma(e) \\ &= \frac{1}{2}\int_{S^{n-1}}\Omega(e)\int_\varepsilon^\infty \frac{f(x-re)-f(x+re)}{r}\,dr\,d\sigma(e) \\ &= \frac{1}{2}\int_{S^{n-1}}\Omega(e)\int_{\{|t|>\varepsilon\}} \frac{f(x-te)}{t}\,dt\,d\sigma(e) \\ &= \frac{\pi}{2}\int_{S^{n-1}}\Omega(e)H_\varepsilon^e f(x)\,d\sigma(e). \end{aligned}$$

□

Exercise 8.2.3. (i) Prove that

$$\|T_\varepsilon f\|_p \leq \left(\frac{\pi}{2}H_p\int_{S^{n-1}}|\Omega(e)|\,d\sigma(e)\right)\|f\|_p$$

$1 < p < \infty$.

(ii) Use (i) to prove that the Riesz transforms satisfy the bound $\|R_j f\|_p \leq H_p\|f\|_p$, $1 < p < \infty$ with H_p as in (i). In particular, if we denote by $\|R_j\|_p$ the L^p–norm of the Riesz transform R_j we have $\|R_j\|_p \leq H_p$ and hence the L^p–norm of the Riesz transforms can be taken to be independent of the dimension n.

Exercise 8.2.4.
Use Proposition 8.6 and the previous exercise to prove that $\|R_j\|_p = H_p$.

Exercise 8.2.5.
Use the representation of R_j in terms of directional Hilbert transforms to prove that for all $1 < p < \infty$

$$\|Rf\|_p \leq C_p\|f\|_p$$

where C_p is independent of the dimension n and as above,

$$Rf(x) = (R_1 f(x), \dots, R_n f(x)).$$

Remark 8.1. Pichorides' result can be found in [Pi]. The first proof showing that the L^p–norms of the Riesz transforms are independent of the dimension was given by Stein [St2] using a corresponding result for Littlewood–Paley square functions. This is discussed in Chapter 10 below. The above argument using the method of rotations was given by Duoandikoetea and Rubio de Francia [DR]. The multiplier argument showing that the L^p norm is actually equal to the one for the Hilbert transform was given in Iwaniec and Martin [IM]. A different and very effective way of obtaining the above results is based on probabilistic argument using methods developed by Burkholder [Bu1] to prove sharp L^p–inequalities for martingales. This approach can be found in [BW] and [BM]. The probabilistic argument, which will not be presented here, has several advantages. First, not only does it give the sharp constant for the single Riesz transform R_j but it also gives better constants than those obtained above for the operator R and it extends to the truly infinite dimensional case of the Riesz transforms for the Orenstein-Uhlenbeck operator; the analogue of the Laplacian in Wiener space. In addition, the probabilistic arguments can be applied to obtain information on the L^p–norms of singular integrals of even kernels where the Calderón–Zygmund method of rotations does not apply. One particular important example of this situation is the Beurling-Ahlfors operator treated in Section 8.4 below. The reader interested in this and related connections, should see [Iw1], [Iw2] and references given there.

The question of whether the Riesz transforms have weak–type constants independent of the dimension remains an interesting and very challenging **open**

problem. Also **open** is the problem of finding the best L^p–constant for the Riesz transforms of order two R_j^2 and R_jR_k. The method of rotations used above does not apply since both of these operators are singular integrals of even kernels. It does follow from Exercise 8.2.4 above that $\|R_j^2 f\|_p \leq H_p^2\|f\|_p$ with the same estimate for R_jR_k. This constant, which is also independent of the dimension, does not however have the correct behavior in p (by Theorem 6.8, Chapter 6) as it is $O(p-1)^2$ as $p \to \infty$ and $O(1/(p-1))^2$ as $p \to 1$. The probabilistic argument in Bañuelos and Wang [BW] proves that $\|R_j^2 f\|_p \leq (p-1)\|f\|_p$ with a similar estimate for R_jR_k. At this point there is no analytic proof of this estimate. Notice however that by Exercise 8.2.1 this last inequality is not sharp even at $p = 2$.

8.3 The Cauchy–Riemann Equations in $\mathbb{R}^{n+1}_+$

Recall that by Exercise 6.1.12, Chapter 6, the function $u_f(x, y)$ in the upper half space

$$\mathbb{R}^{n+1}_+ = \{(x, y) \colon x \in \mathbb{R}^n,\ y > 0\}$$

defined by

$$u_f(x, y) = \int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi} e^{-2\pi|\xi|y} \hat{f}(\xi)\, d\xi, \tag{8.22}$$

for $f \in L^2(\mathbb{R}^n)$, is harmonic with boundary values f. That is,

$$\Delta u_f(x, y) = \frac{\partial^2 u}{\partial y^2}(x, y) + \sum_{j=1}^{n} \frac{\partial^2 u}{\partial x_j^2}(x, y) = 0$$

and $u_f(x, y) \to f(x)$ for almost every $x \in \mathbb{R}^n$ as $y \to 0$. In fact, by the Fourier inversion formula,

$$u_f(x, y) = P_y * f(x) \tag{8.23}$$

where $P_y(x)$ is the Poisson kernel

$$P_y(x) = \frac{C_n y}{(|x|^2 + y^2)^{(n+1)/2}}, \quad C_n = \frac{\Gamma(\frac{n+1}{2})}{\pi^{(n+1)/2}} = \frac{1}{y^n} P(x/y) \tag{8.24}$$

with

$$\hat{P}_y(\xi) = e^{-2\pi|\xi|y}.$$

If $n = 1$ then

$$P_y(x) = \frac{1}{\pi} \frac{y}{x^2 + y^2} \tag{8.25}$$

and if we use the complex notation of $z = x+iy$ for points in the upper half space $\mathbb{R}^2_+$ we see that

$$P_y(x) = \Re \frac{i}{\pi z}. \tag{8.26}$$

Let

$$Q_y(x) = \Im \frac{i}{\pi z} = \frac{1}{\pi} \frac{x}{x^2+y^2}. \tag{8.27}$$

Exercise 8.3.1.
Let $f \in L^p(\mathbb{R})$, $1 \leq p < \infty$, and set

$$F(z) = P_y * f(x) + iQ_y * f(x)$$

Prove that F is the unique analytic function in R^2_+ vanishing at infinity, that is, with

$$\lim_{y\to\infty} \sup_{x\in\mathbb{R}} |F(x+iy)| = 0,$$

and with

$$\lim_{y\to 0} \Re F(x+iy) = f(x),$$

for almost every $x \in \mathbb{R}$.

From this exercise we see that the conjugate harmonic function of $u_f(z)$ in $\mathbb{R}^2_+$ is given by

$$\widetilde{u}_f(z) = Q_y * f(x)$$

and therefore the pair $\{u_f, \widetilde{u}_f\}$ satisfies the Cauchy–Riemann equations

$$\begin{cases} \dfrac{\partial u}{\partial x} & = \dfrac{\partial \widetilde{u}}{\partial y} \\ \dfrac{\partial u}{\partial y} & = -\dfrac{\partial \widetilde{u}}{\partial x}. \end{cases}$$

On the other hand we also have

$$\begin{aligned} \lim_{y\to 0} Q_y * f(x) &= \lim_{y\to 0} \frac{1}{\pi} \int_{\mathbb{R}} \frac{x-\widetilde{x}}{(x-\widetilde{x})^2+y^2} f(\widetilde{x})\, d\tilde{x} \\ &= \lim_{\varepsilon\to 0} \frac{1}{\pi} \int_{\{|x-\widetilde{x}|>\varepsilon\}} \frac{x-\widetilde{x}}{(x-\tilde{x})^2} f(\widetilde{x})\, d\tilde{x} \\ &= \lim_{\varepsilon\to 0} \frac{1}{\pi} \int_{\{|x-\widetilde{x}|>\varepsilon\}} \frac{f(\widetilde{x})}{x-\widetilde{x}}\, d\widetilde{x} = Hf(x). \end{aligned} \tag{8.28}$$

and this gives the classical interpretation of the Hilbert transform: it is given by the boundary values of the conjugate harmonic function of the harmonic function with boundary values f. The $L^p(\mathbb{R})$–inequality for H was first proved by M. Riesz [Ri] in 1927 using this connection to analytic functions. We strongly recommend the beautiful survey article by M. L. Cartwright [Ca] for a historical account of the developments of M. Riesz's proof. The description of the exchange of letters between Hardy and M. Riesz on this subject is particularly interesting. The next exercise outlines M. Riesz original proof, as discussed in [Ca].

Exercise 8.3.2.
Let $f \in \mathcal{S}(\mathbb{R}^n)$ and as in Exercise 8.3.1 set $F(z) = P_y * f(x) + iQ_y * f(x) = u_f(x) + i\widetilde{u}_f(z)$. Use the Cauchy integral formula to prove that for any $y > 0$ and any even integer m

$$\int_{\mathbb{R}} (F(x+iy))^m \, dx = 0 \tag{8.29}$$

Set the real part of the left hand side of (8.29) equal to zero to show that

$$\begin{gathered}\int_{\mathbb{R}} \widetilde{u}_f^m(x+iy)\,dx - \tbinom{m}{2}\int_{\mathbb{R}} \widetilde{u}_f^{m-2}(x+iy)u_f^2(x+iy)\,dx \\ +\tbinom{m}{4}\int_{\mathbb{R}} \widetilde{u}_f^{m-4}(x+iy)u_f^4(x+iy)\,dx - \cdots \pm \int_{\mathbb{R}} u_f^m(x+iy)\,dx \\ = 0.\end{gathered}$$

This gives that

$$\begin{aligned}&\int_{\mathbb{R}} \widetilde{u}_f^m(x+iy)\,dx \\ &\quad\le \binom{m}{2}\int_{\mathbb{R}} \widetilde{u}_f^{m-2}(x+iy)u_f^2(x+iy)\,dx \\ &\qquad +\binom{m}{6}\int_{\mathbb{R}} \widetilde{u}_f^{m-6}(x+iy)u_f^6(x+iy)\,dx + \cdots\end{aligned} \tag{8.30}$$

Use Hölders inequality and (8.30) to conclude that

$$\int_{\mathbb{R}} \widetilde{u}_f^m(x+iy)\,dx \le C_m \int_{\mathbb{R}} u_f^m(x+iy)\,dx \le C_m \int_{\mathbb{R}} f^m(x)\,dx, \tag{8.31}$$

where C_m depends only on m. (8.31) proves the result for any p which is an even integer. The other cases of $2 \le p < \infty$ are obtain by interpolation and the case $1 < p < 2$ is done by duality.

We now wish to have a similar interpretation for the Riesz transforms in $\mathbb{R}^{n+1}_+$ and in particular, to explore the possibility of having Cauchy–Riemann equations in $\mathbb{R}^{n+1}_+$. First we define the functions

$$Q^j_y(x) = C_n \frac{x_j}{(|x|^2 + y^2)^{(n+1)/2}}, \quad j = 1, \ldots, n$$

where C_n is as in the Poisson kernel and notice that when $n = 1$ this just reduces to the function $Q_y(x)$ defined in (8.27). By Exercise 5.1.14 of Chapter 5, for any $f \in L^2(\mathbb{R}^n)$ the functions

$$u_j(x, y) = Q^j_y * f(x) \quad j = 1, \ldots, n$$

are harmonic in $\mathbb{R}^{n+1}_+$. If $n = 1$ this once again reduces to the $\widetilde{u}$ discussed above. As before,

$$\begin{aligned} \lim_{y\to 0} u_j(x, y) &= \lim_{y\to 0} C_n \int_{\mathbb{R}^n} \frac{\widetilde{x}_j}{(|\widetilde{x}|^2 + y^2)^{(n+1)/2}} f(x - \widetilde{x})\, d\widetilde{x} \qquad (8.32) \\ &= \lim_{\varepsilon\to 0} C_n \int_{\{|\widetilde{x}|>\varepsilon\}} \frac{\widetilde{x}_j}{|\widetilde{x}|^{n+1}} f(x - \widetilde{x})\, d\tilde{x} = R_j f(x), \end{aligned}$$

exactly as in (8.28).

Exercise 8.3.3.
Let $f \in L^2(\mathbb{R}^n)$ and let $f_j = R_j f$ be its Riesz transforms which are also in $L^2(\mathbb{R}^n)$. Let $u_j(x, y) = P_y * f_j(x)$. Prove that $u_j(x, y) = Q^j_y * f(x)$.

Theorem 8.8. *Let $f_0, f_1, \ldots, f_n \in L^2(\mathbb{R}^n)$ and set $u_0(x, y) = P_y * f_0(x)$, $u_1(x, y) = P_y * f_1(x), \ldots, u_n(x, y) = P_y * f_n(x)$. Then $f_j(x) = R_j f_0(x)$, $j = 1, \ldots n$, if and only if the Generalized Cauchy Riemann Equations*

$$\begin{cases} \displaystyle\sum_{j=0}^{n} \frac{\partial u_j}{\partial x_j} = 0, \\ \displaystyle\frac{\partial u_j}{\partial x_k} = \frac{\partial u_k}{\partial x_j}, \quad j \neq k, \end{cases} \qquad (8.33)$$

where $x_0 = y$, are satisfied. That is, $f_j(x) = R_j f(x)$ if and only if the $(n+1) \times (n+1)$-matrix of partial derivatives $\left(\frac{\partial u_j}{\partial x_k}\right)$ is symmetric of trace zero.

Proof. Suppose $f_j = R_j f$. Then $\widehat{f_j}(\xi) = \frac{i\xi_j}{|\xi|}\hat{f}(\xi)$ and by (8.22),

$$u_j(x,y) = \int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi} e^{-2\pi|\xi|y} \frac{i\xi_j}{|\xi|}\hat{f}(\xi)\,d\xi.$$

Differentiating under the integral sign we see that

$$\begin{aligned}\frac{\partial u_j}{\partial x_j}(x,y) &= \int_{\mathbb{R}^n} -2\pi i\xi_j e^{-2\pi i x\cdot\xi} e^{-2\pi|\xi|y} \frac{i\xi_i}{|\xi|}\widehat{f}(\xi)\,d\xi \\ &= 2\pi \int_{\mathbb{R}^n} \frac{|\xi_j|^2}{|\xi|} e^{-2\pi i x\cdot\xi} e^{-2\pi|\xi|y}\widehat{f}(\xi)\,d\xi.\end{aligned}$$

Thus

$$\sum_{j=1}^{n} \frac{\partial u_j}{\partial x_j} = 2\pi \int_{\mathbb{R}^n} |\xi|\widehat{f}(\xi) e^{-2\pi i x\cdot\xi} e^{-2\pi|\xi|y}\,d\xi = -\frac{\partial u_0}{\partial y}(x,y) = -\frac{\partial u_0}{\partial x_0}(x,y)$$

or

$$\sum_{j=0}^{n} \frac{\partial u_j}{\partial x_j} = 0,$$

proving the first equality of (8.33). The second is proved exactly the same way.

Next, suppose that (8.33) holds. We would like to show that $\hat{f}_j(\xi) = \frac{i\xi_j}{|\xi|}\widehat{f}(\xi)$. By (8.22) again and differentiation we have

$$\frac{\partial u_0}{\partial x_j}(x,y) = -2\pi \int_{\mathbb{R}^n} i\widehat{f}(\xi)\xi_j e^{-2\pi i\xi\cdot x} e^{-2\pi|\xi|y}\,d\xi$$

and

$$\frac{\partial u_j}{\partial x_0}(x,y) = -2\pi \int_{\mathbb{R}^n} \widehat{f_j}(\xi)|\xi| e^{-2\pi i\xi\cdot x} e^{-2\pi|\xi|y}\,d\xi.$$

Since

$$\frac{\partial u_0}{\partial x_j} = \frac{\partial u_j}{\partial x_0}$$

we see that

$$-2\pi \int_{\mathbb{R}^n} i\widehat{f}(\xi)\xi_j e^{-2\pi i\xi\cdot x} e^{-2\pi|\xi|y}\,d\xi = -2\pi \int_{\mathbb{R}^n} \widehat{f_j}(\xi)|\xi| e^{-2\pi i\xi\cdot x} e^{-2\pi|\xi|y}\,d\xi.$$

Since the functions $\xi_j\widehat{f}(\xi)e^{-2\pi|\xi|y}$ and $|\xi|\widehat{f_j}(\xi)e^{-2\pi|\xi|y}$ are clearly in $L^2(\mathbb{R}^n)$, the Fourier inversion formula gives (as functions in $L^2(\mathbb{R}^n)$)

$$-2\pi i\xi_j\widehat{f}(\xi)e^{-2\pi|\xi|y} = -2\pi|\xi|\widehat{f_j}(\xi)e^{-2\pi|\xi|y}$$

which implies

$$\widehat{f_j}(\xi) = \frac{i\xi_j}{|\xi|}\widehat{f}(\xi),$$

as desired. □

8.4 The Beurling–Ahlfors Operator

Before we discuss the Beurling-Ahlfors operator we present some of the more standard applications of the Riesz transforms. First recall that if $f \in C_0^2(\mathbb{R}^n)$ then

$$\widehat{\frac{\partial^2 f}{\partial x_j \partial x_k}}(\xi) = -4\pi^2 \xi_j \xi_k \widehat{f}(\xi)$$

and

$$\widehat{\Delta f}(\xi) = -4\pi^2 |\xi|^2 \widehat{f}(\xi), \tag{8.34}$$

where

$$\Delta f(x) = \sum_{k=1}^n \frac{\partial^2 f}{\partial x_k^2}(x).$$

Thus

$$(\widehat{R_j R_k \Delta f})(\xi) = \frac{-\xi_j \xi_k}{|\xi|^2}\widehat{\Delta f}(\xi) = -4\pi^2 \xi_j \xi_k \widehat{f}(\xi) = \widehat{\frac{\partial^2 f}{\partial x_j \partial x_k}}(\xi),$$

or

$$\frac{\partial^2 f}{\partial x_j \partial x_k}(x) = -R_j R_k(\Delta f)(x).$$

This gives

Theorem 8.9. *Suppose $f \in C_0^2(\mathbb{R}^n)$. Then*

$$\left\| \frac{\partial^2 f}{\partial x_j \partial x_k} \right\|_p \le A_p \|\Delta f\|_p,$$

for $1 < p < \infty$.

