Lectures in Harmonic Analysis

Joseph Breen

Lectures by: Monica Visan

Last updated: June 29, 2017

Department of Mathematics
University of California, Los Angeles

Contents

Preface

Eventually I'll probably write a better introduction, this is a placeholder for now

The following text covers material from Monica Visan's Math 247A and Math 247B classes, offered at UCLA in the winter and spring quarters of 2017. The content of each chapter is my own adaptation of lecture notes taken throughout the two quarters.

Chapter 1 contains a selection of background material which was not covered in lectures. The rest of the chapters roughly follow the chronological order of lectures given throughout the winter and spring, with some slight rearranging of my own choice. For example, basic results about the Fourier transform on the torus are included in Chapter 2, despite being the subject of lectures in the spring. A strictly chronological (and incomplete) version of these notes can be found on my personal web page at `http://www.math.ucla.edu/~josephbreen/`.

Finally, any mistakes are likely of my own doing and not Professor Visan's. Comments and corrections are welcome.

Chapter 1

Preliminaries

The material covered in these lecture notes is presented at the level of an upper division graduate course in harmonic analysis. As such, we assume that the reader is comfortable with real and complex analysis at a sophisticated level. However, for convenience and completeness we recall some of the basic facts and inequalities that are used throughout the text. We also establish conventions with notation, and discuss a few select topics that the reader may be unfamiliar with (for example, the Riesz-Thorin interpolation theorem). Any standard textbook in real analysis or harmonic analysis is a suitable reference for this material, for example, [3], [6], and [8].

1.1 Notation and Conventions

If $x \leq Cy$ for some constant C, we say $x \lesssim y$. If $x \lesssim y$ and $y \lesssim x$, then $x \sim y$. If the implicit constant depends on additional data, this is manifested as a subscript in the inequality sign. For example, if A denotes the ball of radius $r > 0$ in $\mathbb{R}^d$ and $|\cdot|$ denotes $\mathbb{R}^d$-Lebesgue measure, we have $|A| \sim_d r^d$. However, the dependencies of implicit constants are usually clear from context, or stated otherwise.

FINISH

1.2 L^p Spaces

ADD

Proposition 1.1 (Young's Convolution inequality)**.** *For $1 \leq p, q, r \leq \infty$,*

$$\|f * g\|_{L^r(\mathbb{R}^d)} \lesssim \|f\|_{L^p(\mathbb{R}^d)} \|g\|_{L^q(\mathbb{R}^d)}$$

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

Young's convolution inequality may be proved directly, or by using the Riesz-Thorin interpolation theorem (see Section 1.3). For details on both methods, refer to [3].

It is convenient to record a few special cases of Young's inequality that arise frequently.

Corollary 1.2 (Young's Convolution inequality, special cases)**.** *The following convolution inequalities hold:*

1. *For any $1 \leq p \leq \infty$, $\|f * g\|_{L^\infty} \lesssim \|f\|_{L^p} \|g\|_{L^{p'}}$. In particular, $\|f * g\|_{L^\infty} \lesssim \|f\|_{L^2} \|g\|_{L^2}$ and $\|f * g\|_{L^\infty} \lesssim \|f\|_{L^1} \|g\|_{L^\infty}$.*

2. *For any $1 \leq p \leq \infty$, $\|f * g\|_{L^p} \lesssim \|f\|_{L^1} \|g\|_{L^p}$.*

1.3 Complex Interpolation

An important tool in harmonic analysis is *interpolation*. Broadly speaking, interpolation considers the following question: given estimates of some kind on two different spaces, say L^{p_0} and L^{p_1}, what can be said about the corresponding estimate on L^p for values of p between p_0 and p_1?

There are a number of theorems which answer this kind of question. In Chapter 3, we will prove the Marcinkiewicz interpolation theorem using real analytic methods. Here, we prove a different interpolation theorem using complex analytic methods. Both theorems have their own advantages and disadvantages, and will be used frequently.

Our treatment of the Riesz-Thorin interpolation theorem follows [8]. The main statement is as follows.

Theorem 1.3 (Riesz-Thorin interpolation)**.** *Suppose that T is a bounded linear map from $L^{p_0} + L^{p_1} \to L^{q_0} + L^{q_1}$ and that*

$$\|Tf\|_{L^{q_0}} \leq M_0 \|f\|_{L^{p_0}} \qquad \textit{and} \qquad \|Tf\|_{L^{q_1}} \leq M_1 \|f\|_{L^{p_1}}$$

for some $1 \leq p_0, p_1, q_0, q_1 \leq \infty$. Then

$$\|Tf\|_{L^{q_t}} \leq M_0^{1-t} M_1^t \|f\|_{L^{p_t}}$$

where $\frac{1}{p_t} = \frac{1-t}{p_0} + \frac{t}{p_1}$ and $\frac{1}{q_t} = \frac{1-t}{q_0} + \frac{t}{q_1}$, for any $0 \leq t \leq 1$.

As noted above, the proof of this theorem relies on complex analysis. Specifically, we need the following lemma.

Lemma 1.4 (Three Lines lemma)**.** *Suppose that $\Phi(z)$ is holomorphic in $S = \{0 < \operatorname{Re} z < 1\}$ and continuous and bounded on $\bar{S}$. Let*

$$M_0 := \sup_{y \in \mathbb{R}} |\Phi(iy)| \qquad \textit{and} \qquad M_1 := \sup_{y \in \mathbb{R}} |\Phi(1 + iy)|.$$

Then

$$\sup_{y \in \mathbb{R}} |\Phi(t + iy)| \leq M_0^{1-t} M_1^t$$

for any $0 \leq t \leq 1$.

Proof. We can assume without loss of generality that Φ is nonconstant, otherwise the conclusion is trivial.

We first prove a special case of the lemma. Suppose that $M_0 = M_1 = 1$ and that $\sup_{0 \leq x \leq 1} |\Phi(x + iy)| \to 0$ as $|y| \to \infty$. In this case, ϕ has a global (on $\bar{S}$) supremum of $M > 0$. Let $\{z_n\}$ be a sequence such that $|\Phi(z_n)| \to M$. Because $\sup_{0 \leq x \leq 1} |\Phi(x + iy)| \to 0$ as $|y| \to \infty$, the sequence $\{z_n\}$ is bounded, hence there is a point $z_\infty \in \bar{S}$ such that a subsequence of $\{z_n\}$ converges to z_∞. By continuity, $|\Phi(z_\infty)| = M$. Since Φ is nonconstant, by the maximum modulus principle, z_∞ is on the boundary of $\bar{S}$. This means that $|\Phi(z)|$ is globally bounded by 1. Since $M_0 = M_1 = 1$, this proves the special case.

Next, we remove the decay assumption and only suppose that $M_0 = M_1 = 1$. For $\varepsilon > 0$, define

$$\Phi_\varepsilon(z) := \Phi(z) e^{\varepsilon(z^2 - 1)}.$$

First observe that $|\Phi_\varepsilon(z)| \leq 1$ if $\operatorname{Re} z = 0$ or $\operatorname{Re} z = 1$. Indeed, note that for $y \in \mathbb{R}$ we have

$$|\Phi_\varepsilon(iy)| = |\Phi(iy)| \left| e^{\varepsilon((iy)^2 - 1)} \right| \leq \left| e^{-\varepsilon(y^2 + 1)} \right| \leq 1$$

and

$$|\Phi_\varepsilon(1 + iy)| = |\Phi(1 + iy)| \left| e^{\varepsilon((1+iy)^2 - 1)} \right| \leq \left| e^{\varepsilon(1 + 2iy - y^2 - 1)} \right| \leq \left| e^{2iy\varepsilon} \right| \left| e^{-\varepsilon y^2} \right| \leq 1.$$

Next, we claim that Φ_ε satisfies the decay condition from the previous case. Indeed, note that

$$\begin{aligned}|\Phi_\varepsilon(x+iy)| &= |\Phi(x+iy)|\left|e^{\varepsilon((x+iy)^2-1)}\right| = |\Phi(x+iy)|\left|e^{\varepsilon(x^2-y^2-1+2xyi)}\right| \\ &\leq |\Phi(x+iy)|\left|e^{\varepsilon(x^2-y^2-1)}\right|.\end{aligned}$$

Because $0 \leq x \leq 1$ and because $|\Phi|$ is bounded, as $|y| \to \infty$, $|\Phi_\varepsilon(x+iy)| \to 0$ uniformly in x. By the previous case, $|\Phi_\varepsilon(z)| \leq 1$ uniformly in $\bar{S}$. Since this holds for all $\varepsilon > 0$, letting $\varepsilon \to 0$ gives the desired result for Φ.

Finally we assume that M_0 and M_1 are arbitrary positive numbers. Define $\tilde{\phi}(z) := M_0^{z-1}M_1^{-z}\phi(z)$. Note that if $\operatorname{Re} z = 0$, then

$$|\tilde{\Phi}(z)| \leq |M_0^{iy-1}M_1^{-iy}|M_0 \leq |M_0^{iy}| \cdot |M_1^{-iy}| \leq 1$$

and if $\operatorname{Re} z = 1$, then

$$|\tilde{\Phi}(z)| \leq |M_0^{iy}M_1^{-1-iy}|M_1 \leq |M_0^{iy}| \cdot |M_1^{-iy}| \leq 1.$$

The claim then follows by applying the previous case. $\square$

Proof of Theorem 1.3. By scaling appropriately, we may assume without loss of generality that $\|f\|_{L^{p_t}} = 1$.

First, consider the case when f is a simple function. By duality,

$$\|Tf\|_{L^{q_t}} = \sup\left|\int Tf(x)\,g(x)\,dx\right|$$

where the supremum is taken over simple functions g with $\|g\|_{L^{q_t'}} = 1$. Fix such a g. First, consider the case $p_t < \infty$ and $q_t > 1$. Let

$$\gamma(z) := p_t\left(\frac{1-z}{p_0} + \frac{z}{p_1}\right) \qquad \text{and} \qquad \delta(z) := q_t'\left(\frac{1-z}{q_0'} + \frac{z}{q_1'}\right).$$

Define

$$f_z := |f|^{\gamma(z)}\frac{f}{|f|} \qquad \text{and} \qquad g_z := |g|^{\delta(z)}\frac{g}{|g|}.$$

Note that $\gamma(t) = 1$. Hence, $f_t = f$. Also, if $\operatorname{Re} z = 0$, then

$$\|f_z\|_{L^{p_0}}^{p_0} = \int \left||f|^{p_t\gamma(z)}\right| dx = \int \left||f|^{p_t - z\left(p_t + \frac{p_0}{p_1}\right)}\right| dx = \int |f|^{p_t}\,dx = \|f\|_{L^{p_t}}^{p_t} = 1.$$

A similar computation gives $\|f_z\|_{L^{p_1}} = 1$ if $\operatorname{Re} z = 1$. Also, $\|g_z\|_{L^{q_0'}} = 1$ if $\operatorname{Re} z = 0$ and $\|g_z\|_{L^{q_1'}} = 1$ if $\operatorname{Re} z = 1$.

Next, define

$$\Phi(z) := \int Tf_z(x)\,g_z(x)\,dx.$$

As f and g are simple, we may write $f = \sum_k a_k\chi_{E_k}$ where the sets E_k are disjoint and of finite measure, and likewise $g = \sum_j b_j\chi_{F_j}$. Then

$$f_z = \sum_k |a_k|^{\gamma(z)}\frac{a_k}{|a_k|}\chi_{E_k} \qquad \text{and} \qquad g_z = \sum_j |b_j|^{\delta(z)}\frac{b_j}{|b_j|}\chi_{F_j}.$$

Also with this notation we have

$$\Phi(z) = \sum_{j,k} |a_k|^{\gamma(z)}|b_j|^{\delta(z)}\frac{a_k}{|a_k|}\frac{b_j}{|b_j|}\int T(\chi_{E_k})\,\chi_{F_j}\,dx.$$

From this it is evident that $\Phi(z)$ is holomorphic in $S = \{0 < \operatorname{Re} z < 1\}$ and continuous and bounded on $\bar{S}$. Moreover, if $\operatorname{Re} z = 0$, then Holder's inequality gives

$$|\Phi(z)| \leq \int |Tf_z(x)\, g_z(x)|\, dx \leq \|Tf_z\|_{L^{q_0}} \|g_z\|_{L^{q_0'}} \leq M_0 \|f\|_{L^{p_0}} = M_0.$$

Similarly, if $\operatorname{Re} z = 1$, $|\Phi(z)| \leq M_1$. Therefore, by the Three Lines lemma, if $\operatorname{Re} z = t$ then

$$|\Phi(z)| \leq M_0^{1-t} M_1^t.$$

As $\Phi(t) = \int Tf(x)\, g(x)\, dx$, this is the desired result.

The passage from simple functions to general functions is a straightforward limiting argument, and the remaining cases $p_t = 1$ and $q_t = \infty$ are handled similarly. For more details, consult [8].

□

It is convenient to state a few special cases of Riesz-Thorin interpolation. These occur frequently, and also provide more concrete examples of interpolation at work.

Corollary 1.5 (Riesz-Thorin interpolation, special cases)**.**

1. *Fix $1 \leq p_0 < p_1 \leq \infty$. Suppose that T is bounded as a linear map from $L^{p_0} \to L^{p_0}$ and also from $L^{p_1} \to L^{p_1}$. Then T is bounded from $L^p \to L^p$ for all $p_0 \leq p \leq p_1$.*
2. *Suppose that T is bounded as a linear map from $L^2 \to L^2$ and also from $L^1 \to L^\infty$. Then T is bounded from $L^p \to L^{p'}$ for all $1 \leq p \leq 2$.*

Proof.

1. If T is bounded from $L^{p_0} \to L^{p_0}$ and $L^{p_1} \to L^{p_1}$, then Riesz-Thorin implies that T is bounded from $L^{p_t} \to L^{q_t}$ for all $0 \leq t \leq 1$ with
$$\frac{1}{p_t} = \frac{1-t}{p_0} + \frac{t}{p_1} \quad \text{and} \quad \frac{1}{q_t} = \frac{1-t}{p_0} + \frac{t}{p_1}.$$
Thus, $p_t = q_t$. Moreover, as t ranges from 0 to 1, p_t ranges from p_0 to p_1. Therefore, T is bounded from $L^p \to L^p$ for all $p_0 \leq p \leq p_1$.

2. Riesz-Thorin gives boundedness of $T : L^{p_t} \to L^{q_t}$ for $0 \leq t \leq 1$ with
$$\frac{1}{p_t} = \frac{1-t}{1} + \frac{t}{2} \quad \text{and} \quad \frac{1}{q_t} = \frac{1-t}{\infty} + \frac{t}{2}.$$
Thus, $\frac{1}{p_t} = 1 - \frac{t}{2}$ and $\frac{1}{q_t} = \frac{t}{2}$ so that $q_t = p_t'$. Moreover, $p_t = \frac{2}{2-t}$, so as t ranges from 0 to 1, p_t ranges from 1 to 2.

□

The first substantial application of interpolation is found in found in Chapter 2, where we show that the Fourier transform is well-defined on L^p spaces for $1 \leq p \leq 2$. As a final remark, the Riesz-Thorin interpolation theorem is an interpolation result based on *complex* analytic tools, namely, the maximum principle for holomorphic functions. In Chapter 3, we prove a different interpolation theorem — the Marcinkiewicz interpolation theorem — based on *real* analytic techniques.

1.4 Some General Principles

Throughout the rest of this text, there are a number of straightforward principles that we will employ over and over again, often without explicit comment. For convenience, we include a short discussion of these recurring principles here.

1.4.1 Integrating $|x|^{-\alpha}$

The first concern integration of functions of the form $|x|^{-\alpha}$. In words, integrating $|x|^{-\alpha}$ over a region of $\mathbb{R}^d$ increases the exponent by d. Explicitly, we have the following proposition.

Proposition 1.6. *Fix $R > 0$. Then*

1. *The integral $\int_{|x|>R} |x|^{-\alpha}\, dx$ is finite when $\alpha > d$, in which case*
$$\int_{|x|>R} |x|^{-\alpha}\, dx \;\sim\; R^{-\alpha+d}.$$
2. *The integral $\int_{|x|<R} |x|^{-\alpha}\, dx$ is finite when $\alpha \leq d$, in which case*
$$\int_{|x|<R} |x|^{-\alpha}\, dx \;\sim\; R^{-\alpha+d}.$$

Consequently, $|x|^{-\alpha} \notin L^p(\mathbb{R}^d)$ for $1 \leq p < \infty$, for any choice of α.

Proof. Fix $R > 0$. Set $r = |x|$. Then $dx = r^{d-1}\, dr$. Thus,
$$\int_{|x|>R} \frac{1}{|x|^\alpha}\, dx = \int_{r>R} \frac{1}{r^\alpha} r^{d-1}\, dr = \int_{r=R}^{\infty} r^{-\alpha+d-1}\, dr \;\sim\; r^{-\alpha+d}\Big|_R^\infty .$$
This limit exists precisely when $-\alpha + d < 0$, i.e., when $\alpha > d$, in which case
$$\int_{|x|>R} \frac{1}{|x|^\alpha}\, dx \;\sim\; R^{-\alpha+d}.$$
Similarly,
$$\int_{|x|<R} \frac{1}{|x|^\alpha}\, dx \;\sim\; r^{-\alpha+d}\Big|_0^R .$$
This limit exists precisely when $-\alpha + d \geq 0$, i.e., $\alpha \leq d$, in which case
$$\int_{|x|<R} \frac{1}{|x|^\alpha}\, dx \;\sim\; R^{-\alpha+d}.$$
□

1.4.2 Dyadic Sums

Another computational tool ubiquitous in harmonic analysis is dyadic sums; that is, geometric series with $r = \frac{1}{2}$. The following proposition is easy, but worth recording.