Let us now identify $\mathbb{R}^2$ with the complex plane $\mathbb{C}$ via $z = x + iy$ and denote its Lebesgue measure $dx\,dy$ by $dm(z)$. Recall that the Cauchy-Riemann operators are defined by

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)$$

and

$$\frac{\partial}{\partial \overline{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

This gives

$$\Delta = 4\frac{\partial}{\partial z}\,\frac{\partial}{\partial \overline{z}} = 4\frac{\partial^2}{\partial z \partial \overline{z}}$$

and

$$i\widehat{\frac{\partial f}{\partial z}}(\xi) = \pi\overline{\xi}\widehat{f}(\xi) \text{ and } i\widehat{\frac{\partial f}{\partial \overline{z}}}(\xi) = \pi\xi\widehat{f}(\xi). \tag{8.35}$$

The Beurling-Ahlfors operator (two dimensional Hilbert transform) is the singular integral defined by

$$\begin{aligned} Bf(z) &= -\lim_{\varepsilon\to 0}\frac{1}{\pi}\int_{\{|\zeta|>\varepsilon\}} \frac{f(z-\zeta)}{\zeta^2}\,dm(\zeta) \\ &= \ -\lim_{\varepsilon\to 0}\frac{1}{\pi}\int_{\{|\zeta|>\varepsilon\}} \frac{\overline{\zeta^2}}{|\zeta|^4} f(z-\zeta)\,dm(\zeta) \\ &= \ \lim_{\varepsilon\to 0}\int_{\{|\zeta|>\varepsilon\}} \frac{\Omega(\zeta)}{|\zeta|^2} f(z-\zeta)\,dm(\zeta), \end{aligned}$$

with

$$\Omega(\zeta) = -\frac{1}{\pi}\frac{\overline{\zeta^2}}{|\zeta|^2}.$$

Thus the Beurling-Ahlfors operator, or transform, is a singular integral of the form given by Theorem 6.8 of Chapter 6 and as such it satisfies the inequalities

$$m\{x \in \mathbb{C} : |Bf(z)| > \lambda\} \leq \frac{C}{\lambda}\|f\|_1 \tag{8.36}$$

and

$$\|Bf\|_p \leq B_p\|f\|_p \tag{8.37}$$

with the constant

$$B_p \leq \begin{cases} C(p-1), & \text{for } 2 \leq p < \infty \\ C\frac{1}{p-1} & \text{for } 1 < p \leq 2, \end{cases} \tag{8.38}$$

where C is an absolute constant. The following conjecture has been of interest for many years.

Conjecture 1 (Iwaniec [Iw1])**.** *Let* $1 < p < \infty$*. Set*

$$p^* = \max\left(p, \frac{p}{p-1}\right)$$

and denote the operator norm of B *in* $L^p(\mathbb{C})$ *by* B_p*. Then*

$$B_p = (p^* - 1).$$

The above conjecture can be reformulated in terms of an L^p inequality between the Cauchy–Riemann operators defined above. We shall now explain this. First, notice that

$$\Omega(\xi) = \frac{P(\xi)}{|\xi|^2}$$

with

$$P(\xi) = -\frac{1}{\pi}\left(\xi_1^2 - \xi_2^2 - 2i\pi\xi_1\xi_2\right),$$

which is a harmonic homogeneous polynomial of degree 2. Thus by (8.21) we have

$$\widehat{Bf}(\xi) = \frac{\overline{\xi}^2}{|\xi|^2}\widehat{f}(\xi) = \frac{\overline{\xi}}{\xi}\widehat{f}(\xi). \tag{8.39}$$

This shows that the operator B is an isometry on $L^2(\mathbb{C})$ and that it can be written in terms of second order Riesz transforms as

$$Bf(x) = R_2^2 f(x) - R_1^2 f(x) + 2iR_1R_2 f(x).$$

If we define Δ^{-1} by $\widehat{\Delta^{-1}f}(\xi) = -\frac{1}{4\pi^2|\xi|^2}\widehat{f}(\xi)$ (we will discuss this type of operator in Chapter 9 in much more detail) we see that that the operator B satisfies

$$B = 4\frac{\partial^2}{\partial z^2}\Delta^{-1}. \tag{8.40}$$

Also, using (8.35) and (8.36) we have that

$$B \circ \frac{\partial}{\partial \overline{z}} = \frac{\partial}{\partial z}. \tag{8.41}$$

Thus the inequality (8.37) is equivalent to the following

Proposition 8.10. *For all smooth functions f on $\mathbb{C}$ of compact support*

$$\left\|\frac{\partial f}{\partial z}\right\|_p \le B_p \left\|\frac{\partial f}{\partial \overline{z}}\right\|_p, \tag{8.42}$$

for $1 < p < \infty$.

The conjecture above is equivalent the assertion that (8.42) should hold with $B_p = (p^* - 1)$. There has been a lot of interest on this conjecture primarily because of its many applications to regularity properties of quasiconformal mappings. We will not discuss this topic here but the interested reader can see the excellent lecture notes of Manfredi [Ma] and the classical book of Lehto and Virtanen [LV]. A weaker form of the above conjecture which will also have interesting applications (see Iwaniec [Iw1]) is the following

Open Problem 1. *Prove that*

$$\limsup_{p\to\infty} \frac{\|Bf\|_p}{\|f\|_p} \le 1.$$

We restrict ourselves here to a discussion of what is known, as of now, about conjecture 1. First if we take $p > 2$ and $0 < \theta < 1$, define the functions

$$f_\theta(z) = \begin{cases} z|z|^{-2\theta/p}, & \text{for } |z| < 1 \\ \overline{z}^{-1}, & \text{for } |z| \ge 1. \end{cases}$$

Exercise 8.4.1.
Prove that

$$\left|\frac{\partial f_\theta}{\partial z}\right| = \begin{cases} \left(1 - \frac{\theta}{p}\right)|z|^{-2\theta/p}, & |z| \le 1 \\ 0, & |z| > 1 \end{cases}$$

and that

$$\left|\frac{\partial f_\theta}{\partial \overline{z}}\right| = \begin{cases} \frac{\theta}{p}|z|^{-2\theta/p}, & |z| \le 1 \\ |z|^{-2}, & |z| > 1. \end{cases}$$

Exercise 8.4.2.
Prove, by integrating in polar coordinates, that

$$\frac{\left\|B\frac{\partial f_\theta}{\partial \overline{z}}\right\|_p}{\left\|\frac{\partial f_\theta}{\partial \overline{z}}\right\|_p} = \frac{\left\|\frac{\partial f_\theta}{\partial z}\right\|_p}{\left\|\frac{\partial f_\theta}{\partial \overline{z}}\right\|_p} = \left(\frac{(p-1)(p-\theta)^p}{(p-1)\theta^p + (1-\theta)p^p}\right)^{1/p}. \tag{8.43}$$

By (8.43), as $\theta \uparrow 1$, we see $\|B\|_p \geq p-1$ for $p > 2$. Thus for $p > 2$ the constant cannot be smaller than (p^*-1) and by duality the same is true for $1 < p < 2$. On the other hand, the following explicit estimate for the constant B_p was proved in Bañuelos and Wang [BW].

Theorem 8.11. *(i) For any $f : \mathbb{C} \to \mathbb{C}$, $f \in L^p(\mathbb{C})$, $1 < p \leq \infty$*

$$\|Bf\|_p \leq 4(p^*-1)\|f\|_p \tag{8.44}$$

(ii) If f is real valued, that is, $f : \mathbb{C} \to \mathbb{R}$, then

$$\|Bf\|_p \leq 2\sqrt{2}(p^*-1)\|f\|_p \tag{8.45}$$

The proof of this theorem is via probabilistic techniques using extensions of some martingale inequalities due to Burkholder [Bu1]. The inequalities above were improved by a factor of 2 by Nazarov and Volberg [NV] using the martingale inequalities of Burkholder applied to Haar martingales. More recently, Bañuelos and Hernández–Mendez [BM] used the showed that this improvement can also be obtained using the same martingale techniques employed by in [BW]. The reader can observe that the result, although the best known for arbitrary p, does not give the correct constant at $p = 2$. However, at present there is no analytic proof of this result and we state the following

Open Problem 2. *Give an analytic proof of Theorem 8.11.*

We believe such a proof may produce an improvement and perhaps the full conjecture. Traditionally the analytic approach to obtaining explicit information on constants for singular integrals has been the Calderón–Zygmund method of rotations discussed in Exercise 8.2.2 above. Indeed, this is perhaps the best and easiest method to obtain the best L^p–constants for the Riesz transforms, as we saw in Exercise 8.2.3. This method does not give good information for the Beurlin–Ahlfors operator since it is a singular integral of even kernel. Indeed, the best this method seems to give is the bound $4H_p^2$, where H_p is the constant for the Hilbert transform. Such a constant no longer has the correct behavior as $p \to \infty$ or as $p \to 1$. The probabilistic methods of [BW] also have lead to an analytic approach to the above problems and to (surprising) connections with another well known open problem in the calculus of variations. This topic is also far from the scope of these notes but we shall briefly present the connections here. At this point in the notes the reader has enough background to go and explore this on her/his own if she/he so desires. The probabilistic inequalities proved in [BW], from which Theorem 8.11 is follows, are obtained using the following (famous) function constructed by Burkholder in [Bu1]: For $(z, w) \in \mathbb{C} \times \mathbb{C}$ we define

$$U(z,w) = p\left(1 - \frac{1}{p^*}\right)^{p-1} (|w| - (p^*-1)|z|)\,(|z| + |w|)^{p-1}\,. \tag{8.46}$$

Set

$$V(z,w) = |w|^p - (p^*-1)^p|z|^p.$$

Burkholder [**?**] proved that, among other things, the function U satisfies

$$V(z,w) \leq U(z,w) \text{ for } (z,w) \in \mathbb{C} \times \mathbb{C}. \tag{8.47}$$

Thus Iwaniec's conjecture would follow from an affirmative answer to

Question 8.4.1.
Let $f \in C_0^\infty(\mathbb{C})$ and Bf be its Beurling-Ahlfors transform. Is it true that

$$\int_{\mathbb{C}} U(f(z),\, Bf(z))\, dA \leq 0\,? \tag{8.48}$$

Here, dA is area measure on $\mathbb{C}$.

This question first appear in print in [BL] although it had been first raised in several lectures earlier by Bañuelos and Wang.

Since B is characterized by (8.41), Question 8.4.1 can be stated without any reference to the Beurling-Ahlfors operator. Indeed, Question 8.4.1 is equivalent to

Question 8.4.2.
Let $f \in C_0^\infty(\mathbb{C})$. Is it true that

$$\int_{\mathbb{C}} U\left(\frac{\partial f}{\partial \bar{z}}, \frac{\partial f}{\partial z}\right)\, dA \leq 0\,?$$

It is interesting to test this "conjecture" with the "extremal" family of functions f_θ given above. By Exercise 8.4.1, we have

$$\int_{\mathbb{C}} U\left(\frac{\partial f_\theta}{\partial \bar{z}}, \frac{\partial f_\theta}{\partial z}\right)\, dA = \int_{|z|\leq 1} U\left(\frac{\partial f_\theta}{\partial \bar{z}}, \frac{\partial f_\theta}{\partial z}\right)\, dA$$

$$+ \int_{|z|>1} U\left(\frac{\partial f_\theta}{\partial \bar{z}}, \frac{\partial f_\theta}{\partial z}\right) dA \equiv I + II.$$

With $\alpha_p = p(1 - \frac{1}{p^*})^{p-1}$ we have

$$\begin{aligned} I &= 2\pi\alpha_p \int_0^1 \left[(1 - \tfrac{\theta}{p}) - (p-1)\tfrac{\theta}{p}\right] r^{-2\theta/p}(r^{-2\theta/p})^{p-1}\, r\, dr \\ &= 2\pi\alpha_p(1-\theta)\int_0^1 r^{1-2\theta}\, dr = \pi\alpha_p, \end{aligned}$$

and

$$\begin{aligned} II &= 2\pi\alpha_p \int_1^\infty \left[-(p-1)r^{-2}\right] (r^{-2})^{p-1}\, r dr \\ &= -2\pi\alpha_p(p-1)\int_1^\infty r^{1-2p}\, dr = -\pi\alpha_p. \end{aligned}$$

Hence, for all $0 < \theta < 1$, we have

$$\int_{\mathbb{C}} U\left(\frac{\partial f_\theta}{\partial \bar{z}}, \frac{\partial f_\theta}{\partial z}\right) dA \equiv 0.$$

So, putting the Iwaniec "extremal" into the Burkholder function has the surprising result that the parameters θ and p do not appear, and that moreover, it integrates to zero exactly. For various other related conjectures involving this and other functions of Burkholder, see [BM].

The advantage of working with U in place of V is that this function has several interesting additional properties that the function V does not have. As shown by Burkholder in [Bu1] the function U satisfies the property that for all $z, w, h, k \in \mathbb{C}$ with $|k| \leq |h|$, the mapping

$$t \mapsto U(z + th, w + hk) \tag{8.49}$$

is concave on $\mathbb{R}$. This is a key property in the proof of Burkholder's inequalities for martingales. However, this U is not the least majorant of V with this property. The smallest such function is given by (see [Bu1], p. 81)

$$\tilde{U}(z,w) = \begin{cases} V(z,w), & \text{if } |w| \leq (p^*-1)|z| \\ U(z,w), & \text{if } |w| > (p^*-1)|z| \end{cases}$$

for $2 \leq p < \infty$ and with U and V interchanged for $1 < p \leq 2$.

Exercise 8.4.3.
Prove that

$$\int_{\mathbb{C}} \tilde{U}\left(\frac{\partial f_\theta}{\partial \bar{z}}, \frac{\partial f_\theta}{\partial z}\right) dA = \pi\left[p(1-\tfrac{1}{p})^{p-1} - (p-1)^{p-1}\right] < 0.$$

Thus even if some of the above questions are false for U, they may still hold for the minimal $\widetilde{U}$.

The connection to the calculus of variations problem mentioned above arises as follows. First, by $M(n,\mathbb{R})$ we mean the set of all $n \times n$ matrices with real coefficients. We have the following definitions of convexity. (For more on this, see [Do].)

Definition 8.12.

(i) A function $g : M(n,\mathbb{R}) \to \mathbb{R}$ is convex if for all $A_1, A_2 \in M(n,\mathbb{R})$,

$$g(\lambda A_1 + (1-\lambda)A_2) \leq \lambda g(A_1) + (1-\lambda)g(A_2) \tag{8.50}$$

for every $\lambda \in [0,1]$.

(ii) The function g is said to be *rank-one convex* if (8.50) holds for all $A_1, A_2 \in M(n,\mathbb{R})$ with rank of $\{A_1 - A_2\}$ less than or equal to 1.

(iii) The function g is said to be *quasiconvex* if

$$g(A) \leq \frac{1}{|\Omega|}\int_\Omega g(A + D\phi(x))\,dx \tag{8.51}$$

for every bounded domain $\Omega \subset \mathbb{R}^n$, for every $A \in M(n,\mathbb{R})$ and for all $\phi \in C_0^1(\Omega;\mathbb{R}^n)$, C^1 functions with values in $\mathbb{R}^n$ and compact support in Ω. Here, $D\phi$ is the total differential of $\phi = (\phi_1, \ldots, \phi_n)$:

$$D\phi = \begin{pmatrix} D_1\phi_1 & \cdots & D_n\phi_1 \\ \vdots & \ddots & \vdots \\ D_1\phi_n & \cdots & D_n\phi_n \end{pmatrix}.$$

Clearly, convexity implies rank–one convexity.

Exercise 8.4.4.
Prove that convexity implies quasinconvexity.

It is also known, but much much more difficult to prove ([Do]) that quasiconvexity implies rank–one convexity. The following, however, is a well known open problem.

Question 8.4.3.
Let $n = 2$. Does rank-one convexity imply quasiconvexity?

The functions U and $\widetilde{U}$ give rise to rank-one convex functions for $n = 2$ using (8.49). Indeed, define $\beta \colon M(2, \mathbb{R}) \to \mathbb{C} \times \mathbb{C}$ by $\beta\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = (z, w)$, where $z = \frac{1}{2}\big((a+d)+i(c-b)\big)$, $w = \frac{1}{2}\big((a-d)+i(c+b)\big)$, and define $U^{\#} \colon M(2, \mathbb{R}) \to \mathbb{R}$ by

$$U^{\#} = -U \circ \beta, \tag{8.52}$$

with a similar definition for $\widetilde{U}^{\#}$.

Exercise 8.4.5.
Use (8.49) to prove that the function $U^{\#}$ is rank–one convex.

If $U^{\#}$ were also quasiconvex, then by taking $A = 0$, we would have that

$$0 = U^{\#}(0) \leq \int_{\operatorname{supp}\phi} U^{\#}(D\phi)\, dx. \tag{8.53}$$

Taking $\phi \equiv (\Re f, \Im f)$, gives $U^{\#}(D\phi) = -U(\frac{\partial f}{\partial \bar{z}}, \frac{\partial f}{\partial z})$. Thus (8.53) implies

$$\int_{\mathbb{C}} U\left(\frac{\partial f}{\partial \bar{z}}, \frac{\partial f}{\partial z}\right) dA \leq 0. \tag{8.54}$$

Thus an affirmative answer to Question 8.4.3 would positively answer Question 8.4.2. However, even if the answer to Question 8.4.3 were negative in general, it could still be possible that $U^{\#}$, or $\widetilde{U}^{\#}$, is quasiconvex. In any case, the resolution of the question below would be very interesting. In the positive, it would imply the conjecture of Iwaniec, and in the negative it would settle Question 8.4.3.

Question 8.4.4.
Is $U^{\#}$ (or $\widetilde{U}^{\#}$) quasiconvex?

The paper [BL] also contains a probabilistic study, using martingales, of the version of the Beurling-Ahlfors operator on $\mathbb{R}^n$ and gives the best known estimates for their L^p–norms. This probabilistic representation is also beyond the scope of these lecture notes. However, we do have the tools to present the analytic definition of this operator and we end this chapter by doing so. This operator was first introduce by Danolson and Sulivan [DS] in their study of quasiconformal four–manifolds. As they showed, the Beurling-Ahlfors operator on $\mathbb{R}^n$, $n \geq 3$, acts on differential forms. We begin its description by first recalling a few basic facts about differential forms in $\mathbb{R}^n$ and their L^p–spaces. We assume that the reader is familiar with the the basic notions of linear algebra and differential forms. An excellent reference for the material on differential forms is W. Rudin [Ru1]. Now for $0 \leq k \leq n$, let $i_1, i_2, \dots, i_k$ be integers satisfying $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ and let I be the ordered k–tuple $\{i_1, i_2, \dots, i_k\}$. We call this k–tuple a k–index. For a fixed $0 \leq k \leq n$, the set of all k–indices is denoted by $\mathcal{I}^{n,k}$. We denote the basic k–forms by

$$dx^I = dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_k}$$

for $I = \{i_1, \dots, i_k\}$. Since the cardinality of the set $\mathcal{I}^{n,k}$ is $\nu = \binom{n}{k}$ this is also the number of distinct basic k–forms and every k–form can be represented in terms of these basic k–forms. Fix the standard basis $\{e^1, e^2, \dots, e^n\}$ and for each $I = \{i_1, i_2, \dots, i_k\} \in \mathcal{I}^{n,k}$, let $e^I = e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_k}$. This is a k–vector. We now denote the complex vector space of k-forms spanned by the e^I as I runs over all the $\mathcal{I}^{n,k}$ by Λ^k. That is,

$$\Lambda^k = \operatorname{span}\{e^I \colon I \in \mathcal{I}^{n,k}\}. \tag{8.55}$$

Λ^k inherits the usual complex Hermitian product: for $\alpha = \sum_I \alpha_I e^I$ and $\beta = \sum_I \beta_I e^I$, where $\alpha_I, \beta_I \in \mathbb{C}$, define

$$\langle \alpha, \beta \rangle_k = \sum_I \alpha_I \overline{\beta}_I.$$

Notice that $\Lambda^0 = \mathbb{C}$ and $\Lambda^1 = \mathbb{C}^n$. Now set $\Lambda = \bigoplus_{k=0}^n \Lambda^k$.

Exercise 8.4.6.
Prove that Λ is a vector space of dimension 2^n which is also a graded algebra with respect to the exterior product $\wedge$. That is to say, for $\alpha \in \Lambda^k$ and $\beta \in \Lambda^m$, we have $\alpha \wedge \beta \in \Lambda^{k+m}$ and Λ has the natural inner product obtained from its direct sum decomposition.

We will need the Hodge star operator, $\star \colon \Lambda \to \Lambda$, defined as follows. Let $N = \{1, 2, \dots, n\}$. For a basis element $e^I \in \Lambda^k$, we set $\star e^I = \sigma(I) e^{N \setminus I}$, where

$$\sigma(I) = \sigma(\{i_1, \dots i_k\}) \equiv \operatorname{sign} \begin{pmatrix} 1 & 2 & \dots & k \\ i_1 & i_2 & \dots & i_k \end{pmatrix},$$

and $N \setminus I$ is set difference with the convention that the elements of $N \setminus I$ are increasing. The operator $\star$ is then extended to all of Λ^k by linearity, and finally to all of Λ via its direct sum decomposition.

Exercise 8.4.7.
Prove that $\star$ is characterized by the identity

$$\langle \alpha, \beta \rangle e^1 \wedge \cdots \wedge e^n = \overline{\beta} \wedge \star\alpha$$

for all $\beta \in \Lambda$ and that $\star \star |_{\Lambda^k} = (-1)^{k(n-k)}\,\mathrm{Id}$.

Next define the class of complex-valued k-forms with L^p coefficients by

$$L^p(\mathbb{R}^n, \Lambda^k) = \Big\{ \alpha = \sum_{I \in \mathcal{I}^{n,k}} \alpha_I(x)\, dx^I : \alpha_I(x) \in L^p(\mathbb{R}^n) \Big\}, \tag{8.56}$$

and the (semi-)norm on $L^p(\mathbb{R}^n, \Lambda^k)$ by

$$\|\alpha\|_{L^p(\mathbb{R}^n, \Lambda^k)} = \left\{ \int_{\mathbb{R}^n} \Big(\sum_I |\alpha_I(x)|^2 \Big)^{p/2} dx \right\}^{1/p}. \tag{8.57}$$

We may now define $L^p(\mathbb{R}^n, \Lambda) = \bigoplus_{k=0}^n L^p(\mathbb{R}^n, \Lambda^k)$ and

$$\|\omega\|_{L^p(\mathbb{R}^n, \Lambda)} = \sum_{k=0}^n \|\omega_k\|_{L^p(\mathbb{R}^n, \Lambda^k)}, \tag{8.58}$$

where $\omega = \omega_0 + \cdots + \omega_n$ is such that $\omega_k \in \Lambda^k$.

We should mention here that while (8.58) is a natural choice for the norm, it is not the only norm we could put on this space. Indeed, since by Exercise 8.4.6, $\dim \Lambda = 2^n$, we can identify $L^p(\mathbb{R}^n, \Lambda)$ with the space of L^p functions in $\mathbb{R}^n$ with values in $\mathbb{R}^{2^n}$, that is with $L^p(\mathbb{R}^n, \mathbb{R}^{2^n})$ with the norm of the latter given by

$$\|\vec{f}\|_p = \left\{ \int_{\mathbb{R}^n} \Big(\sum_{k=0}^{2^n} |f_k(x)|^2 \Big)^{p/2} dx \right\}^{1/p}.$$

It is enough, for the purpose of proving the L^p–inequalities below, to restrict ourselves to differential forms which have coefficients which are C^∞ functions of compact support. That is, we may restrict ourselves to differential forms in

$C_0^\infty(\mathbb{R}^n, \Lambda^k) \cap L^2(\mathbb{R}^n, \Lambda^k)$. First recall that the exterior derivative d takes k-forms to $(k+1)$-forms. We define δ as the adjoint of d with respect to the inner product

$$(\alpha, \beta) \equiv \int_{\mathbb{R}^n} \langle \alpha(x), \beta(x) \rangle \, dx = \int_{\mathbb{R}^n} \overline{\beta} \wedge \star\alpha.$$

Note that in fact the statement $(d\alpha, \beta) = -(\alpha, \delta\beta)$ is nothing more than integration by parts.

Exercise 8.4.8.
Prove that

$$\delta = (-1)^{k(n-k)} \star d \star$$

and hence that δ takes k-forms to $(k-1)$-forms.

Our next task is to write the operators d and δ in terms of Fourier multipliers very much as we did above with $\frac{\partial}{\partial x_i}$ and the Laplacian. It will be convenient for our discussion here to denote the Fourier transform by $\mathcal{F}$, as in Chapter 6. Now, given a differential form

$$\alpha = \sum_{I \in \mathcal{I}^{n,k}} \alpha_I(x) \, dx^I$$

we define its Fourier transform by

$$\mathcal{F}\alpha = \sum_{I \in \mathcal{I}^{n,k}} \mathcal{F}\alpha_I(x) \, dx^I.$$

That is, the Fourier transform acts on differential forms by acting on its coefficients.

Exercise 8.4.9.
Use the identity $(\mathcal{F}\frac{\partial f}{\partial x_j})(\xi) = -2\pi i \xi_j (\mathcal{F}f)(\xi)$ to prove that

$$\left(\mathcal{F}(d\alpha)\right)(\xi) = -2\pi i \xi \wedge \left(\mathcal{F}\alpha\right)(\xi). \tag{8.59}$$

Now fix an ordered basis of Λ^k. For each vector $\xi = \xi_1 e^1 + \cdots + \xi_n e^n \in \mathbb{R}^n$, let $[\xi]$ be the matrix of the linear mapping $\omega \mapsto \xi \wedge \omega$. With this notation we can write (8.59) as

$$d = \mathcal{F}^{-1}[-2\pi i \xi]\mathcal{F}. \tag{8.60}$$

Exercise 8.4.10.
Using (8.60) and the fact δ is the adjoint of d and that $\mathcal{F}$ is an isometry on $C_0^\infty(\mathbb{R}^n, \Lambda^k) \cap L^2(\mathbb{R}^n, \Lambda^k)$, prove that

$$\delta = \mathcal{F}^{-1}[-2\pi i\xi]^t\mathcal{F}, \tag{8.61}$$

where $[-2\pi i\xi]^t$ is the transpose of the matrix $[-2\pi i\xi]$.

In the same way, we define the Laplacian, $\boldsymbol{\Delta}$, acting on differential forms by the action of the ordinary Laplacian acting on the coefficients of the forms. That is

$$\boldsymbol{\Delta}\alpha = \sum_{I\in\mathcal{I}^{n,k}} \Delta\alpha_I(x)\,dx^I \tag{8.62}$$

where Δ is just the Laplacian on functions. In the same way the Laplacian inverse, $\boldsymbol{\Delta}^{-1}$, is the operator defined by

$$\boldsymbol{\Delta}^{-1}\alpha = \sum_{I\in\mathcal{I}^{n,k}} \Delta^{-1}\alpha_I(x)\,dx^I, \tag{8.63}$$

where as above Δ^{-1} is defined in terms of the Fourier transform by

$$\mathcal{F}(\Delta^{-1}f)(\xi) = \frac{-1}{4\pi^2|\xi|^2}(\mathcal{F}f)(\xi), \tag{8.64}$$

as in (8.34).

Exercise 8.4.11.
Prove that $\boldsymbol{\Delta} = d\delta + \delta d$ as operators on $C_0^\infty(\mathbb{R}^n, \Lambda^k) \cap L^2(\mathbb{R}^n, \Lambda^k)$.

Definition 8.13. The Beurling–Ahlfors operator for $\omega \in C_0^\infty(\mathbb{R}^n, \Lambda) \cap L^p(\mathbb{R}^n, \Lambda)$ is defined by

$$S\omega = (d\delta - \delta d) \circ \boldsymbol{\Delta}^{-1}\omega. \tag{8.65}$$

Exercise 8.4.12.
Prove that for $n = 2$ and $\omega = f_1dx + f_2dy \in \Lambda^1$, the operator defined by (8.65) reduces to the Bf, with $f = f_1 + if_2$ where B is the Beurling-Ahlforms operator as defined above.

Our next task is to compute the Fourier multiplier of $S_k = S|_{L^p(\mathbb{R}^n,\Lambda^k)}$ using some of the tools developed in the above exercises. Given $\omega \in C_0^\infty(\mathbb{R}^n, \Lambda) \cap L^p(\mathbb{R}^n, \Lambda)$ we solve “Poisson’s” equation

$$\boldsymbol{\Delta}\varphi = (d\delta + \delta d)\varphi = \omega. \tag{8.66}$$

This solution can be written formally as

$$\varphi = \mathcal{F}^{-1}|\xi|^{-2}\mathcal{F}\omega \tag{8.67}$$

or, as we shall see in the next chapter, as

$$\varphi(x) = \frac{\Gamma\big((n-2)/2\big)}{4\pi^{n/2}} \int_{\mathbb{R}^n} \frac{\omega(x)}{|x-y|^{n-2}}\, dy$$

where the integration of differential forms is done by integrating the coefficients and using linearity. From (8.66) and Exercise 8.4.11, we get the Hodge decomposition

$$\omega = d(\delta\varphi) + \delta(d\varphi) = d\alpha + \delta\beta$$

where α is a $(k+1)$–form and β is a $(k-1)$–form with the convention that $\Lambda^{-1} = \Lambda^{(n+1)} = \{0\}$. From (8.65) we have

$$S_k\omega = d\alpha - \delta\beta = d(\delta\varphi) - \delta(d\varphi) = (d\delta - \delta d)\varphi = (d\delta - \delta d) \circ \mathbf{\Delta}^{-1}\omega. \tag{8.68}$$

Because the Hodge decompositions of a 0-form ω_0 and an n-form ω_n are $\omega_0 = \delta\beta_0$ and $\omega_n = d\alpha_n$, respectively, we see $S_0 = -\mathrm{Id}$ and $S_n = \mathrm{Id}$.

Exercise 8.4.13.
Use Exercise 8.4.8 and the fact that $\star$ commutes with $\mathbf{\Delta}$ and $\mathbf{\Delta}^{-1}$ to prove that

$$\star\, S_k = -S_{n-k}\star \tag{8.69}$$

and that

$$\|S_k\|_{L^p(\mathbb{R}^n,\Lambda^k)} = \|S_{n-k}\|_{L^p(\mathbb{R}^n,\Lambda^{n-k})}. \tag{8.70}$$

From (8.60), (8.61), and (8.67) we see that as an operator on $C_0^\infty(\mathbb{R}^n, \Lambda^k) \cap L^2(\mathbb{R}^n, \Lambda^k)$,

$$S_k = \mathcal{F}^{-1}\left(\frac{[\xi]^t[\xi] - [\xi][\xi]^t}{|\xi|^2}\right)\mathcal{F} \equiv \mathcal{F}^{-1}M(\xi)\mathcal{F}. \tag{8.71}$$

Exercise 8.4.14.
Using the fact that the rows and columns of $[\xi]$ are indexed by the set of ordered k-indices $\mathcal{I}^{n,k}$, show that the entries of $M(\xi)$ are given by

$$[M(\xi)]_{I,J} = \begin{cases} \sum_{j\notin I} \frac{\xi_j^2}{|\xi|^2} - \sum_{i\in I} \frac{\xi_i^2}{|\xi|^2}, & \text{if } I = J \\ -2\xi_i\xi_j/|\xi|^2, & \text{if } I\setminus J = \{i\} \\ & \text{and } J\setminus I = \{j\} \\ 0, & \text{otherwise.} \end{cases} \tag{8.72}$$

Recalling that the composition of the Riesz transforms R_i and R_k is given by $R_iR_j = \mathcal{F}^{-1}\left(\frac{-\xi_i\xi_j}{|\xi|^2}\right)\mathcal{F}$, we see that the restriction of S to $L^2(\mathbb{R}^n, \Lambda^k)$ can be represented as the matrix of second ordered Riesz transforms

$$[S_k]_{I,J} = \begin{cases} \sum_{i\in I} R_i^2 - \sum_{j\notin I} R_j^2, & \text{if } I = J \\ 2R_iR_j, & \text{if } I\setminus J = \{i\} \\ & \text{and } J \setminus I = \{j\} \\ 0, & \text{otherwise.} \end{cases} \tag{8.73}$$

We are now able to represent the full operator S as $S = S_0 \otimes \cdots \otimes S_n$, corresponding to the matrix of operators which has $[S_k]$, $k = 0, \ldots, n$ along its diagonal. By the direct sum decomposition of $L^p(\mathbb{R}^n, \Lambda)$ we have

$$\|S\|_{L^p(\mathbb{R}^n,\Lambda)} = \max_{0\le k\le n} \|S_k\|_{L^p(\mathbb{R}^n,\Lambda^k)}.$$

Since by Exercise 8.4.13,

$$\|S_k\|_{L^p(\mathbb{R}^n,\Lambda^k)} = \|S_{n-k}\|_{L^p(\mathbb{R}^n,\Lambda^{n-k})} \tag{8.74}$$

we see that in estimating the norm of S it is enough to estimate

$$\max_{0\le k\le \widehat{n}} \|S_k\|_{L^p(\mathbb{R}^n,\Lambda^k)}, \tag{8.75}$$

where

$$\widehat{n} = \begin{cases} n/2, & \text{if } n \text{ is even} \\ (n-1)/2, & \text{if } n \text{ is odd.} \end{cases}$$

Exercise 8.4.15.
Use (8.73) to prove that for any $0 \le k \le \widehat{n}$, $\|S_k\|_{L^p(\mathbb{R}^n,\Lambda^k)} \le C_p$ for any $1 < p < \infty$. Furthermore, $C_p \le C_{k,n}p^2$ for $p > 2$ where $C_{k,n}$ is independent of p.