Proposition 1.7. *Let $2^{\mathbb{Z}} = \{\, 2^n \;:\; n \in \mathbb{Z} \,\}$ be the set of dyadic numbers. Then*
$$\sum_{N \in 2^{\mathbb{Z}};\, N \geq N_0} \frac{1}{N} = 2 \cdot \frac{1}{N_0}$$
and
$$\sum_{N \in 2^{\mathbb{Z}};\, N \leq N_0} N = 2 \cdot N_0$$

Proof. Fix a dyadic number $N_0 = 2^{n_0} \in 2^{\mathbb{Z}}$. Then

$$\sum_{N \in 2^{\mathbb{Z}};\, N \geq N_0} \frac{1}{N} = \sum_{n \in \mathbb{Z};\, n \geq n_0} 2^{-n} = \sum_{n=n_0}^{\infty} \left(\frac{1}{2}\right)^n = \frac{\left(\frac{1}{2}\right)^{n_0}}{1 - \frac{1}{2}} = 2 \cdot 2^{-n_0} = 2 \cdot \frac{1}{N_0}.$$

Essentially the same computation gives

$$\sum_{N \in 2^{\mathbb{Z}};\, N \leq N_0} N = 2 \cdot N_0.$$

$\square$

Chapter 2

The Fourier Transform

Our study of harmonic analysis naturally begins with the Fourier transform. Historically, the Fourier transform was first introduced by Joseph Fourier in his study of the heat equation. In this context, Fourier showed how a periodic function could be decomposed into a sum of sines and cosines which represent the frequencies of the function. From a modern point of view, the Fourier transform is a transformation which accepts a function and returns a new function, defined via the frequency data of the original function. As a bridge between the physical domain and the frequency domain, the Fourier transform is the main tool of use in harmonic analysis.

The basic theory of the Fourier transform is standard, and there are a wealth of references (for example, [3], [6], and [7]) for the following material.

2.1 The Fourier Transform on $\mathbb{R}^d$

We begin by discussing the Fourier transform of complex-valued functions on the Euclidean space $\mathbb{R}^d$. The Fourier transform is defined as an integral transformation of a function; as such, it is natural to first consider integrable functions.

Definition 2.1. The **Fourier transform** of a function $f \in L^1(\mathbb{R}^d)$ is given by

$$\mathcal{F}(f)(\xi) := \hat{f}(\xi) := (f)\hat{}\,(\xi) := \int_{\mathbb{R}^d} e^{-2\pi i x\cdot\xi} f(x)\,dx.$$

There are a number of conventions for placement of constants when defining the Fourier transform. For example, another common definition of the Fourier transform is

$$\hat{f}(\xi) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x)\,dx.$$

These choices clearly do not alter the theory in any meaningful way; they are simply a matter of notational preference. In this text, we will use either convention when convenient. In this section, we will use the former.

We have two immediate goals. One is to understand the basic properties of the Fourier transform and how it interacts with operations such as differentiation, convolution, etc. The other goal is to understand which classes of functions the Fourier transform can be reasonably defined on. As and aid towards both of these goals, we introduce (or recall) a suitably nice class of functions. Here, we use the following *multi-index* notation: for $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$, define

$$|\alpha| := \alpha_1 + \cdots \alpha_d; \qquad x^\alpha := x_1^{\alpha_1} \cdots x_d^{\alpha_d}; \qquad D^\alpha := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}.$$

Definition 2.2. A C^∞-function $f : \mathbb{R}^d \to \mathbb{C}$ is **Schwartz** if $x^\alpha D^\beta f \in L^\infty_x(\mathbb{R}^d)$ for every multiindex $\alpha, \beta \in \mathbb{N}^d$. The collection of Schwartz functions, **Schwartz space**, is denoted $\mathcal{S}(\mathbb{R}^d)$.

In words, a function is Schwartz if it is smooth, and if all of its derivatives decay faster than any polynomial. An example of a Schwartz function is $e^{-|x|^2}$.

The set of Schwartz functions forms a *Frechet space*, i.e., a locally convex space (a vector space endowed with a family of seminorms $\{\rho_\alpha\}$ that separates points: if $\rho_\alpha(f) = 0$ for all α, then $f = 0$) which is metrizable and complete. In the case of $\mathcal{S}(\mathbb{R}^d)$, the collection of seminorms is given by $\{\rho_{\alpha,\beta}\}_{\alpha,\beta\in\mathbb{N}^d}$, where

$$\rho_{\alpha,\beta}(f) := \left\| x^\alpha D^\beta f \right\|_{L^\infty_x(\mathbb{R}^d)}.$$

The Frechet space structure of $\mathcal{S}(\mathbb{R}^d)$ is not our main concern; rather, density in L^p spaces (indeed, note that $C^\infty_c(\mathbb{R}^d) \subseteq \mathcal{S}(\mathbb{R}^d)$) and its interaction with the Fourier transform are of more importance.

Here, we collect a number of basic and important properties of the Fourier transform of Schwartz functions.

Proposition 2.3. *Let $f \in \mathcal{S}(\mathbb{R}^d)$. Then:*

1. *If $g(x) = f(x-y)$, then $\hat{g}(\xi) = e^{-2\pi i y\cdot\xi}\hat{f}(\xi)$.*
2. *If $g(x) = e^{2\pi i x\cdot\eta} f(x)$, then $\hat{g}(\xi) = \hat{f}(\xi - \eta)$.*
3. *If $g(x) := f(Tx)$ for $T \in GL(\mathbb{R}^d)$, then $\hat{g}(\xi) = |\det T|^{-1}\hat{f}(T^{-t}\xi)$. In particular, if T is a rotation or reflection and $f(Tx) = f(x)$ (i.e. f is rotation or reflection invariant), then $\hat{g}(\xi) = \hat{f}(\xi)$.*
4. *If $g(x) := \overline{f(x)}$, then $\hat{g}(\xi) = \overline{\hat{f}(-\xi)}$.*
5. *If $g(x) := D^\alpha f(x)$, then $\hat{g}(\xi) = (2\pi i\xi)^\alpha \hat{f}(\xi)$.*
6. *If $g(x) := x^\alpha f(x)$, then $\hat{g}(\xi) = \left(\frac{i}{2\pi}\right)^{|\alpha|} D^\alpha_\xi \hat{f}(\xi)$.*
7. *If $g(x) := (k * f)(x)$ for $k \in L^1(\mathbb{R}^d)$, then $\hat{g}(\xi) = \hat{k}(\xi)\cdot\hat{f}(\xi)$.*
8. *We have the basic estimate $\left\|\hat{f}\right\|_{L^\infty_\xi} \leq \|f\|_{L^1_x}$.*

Proof.

1. Making the change of variables $z = x - y$, we have

$$\begin{aligned}\hat{g}(\xi) &= \int e^{-2\pi i x\cdot\xi} f(x-y)\,dx = \int e^{-2\pi i(z+y)\cdot\xi} f(z)\,dz = e^{-2\pi i y}\int e^{-2\pi i z\cdot\xi} f(z)\,dz \\ &= e^{-2\pi i y}\hat{f}(\xi).\end{aligned}$$

2. Computing directly gives:

$$\hat{g}(\xi) = \int e^{-2\pi i x\cdot\xi} e^{2\pi i x\cdot\eta} f(x)\,dx = \int e^{-2\pi i x\cdot(\xi-\eta)} f(x)\,dx = \hat{f}(\xi-\eta).$$

3. For $T \in GL(\mathbb{R}^d)$, make the change of variables $y = Tx$. Then

$$\begin{aligned}\hat{g}(\xi) &= \int e^{-2\pi i x\cdot\xi} f(Tx)\,dx = \int e^{-2\pi i(T^{-1}y)\cdot\xi} f(y)|\det T|^{-1}\,dy \\ &= |\det T|^{-1}\int e^{-2\pi i y\cdot(T^{-t}\xi)} f(y)\,dy \\ &= |\det T|^{-1}\hat{f}(T^{-t}\xi).\end{aligned}$$

If T is a rotation or reflection (or more generally, T is orthogonal), then $T^tT = TT^t = I$ and $\det T = 1$. In this case, $\hat{g}(\xi) = |\det T|^{-1}\hat{f}(T^{-t}\xi) = \hat{f}(\xi)$.

4. Computing directly gives:

$$\hat{g}(\xi) = \int e^{-2\pi i x\cdot\xi}\overline{f(x)}\,dx = \overline{\int \overline{e^{-2\pi i x\cdot\xi}}f(x)\,dx} = \overline{\int e^{-2\pi i x\cdot(-\xi)}f(x)\,dx} = \overline{\hat{f}(-\xi)}.$$

5. Integrating by parts, we have:

$$\begin{aligned}\hat{g}(\xi) &= \int e^{-2\pi i x\cdot\xi}D^\alpha f(x)\,dx = (-1)^{|\alpha|}\int D^\alpha e^{-2\pi i x\cdot\xi}f(x)\,dx \\ &= (-1)^{|\alpha|}\int(-2\pi i\xi)^\alpha e^{-2\pi i x\cdot\xi}f(x)\,dx = (2\pi i\xi)^\alpha\int e^{-2\pi i x\cdot\xi}f(x)\,dx \\ &= (2\pi\xi)^\alpha\hat{f}(\xi).\end{aligned}$$

6. Note that $D_x^\alpha e^{-2\pi i x\cdot\xi} = (-2\pi i x)^\alpha e^{-2\pi i x\cdot\xi}$, so that $x^\alpha e^{-2\pi i x\cdot\xi}$. Thus, integrating by parts again gives:

$$\hat{g}(\xi) = \int e^{-2\pi i x\cdot\xi}x^\alpha f(x)\,dx = \left(\frac{i}{2\pi}\right)^{|\alpha|}\hat{f}(\xi).$$

7. Computing, we have

$$\begin{aligned}\hat{g}(\xi) &= \int e^{-2\pi i x\cdot\xi}(k*f)(x)\,dx = \int e^{-2\pi i x\cdot\xi}\int k(x-y)f(y)\,dy\,dx \\ &= \int\int e^{-2\pi i x\cdot\xi}k(x-y)f(y)\,dx\,dy.\end{aligned}$$

We make a change of variables $z = x - y$ in the inner integral to get

$$\begin{aligned}\hat{g}(\xi) &= \int\int e^{-2\pi i(z+y)\cdot\xi}k(z)f(y)\,dz\,dy = \int e^{-2\pi i z\cdot\xi}k(z)\,dz\cdot\int e^{-2\pi i y\cdot\xi}f(y)\,dy \\ &= \hat{k}(\xi)\cdot\hat{f}(\xi).\end{aligned}$$

8. This follows from the triangle inequality for integrals:

$$\left|\int e^{-2\pi i x\cdot\xi}f(x)\,dx\right| \leq |\int |e^{-2\pi i x\cdot\xi}||f(x)|\,dx = \int |f(x)|\,dx.$$

□

From this proposition, it becomes immediately clear why the Fourier transform is such a powerful computational tool. For example, properties 5 and 7 above describe how the Fourier transform turns complicated function operations like differentiation and convolution into multiplication. This becomes incredibly useful when studying partial differential equations, for instance. Also note that properties 1,2,3,4,7, and 8 extend immediately to functions in $L^1(\mathbb{R}^d)$.

Proposition 2.4. *If $f \in \mathcal{S}(\mathbb{R}^d)$, then $\hat{f} \in \mathcal{S}(\mathbb{R}^d)$. Moreover, if $f_n \to f \in \mathcal{S}(\mathbb{R}^d)$, then $\hat{f}_n \to \hat{f} \in \mathcal{S}(\mathbb{R}^d)$.*

Proof. Suppose that $f \in \mathcal{S}(\mathbb{R}^d)$. Note that

$$|\xi^\alpha D^\beta \hat{f}(\xi)| \sim |\xi^\alpha \widehat{x^\beta f}(\xi)| \sim |\widehat{D^\alpha x^\beta f}(\xi)| = \left| \int e^{-2\pi i x\cdot\xi} D^\alpha(x^\beta f(x))\, dx \right|.$$

Because f is Schwartz, $D^\alpha(x^\beta f(x)) \in L^1(\mathbb{R}^d)$. Therefore,

$$\left\| \xi^\alpha D^\beta \hat{f}(\xi) \right\|_{L^\infty_\xi} \lesssim \left\| D^\alpha(x^\beta f(x)) \right\|_{L^1} < \infty$$

so that $\hat{f} \in \mathcal{S}(\mathbb{R}^d)$. It is clear that $\{f_n\} \to f \in \mathcal{S}(\mathbb{R}^d)$ if and only if $\{\hat{f}_n\} \to \hat{f} \in \mathcal{S}(\mathbb{R}^d)$ by properties 5 and 6 of the previous proposition. □

Corollary 2.5 (Riemann-Lebesgue Lemma)**.** *If $f \in L^1(\mathbb{R}^d)$, then $\hat{f}$ is uniformly continuous and vanishes at ∞.*

Proof. Let $C_0(\mathbb{R}^d)$ denote the space of continuous functions which vanish at ∞.

Let $\{f_n\}$ be a sequence of Schwartz functions with $f_n \to f$ in $L^1(\mathbb{R}^d)$. Since

$$\left\| \widehat{f_n - f} \right\|_{L^\infty} \leq \|f_n - f\|_{L^1}$$

it follows that $\hat{f}_n \to \hat{f}$ in $L^\infty(\mathbb{R}^d)$. As $\hat{f} \in \mathcal{S}(\mathbb{R}^d)$, $\hat{f} \in C_0(\mathbb{R}^d)$. Since $C_0(\mathbb{R}^d)$ is closed under uniform convergence, we are done. □

Aside from simple properties and estimates of the Fourier transform and its interaction with various function operations, we have not computed any Fourier transforms of actual functions.

Lemma 2.6. *Let A be a real, symmetric, positive definite $d \times d$ matrix. Then*

$$\int e^{-x\cdot Ax} e^{-2\pi i x\cdot\xi}\, dx = \pi^{\frac{d}{2}} (\det A)^{-\frac{1}{2}} e^{-\pi^2 \xi\cdot A^{-1}\xi}.$$

Proof. Since A is real, symmetric, and positive definite, it is diagonalizable. Explicitly, there is an orthogonal matrix O and a diagonal matrix $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_d)$, $\lambda_j > 0$, so that $A = O^T D O$. Let $y = Ox$ and $\eta = O\xi$. Then

$$x \cdot Ax = x \cdot O^T DOx = Ox \cdot DOx = y \cdot Dy = \sum_{j=1}^d \lambda_j y_j^2$$

and

$$x \cdot \xi = O^T y \cdot O^T \eta = y \cdot \eta.$$

Thus,

$$\int e^{-x\cdot Ax} e^{-2\pi i x\cdot\xi}\, dx = \int e^{-\sum_{j=1}^d \lambda_j y_j^2} e^{-2\pi i y\cdot\eta}\, dy = \prod_{j=1}^d \int_{\mathbb{R}} e^{-\lambda_j y_j^2 - 2\pi i y_j \eta_j}\, dy_j.$$

Computing each of these one-variable integrals, we have:

$$\int_{\mathbb{R}} e^{-\lambda y^2 - 2\pi i y\eta}\, dy = \int_{\mathbb{R}} e^{-\lambda\left(y - \frac{\pi i \eta}{\lambda}\right)^2 - \frac{\pi^2\eta^2}{\lambda}}\, dy = e^{-\frac{\pi^2\eta^2}{\lambda}} \int_{\mathbb{R}} e^{-\lambda y^2}\, dy = e^{-\frac{\pi^2\eta^2}{\lambda}} \lambda^{-\frac{1}{2}} \pi^{\frac{1}{2}}.$$

So

$$\begin{aligned}
\int e^{-x\cdot Ax} e^{-2\pi i x\cdot\xi}\, dx &= \prod_{j=1}^d e^{-\frac{\pi^2\eta_j^2}{\lambda_j}} \lambda_j^{-\frac{1}{2}} \pi^{\frac{1}{2}} = \pi^{\frac{d}{2}} (\lambda_1 \cdots \lambda_d)^{-\frac{1}{2}} e^{-\sum_{j=1}^d \frac{\pi^2}{\lambda_j}\eta_j^2} \\
&= \pi^{\frac{d}{2}} (\det A)^{-\frac{1}{2}} e^{-\pi^2 \eta\cdot D^{-1}\eta}.
\end{aligned}$$

The result follows from the fact that $\eta \cdot D^{-1}\eta = O\xi \cdot D^{-1} O\xi = \xi \cdot A^{-1}\xi$. □

Corollary 2.7. *The function $e^{-\pi|x|^2}$ is an eigenvalue of the Fourier transform with eigenvalue 1. Explicitly, $\widehat{e^{-\pi|x|^2}} = e^{-\pi|\xi|^2}$.*

Proof. This follows from the previous lemma with $A = \pi I_d$. Indeed,

$$\int e^{-\pi|x|^2} e^{-2\pi i x\cdot\xi}\, dx = \pi^{\frac{d}{2}} \left(\pi^d\right)^{-\frac{1}{2}} e^{-\pi^2 \xi\cdot\frac{1}{\pi}\xi} = e^{-\pi|\xi|^2}.$$

□

Above, we computed that the Fourier transform of a Schwartz function is again Schwartz. In fact, the Fourier transform is an isometry on Schwartz space and extends to a unitary operator on L^2. The next few results summarize these facts.

Theorem 2.8 (Fourier Inversion). *If $f \in \mathcal{S}(\mathbb{R}^d)$, then $(\hat{f})\hat{}(x) = f(-x)$. Equivalently,*

$$f(x) = (\hat{f})\hat{}(-x) =: (\hat{f})\check{}(x) =: \mathcal{F}^{-1}(\hat{f})(x).$$

Proof. For $\varepsilon > 0$, let

$$I_\varepsilon(x) := \int e^{-\pi\varepsilon^2|\xi|^2} e^{2\pi i x\cdot\xi} \hat{f}(\xi)\, d\xi.$$

By the dominated convergence theorem, as $\varepsilon \to 0$, $I_\varepsilon(x) \to \int e^{2\pi i x\cdot\xi}\hat{f}(\xi)\, d\xi$. Thus, to prove the theorem, it suffices to show that $I_\varepsilon(x) \to f(x)$ as $\varepsilon \to 0$. We have

$$\begin{aligned} I_\varepsilon(x) &= \int e^{-\pi\varepsilon^2|\xi|^2} e^{2\pi i x\cdot\xi}\hat{f}(\xi)\, d\xi = \int e^{-\pi\varepsilon^2|\xi|^2} e^{2\pi i x\cdot\xi} \int e^{-2\pi i y\cdot\xi} f(y)\, dy\, d\xi \\ &= \int f(y) \int e^{-\pi\varepsilon^2|\xi|^2} e^{-2\pi i(y-x)\cdot\xi}\, d\xi\, dy. \end{aligned}$$

Note that

$$\begin{aligned} \int e^{-\pi\varepsilon^2|\xi|^2} e^{-2\pi i(y-x)\cdot\xi}\, d\xi &= \left(e^{-\pi\varepsilon^2|\xi|^2}\right)\hat{}\,(y-x) = \pi^{\frac{d}{2}}\left((\pi\varepsilon^2)^d\right)^{-\frac{1}{2}} e^{-\pi^2(y-x)\cdot\frac{1}{\pi\varepsilon^2}(y-x)} \\ &= \varepsilon^{-d} e^{-\pi\frac{|y-x|^2}{\varepsilon^2}}. \end{aligned}$$

Consequently,

$$I_\varepsilon(x) = \int f(y)\, \varepsilon^{-d} e^{-\pi\frac{|y-x|^2}{\varepsilon^2}}\, dy = (f * \phi_\varepsilon)(x)$$

where $\phi_\varepsilon(x) = \varepsilon^{-d}\phi\left(\frac{x}{\varepsilon}\right)$ with $\phi(x) = e^{-\pi|x|^2}$. It is a standard result that $\{\phi_\varepsilon\}$ is an approximation to the identity. Thus, as $\varepsilon \to 0$,

$$I_\varepsilon(x) = (f * \phi_\varepsilon)(x) \to f(x).$$

□

Lemma 2.9. *If $f, g \in \mathcal{S}(\mathbb{R}^d)$, then $\int \hat{f}(\xi)g(\xi)\, d\xi = \int f(x)\hat{g}(x)\, dx$. In particular, $\int \hat{f}(\xi)\overline{\hat{g}(\xi)}\, dx = \int f(x)\overline{g(x)}\, dx$, so that $\left\|\hat{f}\right\|_{L^2} = \|f\|_{L^2}$. Hence, the Fourier transform is an isometry on $\mathcal{S}(\mathbb{R}^d)$.*

Proof. Computing, we have

$$\begin{aligned} \int \hat{f}(\xi)g(\xi)\, d\xi &= \int\int e^{-2\pi i x\cdot\xi} f(x)\, dx\, g(\xi)\, d\xi = \int f(x) \int e^{-2\pi i x\cdot\xi} g(\xi)\, d\xi\, dx \\ &= \int f(x)\hat{g}(x)\, dx. \end{aligned}$$

For the "in particular" statement, let $h = \overline{\hat{g}}$. Then

$$\hat{h}(x) = \hat{\overline{\hat{g}}} = \overline{(\hat{g})\hat{}(-x)} = \overline{g(x)}.$$

□

Theorem 2.10 (Plancharel)**.** *The Fourier transform extends from an operator on $\mathcal{S}(\mathbb{R}^d)$ to a unitary operator on $L^2(\mathbb{R}^d)$.*

Proof. Fix $f \in L^2(\mathbb{R}^d)$. Let $\{f_n\} \subseteq \mathcal{S}(\mathbb{R}^d)$ such that $f_n \xrightarrow{L^2} f$. As the Fourier transform is an isometry on $\mathcal{S}(\mathbb{R}^d)$, we have $\left\|\hat{f}_n - \hat{f}_m\right\|_{L^2} = \|f_n - f_m\|_{L^2}$. This shows that $\{\hat{f}_n\}$ is Cauchy in L^2. Let $\hat{f} := \lim_{n\to\infty} \hat{f}_n$, the limit being taken in the L^2-sense.

We claim that $\hat{f}$ is well-defined. Let $\{f_n\}$ and $\{g_n\}$ be two sequences of Schwartz functions such that $f_n, g_n \xrightarrow{L^2} f$. Define

$$h_n := \begin{cases} f_k & \text{if } n = 2k-1 \\ g_k & \text{if } n = 2k \end{cases}.$$

Then $h_n \xrightarrow{L^2} f$. By the argument above, $\{\hat{h_n}\}$ converges in L^2, which implies by the uniqueness of limits that $\lim_{n\to\infty} \hat{f}_n = \lim_{n\to\infty} \hat{g}_n$.

Next we claim that $f \in L^2(\mathbb{R}^d)$, then $\left\|\hat{f}\right\|_{L^2} = \|f\|_{L^2}$, so the Fourier transform is an isometry on $L^2(\mathbb{R}^d)$. Because the norm function is continuous in the L^2 topology, we have

$$\left\|\hat{f}\right\|_{L^2} = \lim_{n\to\infty} \left\|\hat{f}_n\right\|_{L^2} = \lim_{n\to\infty} \|f_n\|_{L^2} = \left\|\lim_{n\to\infty} f_n\right\|_{L^2} = \|f\|_{L^2}.$$

Before completing the proof, we remark that in infinite dimensions, an isometry is not necessarily a unitary operator. For example, let $T : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ be the right-shift operator given by

$$T(a_0, a_1, a_2, \dots) := (0, a_0, a_1, \dots).$$

Then

$$T^*(a_0, a_1, a_2, \dots) := (a_1, a_2, a_3, \dots).$$

Clearly T is an isometry, but $TT^* \neq I$.

However, to prove that the isometry $\mathcal{F}$ is unitary, it suffices to prove that $\mathcal{F}$ is *surjective.* We claim that $\operatorname{im}\mathcal{F}$ is closed in L^2. From this claim, since $\mathcal{S}(\mathbb{R}^d) \subseteq \operatorname{im}\mathcal{F}$ and $\mathcal{S}(\mathbb{R}^d)$ is dense in $L^2(\mathbb{R}^d)$, the proof is complete.

To demonstrate this claim, let $g \in \overline{\operatorname{im}\mathcal{F}}$. Then there is a sequence $\{f_n\}$ of L^2 functions so that $\hat{f}_n \xrightarrow{L^2} g$. As the Fourier transform is an isometry on L^2, this implies that $\{f_n\}$ converges in L^2. Let $f := \lim_{n\to\infty} f_n$. Then $\left\|\hat{f}_n - \hat{f}\right\|_{L^2} = \|f_n - f\|_{L^2} \to 0$ so that $g = \hat{f}_n$. $\square$

To summarize our progress thus far, we have defined the Fourier transform on $L^1(\mathbb{R}^d)$, investigated its properties on Schwartz functions, and used the class of Schwartz functions to define the Fourier transform on $L^2(\mathbb{R}^d)$. Next, we used Riesz-Thorin interpolation (see Chapter 1) to define the Fourier transform on $L^p(\mathbb{R}^d)$ for $1 \leq p \leq 2$. Even more, we will show that the Fourier transform *cannot* be defined as a bounded operator on $L^p(\mathbb{R}^d)$ spaces for $p > 2$.

Theorem 2.11 (Hausdorff-Young)**.** *If $f \in \mathcal{S}(\mathbb{R}^d)$, then*

$$\left\|\hat{f}\right\|_{L^{p'}} \lesssim \|f\|_{L^p}$$

for $1 \leq p \leq 2$, where $\frac{1}{p} + \frac{1}{p'} = 1$.

Proof. We have the estimates $\left\|\hat{f}\right\|_{L^\infty} \leq \|f\|_{L^1}$ and $\left\|\hat{f}\right\|_{L^2} = \|f\|_{L^2}$. Applying the Riesz-Thorin interpolation theorem with $p_0 = 1$, $p_1 = 2$, $q_0 = \infty$, and $q_1 = 2$, it follows that the Fourier transform is bounded from $L^{p_\theta} \to L^{q_\theta}$, where

$$\frac{1}{p_\theta} = \frac{1-\theta}{1} + \frac{\theta}{2} \quad \text{and} \quad \frac{1}{q_\theta} = \frac{\theta}{2}$$

so that $\frac{1}{p_\theta} = 1 - \frac{\theta}{2}$, hence $\frac{1}{p_\theta} + \frac{1}{q_\theta} = 1$. Since $p_\theta = \frac{2}{2-\theta}$, as $\theta \in [0,1]$ we have $p_\theta \in [1,2]$ as desired. $\square$

Conversely, we have the following. This is one of our first examples of the power of scaling arguments.

Theorem 2.12. *If $\left\|\hat{f}\right\|_{L^q} \lesssim \|f\|_{L^p}$ for all $f \in \mathcal{S}(\mathbb{R}^d)$ for some $1 \leq p, q \leq \infty$, then $1 \leq p \leq 2$ and $q = p'$.*

Proof. Fix $f \neq 0 \in \mathcal{S}(\mathbb{R}^d)$. For $\lambda > 0$, let $f_\lambda(x) := f(x/\lambda)$. Then

$$\hat{f}_\lambda(\xi) = \int e^{-2\pi i x\cdot\xi} f\left(\frac{x}{\lambda}\right) dx = \lambda^d \hat{f}(\lambda\xi).$$

We have $\left\|\hat{f}_\lambda\right\|_{L^q} \lesssim \|f_\lambda\|_{L^p}$, which implies

$$\lambda^d \cdot \lambda^{-\frac{d}{q}} \left\|\hat{f}\right\|_{L^q} \lesssim \lambda^{\frac{d}{p}} \|f\|_{L^p}.$$

Thus, $\lambda^d \cdot \lambda^{-\frac{d}{q}} \lesssim \lambda^{\frac{d}{p}}$, so that $\lambda^{\frac{d}{q'}} \lesssim \lambda^{\frac{d}{p}}$ and hence $\lambda^{\frac{1}{q'}} \lesssim \lambda^{\frac{1}{p}}$ for all $0 < \lambda < \infty$. Letting $\lambda \to \infty$ gives $\frac{1}{q'} \leq \frac{1}{p}$, and letting $\lambda \to 0$ gives $\frac{1}{q'} \geq \frac{1}{p}$. Thus, so $q = p'$.

Next, we show $1 \leq p \leq 2$. It suffices to show $p \leq p'$. Let $\varphi \in C_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp}\varphi \subseteq B(0,1/2)$. Let $\varphi_k(x) = e^{-2\pi i x\cdot\lambda k e_1}\varphi(x - ke_1)$. Using properties 1 and 2 of the Fourier transform, it follows that

$$\hat{\varphi}_k(\xi) = e^{-2\pi i\xi\cdot ke_1}\hat{\varphi}(\xi - \lambda k e_1).$$

Let $f = \sum_{k=1}^N \varphi_k$. Since the supports of φ_k are disjoint, it follows that $\|f\|_{L^p} \sim N^{\frac{1}{p}}$. Next, we compute, expanding out $\left\|\hat{f}\right\|_{L^{p'}}$ as follows:

$$\begin{aligned}
&\left\|\sum_{k=1}^N e^{-2\pi i\xi\cdot ke_1}\hat{\varphi}(\xi - \lambda ke_1)\chi_{B(\lambda ke_1,\lambda/2)}(\xi) + \sum_{k=1}^N e^{-2\pi i\xi\cdot ke_1}\hat{\varphi}(\xi - \lambda ke_1)\chi^C_{B(\lambda ke_1,\lambda/2)}(\xi)\right\|_{L^{p'}} \\
&\quad\geq \left\|\sum_{k=1}^N e^{-2\pi i\xi\cdot ke_1}\hat{\varphi}(\xi - \lambda ke_1)\chi_{B(\lambda ke_1,\lambda/2)}(\xi)\right\|_{L^{p'}} \\
&\qquad - \sum_{k=1}^N \left\|e^{-2\pi i\xi\cdot ke_1}\hat{\varphi}(\xi - \lambda ke_1)\chi^C_{B(\lambda ke_1,\lambda/2)}(\xi)\right\|_{L^{p'}} \\
&\gtrsim N^{\frac{1}{p'}} - N\left\|\hat{\varphi}(\xi)\chi^C_{B(0,\lambda/2)}(\xi)\right\|_{L^{p'}}.
\end{aligned}$$

Because $\hat{\varphi}$ is Schwartz,

$$\left\|\hat{\varphi}(\xi)\chi^C_{B(0,\lambda/2)}(\xi)\right\|_{L^{p'}} \lesssim \left(\int_{|\xi|>\lambda/2} \frac{1}{|\xi|^{2d}}\,d\xi\right)^{\frac{1}{p'}} \lesssim \lambda^{-\frac{d}{p'}}$$

which $\to 0$ as $\lambda \to \infty$. Thus, $\left\|\hat{f}\right\|_{L^{p'}} \lesssim \|f\|_{L^p}$ if and only if $N^{\frac{1}{p'}} \lesssim N^{\frac{1}{p}}$ for all N, which is true if and only if $p \leq p'$. $\square$

2.2 The Fourier Transform on $\mathbb{T}^d$.

The primary focus of this text is harmonic analysis on Euclidean space, and consequently the Fourier transform on $\mathbb{R}^d$ is the most important tool for our purposes. One can also define the Fourier transform on the torus $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$, equivalently, for periodic functions. In the interest of studying dispersive partial differential equations on the torus (see Chapter 9), we include a brief discussion of the Fourier transform in this setting.

Definition 2.13. For a C^∞-function $f : \mathbb{T}^d \to \mathbb{C}$, we define its **Fourier transform** $\hat{f} : \mathbb{Z}^d \to \mathbb{C}$ by

$$\hat{f}(k) = \int_{\mathbb{T}^d} e^{-2\pi i k\cdot x} f(x)\, dx.$$

Then $\hat{f}(k) = \langle f, e_k \rangle$ where $e_k(x) := e^{2\pi i k\cdot x}$.

The characters e_k are orthonormal, since

$$\langle e_k, e_l \rangle = \int_{\mathbb{T}^d} e^{2\pi i (k-l)\cdot x} = \delta_{kl}.$$

We quickly establish some familiar properties of the Fourier transform.

Proposition 2.14 (Bessel's Inequality)**.** *For $f \in C^\infty(\mathbb{T}^d)$,*

$$\sum_{k\in\mathbb{Z}^d} |\langle f, e_k \rangle|^2 \leq \|f\|^2_{L^2(\mathbb{T}^d)}.$$

Proof. Let $S \subseteq \mathbb{Z}^d$ be a finite set. Then

$$\begin{aligned}
0 \leq \left\| f - \sum_{k\in S} \langle f, e_k\rangle e_k \right\|^2_{L^2(\mathbb{T}^d)} &= \left\langle f - \sum_{k\in S} \langle f, e_k \rangle e_k, f - \sum_{l\in S} \langle f, e_l\rangle e_l \right\rangle \\
&= \|f\|^2_{L^2(\mathbb{T}^d)} - 2\sum_{k\in S} \langle f, e_k\rangle \langle e_k, f\rangle + \left\langle \sum_{k\in S} \langle f, e_k\rangle e_k, \sum_{l\in S} \langle f, e_l\rangle e_l \right\rangle \\
&= \|f\|^2_{L^2(\mathbb{T}^d)} - 2\sum_{k\in S} |\langle f, e_k\rangle|^2 + \sum |\langle f, e_k\rangle|^2
\end{aligned}$$

where the last equality follows from orthogonality of the e_k's. Simplifying and rearranging gives

$$\sum_{k\in S} |\langle f, e_k\rangle|^2 \leq \|f\|^2_{L^2(\mathbb{R}^d)}$$

for every finite set S, hence the desired inequality. ☐

This fact shows that if $f \in C^\infty(\mathbb{T}^d)$, then $\sum_{k\in\mathbb{Z}^d} \langle f, e_k\rangle e_k \in L^2(\mathbb{T}^d)$. Unsurprisingly, these two objects can be identified.

Proposition 2.15 (Fourier inversion)**.** *If $f \in C^\infty(\mathbb{T}^d)$, then*

$$f = \sum_{k\in\mathbb{Z}^d} \langle f, e_k\rangle e_k = \sum_{k\in\mathbb{Z}^d} \hat{f}(k) e^{2\pi i k\cdot x}.$$

Proof. Suppose for the sake of contradiction that $f \neq \sum_{k\in\mathbb{Z}^d} \langle f, e_k\rangle e_k$. Then by Stone-Weierstrass, there is a trigonometric polynomial g such that

$$\left\langle f - \sum_{k\in\mathbb{Z}^d} \langle f, e_k\rangle e_k, g \right\rangle \neq 0.$$

For any character e_l,

$$\left\langle f - \sum_{k\in\mathbb{Z}^d} \langle f, e_k\rangle e_k, e_l \right\rangle = \langle f, e_l\rangle - \langle f, e_l\rangle = 0.$$

This is a contradiction. ☐

The following result is unique to the periodic case.

Lemma 2.16 (Poisson summation)**.** *Let* $\varphi \in \mathcal{S}(\mathbb{R})$*. Then*

$$\sum_{n\in\mathbb{Z}} \varphi(x+n) = \sum_{n\in\mathbb{Z}} \hat{\varphi}(n) e^{2\pi i n x}$$

for all $x \in \mathbb{R}$*. In particular,* $\sum_{n\in\mathbb{Z}} \varphi(n) = \sum_{n\in\mathbb{Z}} \hat{\varphi}(n)$*.*

Proof. Define $F_1(x) = \sum_{n\in\mathbb{Z}} \varphi(x+n)$ and $F_2(x) = \sum_{n\in\mathbb{Z}} \hat{\varphi}(n)e^{2\pi i n x}$. Note that both F_1 and F_2 are 1-periodic in x. Also, since φ is Schwartz, each sum converges absolutely in n and converges uniformly in x on compact sets. Therefore, F_1 and F_2 are both continuous functions on $\mathbb{T}$. As such, to give the desired equality it suffices to prove that F_1 and F_2 have the same Fourier coefficients. The definition of F_2 immediately gives $\hat{F_2}(k) = \hat{\varphi}(k)$ for all $k \in \mathbb{Z}$.

Next, we compute:

$$\begin{aligned}
\hat{F}_1(k) &= \sum_{n\in\mathbb{Z}} \int_0^1 \varphi(x+n) e^{-2\pi i k x}\, dx = \sum_{n\in\mathbb{Z}} \int_n^{n+1} \varphi(y) e^{-2\pi i k (y-n)}\, dy \\
&= \sum_{n\in\mathbb{Z}} \int_n^{n+1} \varphi(y) e^{-2\pi i k y}\, dy \\
&= \int_{\mathbb{R}} \varphi(y) e^{-2\pi i k y}\, dy \\
&= \hat{\varphi}(k).
\end{aligned}$$

The in particular statement follows by considering $x = 0$. $\square$

Chapter 3

Lorentz Spaces and Real Interpolation

In this chapter, we introduce a new class of functions space, namely, Lorentz spaces. These spaces measure the size of functions in a more refined manner than the classical L^p spaces. After developing some of the basic theory of Lorentz spaces, we then state and prove the Marcinkiewicz interpolation theorem in its most general setting.

Some references for Lorentz spaces are [5] and [6].

3.1 Lorentz Spaces

Before defining Lorentz spaces in their full generality, we first consider a simpler and more familiar class of functions.