Chapter 9

Fractional Integration

The Riesz potentials are the basic and natural generalizations of the classical Green's potentials. An effective and very natural way to introduce these operators is via the heat kernel.

9.1 Definitions and Boundedness

Recall that the heat kernel in $\mathbb{R}^n$ is given by

$$H_t(x) = \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2/(4t)}$$

with Fourier transform, (Chapter 6, Exercise 6.1.2),

$$\widehat{H}_1(\xi) = e^{-4\pi^2|\xi|^2}.$$

For any $0 < \alpha < n$ we define the operator, called the fractional integral of f, by

$$I_\alpha(f)(x) = \frac{1}{\Gamma(\alpha/2)} \int_0^\infty t^{\alpha/2-1} H_t * f(x)\, dt, \tag{9.1}$$

for $f \in L^1(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$.

We first note that the integral defining $I_\alpha(f)$ is absolutely convergent. To see this observe that for any $1 \le p < \infty$, Jensen's inequality gives

$$|H_t * f(x)|^p = \left| \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-|y|^2/(4t)} f(y-x)\, dy \right|^p \le \frac{1}{(4\pi t)^{n/2}} \|f\|_p^p$$

which is equivalent to

$$|H_t * f(x)| \le \frac{1}{(4\pi t)^{n/(2p)}} \|f\|_p. \tag{9.2}$$

Also recall that if we set

$$f^*(x) = \sup_{t>0} |H_t * f(x)|$$

we have

$$\|f^*\|_p \le C_p \|f\|_p, \tag{9.3}$$

for $1 < p \le \infty$ and

$$m\{x \in \mathbb{R}^n \colon f^*(x) > \lambda\} \le \frac{C}{\lambda} \|f\|_1. \tag{9.4}$$

From (9.2) and the fact that $0 < \alpha < n$, it follows that

$$\begin{aligned} |I_\alpha(f)(x)| &\le \frac{1}{\Gamma(\alpha/2)} \left(\int_0^1 t^{\alpha/2-1}\, dt \right) f^*(x) \\ &\quad + \frac{\|f\|_1}{\Gamma(\alpha/2)(4\pi)^{n/2}} \int_1^\infty t^{\alpha/2-n/2-1}\, dt \\ &= C_{n,\alpha} f^*(x) + C_{n,\alpha} \|f\|_1, \end{aligned} \tag{9.5}$$

This simple argument just presented to prove the absolute convergence of the integral defining I_α can be easily modified to prove the boundedness properties of the operators I_α. We first state and prove the theorem and then explain the reason for the restriction on the exponents.

Theorem 9.1 (Hardy–Littlewood–Sobolev)**.** *Let $0 < \alpha < n$ and set $\frac{1}{q} = \frac{1}{p} - \frac{\alpha}{n}$, for $1 \le p < \infty$. Then*

(i) $\|I_\alpha(f)\|_q \le A_{p,q,\alpha} \|f\|_p$, $1 < p < q < \infty$, *and*

(ii) $m\{x \in \mathbb{R}^n \colon |I_\alpha(f)(x)| > \lambda\} \le A \left(\frac{\|f\|_1}{\lambda} \right)^q$, $p = 1$.

Proof. Let $\delta > 0$ to be chosen later. As in (9.5) we have, using (9.1),

$$|I_\alpha(f)(x)| \le \frac{1}{\Gamma(\alpha/2)} \left(\int_0^\delta t^{\alpha/2-1}\, dt \right) f^*(x)$$

$$
\begin{aligned}
& + \frac{\|f\|_p}{\Gamma(\alpha/2)(4\pi)^{n/(2p)}} \int_\delta^\infty t^{\alpha/2-n/(2p)-1}\,dt \qquad (9.6)\\
= \;& C_{n,\alpha}\delta^{\alpha/2} f^*(x) + C_{n,\alpha}\delta^{\alpha/2-n/(2p)}\|f\|_p,
\end{aligned}
$$

where this time we used the fact that $\alpha - \frac{n}{p} < 0$. Picking

$$\delta = \left(\frac{\|f\|_p}{f^*}\right)^{2p/n}$$

to minimize the right hand side of (9.6) gives

$$|I_\alpha f(x)| \le C_{n,\alpha}\,(f^*)^{1-\alpha p/n}\,\|f\|_p^{\alpha p/n} = C_{n,\alpha}\,(f^*)^{p/q}\,\|f\|_p^{\alpha p/n}. \tag{9.7}$$

Thus for $1 < p < \infty$, (9.3) and (9.7) give

$$
\begin{aligned}
& \int_{\mathbb{R}^n} |I_\alpha f(x)|^q\,dx \le C_{n,p,\alpha}\left(\int_{\mathbb{R}^n} (f^*(x))^p\;dx\right)\|f\|_p^{\alpha pq/n}\\
\le \;& C_{n,p,\alpha}\left(\int_{\mathbb{R}^n} |f(x)|^p\,dx\right)\|f\|_p^{\alpha pq/n} = C_{p,\alpha,n}\left(\int_{\mathbb{R}^n} |f(x)|^p\,dx\right)^{q/p},
\end{aligned}
$$

which proves (i).

If $p = 1$, (9.4) and (9.7) give

$$
\begin{aligned}
& m\{x \in \mathbb{R}^n \colon |I_\alpha(f)(x)| > \lambda\}\\
\le \;& m\left\{x \in \mathbb{R}^n \colon f^*(x) > \left(\frac{\lambda}{C_{n,\alpha}}\|f\|_1^{\alpha/n}\right)^q\right\}\\
\le \;& C_{n,\alpha}\left(\frac{\|f\|_1^{\alpha/n}}{\lambda}\right)^q \|f\|_1 = C_{n,\alpha}\left(\frac{\|f\|_1}{\lambda}\right)^q,
\end{aligned}
$$

and this proves (ii). □

The restriction $\frac{1}{q} = \frac{1}{p} - \frac{\alpha}{n}$ arises from the scaling properties of the operator. Indeed, with $\tau_\delta f(x) = f(\delta x)$ as before, a simple change of variables shows that $\tau_{\delta^{-1}} I_\alpha \tau_\delta = \delta^{-\alpha} I_\alpha$. Since

$$\|\tau_\delta f\|_p = \delta^{-n/p}\|f\|_p, \tag{9.8}$$

we have that

$$\|\tau_{\delta^{-1}} I_\alpha \tau_\delta f\|_q = \delta^{-\alpha}\|I_\alpha f\|_q. \tag{9.9}$$

On the other hand that

$$\|\tau_{\delta^{-1}} I_\alpha \tau_\delta f\|_q = \delta^{n/q} \|I_\alpha(\tau_\delta f)\|_q. \tag{9.10}$$

Thus if we had $\|I_\alpha(f)\|_q \le A_{p,q}\|f\|_p$, (9.8) and (9.9) would imply that

$$\delta^{-\alpha}\|I_\alpha(f)\|_q = \|\tau_{\delta^{-1}} I_\alpha \tau_\delta f\|_q \le A_{p,q}\delta^{n/q-n/p}\|f\|_p$$

or

$$\|I_\alpha(f)\|_q \le A_{p,q}\delta^{\alpha-n/p+n/q}\|f\|_p.$$

Since $\delta > 0$ is arbitrary, this forces $\alpha - \frac{n}{p} + \frac{n}{q} = 0$, or $\frac{1}{q} = \frac{1}{p} - \frac{\alpha}{n}$.

We now show that $I_\alpha(f)$ can also be written as the convolution of f with the kernel $|x|^{-n+\alpha}$. The case $\alpha = 2$ is the classical Green's potential and the others are often referred to as the Riesz potentials. First, note that taking Fourier transforms of both sides of (8.1) gives

$$\begin{aligned}
\widehat{I_\alpha(f)}(\xi) &= \frac{1}{\Gamma(\alpha/2)}\widehat{f}(\xi)\int_0^\infty t^{\alpha/2-1}e^{-4\pi^2 t|\xi|^2}\,dt \\
&= \frac{1}{\Gamma(\alpha/2)}\widehat{f}(\xi)(4\pi^2|\xi|^2)^{-\alpha/2}\int_0^\infty t^{\alpha/2-1}e^{-t}\,dt \\
&= (4\pi^2|\xi|^2)^{-\alpha/2}\widehat{f}(\xi). \qquad (9.11)
\end{aligned}$$

Recalling that for any $f \in C_0^2(\mathbb{R}^n)$, $\widehat{(-\Delta)f}(\xi) = 4\pi^2|\xi|^2\widehat{f}(\xi)$, and formally defining the fractional power of $(-\Delta)$ via the Fourier transform by

$$\widehat{(-\Delta)}^{-\alpha/2} f(\xi) = (4\pi^2|\xi|^2)^{-\alpha/2}\widehat{f}(\xi),$$

we see that

$$I_\alpha(f)(x) = (-\Delta)^{-\alpha/2}f(x). \tag{9.12}$$

We next show that $I_\alpha(f)$ is also given, up to constants, as the convolution of f with the kernel $|x|^{-n+\alpha}$. To see this, apply Fubini's theorem to obtain

$$\begin{aligned}
I_\alpha(f)(x) &= \frac{1}{\Gamma(\alpha/2)}\int_{\mathbb{R}^n}\left(\int_0^\infty \frac{t^{\alpha/2-1}}{(4\pi t)^{n/2}}\,e^{-|x-y|^2/(4t)}\,dt\right)f(y)\,dy \\
&= \frac{1}{(4\pi)^{n/2}\Gamma(\alpha/2)}\int_{\mathbb{R}^n} f(y)I(x,y)\,dy
\end{aligned}$$

with

$$I(x,y) = \int_0^\infty t^{\alpha/2-n/2-1}e^{-|x-y|^2/(4t)}\,dt.$$

The substitution $s = |x-y|^2/(4t)$ leads to

$$I(x,y) = 4^{(n-\alpha)/2}\Gamma\left(\tfrac{n-\alpha}{2}\right)\frac{1}{|x-y|^{n-\alpha}}.$$

Thus

$$\begin{aligned} I_\alpha(f)(x) &= \frac{4^{(n-\alpha)/2}}{(4\pi)^{n/2}}\frac{\Gamma\big((n-\alpha)/2\big)}{\Gamma(\alpha/2)}\int_{\mathbb{R}^n}\frac{f(y)}{|x-y|^{n-\alpha}}\,dy \\ &= \frac{\Gamma\big((n-\alpha)/2\big)}{2^\alpha\pi^{n/2}\Gamma(\alpha/2)}\int_{\mathbb{R}^n}\frac{f(y)}{|x-y|^{n-\alpha}}\,dy, \end{aligned}$$

which is one of the more familiar formulas for I_α.

9.2 Inequalities of Sobolev and Nash

Theorem 9.2 (Sobolev)**.** *Assume $n \geq 3$ and $f \in C_0^1(\mathbb{R}^n)$. Then*

$$\|f\|_{2n/(n-2)} \leq C\|\nabla f\|_2,$$

where C is a constant depending only on n.

Proof. We apply Theorem 9.1 with $\alpha = 1$, $p = 2$ and hence $q = 2n/(n-2)$. Since

$$\begin{aligned} \widehat{I_1(f)}(\xi) &= \left(4\pi^2|\xi|^2\right)^{-1/2}\widehat{f}(\xi), \\ \widehat{\frac{\partial f}{\partial x_j}}(\xi) &= 2\pi i\xi_j\widehat{f}(\xi), \end{aligned}$$

and

$$\widehat{R_j f}(\xi) = \frac{i\xi_j}{|\xi|}\widehat{f}(\xi),$$

we have, after taking Fourier transforms of both sides that,

$$f = -I_1\left(\sum_{j=1}^n R_j\left(\frac{\partial f}{\partial x_j}\right)\right).$$

By Theorem 9.1,

$$\|f\|_{2n/(n-2)} \leq C \left\| \sum_{j=1}^{n} R_j \frac{\partial f}{\partial x_j} \right\|_2 = C \left(\int_{\mathbb{R}^n} |\xi|^2 |\widehat{f}(\xi)|^2 \, d\xi \right)^{1/2} = C \|\nabla f\|_2,$$

where we used Exercise 6.3.3 of Chapter 6 in the last equality. □

We also have the local version of Theorem 9.2.

Corollary 9.3. *Suppose $n \geq 3$ and let Q be a cube in $\mathbb{R}^n$. Suppose f is a smooth (C^2 is enough) function in the interior Q. Then*

$$\left(\int_Q |f(x)|^{2n/(n-2)} \, dx \right)^{(n-2)/(2n)} \leq C \left(\int_Q |\nabla f(x)|^2 + \int_Q |f(x)|^2 \, dx \right)^{1/2}.$$

Proof. As before, let Q^* be the cube with same center as Q but with twice the length. For any function g let $\|g\|_{L^p(Q)}$ denote the L^p–norm of g over the cube Q with a similar definition for Q^*. Extend f, by reflection, to Q^* in such a way that

$$\|f\|_{L^{2n/(n-2)}(Q^*)} \leq C_1 \|f\|_{L^{2n/(n-2)}(Q)}, \tag{9.13}$$

$$\|f\|_{L^2(Q^*)} \leq C_2 \|f\|_{L^2(Q)}, \tag{9.14}$$

$$\|\nabla f\|_{L^2(Q^*)} \leq C_3 \|\nabla f\|_{L^2(Q)}, \tag{9.15}$$

for some constants C_1, C_2, C_3. Now pick a C^∞ function ψ which is 0 on the complement of Q^* and 1 on Q. Simple differentiation gives

$$|\nabla \psi f|^2 \leq 2|\nabla f|^2 + C_1 |f|^2.$$

Applying this and Theorem 9.2 to the C_0^∞ function ψf and using (9.13)–(9.15) gives the result. □

Corollary 9.4 (Logarithmic Sobolev Inequality)**.** *Assume $n \geq 3$ and let $f \in C_0^1(\mathbb{R}^n)$ be nonnegative. Then for all $0 < \varepsilon < 1$,*

$$\int_{\mathbb{R}^n} f^2(x) \log f(x) \, dx \leq \varepsilon \int_{\mathbb{R}^n} |\nabla f(x)|^2 \, dx + \beta(\varepsilon) \|f\|_2^2 + \|f\|_2 \log \|f\|_2,$$

where $\beta(\varepsilon) = C + \frac{n}{4} \log \frac{1}{\varepsilon}$.

Proof. We may assume $\|f\|_2 = 1$. Set $p = \frac{2n}{n-2}$ and apply Jensen's inequality with the measure $d\nu = f^2\,dx$ to obtain

$$\begin{aligned}\int_{\mathbb{R}^n} f^2(x)\log f(x)\,dx &= \frac{1}{p-2}\int_{\mathbb{R}^n}\log(f^{p-2}(x))f^2(x)\,dx\\ \leq\ & \frac{1}{p-2}\log\int_{\mathbb{R}^n} f^p(x)\,dx = \frac{p}{2(p-2)}\log\|f\|_p^2\\ \leq\ & \frac{n}{4}\{\varepsilon\|f\|_p^2 + \log\tfrac{1}{\varepsilon}\} \leq \frac{n}{4}\{C\varepsilon\|\nabla f\|_2^2 + \log\tfrac{1}{\varepsilon}\}\end{aligned}$$

where we have used the elementary inequality $\log(a) \leq \varepsilon a + \log\frac{1}{\varepsilon}$ and Theorem 9.2. □

Theorem 9.5 (Nash's Inequality)**.** *Let $n \geq 2$ and let $f \in C_0^1(\mathbb{R}^n)$. Then*

$$\left(\int_{\mathbb{R}^n}|f(x)|^2\,dx\right)^{1+2/n} \leq C\left(\int_{\mathbb{R}^n}|\nabla f(x)|^2\,dx\right)\left(\int_{\mathbb{R}^n}|f(x)|\,dx\right)^{4/n}$$

Proof. By Plancherel's theorem, and the fact that $|\widehat{f}(\xi)| \leq \|f\|_1$ we obtain for any positive M,

$$\begin{aligned}\int_{\mathbb{R}^n}|f(x)|^2\,dx &= \int_{\mathbb{R}^n}|\widehat{f}(\xi)|^2\,d\xi\\ \leq\ & \int_{\{|\xi|\leq M\}}|\widehat{f}(\xi)|^2\,d\xi + \int_{\{|\xi|>M\}}\frac{|\xi|^2}{M^2}|\widehat{f}(\xi)|^2\,d\xi\\ \leq\ & \gamma_n M^n\|f\|_1^2 + M^{-2}\int_{\mathbb{R}^n}|\nabla f(x)|^2\,dx.\end{aligned}$$

Now we choose M to minimize the right hand side and this gives the result. □

Remark 9.1. The logarithmic Sobolev inequality of Corollary 9.4 also holds for $n = 2$, although our proof above does not apply to this case. What we proved above is that the Sobolev inequality directly implies the logarithmic Sobolev inequality. The converse is also true. Namely, the logarithmic Sobolev inequality implies the Sobolev inequality. There is a deep connection of the above inequalities with the theory of semigroups and Dirichlet forms which makes these implications very clear. The reader interested in these directions and its many fascinating applications can see E. B. Davies [Da] and L. Saloff–Coste [Sa]. Notice also that in the proof of Nash's inequality we used nothing more than the most basic properties of the Fourier transform. It is interesting to note that this inequality can also be proved without the Fourier transform. The proof is outlined in the following three exercises. There are many advantages to this proof since it actually extends to the general setting of semigroups. Again, the reader should see E. B. Davies[Da] for more on this direction.