Definition 3.1. For $1 \leq p \leq \infty$ and $f : \mathbb{R}^d \to \mathbb{C}$ measurable, define

$$\|f\|^*_{L^p_{weak}} := \sup_{\lambda > 0} \lambda \left| \left\{ x \in \mathbb{R}^d \ : \ |f(x)| > \lambda \right\} \right|^{\frac{1}{p}}. \tag{3.1}$$

The **weak-L^p space**, written $L^p_{weak}(\mathbb{R}^d)$, is the family of measurable functions f for which $\|f\|^*_{L^p_{weak}}$ is finite.

The quantity defined in (3.1) contains an asterisk as a superscript because is *not* a norm. However, it is a *quasinorm*. Recall that a quasinorm satisfies the same properties as a norm, except that the triangle inequality is replaced by the *quasi triangle inequality*: $\|f+g\| \leq C\left(\|f\| + \|g\|\right)$ for some universal constant $C > 0$.

To clarify the definition of the weak-L^p (quasi)norm, consider a function $f \in L^p(\mathbb{R}^d)$. Using the fundamental theorem of calculus, we can write

$$\|f\|^p_{L^p(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |f(x)|^p \, dx = \int_{\mathbb{R}^d} \int_0^{|f(x)|} p\lambda^{p-1} \, d\lambda \, dx.$$

By Tonelli's theorem,

$$\begin{aligned} \|f\|^p_{L^p(\mathbb{R}^d)} &= \int_0^\infty p\lambda^{p-1} \int_{\left\{ x \in \mathbb{R}^d \, : \, |f(x)| > \lambda \right\}} dx \, d\lambda = \int_0^\infty p\lambda^{p-1} \left| \left\{ x \in \mathbb{R}^d \ : \ |f(x)| > \lambda \right\} \right| d\lambda \\ &= p \int_0^\infty \left(\lambda \left| \left\{ x \in \mathbb{R}^d \ : \ |f(x)| > \lambda \right\} \right|^{\frac{1}{p}} \right)^p \frac{d\lambda}{\lambda}. \end{aligned}$$

So $\|f\|_{L^p(\mathbb{R}^d)} = p^{\frac{1}{p}} \left\| \lambda \left| \left\{ x \in \mathbb{R}^d \ : \ |f(x)| > \lambda \right\} \right|^{\frac{1}{p}} \right\|_{L^p\left((0,\infty), \frac{d\lambda}{\lambda}\right)}$. Using the convention $p^{\frac{1}{\infty}} = 1$, we can then write

$$\|f\|^*_{L^p_{weak}(\mathbb{R}^d)} = p^{\frac{1}{\infty}} \left\| \lambda \left| \left\{ x \in \mathbb{R}^d \ : \ |f(x)| > \lambda \right\} \right|^{\frac{1}{p}} \right\|_{L^\infty\left((0,\infty), \frac{d\lambda}{\lambda}\right)}.$$

In this way, weak-L^p spaces are a clear generalization of L^p spaces.

Example 3.2. Let $f(x) = |x|^{-\frac{d}{p}}$. Then $f \in L^p_{weak}(\mathbb{R}^d)$, but $f \notin L^p(\mathbb{R}^d)$.

To see this, first note that

$$\int_{\mathbb{R}^d} |f(x)|^p \, dx = \int_{\mathbb{R}^d} \frac{1}{|x|^d} \, dx \sim \int_0^\infty \frac{1}{r^d} \, r^{d-1} \, dr \sim \int_0^\infty \frac{1}{r} \, dr.$$

This latter integral does not converge, hence the L^p-norm of f is not finite, so $f \notin L^p(\mathbb{R}^d)$.

On the other hand,

$$\left| \left\{ x \in \mathbb{R}^d \ : \ |f(x)| > \lambda \right\} \right|^{\frac{1}{p}} = \left| \left\{ x \in \mathbb{R}^d \ : \ \frac{1}{|x|^{\frac{d}{p}}} > \lambda \right\} \right|^{\frac{1}{p}} = \left| \left\{ x \in \mathbb{R}^d \ : \ |x| < \left(\frac{1}{\lambda} \right)^{\frac{p}{d}} \right\} \right|^{\frac{1}{p}}.$$

The set being measured is a ball of radius $(1/\lambda)^{p/d}$. The volume of a such a ball scales according to $\left((1/\lambda)^{p/d}\right)^d = (1/\lambda)^p$. Taking pth roots as above then gives

$$\left| \left\{ x \in \mathbb{R}^d \ : \ |f(x)| > \lambda \right\} \right|^{\frac{1}{p}} \lesssim \frac{1}{\lambda}.$$

Hence,

$$\|f\|^*_{L^p_{weak}(\mathbb{R}^d)} \lesssim \sup_{\lambda > 0} \lambda \, \frac{1}{\lambda} = 1 < \infty$$

so that $f \in L^p_{weak}(\mathbb{R}^d)$.

Lorentz spaces are an even further generalization of L^p and weak-L^p spaces.

Definition 3.3. For $1 \le p \le \infty$ and $1 \le q \le \infty$, the **Lorentz space** $L^{p,q}(\mathbb{R}^d)$ is the space of measurable functions $f : \mathbb{R}^d \to \mathbb{C}$ for which the quantity

$$\|f\|^*_{L^{p,q}(\mathbb{R}^d)} := p^{\frac{1}{q}} \left\| \lambda \left| \left\{ x \in \mathbb{R}^d \ : \ |f(x)| > \lambda \right\} \right|^{\frac{1}{p}} \right\|_{L^q\left((0,\infty), \frac{d\lambda}{\lambda}\right)} \tag{3.2}$$

is finite.

By our previous computation, $L^p(\mathbb{R}^d) = L^{p,p}(\mathbb{R}^d)$, and $L^p_{weak}(\mathbb{R}^d) = L^{p,\infty}(\mathbb{R}^d)$. Also, as with the weak-L^p norm, the $L^{p,q}$-norm defined in (3.2) is not actually a norm.

Lemma 3.4. *The quantity* $\|\cdot\|^*_{L^{p,q}(\mathbb{R}^d)}$ *is a quasinorm.*

Proof. First, note that if $\|f\|^*_{L^{p,q}(\mathbb{R}^d)} = 0$, then

$$\left\| \lambda \left| \left\{ x \in \mathbb{R}^d \ : \ |f(x)| > \lambda \right\} \right|^{\frac{1}{p}} \right\|_{L^q\left((0,\infty), \frac{d\lambda}{\lambda}\right)} = 0$$

which implies that, for almost all $\lambda > 0$, $\left|\left\{ x \in \mathbb{R}^d \ : \ |f(x)| > \lambda \right\}\right| = 0$. It follows that $f(x) = 0$ almost everywhere. Thus, $f = 0$.

Next, let $a \in \mathbb{C}$. Then

$$\begin{aligned} \|af\|^*_{L^{p,q}(\mathbb{R}^d)} &= p^{\frac{1}{q}} \left\| \lambda \left| \left\{ x \in \mathbb{R}^d \ : \ |af(x)| > \lambda \right\} \right|^{\frac{1}{p}} \right\|_{L^q\left((0,\infty), \frac{d\lambda}{\lambda}\right)} \\ &= |a| p^{\frac{1}{q}} \left\| \frac{\lambda}{a} \left| \left\{ x \in \mathbb{R}^d \ : \ |f(x)| > \frac{\lambda}{a} \right\} \right|^{\frac{1}{p}} \right\|_{L^q\left((0,\infty), \frac{d\lambda}{\lambda}\right)}. \end{aligned}$$

By making the change of variables $\eta = \lambda/a$, we have $\|af\|^*_{L^{p,q}(\mathbb{R}^d)} = |a| \, \|f\|^*_{L^{p,q}(\mathbb{R}^d)}$.

It remains to prove that quasi-triangle inequality. To do this, we invoke the fact that for $0 < \alpha < 1$, the map $x \mapsto x^\alpha$ is concave. It is a standard fact that concave functions are subadditive, so that $(x+y)^\alpha \le x^\alpha + y^\alpha$. With this in mind, we compute:

$$\|f+g\|^*_{L^{p,q}(\mathbb{R}^d)} = p^{\frac{1}{q}} \left\| \lambda \left| \left\{ x \in \mathbb{R}^d \ : \ |f(x)+g(x)| > \lambda \right\} \right|^{\frac{1}{p}} \right\|_{L^q\left((0,\infty),\frac{d\lambda}{\lambda}\right)}.$$

Observe that

$$\left\{ x \in \mathbb{R}^d \ : \ |f(x)+g(x)| > \lambda \right\} \subseteq \left\{ x \in \mathbb{R}^d \ : \ |f(x)| > \frac{\lambda}{2} \right\} \cup \left\{ x \in \mathbb{R}^d \ : \ |g(x)| > \frac{\lambda}{2} \right\}.$$

Thus,

$$\|f+g\|^*_{L^{p,q}(\mathbb{R}^d)} \le p^{\frac{1}{q}} \left\| \lambda \left(\left| \left\{ x \ : \ |f(x)| > \frac{\lambda}{2} \right\} \right| + \left| \left\{ x \ : \ |g(x)| > \frac{\lambda}{2} \right\} \right| \right)^{\frac{1}{p}} \right\|_{L^q\left((0,\infty),\frac{d\lambda}{\lambda}\right)}.$$

By subbaditivity of the concave map $x \mapsto x^{\frac{1}{p}}$, we have

$$\|f+g\|^*_{L^{p,q}(\mathbb{R}^d)} \le p^{\frac{1}{q}} \left\| \lambda \left(\left| \left\{ x \ : \ |f(x)| > \frac{\lambda}{2} \right\} \right|^{\frac{1}{p}} + \left| \left\{ x \ : \ |g(x)| > \frac{\lambda}{2} \right\} \right|^{\frac{1}{p}} \right) \right\|_{L^q\left((0,\infty),\frac{d\lambda}{\lambda}\right)}.$$

Distributing the λ, factoring out a 2, and appliyng the usual L^q-Minkowski inequality then yields:

$$\begin{aligned}
\|f+g\|^*_{L^{p,q}(\mathbb{R}^d)} &\le 2p^{\frac{1}{q}} \left\| \frac{\lambda}{2} \left| \left\{ x \ : \ |f(x)| > \frac{\lambda}{2} \right\} \right|^{\frac{1}{p}} + \frac{\lambda}{2} \left| \left\{ x \ : \ |g(x)| > \frac{\lambda}{2} \right\} \right|^{\frac{1}{p}} \right\|_{L^q\left((0,\infty),\frac{d\lambda}{\lambda}\right)} \\
&\le 2p^{\frac{1}{q}} \left(\left\| \frac{\lambda}{2} \left| \left\{ x \ : \ |f(x)| > \frac{\lambda}{2} \right\} \right|^{\frac{1}{p}} \right\|_{L^q\left((0,\infty),\frac{d\lambda}{\lambda}\right)} \right. \\
&\qquad \left. + \left\| \frac{\lambda}{2} \left| \left\{ x \ : \ |g(x)| > \frac{\lambda}{2} \right\} \right|^{\frac{1}{p}} \right\|_{L^q\left((0,\infty),\frac{d\lambda}{\lambda}\right)} \right) \\
&= 2 \left(\|f\|^*_{L^{p,q}(\mathbb{R}^d)} + \|g\|^*_{L^{p,q}(\mathbb{R}^d)} \right).
\end{aligned}$$

□

For $1 < p < \infty$ and $1 \le q \le \infty$, we will show that the quasinorm $\|\cdot\|^*_{L^{p,q}(\mathbb{R}^d)}$ is in fact equivalent to a norm. For $p = 1$, $q \ne 1$, the quasinorm is *not* equivalent to a norm. However, in this case, there does exist a metric that generates the same topology. Thus, in either of these cases, $L^{p,q}(\mathbb{R}^d)$ is a complete metric space.

As another quick remark, we note that if $|g| \le |f|$, then for any $\lambda > 0$,

$$\left\{ x \in \mathbb{R}^d \ : \ |g(x)| > \lambda \right\} \subseteq \left\{ x \in \mathbb{R}^d \ : \ |f(x)| > \lambda \right\}.$$

This implies that $\|g\|^*_{L^{p,q}(\mathbb{R}^d)} \le \|f\|^*_{L^{p,q}(\mathbb{R}^d)}$.

It is natural to ask whether different Lorentz spaces have any simple embedding properties. The next result gives a partial answer to this question.

Proposition 3.5. *Let $1 \le p < \infty$, $1 \le q \le \infty$, and $f \in L^{p,q}(\mathbb{R}^d)$. Write $f = \sum_{m\in\mathbb{Z}} f_m$, where*

$$f_m(x) := f(x)\, \chi_{\left\{ x\in\mathbb{R}^d \ : \ 2^m \le |f(x)| \le 2^{m+1} \right\}}.$$

Then

$$\|f\|^*_{L^{p,q}(\mathbb{R}^d)} \sim_{p,q} \left\| \|f_m\|_{L^p(\mathbb{R}^d)} \right\|_{\ell^q_m(\mathbb{Z})}.$$

In particular, $L^{p,q_1}(\mathbb{R}^d) \subseteq L^{p,q_2}(\mathbb{R}^d)$ whenever $q_1 \le q_2$.

Proof. By the preceding remark, note that

$$\left\| \sum_{m\in\mathbb{Z}} 2^m \, \chi_{\{\, x \,:\, 2^m \le |f| \le 2^{m+1} \,\}} \right\|^*_{L^{p,q}(\mathbb{R}^d)} \le \left\| \sum_{m\in\mathbb{Z}} f_m \right\|^*_{L^{p,q}(\mathbb{R}^d)} \le \left\| \sum_{m\in\mathbb{Z}} 2^{m+1} \, \chi_{\{\, x \,:\, 2^m \le |f| \le 2^{m+1} \,\}} \right\|^*_{L^{p,q}(\mathbb{R}^d)}.$$

Therefore, it suffices to prove the proposition for a function of the form $f(x) = \sum_{m\in\mathbb{Z}} 2^m \chi_{E_m}$ where $\{E_m\}$ is a pairwise disjoint collection of sets.

In this case, we need to show that $\|f\|^*_{L^{p,q}(\mathbb{R}^d)} \sim_{p,q} \left\| 2^m |E_m|^{\frac{1}{p}} \right\|_{\ell^q_m(\mathbb{Z})}$. We compute:

$$\begin{aligned} \left(\|f\|^*_{L^{p,q}(\mathbb{R}^d)} \right)^q &= p \int_0^\infty \lambda^q \, | \{ \, x \;:\; |f(x)| > \lambda \, \} |^{\frac{q}{p}} \, \frac{d\lambda}{\lambda} \\ &= p \sum_{m\in\mathbb{Z}} \int_{2^{m+1}}^{2^m} \lambda^q \, | \{ \, x \;:\; |f(x)| > \lambda \, \} |^{\frac{q}{p}} \, \frac{d\lambda}{\lambda}. \end{aligned}$$

Next, observe that for $\lambda \in [2^{m+1}, 2^m)$, $\{\, x \,:\, |f(x)| > \lambda \,\} = \bigcup_{n\ge m} E_M$. As the E_m's are disjoint, we then have

$$\left(\|f\|^*_{L^{p,q}(\mathbb{R}^d)} \right)^q = p \sum_{m\in\mathbb{Z}} \int_{2^{m+1}}^{2^m} \lambda^q \left(\sum_{n\ge m} |E_m| \right)^{\frac{q}{p}} \frac{d\lambda}{\lambda}.$$

This removes the dependence on λ in the set inside the integral. Because

$$p \int_{2^{m+1}}^{2^m} \lambda^q \, \frac{d\lambda}{\lambda} = \frac{p}{q} (2^{mq} - 2^{(m+1)q}) = \frac{p}{q} \, (1 - 2^q) \, 2^{mq}$$

it follows that

$$\left(\|f\|^*_{L^{p,q}(\mathbb{R}^d)} \right)^q \sim_{p,q} \sum_{m\in\mathbb{Z}} 2^{mq} \left(\sum_{n\ge m} |E_n| \right)^{\frac{q}{p}} = \left\| 2^m \left(\sum_{n\ge m} |E_m| \right)^{\frac{1}{p}} \right\|^q_{\ell^q_m(\mathbb{Z})}.$$

Thus,

$$\|f\|^*_{L^{p,q}(\mathbb{R}^d)} \sim_{p,q} \left\| 2^m \left(\sum_{n\ge m} |E_n| \right)^{\frac{1}{p}} \right\|_{\ell^q_m(\mathbb{Z})} \gtrsim_{p,q} \left\| 2^m \, (|E_m|)^{\frac{1}{p}} \right\|_{\ell^q_m(\mathbb{Z})}.$$

This gives half of the $\sim_{p,q}$ relation that we need.

To get the other inequality, we compute as follows. First, invoking concavity of fractional powers as before,

$$\|f\|^*_{L^{p,q}(\mathbb{R}^d)} \sim_{p,q} \left\| 2^m \left(\sum_{n\ge m} |E_n| \right)^{\frac{1}{p}} \right\|_{\ell^q_m(\mathbb{Z})} \le \left\| 2^m \sum_{n\ge m} |E_n|^{\frac{1}{p}} \right\|_{\ell^q_m(\mathbb{Z})}.$$

Next, making the change of variables $n = m + k$,

$$\begin{aligned} &= \left\| 2^m \sum_{k\ge 0} |E_{m+k}|^{\frac{1}{p}} \right\|_{\ell^q_m(\mathbb{Z})} = \left\| \sum_{k\ge 0} 2^{-k} 2^{m+k} |E_{m+k}|^{\frac{1}{p}} \right\|_{\ell^q_m(\mathbb{Z})} \le \sum_{k\ge 0} 2^{-k} \left\| 2^{m+k} |E_{m+k}|^{\frac{1}{p}} \right\|_{\ell^q_m(\mathbb{Z})} \\ &= \sum_{k\ge 0} 2^{-k} \left\| 2^m |E_m|^{\frac{1}{p}} \right\|_{\ell^q_m(\mathbb{Z})} \\ &= \left\| 2^m |E_m|^{\frac{1}{p}} \right\|_{\ell^q_m(\mathbb{Z})} \end{aligned}$$

so that $\|f\|^*_{L^{p,q}(\mathbb{R}^d)} \lesssim_{p,q} \left\| 2^m |E_m|^{\frac{1}{p}} \right\|_{\ell^q_m(\mathbb{Z})}$. Therefore,

$$\|f\|^*_{L^{p,q}(\mathbb{R}^d)} \sim_{p,q} \left\| 2^m |E_m|^{\frac{1}{p}} \right\|_{\ell^q_m(\mathbb{Z})}$$

as desired.

The "in particular" statement follows from the fact that if $q_1 \leq q_2$, then $\ell^{q_1}(\mathbb{Z}) \subseteq \ell^{q_2}(\mathbb{Z})$. □

As noted in the proof of the above proposition, the monotonicty of the quasinorm allows for the following reduction in many computations involving Lorentz spaces. If $f \in L^{p,q}(\mathbb{R}^d)$ is real-valued and nonnegative, then we can assume without loss of generality that $f = \sum_{m\in\mathbb{Z}} 2^m \chi_{E_m}$, where the sets E_m are pairwise disjoint. In this case,

$$\|f\|^*_{L^{p,q}(\mathbb{R}^d)} = \left\| 2^m |E_m|^{\frac{1}{p}} \right\|_{\ell^q_m(\mathbb{Z})}.$$

We can further decompose a general function into its real and imaginary parts, and then into their corresponding positive and negative parts, to apply this reduction.

3.2 Duality in Lorentz Spaces

One of the most important principles in L^p space theory is that of duality. The analogous fact for Lorentz spaces is given by the following proposition.

Proposition 3.6. *Let $1 < p < \infty$ and $1 \leq q \leq \infty$, and let p' and q' denote the respective Holder conjugates. Then*

$$\|f\|^*_{L^{p,q}(\mathbb{R}^d)} \sim_{p,q} \sup_{\|g\|^*_{L^{p',q'}(\mathbb{R}^d)} \leq 1} \left| \int_{\mathbb{R}^d} f(x)\overline{g(x)}\,dx \right|. \tag{3.3}$$

Proof. As the quasinorm is positively homogeneous, we can scale the function f appropriately and assume without loss of generality that $\|f\|^*_{L^{p,q}(\mathbb{R}^d)} = 1$. Also, as remarked previously, we can assume without loss of generality that f and g are real-valued and nonnegative, hence we can take $f = \sum_{n\in\mathbb{Z}} 2^n \chi_{F_n}$ for disjoint sets F_n and $g = \sum_{m\in\mathbb{Z}} 2^m \chi_{E_m}$ for disjoint sets E_m.

In this case, we have

$$1 = \left(\|f\|^*_{L^{p,q}(\mathbb{R}^d)}\right)^q = \left\| 2^n |F_n|^{\frac{1}{p}} \right\|^q_{\ell^q_n(\mathbb{Z})} = \sum_{n\in\mathbb{Z}} 2^{nq} |F_n|^{\frac{q}{p}}.$$

We decompose the above sum as follows. Let $2^{\mathbb{Z}}$ denote the set of dyadic numbers. Then

$$\sum_{n\in\mathbb{Z}} 2^{nq}|F_n|^{\frac{q}{p}} = \sum_{N\in 2^{\mathbb{Z}}} \sum_{n:|F_n|\sim N} 2^{nq} N^{\frac{q}{p}} = \sum_{N\in 2^{\mathbb{Z}}} N^{\frac{q}{p}} \sum_{n:|F_n|\sim N} 2^{nq} \sim \sum_{N\in 2^{\mathbb{Z}}} N^{\frac{q}{p}} \left(\sum_{n:|F_n|\sim N} 2^n \right)^q.$$

The latter equivalence is a straightforward exercise in dealing with dyadic sums, together with the fact that the ℓ^q-norm is bounded by the ell^1-norm. Thus,

$$1 \sim \sum_{N\in 2^{\mathbb{Z}}} \left(\sum_{n:|F_n|\sim N} 2^n N^{\frac{1}{p}} \right)^q.$$

Similarly, since $\|g\|^*_{L^{p',q'}(\mathbb{R}^d)} \leq 1$, the same computation gives

$$\sum_{M\in 2^{\mathbb{Z}}} \left(\sum_{m:|E_m|\sim M} 2^m |E_m|^{\frac{1}{p'}} \right)^{q'} \lesssim 1.$$

With these computations and our reductions in mind, we demonstrate the equivalence in (3.3). We first show that the right hand side is bounded by the left.

$$\begin{aligned}\int f(x)g(x)\,dx &= \sum_{n,m\in\mathbb{Z}} \int 2^n 2^m \chi_{F_n\cap E_m}(x)\,dx = \sum_{n,m\in\mathbb{Z}} 2^n 2^m |F_n \cap E_m| \\ &\lesssim \sum_{N,M\in 2^{\mathbb{Z}}} \sum_{n:|F_n|\sim N} 2^n \sum_{m:|E_m|\sim M} 2^m \min\{N, M\}.\end{aligned}$$

By our above computations, we eventually want to introduce the quantities $N^{\frac{1}{p}}$ and $|E_m|^{\frac{1}{p'}}$. We force them in the sum as follows:

$$\begin{aligned}&\lesssim \sum_{N,M\in 2^{\mathbb{Z}}} \sum_{n:|F_n|\sim N} 2^n N^{\frac{1}{p}} \sum_{m:|E_m|\sim M} 2^m |E_m|^{\frac{1}{p'}} \min\left\{\frac{N}{N^{\frac{1}{p}} M^{\frac{1}{p'}}}, \frac{M}{N^{\frac{1}{p}} M^{\frac{1}{p'}}}\right\} \\ &= \sum_{N,M\in 2^{\mathbb{Z}}} \sum_{n:|F_n|\sim N} 2^n |F_n|^{\frac{1}{p}} \sum_{m:|E_m|\sim M} 2^m |E_m|^{\frac{1}{p'}} \min\left\{\left(\frac{N}{M}\right)^{\frac{1}{p'}}, \left(\frac{M}{N}\right)^{\frac{1}{p}}\right\}.\end{aligned}$$

Next, we use Holder's inequality on the above sum, viewing the n-sum as living in ℓ^q_N and the m-sum as living in $\ell^{q'}_M$. This yields

$$\begin{aligned}&\lesssim \left[\sum_{N,M\in 2^{\mathbb{Z}}} \left(\sum_{n:|F_n|\sim N} 2^n |F_n|^{\frac{1}{p}}\right)^q \min\left\{\left(\frac{N}{M}\right)^{\frac{1}{p'}}, \left(\frac{M}{N}\right)^{\frac{1}{p}}\right\}\right]^{\frac{1}{q}} \\ &\qquad \cdot \left[\sum_{N,M\in 2^{\mathbb{Z}}} \left(\sum_{m:|E_m|\sim M} 2^M |E_m|^{\frac{1}{p'}}\right)^{q'} \min\left\{\left(\frac{N}{M}\right)^{\frac{1}{p'}}, \left(\frac{M}{N}\right)^{\frac{1}{p}}\right\}\right]^{\frac{1}{q'}}.\end{aligned}$$

In the first term, freeze N and sum over M. Because $p > 1$ and hence $p' < \infty$, we have

$$\sum_{M\in 2^{\mathbb{Z}}} \min\left\{\left(\frac{N}{M}\right)^{\frac{1}{p'}}, \left(\frac{M}{N}\right)^{\frac{1}{p}}\right\} = \sum_{M\leq N} \left(\frac{M}{N}\right)^{\frac{1}{p}} + \sum_{M>N} \left(\frac{N}{M}\right)^{\frac{1}{p'}} \lesssim 1 + 1 \lesssim 1.$$

Similarly, in the second term we freeze M and sum over N. This gives

$$\begin{aligned}\int f(x)g(x)\,dx &\lesssim \left[\sum_{N\in 2^{\mathbb{Z}}} \left(\sum_{n:|F_n|\sim N} 2^n |F_n|^{\frac{1}{p}}\right)^q\right]^{\frac{1}{q}} \cdot \left[\sum_{M\in 2^{\mathbb{Z}}} \left(\sum_{m:|E_m|\sim M} 2^M |E_m|^{\frac{1}{p'}}\right)^{q'}\right]^{\frac{1}{q'}} \\ &\lesssim 1.\end{aligned}$$

This gives one inequality.

Next, we consider the converse inequality. Again we can take f to be of the form $\sum_n 2^n \chi_{F_n}$ with $\|f\|^*_{L^{p,q}(\mathbb{R}^d)} \sim 1$. Let

$$g = \sum_n \left(2^n |F_n|^{\frac{1}{p}}\right)^{q-1} |F_n|^{-\frac{1}{p'}} \chi_{F_n} = \sum_n 2^{n(q-1)} |F_n|^{\frac{q}{p}-1} \chi_{F_n}.$$

We make two claims, from which the desired inequality clearly follows:

1. $\int f(x)g(x)\,dx \gtrsim 1$;
2. $\|g\|^*_{L^{p',q'}(\mathbb{R}^d)} \lesssim 1$.

The first claim follows from

$$\begin{aligned}\int f(x)g(x)\,dx &= \int \sum_n 2^{n(q-1)}2^n|F_n|^{\frac{q}{p}-1}\chi_{F_n} = \sum_n 2^{nq}|F_n|^{\frac{q}{p}} \\ &= \left\|2^n|F_n|^{\frac{1}{p}}\right\|^q_{\ell^q_n(\mathbb{Z})} \\ &\sim 1.\end{aligned}$$

To see the second claim, write

$$g \sim \sum_{N\in 2^{\mathbb{Z}}} N\chi_{\bigcup_{n\in S_n} F_n}$$

where $S_n = \left\{ n \ : \ 2^{n(q-1)}|F_n|^{\frac{q}{p}-1} \sim N \right\}$. Note that the S_n's are disjoint sets. Then

$$\begin{aligned}\left(\|g\|^*_{L^{p',q'}(\mathbb{R}^d)}\right)^{q'} &\sim \sum_{N\in 2^{\mathbb{Z}}} N^{q'}\left(\sum_{n\in S_n}|F_n|\right)^{\frac{q'}{p'}} \\ &\sim \sum_{N\in 2^{\mathbb{Z}}} N^{q'}\left(\sum_{n\in S_n}\left(N2^{-n(q-1)}\right)^{\frac{p}{q-p}}\right)^{\frac{q'}{p'}} \\ &\sim \sum_{N\in 2^{\mathbb{Z}}} N^{q'}\sum_{n\in S_n} N^{\frac{p}{q-p}\cdot\frac{q'}{p'}}2^{-nq\frac{p-1}{q-p}}\end{aligned}$$

where in the last equivalence we have used the fact about dyadic sums and ℓ^q norms from above and the fact that $q' = \frac{q}{q-1}$. It remains to compute and simplify all of the Holder exponents. Doing this gives

$$\begin{aligned}\left(\|g\|^*_{L^{p',q'}(\mathbb{R}^d)}\right)^{q'} &\sim \sum_{n\in\mathbb{Z}}\left(2^{n(q-1)}|F_n|^{\frac{q-p}{p}}\right)^{\frac{q}{q-p}}2^{-nq\frac{p-q}{q-p}} \\ &\sim \sum_{n\in\mathbb{Z}}|F_n|^{\frac{q}{p}}2^{nq} \\ &\lesssim 1.\end{aligned}$$

This proves both claims, the the converse inequality, and hence completes the proof.

□

The right hand side of (3.3) defines a norm which is equivalent to the quasinorm Lorentz quasinorm. With respect to this norm, $L^{p,q}(\mathbb{R}^d)$ for $1 < p < \infty$, $1 \le q \le \infty$, is a Banach space. The proof of completeness is the same as the proof of completeness in L^p spaces. The dual space in this case is $L^{p',q'}(\mathbb{R}^d)$.

As remarked above, if $p = 1$ and $q \neq 1$, there is no norm which is equivalent to the quasinorm. The following example demonstrates this explicitly.

Example 3.7. Consider the case $p = 1, q = \infty, d = 1$. Let $f(x) = \sum_{n=1}^N \frac{1}{|x-n|}$. We saw in a previous example that $\left\|\frac{1}{|x-n|}\right\|^*_{L^{1,\infty}(\mathbb{R})} = \left\|\frac{1}{|x-n|}\right\|^*_{L^1_{weak}(\mathbb{R})} \lesssim 1$. This implies that $\sum_{n=1}^N \left\|\frac{1}{|x-n|}\right\|^*_{L^{1,\infty}(\mathbb{R})} \lesssim N$.

We claim that $\left\|\sum_{n=1}^N \frac{1}{|x-n|}\right\|^*_{L^{1,\infty}(\mathbb{R})} \gtrsim N\log N$. We have

$$\left\|\sum_{n=1}^N \frac{1}{|x-n|}\right\|^*_{L^{1,\infty}(\mathbb{R})} = \sup_{\lambda>0}\lambda\left|\left\{x\in\mathbb{R}^d \ : \ \sum_{n=1}^N\frac{1}{|x-n|} > \lambda\right\}\right|.$$

Fix $x \in [0, N]$. Note that, for $n \geq x$, $|x-n| < n$, so that $\frac{1}{|x-n|} > \frac{1}{n}$. For $n < x$, we can consider each finite number and rearrange them so that their sum is comparable to $\sum_{n<x} \frac{1}{n}$. Thus, for $x \in [0, N]$ we have $f(x) \gtrsim \sum_{n=1}^{N} \frac{1}{n} \sim \log N$. Choose $\lambda = C \log N$ for some appropriate constant C, we then have $\left\{ x \in \mathbb{R}^d \ : \ \sum_{n=1}^{N} \frac{1}{|x-n|} > C \log N \right\} \supseteq [0, N]$. Therefore,

$$\|f\|_{L^{1,\infty}(\mathbb{R})} \geq C \log N \left| \left\{ x \in \mathbb{R}^d \ : \ \sum_{n=1}^{N} \frac{1}{|x-n|} > C \log N \right\} \right| \geq CN \log N$$

as claimed.

Now, suppose that the quasinorm was equivalent to a norm. Then by the triangle inequality,

$$N \log N \lesssim \left\| \sum_{n=1}^{N} \frac{1}{|x-n|} \right\|^{*}_{L^{1,\infty}(\mathbb{R})} \lesssim \sum_{n=1}^{N} \left\| \frac{1}{|x-n|} \right\|^{*}_{L^{1,\infty}(\mathbb{R})} \lesssim N,$$

which is a contradiction.

3.3 The Marcinkiewicz Interpolation Theorem

Finally, we arrive at the Marcinkiewicz interpolation theorem. Before stating and proving the theorem, we recall the notion of sublinearity and introduce some language that will be used throughout the remainder of this text.

Definition 3.8. A mapping T on a class of measurable functions is **sublinear** if $|T(cf)| \leq |c||T(f)|$ and $|T(f+g)| \leq |T(f)| + |T(g)|$ for all $c \in \mathbb{C}$ and f, g in the support of T.

Any linear operator is obviously sublinear. A more interesting class of examples, and perhaps the primary motivation for defining sublinearity, is the following.

Example 3.9. Given a family of linear mappings $\{T_t\}_{t \in I}$, the mapping defined by

$$T(f)(x) := \|T_t(f)(x)\|_{L^q_t}$$

is sublinear. When $q = \infty$, operators of this form are called *maximal functions.* When $q = 2$, the term *square function* is used.

Definition 3.10. Let $1 \leq p, q \leq \infty$. A mapping of functions T is of **(strong) type** (p, q) if $\|Tf\|_{L^q(\mathbb{R}^d)} \lesssim_T \|f\|_{L^p(\mathbb{R}^d)}$. That is, T is of type (p, q) if it is bounded as a mapping from $L^p(\mathbb{R}^d) \to L^q(\mathbb{R}^d)$. Similarly, for $q < \infty$, T is of **weak type** (p, q) if $\|Tf\|^{*}_{L^{q,\infty}(\mathbb{R}^d)} \lesssim_T \|f\|_{L^p(\mathbb{R}^d)}$, and is of **restricted weak type** (p, q) if $\|T\chi_F\|^{*}_{L^{q,\infty}(\mathbb{R}^d)} \lesssim_T |F|^{\frac{1}{p}}$ for all finite measure sets F.

Note that type (p, q) implies weak type (p, q), which implies restricted weak type (p, q). An alternative characterization of restricted weak type operators is given by the following proposition.

Proposition 3.11. *A mapping T is of restricted weak type (p, q) if and only if*

$$\int |T\chi_F||\chi_E| \, dx \lesssim |E|^{\frac{1}{p}} |F|^{\frac{1}{q}}$$

for all finite measure sets E, F.

Proof. First, suppose that T is of restricted type (p,q). Then $\|T\chi_F\|^*_{L^{q,\infty}(\mathbb{R}^d)} \lesssim_T |F|^{\frac{1}{p}}$. By the duality of Lorentz quasinorms, this implies

$$\int |T\chi_F||\chi_E|\, dx \lesssim \|T\chi_F\|^*_{L^{q,\infty}(\mathbb{R}^d)} \|\chi_E\|^*_{L^{q',1}(\mathbb{R}^d)}.$$

Note that

$$\left(\|\chi_E\|^*_{L^{q',1}(\mathbb{R}^d)}\right)^{q'} = q' \int_0^1 \lambda^{q'} |\{x \,:\, \chi_E(x) > \lambda\}| \frac{d\lambda}{\lambda} \sim |E|.$$

Hence,

$$\int |T\chi_F||\chi_E|\, dx \lesssim |E|^{\frac{1}{p}}|F|^{\frac{1}{q}}.$$

Conversely, suppose that the above inequality holds for all finite measure sets E and F. We again invoke duality to write

$$\|T\chi_F\|^*_{L^{q,\infty}(\mathbb{R}^d)} \sim \sup_{\|g\|^*_{L^{q',1}} \le 1} \left| \int T\chi_F\, \overline{g}\, dx \right| \lesssim \sup_{\|g\|^*_{L^{q',1}} \le 1} \int |T\chi_F||g|\, dx.$$

Consider $g := \sum_m 2^m E_m$ for an appropriate disjoint collection of sets E_m. Then

$$\begin{aligned}\int |T\chi_F||g|\, dx &\lesssim \sum_m 2^m |F|^{\frac{1}{p}}|E|^{\frac{1}{q'}} \lesssim |F|^{\frac{1}{p}} \|g\|^*_{L^{q',1}(\mathbb{R}^d)} \\ &\lesssim |F|^{\frac{1}{p}}\end{aligned}$$

as desired.

□

The formulation of the Marcinkiewicz interpolation theorem that we state below is due to Hunt (see [5]). The classical statement of the theorem is given as Corollary 3.14 below.