Exercise 9.2.1.
Let $H_t(x)$ be the heat kernel as above and define the operators in $T_t\colon L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ by $T_t f(x) = H_t * f(x)$. By Exercise 5.1.13, Chapter 5,

$$\frac{dT_t f}{dt}(x,t) = \Delta_x T_t f(x) \tag{9.16}$$

and

$$\lim_{t\to 0} T_t f(x) = f(x) \tag{9.17}$$

almost everywhere and in L^2. Prove that the operators T_t are self-adjoint. That is

$$\int_{\mathbb{R}^n} g(x) T_t f(x)\, dx = \int_{\mathbb{R}^n} T_t g(x) f(x)\, dx, \tag{9.18}$$

and that they form a semigroup in the sense that $T_t(T_s f) = T_{t+s} f$, for all $t, s > 0$ and $T_0 = I$ where I is the identity operator and this is interpreted as (9.17).

Exercise 9.2.2.
Let $f \in L^2(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$. Use duality, Exercise 9.2.1 and (9.2) to prove that

$$\int_{\mathbb{R}^n} |T_t f(x)|^2\, dx \le C t^{-n/2} \|f\|_1^2$$

Exercise 9.2.3.
Let $f \in C_0^2(\mathbb{R}^n)$. Use

$$T_{2t} f(x) = f(x) + \int_0^{2t} \frac{dT_s f(x)}{ds}\, ds,$$

exercises 9.2.1, 9.2.2, and integration by parts to conclude that for any $t > 0$,

$$\int_{\mathbb{R}^n} |f(x)|^2\, dx \le 2t \int_{\mathbb{R}^n} |\nabla f(x)|^2\, dx + C t^{-n/2} \left(\int_{\mathbb{R}^n} |f(x)|\, dx \right)^2. \tag{9.19}$$

Minimize the right hand side of (9.19) with respect to t to obtain another proof of Nash's inequality.

We end this Chapter by proving a more general version of Theorem 9.2

Theorem 9.6. *Let $f \in C_0^1(\mathbb{R}^n)$, $n \geq 2$ and let $1 \leq p < n$. Then*

$$\|f\|_{np/(n-p)} \leq C\|\nabla f\|_p$$

Proof. Let us assume that $p = 1$. By the fundamental theorem of calculus,

$$|f(x)| \leq \int_{\infty}^{x_1} \left|\frac{\partial f}{\partial x_i}\right| \, dx_1 \tag{9.20}$$

with a similar expression for $x_2, \dots, x_n$. From this it follows that

$$|f(x)|^{n/(n-1)} \leq \left(\prod_{k=1}^{n} \int_{\mathbb{R}} |\frac{\partial f}{\partial x_k}| \, dx_k\right)^{1/(n-1)}$$

Integrating this inequality, applying Hölders inequality, Fubini's theorem and the geometric mean inequality we arrive at

$$\|f\|_{n/(n-1)} \leq \frac{1}{\sqrt{n}}\|\nabla f\|_1. \tag{9.21}$$

Now assume $p > 1$. Since $p < n$ the number $\alpha = \frac{(n-1)p}{n-p}$ is larger than one and we can apply (9.21) to the function f^α to obtain that

$$\|f^\alpha\|_{n/(n-1)} \leq \frac{\alpha}{\sqrt{n}} \int_{\mathbb{R}^n} |f(x)|^{\alpha-1} |\nabla f(x)| \, dx \leq \frac{\alpha}{\sqrt{n}} \|f^{\alpha-1}\|_q \|\nabla f\|_p, \tag{9.22}$$

where $q = p/(p-1)$ is the conjugate exponent of p. However, $(\alpha - 1)q = \frac{np}{n-p}$ and $\alpha\frac{n}{n-1} = \frac{np}{n-p}$, and (9.22) gives the result. $\square$

Chapter 10

Littlewood–Paley and Lusin Square Functions

In this chapter we present the $L^p(\mathbb{R}^n)$ boundedness of the classical Littlewood–Paley and the Lusin square functions. We then give an application of these results to the Hörmander multiplier Theorem. Other important applications to, for example, the study of boundary behavior of harmonic functions and Hardy (H^p) spaces will not be given here. For these and other connections we refer the reader to Stein [St2] or Bañuelos and Moore [BaMo]. Our strategy here will be to interpret square functions as vector valued singular integrals and obtain their L^p–boundedness as a consequence of the results of Chapter 7. While this is certainly not the only way to obtain their L^p–boundedness, it is certainly the most direct method given the tools we have developed in the earlier chapter. In addition, this method has the advantage of clarifying the role played in these type of results by the Poisson kernel.

10.1 Definitions, L^2–Properties and Pointwise Comparisons

Let us recall that $\mathbb{R}^{n+1}_+ = \{(x,y)\colon x \in \mathbb{R}^n,\ y > 0\}$ is the upper half space. For u harmonic in $\mathbb{R}^{n+1}_+$ we write

$$\nabla u = \left(\tfrac{\partial u}{\partial y},\ \tfrac{\partial u}{\partial x_1},\ \tfrac{\partial u}{\partial x_2}, \dots, \tfrac{\partial u}{\partial x_n}\right)$$

for its full gradient. Also recall that

$$\nabla_x u = \left(\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \dots, \frac{\partial u}{\partial x_n}\right).$$

The Littlewood–Paley square function of u is defined by

$$g(u)(x) = \left(\int_0^\infty |\nabla u(x,y)|^2 y\,dy\right)^{1/2}. \tag{10.1}$$

In the case when u_f is the Poisson integral of a function $f \in L^p(\mathbb{R}^n)$, $1 \le p \le \infty$, that is

$$u_f(x,y) = P_y * f(x) = \int_{\mathbb{R}^n} \frac{c_n y}{(|x-s|^2 + y^2)^{(n+1)/2}} f(s)\,ds,$$

we simply write $g(f)(x)$ for $g(u_f)(x)$. Notice that

$$g^2(f)(x) = g_v^2(f)(x) + g_h^2(f)(x) \tag{10.2}$$

where $g_v(f)$ and $g_h(f)$ are, respectively, the vertical and horizontal g–functions as defined in Exercise 6.3.4 of Chapter 6. It follows from that exercise that for $f \in L^2(\mathbb{R}^n)$,

$$\|g(f)\|_2 = \frac{1}{\sqrt{2}}\|f\|_2. \tag{10.3}$$

Let us recall the proof of (10.3) and the tools used in it. These will be useful below as well. First, since

$$\hat{P}_y(\xi) = e^{-2\pi y|\xi|}$$

we have

$$\widehat{\frac{\partial P_y}{\partial y}}(\xi) = -2\pi|\xi|e^{-2\pi y|\xi|}$$

and

$$\widehat{\frac{\partial P_y}{\partial x_i}}(\xi) = -2\pi i\xi_i e^{-2\pi y|\xi|}.$$

By Plancherel's theorem,

$$\begin{aligned}\int_{\mathbb{R}^n} g^2(f)(x)\,dx &= \int_0^\infty \int_{\mathbb{R}^n} y|\nabla u_f(x,y)|^2\,dx\,dy \\ &= \int_0^\infty y \int_{\mathbb{R}^n} 8\pi^2|\xi|^2 e^{-4\pi y|\xi|}|\hat{f}(\xi)|^2\,d\xi\,dy\end{aligned}$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 |\xi|^2 \left(\int_0^\infty 8\pi^2 y e^{-4\pi y|\xi|}\, dy \right) d\xi \\
&= \frac{1}{2} \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 \, d\xi = \frac{1}{2} \|f\|_2^2,
\end{aligned}
$$

hence (1.3).

For $\alpha > 0$ and $x \in \mathbb{R}^n$, denote the cone in $\mathbb{R}^{n+1}_+$ with vertical axes, vertex at x and opening angle $2\arctan(\alpha)$ by $\Gamma_\alpha(x)$. That is

$$\Gamma_\alpha(x) = \{(w, y) \colon w \in \mathbb{R}^n,\ y > 0,\ |x - w| < \alpha y\}.$$

If u is harmonic in $\mathbb{R}^{n+1}_+$ the Lusin square function of u is defined by

$$A_\alpha(u)(x) = \left(\int_{\Gamma_\alpha(x)} |\nabla u(w, y)|^2 y^{1-n} \, dw \, dy \right)^{1/2}. \tag{10.4}$$

As before, if $u(x, y) = u_f(x, y) = P_y * f(x)$ for some $f \in L^p(\mathbb{R}^n)$, $1 \le p \le \infty$, we simply write $A_\alpha(f)(x)$ for $A_\alpha(u_f)(x)$. We will momentarily see that this operator is also, up to a constant, an isometry on $L^2(\mathbb{R}^n)$. Before we do this, however, let us give a little geometric insight into $A_\alpha(u)$ in the case when $n = 1$ and u is the real part of an analytic function in $\mathbb{R}^2_+$. That is, suppose $F = u + iv$ is analytic in $\mathbb{R}^2_+$. Then

$$|\nabla u|^2 = |F'|^2 = \frac{\partial(u, v)}{\partial(x, y)}$$

is the Jacobian of the mapping from $\mathbb{R}^2_+$ into the complex plane. By the change of variables formula $A^2_\alpha(u)(x)$ is simply the area, counting multiplicities, of the image of $\Gamma_\alpha(x)$ under the analytic function F. If in addition F is conformal (one–to–one), then $A^2_\alpha(u)(x) = \mathrm{Area}(F(\Gamma_\alpha(x)))$. For this reason, $A_\alpha(u)(x)$ is often called the Lusin area function of u.

As before, the boundedness of $A_\alpha(f)$ in $L^2(\mathbb{R}^n)$ follows from Plancherel's theorem. Indeed, let $\chi_{(0,1)}(r)$ be the characteristic function of the interval $(0, 1)$. By Fubini's theorem

$$
\begin{aligned}
\|A_\alpha(f)\|_2^2 &= \int_{\mathbb{R}^n} \left(\int_0^\infty \int_{\mathbb{R}^n} \chi_{(0,1)} \left(\frac{|x - w|}{\alpha y} \right) |\nabla u(w, y)|^2 y^{1-n} \, dw \, dy \right) dx \\
&= \int_{\mathbb{R}^n} \int_0^\infty |\nabla u(w, y)|^2 y^{1-n} \left(\int_{\mathbb{R}^n} \chi_{(0,1)} \left(\frac{|x - w|}{\alpha y} \right) dx \right) dy \, dw
\end{aligned}
$$

$$= \gamma_n \alpha^n \int_{\mathbb{R}^n} \int_0^\infty y |\nabla u(w,y)|^2 \, dy \, dw = \gamma_n \alpha^n \|g(f)\|_2^2 = \frac{\gamma_n \alpha^n}{2} \|f\|_2^2,$$

where γ_n is the volume of the unit ball in $\mathbb{R}^n$.

As in the case of the Littlewood–Paley g–function, we also have the vertical and horizontal Lusin area functions $A_{\alpha,v}$ and $A_{\alpha,h}$ defined by

$$A_{\alpha,v}(u)(x) = \left(\int_{\Gamma_\alpha(x)} \left| \frac{\partial u}{\partial y}(w,y) \right|^2 y^{1-n} \, dw \, dy \right)^{1/2}$$

and

$$A_{\alpha,h}(u)(x) = \left(\int_{\Gamma_\alpha(x)} |\nabla_w u(w,y)|^2 y^{1-n} \, dw \, dy \right)^{1/2}.$$

It follows from the argument just presented that

$$\|A_{\alpha,v}(f)\|_2^2 = \gamma_n \alpha^n \|g_1(f)\|_2^2 = \frac{\gamma_n \alpha^n}{4} \|f\|_2^2$$

and

$$\|A_{\alpha,h}(f)\|_2^2 = \gamma_n \alpha^n \|g_2(f)\|_2^2 = \frac{\gamma_n \alpha^n}{4} \|f\|_2^2.$$

The following exercises will be useful later.

Exercise 10.1.1.
Let $E \subset \mathbb{R}^n$ and set $G_\alpha(E) = \bigcup_{x \in E} \Gamma_\alpha(x)$. Prove that

$$\begin{aligned} &\int_E A_\alpha^2(u)(x) \, dx \\ &= \int_{G_\alpha(E)} |\nabla u(w,y)|^2 y^{1-n} \, |\{x \in E \colon (w,y) \in \Gamma_\alpha(x)\}| \, dw \, dy \\ &= \int_{G_\alpha(E)} |\nabla u(w,y)|^2 y^{1-n} \, |\{E \cap B(w, \alpha y)\}| \, dw \, dy. \end{aligned}$$

The set $G_\alpha(E)$ is often called a saw tooth region with base E. To see why, take $E \subset \mathbb{R}$ to be a collection of disjoint intervals and draw a picture of the set G_α.

Exercise 10.1.2.
Fix $\alpha > 1$ Let E be an open set of finite measure and set

$$\widetilde{E} = \left\{ x \in \mathbb{R}^n \colon M\chi_E(x) > \frac{1}{2\alpha^n} \right\}.$$

Prove that

(i) $|\widetilde{E}| \le C_1 \alpha^n |E|$,

(ii) $G_\alpha(\widetilde{E}^c) \subset G_1(E^c)$, and

(iii) for all $(w, y) \in G_\alpha(E^c)$, $\left|B(w, \alpha y) \cap \widetilde{E}\right| \le C_2 \alpha^n \left|B(w, y) \cap E\right|$, with C_1 and C_2 depending only on n.

Exercise 10.1.3.
Let E and $\widetilde{E}$ be as in Exercise 10.1.2. Prove that there is a constant C_n depending only on n such that for all $\alpha \ge 1$,

$$\int_{\widetilde{E}^c} A_\alpha^2(u)(x)\,dx \le C_n \alpha^n \int_{E^c} A_1^2(u)(x)\,dx.$$

Exercise 10.1.4.
Let $\beta > 0$ and suppose there is a $x_0 \in \mathbb{R}^n$ such that $A_\beta u(x_0) < \infty$. Let $0 < \alpha < \beta$. Prove that the set $\{x \in \mathbb{R}^n \colon A_\alpha u(x) > \lambda\}$ is open.

Another very useful operator, as we shall see below, is the Littlewood–Paley square function g_λ^* defined by

$$g_\lambda^*(u)(x) = \left(\int_{\mathbb{R}_+^{n+1}} \left(\frac{y}{|x - w| + y}\right)^{\lambda n} |\nabla u(w, y)|^2 y^{1-n}\,dw\,dy\right)^{1/2} \tag{10.5}$$

for any harmonic function u and any $\lambda > 0$. As before, if $u(x, y) = P_y * f(x)$ for some $f \in L^p(\mathbb{R}^n)$, $1 \le p \le \infty$, we simply write $g_\lambda^*(f)$. As with the previous two operators we also have the corresponding vertical and horizontal functions defined by

$$g_{\lambda,v}^*(u)(x) = \left(\int_{\mathbb{R}_+^{n+1}} \left(\frac{y}{|x - w| + y}\right)^{\lambda n} \left|\tfrac{\partial u}{\partial y}(w, y)\right|^2 y^{1-n}\,dw\,dy\right)^{1/2}$$

and

$$g_{\lambda,h}^*(u)(x) = \left(\int_{\mathbb{R}_+^{n+1}} \left(\frac{y}{|x - w| + y}\right)^{\lambda n} |\nabla_w u(w, y)|^2 y^{1-n}\,dw\,dy\right)^{1/2},$$

respectively. Thus,

$$\left(g_\lambda^* u(x)\right)^2 = \left(g_{\lambda,v}^* u(x)\right)^2 + \left(g_{\lambda,h}^* u(x)\right)^2.$$

There is a more convenient way to write the g^* functions which makes some of their properties very transparent. For this set

$$\varphi^\lambda(x) = \frac{1}{(|x|+1)^{n\lambda}} \tag{10.6}$$

and notice that

$$g_\lambda^*(u)(x) = \left(\int_{\mathbb{R}_+^{n+1}} \varphi_y^\lambda(x-w)|\nabla u(w,y)|^2 y\, dw\, dy \right)^{1/2}. \tag{10.7}$$

If $\lambda > 1$ the function φ^λ is in $L^1(\mathbb{R}^n)$. Set $C_\lambda = \|\varphi^\lambda\|_1$ and recall (Lemma 4.7 of chapter 4) that with $\varphi_y^\lambda(x) = \frac{1}{y^n}\varphi^\lambda(x/y)$ we also have $\|\varphi_y^\lambda\|_1 = C_\lambda$. Hence by Fubini's theorem we have

$$\|g_\lambda^*(u)\|_2^2 = C_\lambda \|g(u)\|_2^2.$$

We conclude from (10.3) that for $\lambda > 1$,

$$\|g_\lambda^*(f)\|_2 = \frac{\sqrt{C_\lambda}}{2}\|f\|_2.$$

Our next Proposition gives pointwise comparisons between the above three functions. Here we do not even need to assume that $\lambda > 1$ nor that u is the Poisson integral of an $L^p(\mathbb{R}^n)$ function.

Proposition 10.1. *There is a constant C_1 depending on α and n, and a constant C_2 depending on α, λ and n, such that*

$$C_1 g(u)(x) \le A_\alpha(u)(x) \le C_2 g_\lambda^*(u)(x).$$

Proof. If $(w,y) \in \Gamma_\alpha(x)$, then $|w-x| < \alpha y$ and we see that

$$\frac{1}{(\alpha+1)^{\lambda n}} \le \left(\frac{y}{|x-w|+y} \right)^{\lambda n}.$$

Thus, the right hand side inequality follows with $C_2 = C_{\alpha,n} = \sqrt{(\alpha+1)^{\lambda n}}$.

For the proof of the left hand–side inequality recall that if u is a harmonic function in a region of $\mathbb{R}^n$ and $B(x_0,r)$ is a ball centered at x_0, radius r and contained in this region, then

$$u(x_0) = \frac{1}{m(B(x_0,r))} \int_{B(x_0,r)} u(x)\, dx. \tag{10.8}$$

(This can be found in Bañuelos and Moore [BM] Chapter 1, Theorem 1.1.3). Now, let $u(x,y)$ be harmonic in $\mathbb{R}^{n+1}_+$. Fix $x_0 \in \mathbb{R}^n$ and consider the cone $\Gamma_\alpha(x_0)$. For $(x_0, y) \in \Gamma_\alpha(x_0)$, let $B\big((x_0,y), r\big)$ be the ball in $\mathbb{R}^{n+1}_+$ centered at (x_0, y) contained in the cone and tangent to its boundary. Since $(x,t) \in \Gamma_\alpha(x_0)$ if and only if $|x - x_0| < \alpha t$, we see that $r = C_\alpha y$ where C_α is a constant depending only on α. Thus applying (10.8) to the derivatives of u, which are also harmonic in $\mathbb{R}^{n+1}_+$, we see that

$$\frac{\partial u}{\partial y}(x_0, y) = \frac{C_{\alpha,n}}{y^{n+1}} \int\limits_{(B(x_0,y),C_\alpha y)} \frac{\partial u}{\partial y}(x,t)\,dx\,dt$$

and

$$\frac{\partial u}{\partial x_i}(x_0, y) = \frac{C_{\alpha,n}}{y^{n+1}} \int\limits_{B((x_0,y),C_\alpha y)} \frac{\partial u}{\partial x_i}(x,t)\,dx\,dt.$$

Applying Jensen's inequality we get

$$|\nabla u(x_0,y)|^2 \le C_{\alpha,n} y^{-n-1} \int\limits_{B((x_0,y),C_\alpha y)} |\nabla u(x,t)|^2\,dx\,dt.$$

Multiplying both sides by y and integrating we obtain

$$g^2(u)(x_0) = \int_0^\infty y|\nabla u(x_0,y)|^2\,dy \le C \int_0^\infty y^{-n} \int\limits_{B((x_0,y),C_\alpha y)} |\nabla u(x,t)|^2\,dx\,dt\,dy.$$

However, for $(x,t) \in B((x_0,y), C_\alpha y) \subset \Gamma_\alpha(x_0)$ we clearly have $C'_\alpha t \le y \le C^2_\alpha t$ and therefore the above integral is dominated by

$$\begin{aligned} & C_{\alpha,n} \int_0^\infty \int\limits_{|x_0-x|<\alpha t} \left(\int_{C^1_\alpha t}^{C^2_\alpha t} y^{-n}\,dy \right) |\nabla u(x,t)|^2\,dx\,dt \\ = \;& C_{\alpha,n} \int_0^\infty \int\limits_{|x-x_0|<\alpha t} |\nabla u(x,t)|^2 t^{1-n}\,dx\,dt \\ = \;& C_{\alpha,n} A_\alpha(u)(x_0), \end{aligned}$$

which proves the Proposition. □

There are other useful generalizations of the vertical functions g_v, $A_{\alpha,v}$ and $g^*_{\lambda,v}$. For any harmonic function u in the upper half space $\mathbb{R}^{n+1}$ and any integer $k \geq 1$, we define

$$g_k u(x) = \left(\int_0^\infty \left|\frac{\partial^k u}{\partial y^k}(x,y)\right|^2 y^{2k-1}\,dy\right)^{1/2} \tag{10.9}$$

$$A_{\alpha,k} u(x) = \left(\int_{\Gamma_{\alpha(x)}} \left|\frac{\partial^k u}{\partial y^k}(w,y)\right|^2 y^{2k-1-n}\,dy\,dw\right)^{1/2} \tag{10.10}$$

and

$$g^*_{\lambda,k}(u)(x) = \left(\int_{\mathbb{R}^{n+1}_+} \varphi^\lambda_y(x-w)\left|\frac{\partial^k u}{\partial y^k}(w,y)\right|^2 y^{2k-1}\,dw\,dy\right)^{1/2}$$

so that $g_1 = g_v$, $A_{\alpha,1} = A_{\alpha,v}$ and $g^*_{\lambda,1} = g^*_{\lambda,v}$. As usual, if $f \in L^p(\mathbb{R}^n)$ we simply write $g_k(f)(x)$ for $g_k u(x)$ and similarly for the other two operators.

Exercise 10.1.5.
Prove that there are constants C_1, C_2, depending on k, n, λ, α, such that

$$C_1 g_k u(x) \leq A_{\alpha,k} u(x) \leq C_2 g^*_{\lambda,k}(u)(x).$$

Exercise 10.1.6.
Let $f \in L^2(\mathbb{R}^n)$. Prove that for any $k \geq 1$ and $\lambda > 1$,

$$\|g_k(f)\|_2 = \sqrt{\frac{\Gamma(2k)}{2^{2k-1}}}\|f\|_2,$$
$$\|A_{\alpha,k} f\|_2 = \sqrt{\gamma_n \alpha^n \frac{\Gamma(2k)}{2^{2k-1}}}\|f\|_2,$$

and

$$\|g^*_{\lambda,k} f\|_p = \sqrt{C_\lambda \frac{\Gamma(2k)}{2^{2k-1}}}\|f\|_2,$$

where γ_n is the volume of the unit ball in $\mathbb{R}^n$ and $C_\lambda = \|\varphi^\lambda\|_1$

Exercise 10.1.7.
Use the fundamental theorem of calculus to prove that for any $f \in C_0^\infty(\mathbb{R}^n)$ and any $k \geq 1$,

$$g_k(f)(x) \leq \frac{1}{\sqrt{(2k-1)}} g_{k+1}(f)(x)$$

and conclude that

$$g_1(f)(x) \leq C_k^1 g_{k+1}(f)(x).$$

Similarly prove that

$$A_{\alpha,1} f(x) \leq C_k^2 A_{\alpha,k} f(x)$$

and that

$$g_{\lambda,1}^* f(x) \leq C_k^3 g_{\lambda,k}^* f(x),$$

for some constants C_k^1, C_k^2, C_k^3 depending only on k and n.

Exercise 10.1.8.
Let u be a harmonic function in $\mathbb{R}^n$ and fix $r > 0$. Let $x, w \in B(x_0, 3r/4)$ with $|x - w|$ small enough. Use (10.8) to prove that

$$\frac{|u(x) - u(w)|}{|x - w|} \leq \frac{C_n}{r} \sup_{z \in B(x_0, r)} |u(z)|, \tag{10.11}$$

where C_n depends only on n. Conclude from (10.9) that

$$r \sup_{x \in B(x_0, 3r/4)} |\nabla u(x)| \leq C_n \sup_{x \in B(x_0, r)} |u(x)|, \tag{10.12}$$

with C_n depending only on n.

10.2 L^p Properties

Theorem 10.2. *Let $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$. There are constants A_1 and A_2 depending only on p and n such that*

$$A_1 \|f\|_p \leq \|g(f)\|_p \leq A_2 \|f\|_p. \tag{10.13}$$

Theorem 10.3. *Let $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$. There are constants A_3 and A_4 depending on p, α and n, such that*

$$A_3\|f\|_p \leq \|A_\alpha(f)\|_p \leq A_4\|f\|_p. \tag{10.14}$$

By Proposition 10.1 and either Theorem 10.2 or 10.3, $A\|f\|_p \leq \|g_\lambda^*(f)\|_p$, for all $1 < p < \infty$ for some constant A independent of f. Unlike the Littlewood–Paley g–function and the Lusin area function A_α, however, the g_λ^*–function is not bounded for the full range of $1 < p < \infty$. For this operator we have

Theorem 10.4. *Let $\lambda > 1$ and $1 < p < \infty$ with $p > 2/\lambda$. There is a constant A_5 depending on p, λ and n, such that*

$$\|g_\lambda^*(f)\|_p \leq A_5\|f\|_p. \tag{10.15}$$

As we shall see, Theorem 10.4 follows from Theorems 10.2 and 10.3. By Proposition 10.1, obviously, the right hand side of Theorem 10.2 also follows from the right hand side of Theorem 10.3. However, from our approach to Theorems 10.2 and 10.3 via vector valued singular integrals, there is very little difference between the proofs of the right hand side of (10.13) and (10.14). Since the proof of (10.13) illustrates these techniques a little more clearly, we present this with full details, leaving some of the details of the proof of (10.14) to the reader.

Proof of Theorem 10.2. Consider the Hilbert space

$$\begin{aligned} H_2 &= L^2(\mathbb{R}^+, \mathbb{R}^{n+1}) \\ &= \left\{ F(y) = \big(f_0(y), f_1(y), \ldots, f_n(y)\big) : \underbrace{\left(\int_0^\infty |F(y)|^2 y\, dy\right)^{1/2}}_{\|F\|_{H_2}} < \infty \right\} \end{aligned}$$

with the obvious L^2 inner product. For any $\varepsilon > 0$, $y > 0$, define

$$K_\varepsilon(x) = \nabla P_{y+\varepsilon}(x) \tag{10.16}$$

and

$$T_\varepsilon f(x) = \int_{\mathbb{R}^n} K_\varepsilon(w) f(x-w)\, dw = \nabla u_f(x, y+\varepsilon).$$

Then for each $x \in \mathbb{R}^n$,

$$\|T_\varepsilon f(x)\|_{H_2} = \left(\int_0^\infty |\nabla u_f(x, y+\varepsilon)|^2 y\, dy\right)^{1/2}$$

and what we like to prove is that

$$\int_{\mathbb{R}^n} \|T_\varepsilon f(x)\|_{H_2}^p \, dx \le A_{p,n} \int_{\mathbb{R}^n} |f(x)|^p \, dx \tag{10.17}$$

with $A_{p,n}$ independent of ε. (After this, simply apply Fatou's Theorem). We apply Theorem 7.23 of Chapter 7. In order to do that we need to verify that $K_\varepsilon(x)$ satisfies

(i) $\|K_\varepsilon(x)\|_{H_2} \in L^2(\mathbb{R}^n)$,

(ii) $\|\nabla K_\varepsilon(x)\|_{H_2} \le \dfrac{B_1}{|x|^{n+1}}$, $x \in \mathbb{R}^n \setminus \{0\}$, and

(iii) $\|\hat{K}_\varepsilon(\xi)\|_{H_2} \le B_2$, with B_1 and B_2 independent of ε.

(Notice that as in the scalar case, (ii) implies the Hörmander condition of Theorem 7.23 of Chapter 7). With

$$P_y(x) = \frac{C_n y}{(|x|^2 + y^2)^{n+1/2}}$$

a straightforward differentiation gives

$$\frac{\partial P_y}{\partial y}(x) = \frac{C_n\big[(|x|^2+y^2)^{(n+1)/2} - (n+1)y^2(|x|^2+y^2)^{(n-1)/2}\big]}{(|x|^2+y^2)^{n+1}}$$

and

$$\frac{\partial P_y}{\partial x_i}(x) = \frac{-(n+1)C_n y x_i}{(|x|^2+y^2)^{(n+3)/2}}.$$

Hence for some constant C_n' depending only on n,

$$|K_\varepsilon(x)| \le \frac{C_n'}{(|x|^2 + (y+\varepsilon)^2)^{(n+1)/2}}.$$

Changing variables we find that

$$\int_0^\infty \frac{y}{(|x|^2 + (y+\varepsilon)^2)^{n+1}} \, dy \le \int_{|x|^2}^\infty \frac{dy}{y^{n+1}} = \frac{C_n'}{|x|^{2n}}$$

and in the same way that

$$\int_0^\infty \frac{y}{(|x|^2 + (y+\varepsilon)^2)^{n+1}} \, dy \le \int_{\varepsilon^2}^\infty \frac{dy}{y^{n+1}} = \frac{C_n''}{\varepsilon^{2n}}.$$

Therefore

$$\int_{\mathbb{R}^n} \|K_\varepsilon(x)\|_{H_2}^2 \, dx \le C_{\varepsilon,n} < \infty,$$

completing the proof of (i). Computing the second derivatives of $P_y(x)$, (ii) follows exactly in the same way. For (iii) we have as above that

$$\begin{aligned} \|\widehat{K}_\varepsilon(\xi)\|_{H_2}^2 &= \int_0^\infty |\widehat{\nabla} P_{y+\varepsilon}(\xi)|^2 y \, dy \\ &= 8\pi^2 \int_0^\infty y|\xi|^2 e^{-4\pi(y+\varepsilon)|\xi|} \, dy \\ &\le 8\pi^2 |\xi|^2 \int_0^\infty y e^{-4\pi y|\xi|} \, dy = \frac{1}{2}. \end{aligned}$$

This gives the right hand side of (10.13).

The left hand side of (10.13) follows from this and duality. Indeed, by Exercise 6.3.5 of Chapter 6, for all $f,\ h \in L^2(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} f\overline{h} \, dx = 2\int_0^\infty \int_{\mathbb{R}^n} y \nabla u_f(x,y) \cdot \overline{\nabla u_h(x,y)} \, dx \, dy$$

and by Cauchy–Schwartz

$$\left| \int_{\mathbb{R}^n} f\overline{h} \, dx \right| \le 2 \int_{\mathbb{R}^n} g(f)(x) g(h)(x) \, dx.$$

If in addition, $f \in L^p(\mathbb{R}^n)$ and $h \in L^q(\mathbb{R}^n)$ with $1/p + 1/q = 1$ and $\|h\|_q \le 1$, the right hand side of (10.13) gives

$$\left| \int_{\mathbb{R}^n} f\overline{h} \, dx \right| \le A_{q,n} \|g(f)\|_p.$$

Taking supremum over all $h \in L^q(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ we obtain

$$\|f\|_p \le A_{p,n} \|g(f)\|_p \tag{10.18}$$

for all $f \in L^p(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$.

To remove the L^2–restriction in (10.18), let $f_m \in L^p(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ with $f_m \to f$ in $L^p(\mathbb{R}^n)$. A simple calculation shows that

$$|g(f_m) - g(f)|^2 = g^2(f_m) - 2g(f_m)g(f) + g^2(f)$$

$$
\begin{aligned}
&\le\; g^2(f_m) - 2\int_0^\infty y|\nabla u_{f_m}(x,y)|\,|\nabla u_f(x,y)|\,dy + g^2(f)\\
&\le\; |g(f_m - f)|^2.
\end{aligned}
$$

That is, $|g(f_m) - g(f)| \le |g(f_m - f)|$ and from the right hand side of (10.13) we again conclude that $g(f_m) \to g(f)$ in $L^p(\mathbb{R}^n)$. This gives (10.18) for all $f \in L^p(\mathbb{R}^n)$ and completes the proof of Theorem 2.1. □

If instead of using the full gradient above we use the identities of Exercise 6.3.5, Chapter 6 with the vertical and horizontal derivatives, we arrive at the following

Corollary 10.5. *There are constants A_1 and A_2 depending only on p and n such that*

(i) $A_1\|f\|_p \le \|g_v(f)\|_p$, *and*

(ii) $A_2\|f\|_p \le \|g_h(f)\|_p$.

The following exercise gives an application of Theorem 10.2.1 and its corollary. This type of idea concerning the "invariance" of Littlewood-Paley functions under certain singular integrals is a powerful tool in applications and it will be used below when we prove the Hörmander multiplier theorem.

Exercise 10.2.1.
Let $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$ and let R_jf, $j = 1,\dots,n$, be its Riesz transforms. Use the Cauchy–Riemann equations to prove that

$$g_h(f)(x) = \sqrt{g_v^2(R_1f)(x) + \cdots + g_v^2(R_nf)(x)}$$

and hence

$$g_v(R_jf)(x) \le g(f)(x),$$

for all $j = 1,\dots,n$.

Proof of Theorem 10.3. First, by Proposition 10.1, the left hand side of (10.14) follows from the left hand side of (10.13). The proof of the right hand side of (10.14) follows exactly the proof of Theorem 10.2 except that we need to adjust the Hilbert space H_2 and the operator T, accordingly. First, note that

$$A_\alpha^2(f)(x) = \int_0^\infty \int_{\{|x-w|<\alpha y\}} |\nabla u(w,y)|^2 y^{1-n}\,dw\,dy$$

$$= \int_0^\infty \int_{\{|w|<\alpha\}} |\nabla u(x - wy, y)|^2 y \, dw \, dy,$$

by a simple change of variables. Therefore it seems natural this time to consider the Hilbert space H_2 consisting of all functions

$$F(w, y) = (f_0(w, y), f_1(w, y), \dots, f_n(w, y)) \colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$$

with the norm

$$\|F\|_{H_2} = \left(\int_0^\infty \int_{\{|w|<\alpha\}} |F(w, y)|^2 y \, dw \, dy \right)^{1/2} < \infty,$$

and with its obvious L^2–inner product.