Theorem 3.12 (Marcinkiewicz interpolation)**.** *Fix $1 \le p_1, p_2, q_1, q_2 \le \infty$ such that $p_1 \ne p_2$ and $q_1 \ne q_2$. Let T be a sublinear operator of restricted weak type (p_1, q_1) and of restricted weak type (p_2, q_2). Then for any $1 \le r \le \infty$ and $0 < \theta < 1$,*

$$\|Tf\|^*_{L^{q_\theta,r}} \lesssim \|f\|^*_{L^{p_\theta,r}}$$

where $\frac{1}{p_\theta} = \frac{\theta}{p_1} + \frac{1-\theta}{p_2}$ and $\frac{1}{q_\theta} = \frac{\theta}{q_1} + \frac{1-\theta}{q_2}$.

Before proving the theorem, we make a few remarks. First, if $p_\theta \le q_\theta$, taking $r = q_\theta$ gives

$$\|Tf\|_{L^{q_\theta}} \lesssim \|f\|^*_{L^{p_\theta,q_\theta}} \lesssim \|f\|_{L^{p_\theta}}$$

so that T is of strong type (p_θ, q_θ). The requirement that $p_\theta \le q_\theta$ is essential for the strong type conclusion, as evidenced by the following example.

Example 3.13. Define $T(f)(x) := \frac{f(x)}{|x|^{\frac{1}{2}}}$. Then T is bounded as an operator from $L^p(\mathbb{R}) \to L^{\frac{2p}{p+2},\infty}(\mathbb{R})$ for any $2 \le p \le \infty$, but is not bounded as an operator from $L^p(\mathbb{R}) \to L^{\frac{2p}{p+2}}(\mathbb{R})$.

To prove this, we invoke *Holder's inequality in Lorentz spaces*: for $1 \le p_1, p_2, p < \infty$ and $1 \le q_1, q_2, q \le \infty$ satisfying $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$ and $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2}$, we have $\|fg\|^*_{L^{p,q}} \lesssim \|f\|^*_{L^{p_1,q_1}} \|g\|^*_{L^{p_2,q_2}}$. By this inequality, we have:

$$\|Tf\|^*_{L^{\frac{2p}{p+2},\infty}(\mathbb{R})} \lesssim \|f\|^*_{L^{p,\infty}(\mathbb{R})} \left\| |x|^{-\frac{1}{2}} \right\|^*_{L^{2,\infty}(\mathbb{R})} \lesssim \|f\|_{L^p(\mathbb{R})}.$$

On the other hand, consider the function $f(x) = |x|^{-\frac{1}{p}} \left|\log\left(|x| + \frac{1}{|x|}\right)\right|^{-\frac{p+2}{2p}}$. We first claim that $f \in L^p(\mathbb{R})$. Indeed, we can write

$$\begin{aligned}
\|f\|^p_{L^p(\mathbb{R})} &= \int \frac{1}{|x|}\left|\log\left(|x| + \frac{1}{|x|}\right)\right|^{-\frac{p+2}{2}} dx \\
&= 2\left(\int_0^2 \frac{1}{|x|}\left|\log\left(|x| + \frac{1}{|x|}\right)\right|^{-\frac{p+2}{2}} dx + \int_2^\infty \frac{1}{|x|}\left|\log\left(|x| + \frac{1}{|x|}\right)\right|^{-\frac{p+2}{2}} dx\right).
\end{aligned}$$

Consider the second integral. Since $p \geq 2$, we have

$$\begin{aligned}
\int_2^\infty \frac{1}{|x|}\left|\log\left(|x| + \frac{1}{|x|}\right)\right|^{-\frac{p+2}{2}} dx \leq \int_2^\infty \frac{1}{|x|}\left|\log\left(|x| + \frac{1}{|x|}\right)\right|^{-2} dx &\leq \int_2^\infty \frac{1}{x(\log x)^2}\, dx \\
&= \int_{\log 2}^\infty \frac{1}{u^2}\, du \\
&< \infty.
\end{aligned}$$

By symmetry, the singularity at 0 behaves similarly. Hence, $f \in L^p(\mathbb{R})$. However, it is not difficult to show that the quantity

$$\|Tf\|^{\frac{2p}{p+2}}_{L^{\frac{2p}{p+2}}} = \int |x|^{-\frac{2}{p+2}} \left|\log\left(|x| + \frac{1}{|x|}\right)\right|^{-1} dx$$

is not finite.

We now prove Theorem 3.12.

Proof. □

Corollary 3.14 (Marcinkiewicz interpolation)**.**

Chapter 4

Maximal Functions

The notion of *averaging* is one which is ubiquitous in analysis. For example, convolution can be viewed as a sort of averaging against a given function. Convolving a function with a smooth function then smooths the original function, a fact which is immensely useful.

In this chapter, we study averaging in more depth via *maximal functions*. We begin by recalling the classical Hardy-Littlewood maximal inequality, and spend the rest of the chapter developing generalizations.

4.1 The Hardy-Littlewood Maximal Function

The reader may already be familiar with the **Hardy-Littlewood maximal function**, defined as

$$M(f)(x) := \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)|\, dy. \tag{4.1}$$

The quantity $\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)|\, dy$ represents the average value of the function on a ball centered at x. Thus, the maximal function measures the largest possible average values for f on various balls.

The following theorem is standard.

Theorem 4.1 (Hardy-Littlewood maximal inequality). *Let $M(f)$ denote the Hardy-Littlewood maximal function as in* (4.1). *Then*

1. *For $f \in L^p(\mathbb{R}^d)$, $1 \le p \le \infty$, $M(f)$ is finite almost everywhere.*

2. *The operator M is of weak type $(1,1)$, and of strong type (p,p) for $1 < p \le \infty$.*

Later in this chapter, we will prove a generalization of Theorem 4.1, and so we defer a proof until then.

Unraveling statement 2 of Theorem 4.1, we see that M being of weak type $(1,1)$ means

$$\|Mf\|^*_{L^{1,\infty}(\mathbb{R}^d)} \lesssim \|f\|_{L^1(\mathbb{R}^d)}$$

and hence

$$|\{ x \,:\, (Mf)(x) > \lambda \}| \le \frac{C}{\lambda} \|f\|_{L^1(\mathbb{R}^d)}$$

for all $\lambda > 0$. This is the usual formulation of the Hardy-Littlewood maximal inequality. We remark that the $p > 1$ condition in statement 2 is necessary, because M is not of strong type $(1,1)$. To see this, let $\phi \in C_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp}\phi \subseteq B_{1/2}(0)$. Fix $|x| > 1$. If $r < |x| - 1/2$, then $\int_{B_r(x)} \phi(y)\, dy = 0$. If $r > |x| + 1/2$, then

$$\frac{1}{|B_r(x)|} \int_{B_r(x)} \phi(y)\, dy = \frac{1}{|B_r(x)|} \int_{B_{|x|+1/2}(x)} \phi(y)\, dy \le \frac{1}{|B_{|x|+1/2}(x)|} \int_{B_{|x|+1/2}(x)} \phi(y)\, dy.$$

Thus,
$$(M\phi)(x) = \sup_{|x|-1/2\le r\le |x|+1/2} \frac{1}{|B_r(x)|}\int_{B_r(x)} \phi(y)\,dy \gtrsim \frac{1}{|x|^d}.$$
But $\frac{1}{|x|^d}$ is not in $L^1(\mathbb{R}^d)$.

We close this section with a classic application of the Hardy-Littlewood maximal inequality.

Theorem 4.2 (Lebesgue differentiation)**.** *Let $f \in L^1_{loc}(\mathbb{R}^d)$. Then*
$$\lim_{r\to 0} \frac{1}{|B_r(x)|}\int_{B_r(x)} f(y)\,dy = f(x)$$
for almost every $x \in \mathbb{R}^d$.

Proof. ADD IN □

More details on the classical Hardy-Littlewood maximal function can be found in [9].

4.2 A_p Weights and the Weighted Maximal Inequality

The main theorem of this section, which is a generalization of Theorem 4.1, is the following.

Theorem 4.3 (Weighted maximal inequality)**.** *Let $\omega : \mathbb{R}^d \to [0,\infty)$ be locally integrable. Associate to ω the measure defined by $\omega(E) := \int_E \omega(y)\,dy$. Then*
$$M : L^1(M(\omega)\,dx) \to L^{1,\infty}(\omega\,dx)$$
and
$$M : L^p(M(\omega)\,dx) \to L^p(\omega\,dx)$$
for $1 < p \le \infty$. Explicitly,
$$\omega\left(\{\, x \;:\; |(M\omega)(x)| > \lambda \,\}\right) \lesssim \frac{1}{\lambda}\int_{\mathbb{R}^d} |f(x)|\,(M\omega)(x)\,dx$$
and
$$\int_{\mathbb{R}^d} |(Mf)(x)|^p \omega(x)\,dx \lesssim \int_{\mathbb{R}^d} |f(x)|^p\,(M\omega)(x)\,dx.$$

Note that if $\omega \equiv 1$, then $M(\omega) \equiv 1$, and we recover the classical Hardy-Littlewood maximal inequality in Theorem 4.1.

Before proving the weighted maximal inequality, we establish some preliminary facts on A_p weights.

Definition 4.4. A function ω satisfies the A_1 **condition**, written $\omega \in A_1$, if $M(\omega) \lesssim \omega$ almost everywhere.

Note that if $\omega \in A_1$, then Theorem 4.3 gives $M : L^1(\omega\,dx) \to L^{1,\infty}(\omega\,dx)$ and $M : L^p(\omega\,dx) \to L^p(\omega\,dx)$ for $1 < p \le \infty$.

Lemma 4.5. *The following are equivalent:*

1. $\omega \in A_1$;
2. $\frac{\omega(B)}{|B|} \lesssim \omega(x)$ *for almost all $x \in B$ and all balls B;*
3. $\frac{1}{|B|}\int_B f(y)\,dy \lesssim \frac{1}{\omega(B)}\int_B f(y)\omega(y)\,dx$ *for all $f \ge 0$ and all balls B.*

Proof. We first show (1) $\Rightarrow$ (2). Fix a ball B_0 of radius r_0 and let $x \in B_0$. Then

$$\omega(x) \gtrsim (M\omega)(x) = \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} \omega(y)\, dy \geq \frac{1}{|B_{2r_0}(x)|} \int_{B_{2r_0}(x)} \omega(y)\, dy.$$

Because $B_{2r_0}(x) \supseteq B_0$ (drawing a picture makes this clearer) and because $|B_{2r_0}(x)| \lesssim_d |B_0|$,

$$\omega(x) \gtrsim \frac{1}{|B_0|} \int_{B_0} \omega(y)\, dy = \frac{\omega(B_0)}{|B_0|}$$

as desired.

The implication (2) $\Rightarrow$ (1) follows immediately from the definition of the maximal function.

Next, we show (2) $\Rightarrow$ (3). Fix a ball B and $f \geq 0$. Then

$$\frac{1}{\omega(B)} \int_B f(y)\omega(y)\, dy \gtrsim \frac{1}{\omega(B)} \int_B f(y) \frac{\omega(B)}{|B|}\, dy = \frac{1}{|B|} \int_B f(y)\, dy.$$

Finally (3) $\Rightarrow$ (2). Fix a ball B and let $x \in B$ be a Lebesgue point of B (i.e., a point for which the conclusion of the Lebesgue differentiation theorem holds). Choose $r << 1$ so that $B_r(x) \subseteq B$. Let $f = \chi_{B_r(x)}$. Then by (3),

$$\frac{1}{|B|} \int_B \chi_{B_r(x)}(y)\, dy \lesssim \frac{1}{\omega(B)} \int_B \chi_{B_r(x)}(y)\omega(y)\, dx \quad \Rightarrow \quad \frac{|B_r(x)|}{|B|} \lesssim \frac{1}{\omega(B)} \int_{B_r(x)} \omega(y)\, dy.$$

Since x is a Lebesgue point,

$$\frac{\omega(B)}{|B|} \lesssim \frac{1}{|B_r(x)|} \int_{B_r(x)} \omega(y)\, dy \to \omega(x)$$

as $r \to 0$. $\square$

Definition 4.6. A function $\omega : \mathbb{R}^d \to [0, \infty)$ satisfies the A_p **condition**, written $\omega \in A_p$, if

$$\sup_B \frac{\omega(B)}{|B|} \left(\frac{1}{|B|} \int_B \omega(y)^{-\frac{p'}{p}}\, dy \right)^{\frac{p}{p'}} \lesssim 1,$$

where the supremum is taken over all open balls in $\mathbb{R}^d$.

We note for convenience that the above condition is equivalent to

$$\sup_B \frac{\omega(B)}{|B|^p} \left(\int_B \omega(y)^{-\frac{p'}{p}}\, dy \right)^{\frac{p}{p'}} \lesssim 1.$$

Also, if $1 < p < q < \infty$, then $A_p \subseteq A_q$. This follows from Holder's inequality.

The importance of the class of A_p weights is the following theorem, which characterizes when the maximal function is a bounded operator on a weighted L^p space.

Theorem 4.7. *Fix $1 < p < \infty$. Then $M : L^p(\omega\, dx) \to L^p(\omega\, dx)$ if and only if $\omega \in A_p$.*

Proof. Here, we only explicitly prove one direction ($\Rightarrow$) of this statement. For the other direction, we will provide an outline of the argument and leave the details as an exercise for the reader.

Suppose that $M : L^p(\omega\, dx) \to L^p(\omega\, dx)$. Fix a ball B and let $f = (\omega + \varepsilon)^{-\frac{p'}{p}} \chi_B$. For $x \in B$,

$$\begin{aligned}(Mf)(x) &= \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} (\omega(y) + \varepsilon)^{-\frac{p'}{p}} \chi_B(y)\, dy \geq \frac{1}{|2B|} \int_B (\omega(y) + \varepsilon)^{-\frac{p'}{p}}\, dy \\ &\gtrsim \frac{1}{|B|} \int_B (\omega(y) + \varepsilon)^{-\frac{p'}{p}}\, dy =: \lambda.\end{aligned}$$

Here our definition of λ includes any implicit constants from the above inequality. Then

$$\omega(B) \le \omega\left(\left\{ x \;:\; (Mf)(x) > \frac{\lambda}{2} \right\}\right) \lesssim \frac{1}{\lambda^p} \int |f(y)|^p \omega(y)\, dy$$

where the final inequality follows from the fact that $M : L^p(M(\omega)\, dx) \to L^p(\omega\, dx)$, and hence

$$M : L^p(M(\omega)\, dx) \to L^{p,\infty}(\omega\, dx).$$

Plugging in our definition of λ gives

$$\begin{aligned}
\omega(B) &\lesssim |B|^p \left(\int_B (\omega(y)+\varepsilon)^{-\frac{p'}{p}}\, dy\right)^{-p} \int_B (\omega(y)+\varepsilon)^{-p'} \omega(y)\, dy \\
&\lesssim |B|^p \left(\int_B (\omega(y)+\varepsilon)^{-\frac{p'}{p}}\, dy\right)^{-p} \int_B (\omega(y)+\varepsilon)^{-p'+1}\, dy \\
&= |B|^p \left(\int_B (\omega(y)+\varepsilon)^{-\frac{p'}{p}}\, dy\right)^{-p} \int_B (\omega(y)+\varepsilon)^{-\frac{p'}{p}}\, dy \\
&= |B|^p \left(\int_B (\omega(y)+\varepsilon)^{-\frac{p'}{p}}\, dy\right)^{-(p-1)}.
\end{aligned}$$

Thus,

$$\frac{\omega(B)}{|B|^p} \left(\int_B (\omega(y)+\varepsilon)^{-\frac{p'}{p}}\, dy\right)^{\frac{p}{p'}} \lesssim 1.$$

Letting $\varepsilon \to 0$ gives $\omega \in A_p$.

Now, we sketch the proof of the converse direction. Suppose that $\omega \in A_p$. Define the following weighted maximal function:

$$(M_\omega f)(x) := \sup_{r>0} \frac{1}{\omega(B(x,r))} \int_{B(x,r)} |f(y)| \omega(y)\, dy.$$

It can be shown (via techniques from this section) that $M_\omega : L^1(\omega\, dx) \to L^{1,\infty}(\omega\, dx)$, and moreover $(Mf)^p \lesssim M_\omega(f^p)$.

Fix $f \in L^p(\omega\, dx)$. Then

$$\begin{aligned}
\left(\|Mf\|^*_{L^{p,\infty}(\omega\, dx)}\right)^p &= \sup_{\lambda>0} \lambda^p \omega\left(\{ x \;:\; (Mf)(x) > \lambda \}\right) = \sup_{\lambda>0} \lambda^p \omega\left(\{ x \;:\; (Mf)^p(x) > \lambda^p \}\right) \\
&= \sup_{\eta>0} \eta\, \omega\left(\{ x \;:\; (Mf)^p(x) > \eta \}\right).
\end{aligned}$$

Since $(Mf)^p(x) \lesssim M_\omega(|f|^p)(x)$,

$$\omega\left(\{ x \;:\; (Mf)^p(x) > \eta \}\right) \lesssim \omega\left(\{ x \;:\; M_\omega(|f|^p)(x) > \eta \}\right).$$

Thus,

$$\left(\|Mf\|^*_{L^{p,\infty}(\omega\, dx}\right)^p \lesssim \sup_{\eta>0} \eta\, \omega\left(\{ x \;:\; M_\omega(|f|^p)(x) > \eta \}\right) = \|M_\omega(|f|^p)\|_{L^{1,\infty}(\omega\, dx)}.$$

Since $M_\omega : L^1(\omega\, dx) \to L^{1,\infty}(\omega\, dx)$,

$$\|M_\omega(|f|^p)\|_{L^{1,\infty}(\omega\, dx)} \lesssim \||f|^p\|_{L^1(\omega\, dx)} = \int |f(x)|^p \omega(x)\, dx = \|f\|^p_{L^p(\omega\, dx)}.$$

Therefore,

$$\|Mf\|^*_{L^{p,\infty}(\omega\, dx} \lesssim \|f\|_{L^p(\omega\, dx)}$$

so that $M : L^p(\omega\, dx) \to L^{p,\infty}(\omega\, dx)$. □

Theorem 4.8. *Fix $1 \leq p \leq \infty$, and let $d\mu$ be a nonnegative Borel measure. If $M : L^p(d\mu) \to L^{p,\infty}(d\mu)$, then $d\mu = \omega\, dx$ for some $\omega \in A_p$.*

Proof. If we can prove that $d\mu$ is absolutely continuous, then in light of the previous proof, we are done.

Decompose $d\mu = \omega(x)\, dx + d\nu$ where $d\nu$ is the singular part of $d\mu$, i.e., there exists a compact set K with $|K| = 0$ but $\nu(K) > 0$. Define $U_n = \{ x : d(x, K) < 1/n \}$ and $f_n = \chi_{U_n \setminus K}$. As $U_n \setminus K \supseteq U_{n+1} \setminus K$, and $\bigcap (U_n \setminus K) = \emptyset$, it follows that $f_n \to 0$ pointwise.

We claim that $d\mu$ is finite on compact sets. To see this, pick a measurable set E with $0 < \mu(E) < \infty$; this possible since μ is nontrivial. We may assume that E is compact by inner regularity. Then

$$(M\chi_E)(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} \chi_E(y)\, dy.$$

Let $r = d(x, E) + \operatorname{diam} E$. Then

$$(M\chi_E)(x) \gtrsim \frac{|E|}{r^d}.$$

Since E is compact, if we restrict x to a compact set then $d(x, E)$ is bounded below and above. Thus, $M\chi_E \gtrsim_{E,F} 1$ uniformly for $x \in F$ with F compact. So suppose for the sake of contradiction that there is a compact set F with $\mu(F) = \infty$. Then

$$\infty = \mu(F) \leq \mu\left(\{ x : M\chi_E(x) \gtrsim_{E,F} 1 \}\right) \lesssim_{E,F} \int_E d\mu = \mu(E) < \infty$$

which is a contradiction.

With this claim proven, next, note that $\int |f_n|^p\, d\mu \to 0$ by the dominated convergence theorem. Fix $x \in K$. Then since $B(x, 1/n) \subseteq U_n$ and since $|K| = 0$,

$$\begin{aligned}
(Mf)(x) &= \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} \chi_{U_n \setminus K}(y)\, dy \geq \frac{1}{|B(x,1/n)|} \int_{B(x,1/n)} \chi_{U_n \setminus K}(y)\, dy \\
&= \frac{1}{|B(x,1/n)|} \int_{K^C} \chi_{B(x,1/n)}(y)\, dy \\
&= \frac{1}{\mathbb{R}^d} \int_{K^C} \chi_{B(x,1/n)}(y)\, dy \\
&= 1.
\end{aligned}$$

So $x \in \{ x : (Mf_n)(x) > 1/2 \}$. Thus, $K \subseteq \{ x : (Mf_n)(x) > 1/2 \}$. So

$$0 < \nu(K) < \mu(K) \leq \mu\left(\{ x : (Mf_n)(x) > 1/2 \}\right) \lesssim \int |f_n|^p\, d\mu \to 0$$

which is a contradiction. Thus, $d\mu = \omega(x)\, dx$, and by the previous theorem, $\omega \in A_p$.

□

Lemma 4.9. *We have $\omega \in A_p$ if and only if*

$$\left(\frac{1}{|B|} \int_B f(y)\, dy \right)^p \lesssim \frac{1}{\omega(B)} \int_B f(y)^p \omega(y)\, dy$$

uniformly for all $f \geq 0$ and all balls B.

Proof. First, suppose that $\omega \in A_p$. Then

$$\sup_B \frac{\omega(B)}{|B|^p} \left(\int_B w(y)^{-\frac{p'}{p}} \right)^{\frac{p}{p'}} \lesssim 1.$$

Then, by Holder,

$$\begin{aligned}
\frac{1}{|B|^p}\left(\int_B f(y)\,dy\right)^p &= \frac{1}{|B|^p}\left(\int_B f(y)\omega(y)^{\frac{1}{p}}\omega(y)^{-\frac{1}{p}}\,dy\right)^p \\
&\lesssim \frac{1}{|B|^p}\int_B |f(y)|^p\omega(y)\,dy\,\left(\int_B \omega(y)^{-\frac{p'}{p}}\right)^{\frac{p}{p'}} \\
&\lesssim \frac{1}{|B|^p}\int_B |f(y)|^p\omega(y)\,dy\,\frac{|B|^p}{\omega(B)} \\
&= \frac{1}{\omega(B)}\int_B f(y)^p\omega(y)\,dy.
\end{aligned}$$

To see the converse direction, suppose that

$$\left(\frac{1}{|B|}\int_B f(y)\,dy\right)^p \lesssim \frac{1}{\omega(B)}\int_B f(y)^p\omega(y)\,dy$$

uniformly for all $f \geq 0$ and all balls B. Let $f = (\omega+\varepsilon)^{-\frac{p'}{p}}$. Then

$$\begin{aligned}
\frac{1}{|B|^p}\left(\int_B (\omega(y)+\varepsilon)^{-\frac{p'}{p}}\,dy\right)^p &\lesssim \frac{1}{\omega(B)}\int_B (\omega(y)+\varepsilon)^{-p'}\omega(y)\,dy \\
&\leq \frac{1}{\omega(B)}\int_B (\omega(y)+\varepsilon)^{-p'+1}\,dy \\
&= \frac{1}{\omega(B)}\int_B (\omega(y)+\varepsilon)^{\frac{p'}{p}}\,dy.
\end{aligned}$$

So

$$\frac{\omega(B)}{|B|^p}\left(\int_B (\omega(y)+\varepsilon)^{-\frac{p'}{p}}\,dy\right)^{p-1} \lesssim 1.$$

Letting $\varepsilon \to 0$ gives the result. □

The final step before proving the weighted maximal inequality is a version of a covering lemma, due to Vitali.

Lemma 4.10 (Vitali)**.** *Let $\mathcal{F}$ be a finite collection of open balls in $\mathbb{R}^d$. Then there exists a subcollection S of $\mathcal{F}$ such that distinct balls in S are disjoint, and $\bigcup_{B\in\mathcal{F}} \subseteq \bigcup_{B\in S} 3B$.*

Proof. Run the following algorithm.

1. Set $S := \emptyset$.

2. Choose a ball in $\mathcal{F}$ of largest radius, and add it to S.

3. Discard all of the balls in $\mathcal{F}$ which intersect balls in S.

4. If all balls in $\mathcal{F}$ are removed, stop. Otherwise, return to step 2.

This algorithm terminates because at least one ball in $\mathcal{F}$ is discarded at every step. By construction, distinct balls in S are disjoint. Finally, if B is a ball in $\mathcal{F}$ which is not in S, it necessarily intersects some ball $B' \in S$ of larger radius. By the triangle inequality (draw a picture), $3B'$ contains B. Thus, $\bigcup_{B\in\mathcal{F}} \subseteq \bigcup_{B\in S} 3B$.

□

At last, we prove the weighted maximal inequality.

Proof of Theorem 4.3. First, we claim that $M : L^\infty(M(\omega)\,dx) \to L^\infty(\omega\,dx)$. To see this, let $f \in L^\infty(M(\omega)\,dx)$. Then

$$\|Mf\|_{L^\infty(\omega\,dx)} = \inf_{\omega(E)=0} \sup_{x\in E^C} |(Mf)(x)| \le \inf_{|E|=0} \sup_{x\in E^C} |(Mf)(x)|$$

because $|E| = 0$ implies $\omega(E) = 0$, as$\omega(E) = \int_E w(y)\,dy$. We also have the trivial estimate that $\|Mf(x)\|_{L^\infty(dx)} \le \|f\|_{L^\infty(dx)}$, which follows from moving absolute value signs inside the integral definition of Mf. Then

$$\|Mf\|_{L^\infty(\omega\,dx)} \le \inf_{|E|=0} \sup_{x\in E^C} |f(x)|.$$

Note that, unless $\omega \equiv 0$, $M\omega(x) > 0$ for all x. Thus,

$$\|Mf\|_{L^\infty(\omega\,dx)} \le \inf_{|E|=0} \sup_{x\in E^C} |f(x)| = \inf_{(M\omega)(E)=0} \sup_{x\in E^C} |f(x)| = \|f\|_{L^\infty(M\omega\,dx)}$$

as desired.

Since M is bounded from $L^\infty(M(\omega)\,dx) \to L^\infty(\omega\,dx)$, by the Marcinkiewicz interpolation theorem, it suffices to prove that M is bounded from $L^1(M(\omega)\,dx) \to L^{1,\infty}(\omega\,dx)$. Indeed, set $p_1 = 1$, $q_1 = 1$ and $p_2 = \infty$, $q_2 = \infty$, and for $\theta \in (0,1)$, define

$$\frac{1}{p_\theta} = \frac{\theta}{p_1} + \frac{1-\theta}{p_2} = \theta \qquad \Rightarrow p_\theta = \frac{1}{\theta}$$

and

$$\frac{1}{q_\theta} = \frac{\theta}{q_1} + \frac{1-\theta}{q_2} = \theta \qquad \Rightarrow q_\theta = \frac{1}{\theta}.$$

For $r = q_\theta = p_\theta = \frac{1}{\theta}$, the interpolation theorem gives boundedness of M as an operator from $L^{\frac{1}{\theta}}(M(\omega)\,dx) \to L^{\frac{1}{\theta}}(\omega\,dx)$. As $0 < \theta < 1$, $1 < \frac{1}{\theta} < \infty$, hence the theorem is proved.

So it remains to prove $M : L^1(M(\omega)\,dx) \to L^{1,\infty}(\omega\,dx)$. For $f \in L^1(M(\omega)\,dx)$, we need to show that $\|Mf\|^*_{L^{1,\infty}(\omega\,dx)} \lesssim \|f\|_{L^1(M(\omega)\,dx)}$, i.e.,

$$\sup_{\lambda>0} \lambda\,\omega\left(\{\,x \;:\; |Mf(x)| > \lambda\,\}\right) \lesssim \int |f(x)| M(\omega)(x)\,dx.$$

Fix $\lambda > 0$ and consider $\{\,x : |Mf(x)| > \lambda\,\}$. Let $K \subseteq \{\,x : |Mf(x)| > \lambda\,\}$ be compact. For $x \in K$, $Mf(x) > \lambda$, so by definition of the maximal function there exists $r_x > 0$ so that

$$\frac{1}{|B_r(x)|} \int_{B_{r_x}(x)} |f(y)|\,dy > \lambda.$$

Then $K \subseteq \bigcup_{x\in K} B_{r_x}(x)$, and compactness then gives a finite subcover of such balls. By Vitali's covering lemma, there exists a subcollection S of these balls such that distinct balls in S are disjoint, and $K \subseteq \bigcup_{B\in S} 3B$. Then $\omega(K) \le \sum_{B\in S} \omega(3B)$.

Consider a ball $B_j \in S$ with radius r_j. For $y \in B_j$, note that $B_{4r_j}(y) \supseteq 3B_j$. Thus,

$$\omega(3B_j) = \int_{3B_j} \omega(x)\,dx \le \int_{B_{4r_j}(y)} \omega(x)\,dx \le (M\omega)(y)\,|B_{4r_j}(y)| = (M\omega)(y)\,4^d\,|B_j|.$$

Therefore,

$$\omega(3B_j) \cdot \int_{B_j} |f(y)|\,dy \le 4^d |B_j| \int_{B_j} (M\omega)(y)|f(y)|\,dy$$

so that

$$\omega(3B_j) \cdot \frac{1}{|B_j|} \int_{B_j} |f(y)|\,dy \le 4^d \int_{B_j} (M\omega)(y)|f(y)|\,dy.$$

But $\frac{1}{|B_j|}\int_{B_j}|f(y)|\,dy > \lambda$, so

$$\omega(3B_j) \lesssim_d \frac{1}{\lambda}\int_{B_j}(M\omega)(y)|f(y)|\,dy.$$

Since the balls B_j are all disjoint, we then have

$$\omega(K) \leq \sum_{B_j \in S}\omega(3B_j) \lesssim \frac{1}{\lambda}\int_{\mathbb{R}^d}(M\omega)(y)|f(y)|\,dy = \frac{1}{\lambda}\|f\|_{L^1(M(\omega)\,dx)}.$$

As this holds for all compact $K \subseteq \{\, x \,:\, |Mf(x)| > \lambda \,\}$, by the inner regularity of the measure ω, we are done. $\square$

4.3 The Vector-Valued Maximal Inequality

Our first application of the weighted maximal inequality is a similar estimate for vector-valued maximal functions.

Definition 4.11. For $f : \mathbb{R}^d \to \ell^2(\mathbb{N})$ given by $f(x) = \{f_n(x)\}_{n\geq 1}$, define

$$\|f\|_{L^p} := \left\|\left\|\{f_n(x)\}_{\ell^2_n}\right\|\right\|_{L^p_x}.$$

The **vector-valued maximal function** $\bar{M}$ is given by

$$\bar{M}(f)(x) := \|\{Mf_n(x)\}\|_{\ell^2_n}.$$

Note that $\bar{M}f : \mathbb{R}^d \to [0,\infty]$, so that $\bar{M}f$ is not itself vector-valued. Rather, the phrase "vector-valued" refers only to the input functions f.

The main estimate on $\bar{M}$ is the following vector-valued version of the classical Hardy-Littlewood maximal inequality.

Theorem 4.12 (Vector-valued maximal inequality)**.**

1. *The operator $\bar{M}$ is of weak type $(1,1)$, that is,*

$$\left\|\bar{M}f\right\|^*_{L^{1,\infty}} \lesssim \|f\|_{L^1}$$

for $f : \mathbb{R}^d \to \ell^2(\mathbb{N})$;

2. *$\bar{M}$ is of strong type (p,p) for all $1 < p < \infty$, that is,*

$$\left\|\bar{M}f\right\|_{L^p} \lesssim \|f\|_{L^p}$$

for $f : \mathbb{R}^d \to \ell^2(\mathbb{N})$.

Before proving Theorem 4.12, we give an outline of the argument. The case $p = 2$ is a straightforward computation, and the case $p > 2$ will follow from the weighted maximal inequality. Having established the case $p = 2$, by the Marcinkiewicz Interpolation theorem it will suffice to prove the weak type estimate for $p = 1$ to prove the entire theorem. To prove this case we employ a *Calderon-Zygmund* decomposition technique.

We adopt the convention that for $f : \mathbb{R}^d \to \ell^2(\mathbb{N})$, $|f(x)| := \|\{f_n(x)\}\|_{\ell^2_n}$.

Lemma 4.13 (Calderon-Zygmund decomposition)**.** *Given $f \in L^1(\mathbb{R}^d)$ (possibly vector valued) and $\lambda > 0$, there is a decomposition $f = g + b$ such that:*

1. *$|g(x)| \leq \lambda$ for almost all $x \in \mathbb{R}^d$;*

2. $b = f\chi_{Q_k}$, *where* $\{Q_k\}$ *is a collection of cubes whose interiors are disjoint and such that*

$$\lambda < \frac{1}{|Q_k|}\int_{Q_k} |b(y)|\, dy \leq 2^d\lambda.$$

In this decomposition, g represents the *good* part of f, i.e., the part which is essentially uniformly bounded by λ, and b represents the *bad* part. The requirement of b in the decomposition says that the average of b on any given cube is on the order of λ.

Proof. Decompose $\mathbb{R}^d$ into dyadic cubes of the form:

$$Q_k = [2^n k_1, 2^n(k_1+1) \times \cdots \times [2^n k_d, 2^n(k_d+1))$$

where the diameter of each cube is chosen to be large enough so that

$$\frac{1}{Q_k}\int_{Q_k} |f(y)|\, dy \leq \lambda.$$

Because $\int_{\mathbb{R}^d} |f(y)|\, dy < \infty$, this is certainly possible.

Run the following algorithm. Fix a cube Q from the above decomposition. Divide Q into 2^d equal sized cubes. Let Q' denote one of these smaller cubes.

If Q satisfies

$$\frac{1}{Q'}\int_{Q'} |f(y)|\, dy > \lambda,$$

then stop and add Q' to the collection of cubes which define the support of b. Note that, for such a cube,

$$\lambda < \frac{1}{|Q'|}\int_{Q'} |f(y)|\, dy \leq \frac{2^d}{|Q|}\int_{Q} |f(y)|\, dy \leq 2^d\lambda$$

as required by the definition of b.

If Q satisfies

$$\frac{1}{Q'}\int_{Q'} |f(y)|\, dy \leq \lambda,$$

then subdivide Q' as we did with Q and continue until, if ever, we land in the previous case.

Having run this algorithm, let $b = f\chi_{\bigcup Q_k}$, where $\{Q_k\}$ is the collection of cubes from the above algorithm. Set $g = f - b$. By construction, the cubes Q_k have disjoint interiors.

It only remains to check that $|g| \leq \lambda$ almost everywhere. If $x \notin \bigcup Q_k$, then by our application there is a sequence of cubes, each containing x, with diameters shrinking to 0, such that the average value of $|f|$ on these cubes is all $\leq \lambda$. By the Lebesgue differentiation theorem, $|g| \leq \lambda$ almost everywhere. □

Now, we prove the vector-valued maximal inequality.

Proof Theorem 4.12. First consider the case $p = 2$. For a vector-valued $f = \{f_n\} \in L^2(\mathbb{R}^d)$, we need to show that $\|\bar{M}f\|_{L^2} \lesssim \|f\|_{L^2}$. By definition,

$$\|\bar{M}f\|_{L^2}^2 = \int_{\mathbb{R}^d} \|\{Mf_n(x)\}\|_{\ell_n^2}^2\, dx = \int_{\mathbb{R}^d} \sum_{n\geq 1} |Mf_n(x)|^2\, dx.$$

Since all the quantities in question are nonnegative, we can interchange the integral and summation (by Tonelli's theorem) to get

$$\|\bar{M}f\|_{L^2}^2 = \sum_{n\geq 1}\int_{\mathbb{R}^d} |Mf_n(x)|^2\, dx.$$

Previously, we proved that M is of strong type $(2,2)$. So

$$\left\|\bar{M}f\right\|_{L^2}^2 = \sum_{n\geq 1}\int_{\mathbb{R}^d} |Mf_n(x)|^2\,dx \lesssim \sum_{n\geq 1}\int_{\mathbb{R}^d} |f_n(x)|^2\,dx = \int_{\mathbb{R}^d}\sum_{n\geq 1} |f_n(x)|^2\,dx = \|f\|_{L^2}^2$$

as desired.