Now, for each $\alpha > 0$ and $\varepsilon > 0$, we define the function $T_\varepsilon f \colon \mathbb{R}^n \to H_2$ by

$$T_\varepsilon f(x)(w, y) = \nabla u(x - w(y + \varepsilon), \ (y + \varepsilon))$$

and observe that

$$\|T_\varepsilon f(x)\|_{H_2} = \Big(\int_\varepsilon^\infty \int_{\{|x-w|<\alpha y\}} |\nabla u(w, y)|^2 y^{1-n} \, dw \, dy \Big)^{1/2}.$$

It is enough to prove that

$$\int_{\mathbb{R}^n} \|T_\varepsilon f(x)\|_{H_2}^p \, dx \le A_{p,\alpha,n} \int_{\mathbb{R}^n} |f(x)|^p \, dx,$$

with the constant $A_{p,\alpha,n}$ independent of ε. Fix y and w. This time we take our kernel function to be

$$K_\varepsilon(x) = \nabla P_{y+\varepsilon}(x - w(y + \varepsilon))$$

and as before we must verify that

(i) $\|K_\varepsilon(x)\|_{H_2} \in L^2(\mathbb{R}^n)$,

(ii) $\|\nabla K_\varepsilon(x)\|_{H_2} \le \dfrac{B_1}{|x|^{n+1}}$, $x \in \mathbb{R}^n \setminus \{0\}$, and

(iii) $\|\widehat{K}_\varepsilon(\xi)\|_{H_2} \le B_2$, with B_1 and B_2 independent of ε and y and w.

(i), (ii) and (iii) follow exactly as in the proof of Theorem 10.2 using the explicit bounds for the derivatives of the Poisson kernel. We leave the details as an exercise. □

As with many of our earlier results concerning approximations to the identity and maximal functions, the exact formula for the Poisson kernel is not important here either. What is essential in the above proofs is certain decay properties of its derivatives. First, if $\psi\colon \mathbb{R}^n \to \mathbb{R}^m$, $\psi(x) = \big(\psi_1(x), \psi_2(x), \ldots, \psi_m(x)\big)$ is a smooth function we set

$$|\nabla\psi| = \sqrt{|\nabla\psi_1|^2 + |\nabla\psi_2|^2 + \cdots + |\nabla\psi_m|^2}.$$

Definition 10.6. We will say that a function $\psi\colon \mathbb{R}^n \to \mathbb{R}^m$ is a Littlewood–Paley function if

(i) $|\psi(x)| \le \dfrac{C_1}{(1+|x|)^{n+1}}$

(ii) $|\nabla\psi(x)| \le \dfrac{C_2}{(1+|x|)^{n+2}}$, and

(iii) $|\widehat{\psi}|$ is a radial function with $\int_0^\infty |\hat{\psi}(y)|^2 \frac{dy}{y} = C_3 < \infty$.

Let us recall that if $\beta = (\beta_1, \ldots, \beta_n)$ is a multi-index (the $\beta_i's$ are nonnegative integers) with $|\beta| = \beta_1 + \cdots + \beta_n$ and $u \in C^{|\beta|}(\mathbb{R}^n)$, we set

$$D^\beta u(x) = \frac{\partial^{|\beta|} u(x)}{\partial x_1^{\beta_1} \cdots \partial x_n^{\beta_n}}. \tag{10.19}$$

Exercise 10.2.2.
Let $P(x) = C_n(1+|x|^2)^{-(n+1)/2}$ be the Poisson kernel. Prove that

$$|D^\beta P(x)| \le \frac{C_{\beta,n}}{(1+|x|)^{n+|\beta|}}$$

for any multi-index β.

Exercise 10.2.3.
With P as in Exercise 10.2.2, let $\psi^1(x) = \nabla_x P(x)$ and $\psi^2(x) = -nP(x) - \sum_{j=1}^n x_j \frac{\partial P}{\partial x_j}$. Prove that ψ^1 and ψ^2 are both Littlewood–Paley functions and that $\psi_y^2(x) = \frac{1}{y^n}\psi^2(x/y) = y\frac{\partial}{\partial y}P_y(x)$.

Exercise 10.2.4.
Let ψ be a Littlewood–Paley function and define

$$Tf(x) = \left(\int_0^\infty |\psi_y * f(x)|^2 \tfrac{dy}{y} \right)^{1/2},$$

for any $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$. Prove, following the argument of Theorem 10.2, that

$$A_1 \|f\|_p \le \|Tf\|_p \le A_2 \|f\|_p$$

where the constants A_1, A_2 depend on p, n and the Littlewood–Paley constants C_1, C_2 and C_3.
Hint: Consider the Hilbert space

$$\begin{aligned} H_2 &= L^2(\mathbb{R}^+, \mathbb{R}^m) \\ &= \left\{ F(y) = (f_1(y), f_2(y), \dots, f_m(y)) \colon \underbrace{\left(\int_0^\infty |F(y)|^2 \tfrac{dy}{y} \right)^{1/2}}_{\|F\|_{H_2}} < \infty \right\}. \end{aligned}$$

Exercise 10.2.5.
Let ψ^1 and ψ^2 be Littlewood–Paley functions and define the corresponding T_1 and T_2 as in Exercise 10.2.4. Prove that $T_1 f = g_v(f)$ and $T_2(f) = g_h(f)$ where g_v and g_h are, as in Exercise 6.3.4 of Chapter 6, the horizontal and vertical Littlewood–Paley functions, respectively.

Exercise 10.2.6.
Let ψ be a Littlewood–Paley function and for $|w| < \alpha$ define $\psi^w(x) = \psi(x - w)$ and let $\psi^w_y(x) = \psi^w(x/y)/y^n$. For $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$, define the operator

$$T_\alpha f(x) = \left(\int_0^\infty \int_{\{|w|<\alpha\}} |\psi^w_y * f(x)|^2 \, dw \tfrac{dy}{y} \right)^{1/2}.$$

Prove that $\|T_\alpha f\|_p \le A_p \|f\|_p$.

Exercise 10.2.7.
Let ψ^1 and ψ^2 be as in Exercise 10.2.3 and let $T_{\alpha,1}$ and $T_{\alpha,2}$ be the corresponding operators. Prove that $T_{\alpha,1} = A_{\alpha,v}$, $T_{\alpha,2} = A_{\alpha,h}$.

The following proposition will play an important role in the proof of Theorem 10.4 but it is also of independent interest. We return to its proof later.

Proposition 10.7. *Let $1 < p < 2$ and $\alpha \geq 1$. There is a constant C independent of α such that*

$$\|A_\alpha(f)\|_p \leq C\alpha^{n/p}\|A_1(f)\|_p \tag{10.20}$$

Proof of Theorem 10.4. Let $1 < q < \infty$ be the conjugate exponent of p and let $h \in L^q(\mathbb{R}^n)$. Let $\varphi^\lambda(x)$ be as in (10.6) and recall, by Corollary 5.5 of Chapter 5, that if $h^*(x) = \sup_{y>0}\{\varphi_y^\lambda * |h|\}(x)$ then $\|h^*\|_q \leq C_q\|h\|_q$. Using this, the representation for g_λ^* given in (10.7), Fubini's theorem, and Hölder's inequality we see that

$$\begin{aligned}\left|\int_{\mathbb{R}^n}\big(g_\lambda^*(f)(x)\big)^2 h(x)\,dx\right| &\leq \int_0^\infty\int_{\mathbb{R}^n}\varphi_y^\lambda * |h|(x)|\nabla u_f(x,y)|^2 y\,dx\,dy\\ &\leq \int_{\mathbb{R}^n} h^*(x)g^2(f)(x)dx \leq \|h^*\|_q\|g^2(f)\|_p \leq C_q\|h\|_q\|g^2(f)\|_p.\end{aligned}$$

Taking supremum over all $h \in L^q(\mathbb{R}^n)$ with $\|h\|_q \leq 1$ gives that

$$\|\big(g_\lambda^*(f)(x)\big)^2\|_p \leq C_p\|g^2(f)\|_p,$$

for all $1 < p < \infty$. Thus Theorem 10.2 gives

$$\|g_\lambda^*(f)(x)\|_p \leq C_p\|f\|_p, \tag{10.21}$$

for all $2 < p < \infty$.

Next we have,

$$\begin{aligned}\big(g_\lambda^*(f)(x)\big)^2 &= \int_0^\infty\int_{\{|x-w|<y\}}\left(\frac{y}{|x-w|+y}\right)^{\lambda n} y^{1-n}|\nabla u_f(w,y)|^2\,dw\,dy\\ &+ \sum_{j=1}^\infty\int_0^\infty\int_{\{2^{j-1}y<|x-w|<2^j y\}}\left(\frac{y}{|x-w|+y}\right)^{\lambda n} y^{1-n}|\nabla u_f(w,y)|^2\,dw\,dy\\ &\leq C\sum_{j=0}^\infty 2^{-jn\lambda}A_{2^j}^2(f)(x).\end{aligned}$$

Hence, if $1 < p < 2$ we obtain

$$\big(g_\lambda^*(f)(x)\big)^p \leq C'\sum_{j=0}^\infty 2^{-jn\lambda p/2}A_{2^j}^p(f)(x).$$

Integrating both sides of this inequality and applying Proposition 10.7 we arrive at

$$\|g_\lambda^*(f)\|_p \leq C_p\Big(\sum_{j=0}^{\infty} 2^{-jn(\lambda p/2-1)}\Big)^{1/p}\|A_1(f)\|_p \tag{10.22}$$

and the sum is finite since by assumption $p > 2/\lambda$. The right hand side (10.14) and the inequality (10.21) proves that

$$\|g_\lambda^*(f)(x)\|_p \leq C_p\|f\|_p, \tag{10.23}$$

for all $1 < p < 2$ and Theorem 10.4 follows. □

It remains to prove Proposition 10.7. Since we already know that $f \in L^p(\mathbb{R}^n)$ implies the same for the area function associated to any cone, we have the hypothesis of Exercise 10.1.4 and therefore the set $E_\lambda = \{x \in \mathbb{R}^n \colon A_1(f)(x) > \lambda\}$ is open. Recall that $\widetilde{E}_\lambda = \big\{x \in \mathbb{R}^n \colon M\chi_{E_\lambda}(x) > \frac{1}{\alpha^n}\big\}$ and by Exercise 10.1.2(i), $|\widetilde{E}_\lambda| \leq C_n\alpha^n|E_\lambda|$ with C_n depending only on n. From this, Chebychev's inequality and Exercise 10.1.3,

$$\begin{aligned}
\int_{\mathbb{R}^n} A_\alpha^p f(x)\,dx &= p\int_0^\infty \lambda^{p-1}\,|\{x \in \mathbb{R}^n \colon A_\alpha(f)(x) > \lambda\}|\,d\lambda \\
&\leq\; p\int_0^\infty \lambda^{p-1}|\widetilde{E}_\lambda|\,d\lambda + p\int_0^\infty \lambda^{p-1}\left|\{x \in \widetilde{E}_\lambda^c \colon A_\alpha f(x) > \lambda\}\right|\,d\lambda \\
&\leq\; pC_n\alpha^n\int_0^\infty \lambda^{p-1}|E_\lambda|\,d\lambda + p\int_0^\infty \lambda^{p-1}\Big\{\frac{1}{\lambda^2}\int_{\widetilde{E}_\lambda^c} A_\alpha^2 f(x)\,dx\Big\}\,d\lambda \\
&\leq\; pC_n\alpha^n\int_0^\infty \lambda^{p-1}|E_\lambda|\,d\lambda + pC_n'\alpha^n\int_0^\infty \lambda^{p-1}\Big\{\frac{1}{\lambda^2}\int_{E_\lambda^c} A_1^2 f(x)\,dx\Big\}\,d\lambda \\
&=\; C_n\alpha^n\int_{\mathbb{R}^n} A_1^p f(x)\,dx + pC_n'\alpha^n\int_{\mathbb{R}^n} A_1^2 f(x)\Big\{\int_{A_1 f(x)}^\infty \lambda^{p-3}\,d\lambda\Big\}\,dx \\
&=\; C_n\alpha^n\int_{\mathbb{R}^n} A_1^p f(x)\,dx + \frac{p}{p-2}C_n'\alpha^n\int_{\mathbb{R}^n} A_1^p f(x)\,dx \\
&=\; C_pC_n\alpha^n\int_{\mathbb{R}^n} A_1^p f(x)\,dx,
\end{aligned}$$

and the Proposition follows.

Exercise 10.2.8.
Let $T_\alpha, \alpha \geq 1$, be as in Exercise 10.2.5 above. Prove that the conclusion of Proposition 10.7 holds with A_α and A_1 replaced by T_α and T_1, respectively.

Finally, associated with any Littlewood–Paley function ψ we also have a corresponding g_λ^*–function. Indeed, let ψ be a Littlewood–Paley function and let ϕ^λ be as above. Define

$$\begin{aligned} T_\lambda^*(f)(x) & \\ &= \left(\int_{\mathbb{R}_+^{n+1}} \left(\frac{y}{|x-w|+y}\right)^{\lambda n} |\psi_y * f(w)|^2 y^{-1-n}\, dw\, dy\right)^{1/2} \\ &= \left(\int_{\mathbb{R}_+^{n+1}} \varphi_y^\lambda(x-w)\, |\psi_y * f(w)|^2\, dw \tfrac{dy}{y}\right)^{1/2} \end{aligned} \tag{10.24}$$

Exercise 10.2.9.
Prove, using Exercises 10.2.3, 10.2.6 and the argument of Theorem 10.4, that for $\lambda > 1$, $1 < p < \infty$ and $p > 2/\lambda$,

$$\|T_\lambda^*(f)\|_p \leq A_p \|f\|_p.$$

Exercise 10.2.10.
Let ψ be define as the inverse Fourier transform of

$$\widehat{\psi}(\xi) = \left(\frac{2}{\pi}\right)^{k+1} |\xi|^{k+1} e^{-2\pi|\xi|}.$$

Prove that ψ is a Littlewood–Paley function and that with $g_{k+1}(f)(x)$ and $A_{\alpha,k+1}f(x)$ as in (10.9) and (10.10), we have

$$g_{k+1}(f)(x) = \left(\int_0^\infty |\psi_y * f(x)|^2 \frac{dy}{y}\right)^{1/2}$$

and

$$A_{\alpha,k+1}f(x) = \left(\int_0^\infty \int_{\{|w|<\alpha\}} |\psi_y^w * f(x)|^2\, dw \frac{dy}{y}\right)^{1/2}.$$

Conclude that

$$C_{k,p}\|f\|_p \leq \|g_{k+1}(f)(x)\|_p \leq C'_{k,p}\|f\|_p, \tag{10.25}$$

and that

$$C_{\alpha,k,p}\|f\|_p \leq \|A_{\alpha,k+1}(f)(x)\|_p \leq C'_{\alpha,k,p}\|f\|_p, \tag{10.26}$$

for any $1 < p < \infty$.

Remark 10.1. The Littlewood–Paley operators were originally studied for the Poisson kernel because of their connections in two dimensions to analytic functions. However, as Exercises 10.2.3, 10.2.5 and 10.2.7 show, the exact formula for the Poisson kernel is really not important when studying their L^p–properties. What is essential is that the derivatives of $P(x)$ satisfy the assumptions of Definition 10.6 and certain properties of its Fourier Transform. This turns out to be very important in some applications, as we shall see in the next section.

10.3 The Hörmander multiplier theorem

Definition 10.8. Fix $1 \leq p < \infty$. The function $m \in L^\infty(\mathbb{R}^n)$ is said to be an L^p–multiplier if the operator T_m defined in $L^2(\mathbb{R}^n)$ by

$$\widehat{T_m f}(\xi) = m(\xi)\widehat{f}(\xi)$$

has the property that

$$\|T_m f\|_p \leq A_{p,m}\|f\|_p \tag{10.27}$$

for all $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$ with some constant $A_{p,m}$ independent of f.

When (10.27) holds the operator T_m has a continuous extension to $L^p(\mathbb{R}^n)$ and we denote its operator norm by $\|T_m\|_{p,p}$. By Plancherel's theorem, $\|T_m\|_{2,2} = \|m\|_\infty$. Theorem 7.14 in Chapter 7 provides several examples of L^p–multipliers for $1 < p < \infty$. If m is the Fourier transform of an $L^1(\mathbb{R}^n)$ function (or even of a finite Borel measure in $\mathbb{R}^n$), then m is an L^p–multiplier. We will denote, as it is customary, the collection of $L^p(\mathbb{R}^n)$ multipliers by $\mathcal{M}_p$. The following exercises describe some of the elementary properties of this space.

Exercise 10.3.1. (i) Let $1 < p < \infty$ and let q be its conjugate exponent. Prove that $m \in \mathcal{M}_p$ if and only if $m \in \mathcal{M}_q$.

(ii) Let $m_1, m_2 \in \mathcal{M}_p$. Prove that $m_1 m_2 \in \mathcal{M}_p$.

(iii) With $2 \leq p < \infty$ and q as in (ii), let $q \leq r \leq p$. Prove that $m \in \mathcal{M}_r$ and furthermore, $\|T\|_{r,r} \leq \|T\|_{p,p}$. In particular, $\|m\|_\infty \leq \|T\|_{p,p}$.

Hint: For (ii) use Plancherel's theorem and duality. For (iii) apply the appropriate interpolation theorem from Chapter 5.

The Hörmander multiplier theorem (for $n = 1$ due to Mihlin) provides further examples of L^p–multipliers.

Definition 10.9. We shall say that the function $m \in C^k\big(\mathbb{R}^n \setminus \{0\}\big)$ satisfies the Hörmander multiplier condition of order k if there is a constant B_k such that $\|m\|_\infty \leq B_k$ and

$$\sup_{R>0} \left(\frac{1}{R^n} \int_{\{R \leq |x| \leq 2R\}} \Big(|x|^{|\beta|} |D^\beta m(x)| \Big)^2 \, dx \right) \leq B_k, \tag{10.28}$$

for all multi-index β with $|\beta| \leq k$.

In many applications the functions m satisfy the stronger condition

$$|D^\beta m(x)| \leq C_k |x|^{-|\beta|},$$

for $|\beta| \leq k$. This is satisfied if, for example, m is homogeneous of degree zero and of class C^k on the sphere, or more generally, if $m(rx) = r^{it} m(x)$, $t \in \mathbb{R}$ and m of is class C^k.

Theorem 10.10. *Let m satisfy the Hörmander multiplier condition of order k. Let $f \in L^2(\mathbb{R}^n)$ and set $F(x) = T_m f(x)$. The function $F \in L^2(\mathbb{R}^n)$ and with $\lambda = 2k/n$ we have*

$$g_v(F)(x) \leq C g_\lambda^*(f)(x), \tag{10.29}$$

where C is a constant depending only on k, n and B_k.

Assume this theorem for the moment. If we take $2 < p < \infty$ and assume the positive integer k is such that $k > n/2$, we see that the hypotheses of Theorem 10.4 are satisfied and hence

$$\|T_m f\|_p \leq C_p \|g_v(T_m f)\|_p \leq C_p' \|g_\lambda^*(f)\|_p \leq C_p'' \|f\|_p, \tag{10.30}$$

for $f \in L^2(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$ with the constants C_p' and C_p'' depending only on p, n, k and B_k. (The first inequality follows from Theorem 10.2). Thus $m \in \mathcal{M}_p$ for $2 < p < \infty$ and hence also for $1 < p < 2$, by Exercise 10.3.1 above. We record this result as

Theorem 10.11 (Hörmander Multiplier Theorem)**.** *Let $k > n/2$ and suppose the function m satisfies the Hörmander multiplier condition of order k. Then $m \in \mathcal{M}_p$ for every $1 < p < \infty$.*

Proof of Theorem 10.10. To simplify notation a little, let us set $U(x,y) = P_y * F(x)$ and $u(x,y) = P_y * f(x)$. Then

$$\widehat{\tfrac{\partial^{k+1}U}{\partial y^{k+1}}}(\xi,y) = (2\pi|\xi|)^{k+1}e^{-2\pi y|\xi|}m(\xi)\widehat{f}(\xi)$$

and

$$\widehat{\frac{\partial u}{\partial y}}(\xi,y) = 2\pi|\xi|e^{-2\pi y|\xi|}\widehat{f}(\xi).$$

If we now set

$$\widehat{M}(\xi,y) = (y|\xi|)^k e^{-2\pi y|\xi|}m(\xi) \tag{10.31}$$

we see that

$$y^k\widehat{\tfrac{\partial^{k+1}U}{\partial y^{k+1}}}(\xi,y) = C_k\widehat{\frac{\partial u}{\partial y}}(\xi,y/2)\widehat{M}(\xi,y/2),$$

where C_k depends only on k. Taking the inverse Fourier transform we obtain the convolution representation

$$y^k\tfrac{\partial^{k+1}U}{\partial y^{k+1}}(x,y) = \int_{\mathbb{R}^n}\frac{\partial u}{\partial y}(x-w,y/2)M(w,y/2)\,dw$$

or

$$\begin{aligned}
&y^k\tfrac{\partial^{k+1}U}{\partial y^{k+1}}(x+x_0,y)\\
&\quad= \int_{\mathbb{R}^n}\frac{\partial u}{\partial y}(x+x_0-w,y/2)M(w,y/2)\,dw\\
&\quad= \int_{\mathbb{R}^n}\frac{\partial u}{\partial y}(x_0-w,y/2)M(w+x,y/2)\,dw.
\end{aligned}$$

From this we obtain

$$y^{2k}\left|\tfrac{\partial^{k+1}U}{\partial y^{k+1}}(x+x_0,y)\right|^2 \le I\cdot II \tag{10.32}$$

where

$$I = \int_{\mathbb{R}^n}\left|\frac{\partial u}{\partial y}(x_0-w,y/2)\right|^2\Big(\frac{y}{y+|w|}\Big)^{2k}\,dw$$

and

$$II = \int_{\mathbb{R}^n}|M(w+x,y/2)|^2\Big(\frac{y+|w|}{y}\Big)^{2k}\,dw.$$

We claim that

$$II \le C_{k,n} y^{-n}. \tag{10.33}$$

Assume (10.33). We have

$$\begin{aligned}
A^2_{1,k+1}(F)(x_0) &= \int_{\Gamma_1(x_0)} \left|\tfrac{\partial^{k+1}U}{\partial y^{k+1}}(x,y)\right|^2 y^{2k+1-n}\,dx\,dy \\
&= \int_{\Gamma_1(0)} \left|\tfrac{\partial^{k+1}U}{\partial y^{k+1}}(x+x_0,y)\right|^2 y^{2k+1-n}\,dx\,dy \\
&\le C_{k,n}\int_0^\infty \int_{|x|<y} y^{-n}\int_{\mathbb{R}^n} \left|\tfrac{\partial u}{\partial y}(x_0-w,y/2)\right|^2 y^{1-n}\Big(\frac{y}{y+|w|}\Big)^{2k}\,dw\,dx\,dy \\
&= C_{k,n}\int_0^\infty \int_{\mathbb{R}^n} \left|\tfrac{\partial u}{\partial y}(x_0-w,y/2)\right|^2 y^{1-n}\Big(\frac{y}{y+|w|}\Big)^{2k}\,dw\,dy \\
&= C_{k,n}\int_0^\infty \int_{\mathbb{R}^n} \left|\tfrac{\partial u}{\partial y}(w,y/2)\right|^2 y^{1-n}\Big(\frac{y}{y+|w-x_0|}\Big)^{2k}\,dw\,dy \\
&= C_{k,n}\big(g^*_{2k/n}f(x_0)\big)^2.
\end{aligned}$$

The inequality (10.29) now follows from this an Exercises 10.1.5 and 10.1.7.

It remains to prove (10.33). Using the triangle inequality and the fact that $\big(|a|+|b|\big)^{2k} \le C\big(|a|^{2k}+|b|^{2k}\big)$, we see that

$$\Big(\frac{y+|w|}{y}\Big)^{2k} \le C\Big(1+\frac{|w+x|^{2k}}{y^{2k}}\Big).$$

if $|x| \le |y|$. From this it follows that

$$\begin{aligned}
II &\le C\int_{\mathbb{R}^n} |M(w+x,y/2)|^2\,dw + Cy^{-2k}\int_{\mathbb{R}^n} \big(|M(w+x,y/2)|\,|w+x|^k\big)^2\,dw \\
&= C\int_{\mathbb{R}^n} |M(w,y/2)|^2\,dw + Cy^{-2k}\int_{\mathbb{R}^n} \big(|M(w,y/2)|\,|w|^k\big)^2\,dw \\
&= C\int_{\mathbb{R}^n} |\widehat{M}(\xi,y)|^2\,d\xi + Cy^{-2k}\int_{\mathbb{R}^n} \big(|M(w,y/2)|\,|w|^k\big)^2\,dw \\
&= C\int_{\mathbb{R}^n} y^{2k}|\xi|^{2k}e^{-4\pi y|\xi|}|m(\xi)|^2\,d\xi + Cy^{-2k}\int_{\mathbb{R}^n} \big(|M(w,y/2)|\,|w|^k\big)^2\,dw \\
&= C\int_{\mathbb{R}^n} y^{-n}|\xi|^{2k}e^{-4\pi|\xi|}|m(\xi/y)|^2\,d\xi + Cy^{-2k}\int_{\mathbb{R}^n} \big(|M(w,y/2)|\,|w|^k\big)^2\,dw \\
&\le Cy^{-n} + Cy^{-2k}\int_{\mathbb{R}^n} \big(|M(w,y/2)|\,|w|^k\big)^2\,dw,
\end{aligned}$$

where in the last inequality we have used the fact that m is bounded. It remains to estimate the second expression in the previous inequality. For this observe that it is enough to estimate

$$y^{-2k}\int_{\mathbb{R}^n}|x^\beta M(x,y)|^2\,dx \tag{10.34}$$

where $\beta=(\beta_1,\ldots,\beta_n)$ is any multi-index with $|\beta|=k$. (Recall that $x^\beta = x_1^{\beta_1}x_2^{\beta_2}\cdots x_n^{\beta_n}$). By (10.31) and Plancherel's theorem, (10.34) equals

$$y^{-2k}\int_{\mathbb{R}^n}\Big|D^\beta\big((y|\xi|)^k e^{-2\pi y|\xi|}m(\xi)\big)\Big|^2\,d\xi=\int_{\mathbb{R}^n}\Big|D^\beta\big(|\xi|^k e^{-2\pi y|\xi|}m(\xi)\big)\Big|^2\,d\xi$$

Writing $\beta=\beta_1+\beta_2+\beta_3$ with $|\beta_1|+|\beta_2|+|\beta_3|=k$ we see by Leibnitz's rule that the derivatives in the previous expression are dominated by sums involving terms of the form $D^{\beta_1}\big(|\xi|^k\big)D^{\beta_2}\big(e^{-2\pi y|\xi|}\big)D^{\beta_3}m(\xi)$ which in term is dominated by a sum involving terms of the form $|\xi|^{k-|\beta_1|}y^{|\beta_2|}\big|D^{\beta_3}m(\xi)\big|e^{-2\pi y|\xi|}=|\xi|^{|\beta_2|+|\beta_3|}y^{|\beta_2|}\big|D^{\beta_3}m(\xi)\big|e^{-2\pi y|\xi|}=\big(y|\xi|\big)^{|\beta_2|}\big(|\xi|^{\beta_3}\big|D^{\beta_3}m(\xi)\big|\big)e^{-2\pi y|\xi|}$.

Let $m_0(x)=\big(|x|^{\beta_3}\big|D^{\beta_3}m(x)\big|\big)^2$. By (10.28)

$$\sup_{R>0}\Big(\frac{1}{R^n}\int_{\{R<|x|<2R\}}m_0(x)\,dx\Big)=\sup_{R>0}\Big(\int_{\{1<|x|<2\}}m_0(Rx)\,dx\Big)\le B_k$$

However,

$$\begin{aligned}
&y^{2\beta_2}\int_{\mathbb{R}^n}|x|^{2\beta_2}m_0(x)e^{-4\pi y|x|}\,dx\\
&=\ C\frac{1}{y^n}\int_{\mathbb{R}^n}|x|^{2\beta_2}m_0(x/y)e^{-|x|}\,dx\\
&=\ C\frac{1}{y^n}\sum_{j=-\infty}^{\infty}\int_{\{2^j<|x|<2^{j+1}\}}|x|^{2\beta_2}m_0(x/y)e^{-|x|}\,dx\\
&\le\ C\frac{1}{y^n}\sum_{j=-\infty}^{\infty}2^{2(j+1)\beta_2}e^{-2^j}\int_{\{2^j<|x|<2^{j+1}\}}m_0(x/y)\,dx\\
&=\ C\frac{1}{y^n}\sum_{j=-\infty}^{\infty}2^{2j\beta_2}e^{-2^j}2^{jn}\int_{\{1<|x|<2\}}m_0(2^jx/y)\,dx\\
&\le\ CB_k\frac{1}{y^n}\sum_{j=-\infty}^{\infty}2^{j(2\beta_2+n)}e^{-2^j}
\end{aligned}$$

$$= \quad C(k, n, B_k)\frac{1}{y^n}.$$

This completes the proof of Theorem 10.10. □

Bibliography

[BL] R. Bañuelos and A.J. Lindeman. *A martingale study of the Beurling-Ahlfors transform in R^n*. Journal of Functional Analysis. **145**, 1997, pp 224–265.

[BM] R. Bañuelos and P.J. Méndez-Hernández. *Space–Time Brownian Motion and the Beurling-Ahlfors Transform.* Indiana University Math J. **52**, 2003, pp. 981–990.

[BaMo] R. Bañuelos and M. Moore. *Probabilistic Behavior of Harmonic Functions* Birkhäuser. 1999.

[BC] R. Bañuelos and T. Carroll. *Brownian motion and the fundamental frequency of a drum.* Duke Math J. **75**, 1994, pp. 575–602.

[BW] R. Bañuelos and G. Wang. *Sharp inequalities for martingales with applications to the Beurling-Ahlfors and Riesz transforms.* Duke Math Journal. **80**, 1995, pp. 575–600.

[Bas] R. Bass. *Probabilistic Techniques in Analysis*. Springer–Verlag, New York, 1995.

[BM] A. Baernstein II and S. J. Montgomery-Smith. *Some Conjectures about integral means of ∂f and $\overline{\partial} f$* Acta Universitatis Upsaliensis (Festschrift for M.Essén, edited by C. Kiselman). **64**, 1999, pp. 92–109.

[Bu1] D.L. Burkholder. *Sharp inequalities for martingales and stochastic integrals, Colloque Paul Lévy (Palaiseau)*. Astérique. 1987, pp. 75–94.

[Bu2] D.L. Burkholder. *Martingales and singular integrals in Banach spaces in Handbook on the geometry of Banach spaces, W. B. Johnson and J. Lindenstrauss, editors*. North Holland Pub. Co. 2001.

[Ca] M.L. Cartright. *Manuscripts of Hardy, Littlewood, Marcel Riesz and Titchmarsh.* Bull. London Math. Soc. **14**, 1982, pp. 472–532.

[DC] R.A. DeVore and R.C .Sharpley. *Maximal functions measuring smoothness*. American Mathematical Society. 1984.

[Da] E.B. Davies. *Heat Kenrels and Spectral Theory* Cambridge University Press, Cambridge. 1989.

[Do] B. Dacorogna. *Direct Methods in the Calculus of Variations*. Springer. 1989.

[DR] J. Duoandikoetxea and J.L. Rubio de Francia. *Estimations independantes de la dimension pour les transformees de Riesz.* C. R. Acad. Sci. **300, Serie I**, 1985, pp. 193–196.

[DS] S. Donalson and D. Sullivan. *Quasiconformal 4–manifolds* Acta Math. **163**, 1987, pp. 181–252.

[Dur1] R. Durrett. *Brownian Motion and Martingales in Analysis*. Wadsworth, Belmont, CA. 1984.

[Dur2] R. Durrett. *Probability: Theory and Examples, Second Edition*. Duxbury Press. 1996.

[Gu] M. de Guzman. *Real variable methods in Fourier analysis*. North-Holland Pub. Co. 1981.

[IM] T. Iwaniec and G. Martin. *The Beurling-Ahlfors transform in* $\mathbb{R}^n$ *and related singular integrals, preprint*.

[Iw1] T. Iwaniec. *Extremal inequalities in Sobolev spaces and quasiconformal mappings*. Z. Anal. Anwendungen **6**, 1982 pp. 1–16.

[Iw2] T. Iwaniec. *An approach to Cauchy–Riemann operators in* $\mathbb{R}^n$*, preprint*.

[LV] O. Lehto and K.I. Virtanen. *Quasiconformal Mappings in the Plane*. Springer. 1973.

[Ma] J. Manfredi. *Quasiregular Mappings from the Multilinear Point of View, Lecture Notes*. University of Jyvskyl, 1994.

[NV] F. Nazarov and A. Volberg. *Heating of the Beurling operator and the estimates of its norm, (preprint)*.

[Pi] S. Pichorides. *A note on the Littlewood–Paley square function inequality*. Colloq. Math. **60/61**, 1990 pp. 687–691.

[Ri] M. Riesz. *Sur les functions conjuguées*. Math. Zeits. **27**, 1927, pp. 218–244.

[Ro] H.L. Royden. *Real Analysis (second edition)*. MacMillan Publishing Co., Inc. 1968.

[Ru1] W. Rudin. *Principles of Mathematical Analysis (third edition)*. McGraw–Hill, New York. 1976.

[Ru2] W. Rudin. *Real and Complex Analysis (third edition)*. McGraw–Hill. New York. 1987.

[Sa] L. Saloff–Coste. *Aspects of Sobolev-type inequalities*. Cambridge ; New York : Cambridge University Press. 2002.

[St-St] E.M. Stein and J.-O. Strömberg. *Behavior of maximal functions in* $\mathbb{R}^n$ *for large* n. Ark. Mat. **21**, 1983, pp. 259–269

[St1] E.M. Stein. *Singular Integrals and Differentiability Properties of Functions*. Princeton Univ. Press, Princeton, NJ. 1970

[St2] E.M. Stein. *Problems in harmonic analysis related to curvature and oscillatory integrals*. Proceedings of the International Congress of Mathematicians, 1986, Berkeley, CA.

[SW] E.M. Stein and G. Weiss *On the theory of harmonic functions of several variables, I.* Acta Math. **103**, 1960, pp. 25–62

[To1] A. Torchinsky. *Real Variable Methods in Harmonic Analysis* Academic Press, Inc. Orlando, FL. 1986

[To2] A. Torchinsky. *Real Variables*. Addison–Wesley. 1988

Index of Notation

Index