Next, suppose that $2 < p < \infty$. We wish to show that $\left\|\bar{M}f\right\|_{L^p} \lesssim \|f\|_{L^p}$. Observe that

$$\left\|\bar{M}f\right\|_{L^p} = \left\|\|Mf_n(x)\|_{\ell_n^2}\right\|_{L_x^p} = \left\|(\|Mf_n(x)\|_{\ell_n^2})^2\right\|_{L_x^{p/2}}^{\frac{1}{2}}$$

so that $\left\|\bar{M}f\right\|_{L^p}^2 = \left\|(\bar{M}f(x))^2\right\|_{L^{p/2}}$. By the duality characterization of norms,

$$\begin{aligned}\left\|\bar{M}f\right\|_{L^p}^2 = \left\|(\bar{M}f(x))^2\right\|_{L^{p/2}} &= \sup_{\|\omega\|_{L^{(p/2)'}}=1}\int (\bar{M}f(x))^2\,\omega(x)\,dx\\ &= \sup_{\|\omega\|_{L^{(p/2)'}}=1}\int \sum_{n\geq 1}|Mf_n(x)|^2\,\omega(x)\,dx\\ &= \sup_{\|\omega\|_{L^{(p/2)'}}=1}\sum_{n\geq 1}\int |Mf_n(x)|^2\,\omega(x)\,dx.\end{aligned}$$

For $\omega \geq 0$, the weighted maximal inequality that we proved earlier tells us that $M : L^2(M(\omega)\,dx) \to L^2(\omega\,dx)$. Thus, supposing without loss of generality that the supremum in the above expression is taken over $\omega \geq 0$, we have

$$\begin{aligned}\left\|\bar{M}f\right\|_{L^p}^2 &= \sup_{\|\omega\|_{L^{(p/2)'}}=1}\sum_{n\geq 1}\int |Mf_n(x)|^2\,\omega(x)\,dx \lesssim \sup_{\|\omega\|_{L^{(p/2)'}}=1}\sum_{n\geq 1}\int |f_n(x)|^2\,(M\omega)(x)\,dx\\ &= \sup_{\|\omega\|_{L^{(p/2)'}}=1}\int\sum_{n\geq 1} |f_n(x)|^2\,(M\omega)(x)\,dx.\end{aligned}$$

Applying Holder's inequality,

$$\left\|\bar{M}f\right\|_{L^p}^2 \lesssim \sup_{\|\omega\|_{L^{(p/2)'}}=1}\left\|\sum_{n\geq 1}|f_n|^2\right\|_{L^{(p/2)}}\|M\omega\|_{L^{(p/2)'}}.$$

By the standard Hardy-Little maximal inequality theorem, $\|M\omega\|_{L^{(p/2)'}} \lesssim \|\omega\|_{L^{(p/2)'}} = 1$. Finally, note that

$$\left\|\sum_{n\geq 1}|f_n|^2\right\|_{L^{(p/2)}} = \left\|\|f_n\|_{\ell_n^2}^2\right\|_{L^{(p/2)}} = \|f\|_{L^p}^2$$

so that $\left\|\bar{M}f\right\|_{L^p} \lesssim \|f\|_{L^p}$ as desired.

As noted above, it remains to prove the weak type $(1,1)$ claim. Explicitly, given $f = \{f_n\} \in L^1(\mathbb{R}^d)$, we need to show that

$$\sup_{\lambda>0}\lambda\left|\left\{\,x\;:\;(\bar{M}f)(x) > \lambda\,\right\}\right| \lesssim \|f\|_{L^1}.$$

Using the Calderon-Zygmund decomposition, write $f = g + b$ with $|g| \leq \lambda$ almost everywhere and $b = f\chi_{\bigcup Q_k}$ where Q_k are cubes with disjoint interiors satisfying

$$\frac{1}{|Q_k|}\int_{Q_k}|b(y)|\,dy \sim \lambda.$$

Because $\bar{M}$ is a sublinear operator,

$$\{\, x \,:\, (\bar{M}f)(x) > \lambda \,\} \subseteq \left\{\, x \,:\, (\bar{M}g)(x) > \frac{\lambda}{2} \,\right\} \cup \left\{\, x \,:\, (\bar{M}b)(x) > \frac{\lambda}{2} \,\right\}.$$

Consider the first set. We have already shown that $\bar{M}$ is of strong type $(2,2)$, so in particular it is of weak type $(2,2)$. Thus,

$$\left|\left\{\, x \,:\, (\bar{M}g)(x) > \frac{\lambda}{2} \,\right\}\right| \lesssim \frac{1}{\lambda}\, \|g\|_{L^2}^2 = \frac{1}{\lambda^2}\, \left\||g|^2\right\|_{L^1} \leq \frac{1}{\lambda^2}\, \|\lambda|g|\|_{L^1} = \frac{1}{\lambda}\, \|g\|_{L^1} \leq \frac{1}{\lambda}\, \|f\|_{L^1}\,.$$

Therefore, it remains to prove the same type of estimate for the set $\{\, x \,:\, (\bar{M}b)(x) > \frac{\lambda}{2} \,\}$.

Let $2Q_k$ denote the cube with same center as Q_k and twice the side length. Then, because the interior of the Q_k's are disjoint, we have:

$$\begin{aligned}\left|\bigcup 2Q_k\right| &\leq 2^d \sum_k |Q_k| \lesssim \sum_k \frac{1}{\lambda} \int_{Q_k} |b(y)|\, dy \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^d} |b(y)|\, dy \\ &\leq \frac{1}{\lambda}\, \|f\|_{L^1}\,.\end{aligned}$$

With this estimate, it remains to show that

$$\left|\left\{\, x \notin \bigcup 2Q_k \,:\, (\bar{M}b)(x) > \frac{\lambda}{2} \,\right\}\right| \lesssim \frac{1}{\lambda}\, \|f\|_{L^1}\,.$$

Define $b_n^{avg}(x) = \sum_k \chi_{Q_k}(x) \frac{1}{|Q_k|} \int_{Q_k} |b_n(y)|\, dy$. If $x \in Q_k$, then

$$\|b_n^{avg}(x)\|_{\ell_n^2} = \frac{1}{|Q_k|} \left\| \int_{Q_k} |b_n(y)|\, dy \right\|_{\ell_n^2} \leq \frac{1}{|Q_k|} \int_{Q_k} \|b_n(y)\|_{\ell_n^2}\, dy = \frac{1}{|Q_k|} \int_{Q_k} |b(y)|\, dy \lesssim \lambda$$

invoking Minkowski's integral inequality to move the ℓ_n^2 norm inside the integral. Thus,

$$\|b^{avg}\|_{L^1} = \left\| \|b_n^{avg}(x)\|_{\ell_n^2} \right\|_{L_x^1} \lesssim \lambda \left|\bigcup Q_k\right| \lesssim \lambda \cdot \frac{1}{\lambda}\, \|f\|_{L^1} = \|f\|_{L^1}\,.$$

Now, fix $x \notin \bigcup 2Q_k$. Then

$$Mb_n(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |b_n(y)|\, dy = \sup_{r>0} \frac{1}{|B(x,r)|} \sum_k \int_{B(x,r)\cap Q_k} |b_n(y)|\, dy.$$

Suppose that $x \notin \bigcup 2Q_k$ but that $B(x,r) \cap Q_k \neq \emptyset$. Let l be the side length of Q_k. Then necessarily $r \geq l/2$. Note that the diameter of Q_k is $\sqrt{d}l \leq 2r\sqrt{d}$. This implies that $B(x, r(1+2\sqrt{d})) \supseteq Q_k$. Thus,

$$\begin{aligned} Mb_n(x) &\lesssim_d \sup_{r>0} \frac{1}{|B(x, r(1+2\sqrt{d})|} \sum_k \int_{Q_k} |b_n(y)|\, dy \\ &= \sup_{r>0} \frac{1}{|B(x, r(1+2\sqrt{d})|} \sum_k \left(\int_{B(x,r(1+2\sqrt{d}))} \chi_{Q_k}(z)\, dz \right) \left(\int_{Q_k} |b_n(y)|\, dy \right) \\ &= \sup_{r>0} \frac{1}{|B(x, r(1+2\sqrt{d})|} \int_{B(x,r(1+2\sqrt{d}))} \sum_k \chi_{Q_k}(z) \cdot \frac{1}{|Q_k|} \int_{Q_k} |b_n(y)|\, dy\, dz \\ &= Mb_n^{avg}(x). \end{aligned}$$

It follows that $\bar{M}b \lesssim \bar{M}b^{avg}$. So we again apply the $(2,2)$ estimate of $\bar{M}$ to get

$$\begin{aligned}\left|\left\{ x \notin \bigcup 2Q_k \,:\, (\bar{M}b)(x) > \frac{\lambda}{2} \right\}\right| &\le \left|\left\{ x \notin \bigcup 2Q_k \,:\, (\bar{M}b^{avg})(x) \gtrsim \lambda \right\}\right| \lesssim \frac{1}{\lambda^2} \left\| b^{avg} \right\|_{L^2}^2 \\ &= \frac{1}{\lambda^2} \left\| \| b_n^{avg}(x) \|_{\ell_n^2} \right\|_{L_x^2}^2 = \frac{1}{\lambda^2} \left\| \| b_n^{avg}(x) \|_{\ell_n^2}^2 \right\|_{L_x^1} \\ &\lesssim \frac{1}{\lambda} \left\| \| b_n^{avg}(x) \|_{\ell_n^2} \right\|_{L_x^1} = \frac{1}{\lambda} \left\| b^{avg} \right\|_{L^1} \\ &\lesssim \frac{1}{\lambda} \| f \|_{L^1} .\end{aligned}$$

$\square$

Chapter 5

Sobolev Inequalities

This chapter contains a brief foray into the world of so-called *Sobolev inequalities*. A typical inequality of this type bounds a certain norm of a function by a norm of a derivative of the function. In vague terms, this kind of estimate bounds the average gradient (or energy) of a function below by its overall spread. This turns out to be a useful fact.

The inequalities we prove in this chapter will appear again throughout the rest of the book in different contexts. For example. we will provide an alternate proof for the Gagliardo-Nirenberg inequality after we develop the machinery of Littlewood-Paley theory. Also, we will discuss the optimal constants of the Gagliardo-Nirenberg and Sobolev embedding inequalities in the future.

5.1 The Hardy-Littlewood-Sobolev Inequality

As a precursor to the Sobolev inequalities we will prove, we begin with an important convolution estimate called the Hardy-Littlewood-Sobolev inequality. We actually prove two such inequalities in this section, the latter a generalization of the first.

There are many ways to prove the following theorem. The approach we take is due to Hedberg [4] and is amenable to proving certain *inverse* inequalities.

Theorem 5.1 (Hardy-Littlewood-Sobolev I). *Let $f \in \mathcal{S}(\mathbb{R}^d)$. Then*

$$\left\| f * \frac{1}{|x|^\alpha} \right\|_{L^r} \lesssim \|f\|_{L^p}$$

whenever $1 + \frac{1}{r} = \frac{1}{p} + \frac{\alpha}{d}$ for $1 < p < r < \infty$ and $0 < \alpha < d$.

Proof. We remark that requiring $f \in \mathcal{S}(\mathbb{R}^d)$ ensures that $f \in L^p(\mathbb{R}^d)$ for all necessary p.

Decompose the convolution as follows:

$$\left(f * \frac{1}{|x|^\alpha} \right)(x) = \int \frac{f(y)}{|x-y|^\alpha}\, dy = \int_{|x-y| \le R} \frac{f(y)}{|x-y|^\alpha}\, dy + \int_{|x-y| > R} \frac{f(y)}{|x-y|^\alpha}\, dy$$

for some $R > 0$. We will estimate each integral and then optimize in R.

Consider the first integral.

$$\begin{aligned} \left| \int_{|x-y| \le R} \frac{f(y)}{|x-y|^\alpha}\, dy \right| &\le \int_{|x-y| \le R} \frac{|f(y)|}{|x-y|^\alpha}\, dy \le \sum_{r \in 2^{\mathbb{Z}}; r \le R} \int_{r < |x-y| < 2r} \frac{|f(y)|}{|x-y|^\alpha}\, dy \\ &\lesssim \sum_{r \in 2^{\mathbb{Z}}; r \le R} r^{-\alpha} \int_{|x-y| < 2r} |f(y)|\, dy \\ &\lesssim \sum_{r \in 2^{\mathbb{Z}}; r \le R} \frac{r^{d-\alpha}}{|B(x, 2r)|} \int_{|x-y| < 2r} |f(y)|\, dy. \end{aligned}$$

The quantity $\frac{1}{|B(x,2r)|}\int_{|x-y|<2r}|f(y)|\,dy$ is bounded by $Mf(x)$, so

$$\left|\int_{|x-y|\le R}\frac{f(y)}{|x-y|^\alpha}\,dy\right|\le Mf(x)\sum_{r\in 2^{\mathbb{Z}};r\le R}r^{d-\alpha}.$$

Because $d-\alpha>0$ and because the r's are dyadic numbers, the sum is summable and is bounded (up to a constant) by the largest term. Thus,

$$\left|\int_{|x-y|\le R}\frac{f(y)}{|x-y|^\alpha}\,dy\right|\le R^{d-\alpha}Mf(x).$$

Now we consider the second integral. We have

$$\int_{|x-y|\le R}\frac{f(y)}{|x-y|^\alpha}\,dy=\left(f*\left(\frac{1}{|x|^\alpha}\chi_{\{|x|>R\}}\right)\right)(x).$$

By Young's convolution inequality (or more directly, Holder's inequality),

$$\left\|f*\left(\frac{1}{|x|^\alpha}\chi_{\{|x|>R\}}\right)\right\|_{L^\infty}\lesssim\|f\|_{L^p}\left\|\frac{1}{|x|^\alpha}\chi_{\{|x|>R\}}\right\|_{L^{p'}}=\|f\|_{L^p}\left(\int_{|x|>R}\frac{1}{|x|^{\alpha p'}}\,dx\right)^{\frac{1}{p'}}.$$

If $\alpha p'>d$, then the integral quantity is finite, and in particular $\lesssim R^{\frac{d-\alpha p'}{p'}}$. Since

$$\frac{\alpha}{d}+\frac{1}{p}>1\qquad\Rightarrow\qquad\frac{\alpha}{d}>1-\frac{1}{p}=\frac{1}{p'}$$

this is indeed the case. Thus,

$$\left|\int_{|x-y|\le R}\frac{f(y)}{|x-y|^\alpha}\,dy\right|\le\left\|f*\left(\frac{1}{|x|^\alpha}\chi_{\{|x|>R\}}\right)\right\|_{L^\infty}\lesssim\|f\|_{L^p}R^{\frac{d}{p'}-\alpha}.$$

We optimize our choice of R by requiring the two estimates to be comparable, so that $R^{d-\alpha}Mf(x)\sim R^{\frac{d}{p'}-\alpha}\|f\|_p$. Choose R so that $R\sim\left(\frac{\|f\|_{L^p}}{Mf(x)}\right)^{\frac{p}{d}}$.

With this choice,

$$\left|f*\frac{1}{|x|^\alpha}\right|(x)\lesssim\left(\left(\frac{\|f\|_p}{Mf(x)}\right)^{\frac{p}{d}}\right)^{d-\alpha}Mf(x)=\|f\|_{L^p}^{\frac{d-\alpha}{d}p}(Mf(x))^{1-\frac{p}{d}(d-\alpha)}.$$

Simplifying the exponents,

$$1-\frac{p}{d}(d-\alpha)=p\left(\frac{1}{p}-1+\frac{\alpha}{d}\right)=\frac{p}{r}.$$

So

$$\left|f*\frac{1}{|x|^\alpha}\right|(x)\lesssim\|f\|_{L^p}^{1-\frac{p}{r}}(Mf(x))^{\frac{p}{r}}.$$

Taking the L^r norm and using the fact that M is of type (p,p) then gives

$$\left\|f*\frac{1}{|x|^\alpha}\right\|_{L^r}\lesssim\|f\|_{L^p}^{1-\frac{p}{r}}\left\|(Mf(x))^{\frac{p}{r}}\right\|_{L^r}=\|f\|_{L^p}^{1-\frac{p}{r}}\|Mf\|_{L^p}^{\frac{p}{r}}\lesssim\|f\|_{L^p}^{1-\frac{p}{r}}\|f\|_{L^p}^{\frac{p}{r}}=\|f\|_{L^p}$$

as desired.

$\square$

It is tempting to view the Hardy-Littlewood-Sobolev inequality as having more or less the same content as Young's convolution inequality, as the statements of the estimates are similar. However, the fact that $|x|^{-\alpha}$ is not in any L^p spaces makes Hardy-Littlewood-Sobolev much more subtle.

Note, however, that the function $|x|^{-\alpha}$ lives in the weak space $L^{\frac{d}{\alpha},\infty}(\mathbb{R}^d)$. This suggests the following generalization of the previous theorem.

Theorem 5.2 (Hardy-Littlewood-Sobolev II). *For $1 < p < r < \infty$ and $1 < q < \infty$, we have*

$$\|f * g\|_{L^r} \lesssim \|f\|_{L^p} \|g\|^*_{L^{q,\infty}}$$

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

Proof. We begin with a few reductions. By rescaling, we can assume without loss of generality that $\|f\|_{L^p} = \|g\|^*_{L^{q,\infty}} = 1$. Furthermore, note that for a fixed such g it suffices to prove that the operator $f \mapsto f * g$ is of strong type (p, r). In fact, by the Marcinkiewicz interpolation theorem, it suffices to prove that this operator is of *weak* type (p, r). This is because the condition $1 < p < r < \infty$ is an *open* condition, so that we can always find necessary (p_1, r_1) and (p_2, r_2) satisfying the relevant conditions to yield the strong type estimate in the interpolation theorem.

So fix $\lambda > 0$. We need to show that $\|f * g\|^*_{L^{r,\infty}} \lesssim 1$, i.e., $|\{\, x \;:\; |f * g|(x) > \lambda \,\}| \lesssim \lambda^{-r}$. As in the previous proof, we decompose $g = g\chi_{|g|\le R} + g\chi_{|g|>R} := g_1 + g_2$. Then

$$|\{\, x \;:\; |f * g|(x) > \lambda \,\}| \le \left|\left\{\, x \;:\; |f * g_1|(x) > \frac{\lambda}{2} \,\right\}\right| + \left|\left\{\, x \;:\; |f * g_2|(x) > \frac{\lambda}{2} \,\right\}\right|.$$

We will show that the first quantity on the right is 0 for an appropriately chosen R. Intuitively, since g_1 is bounded by R, the convolution of f with g cannot be too large. Computing,

$$\begin{aligned}\|f * g_1\|_{L^\infty} \lesssim \|f\|_{L^p} \|g_1\|_{L^{p'}} = \|g_1\|_{L^{p'}} &\lesssim \left(\int_0^\infty \alpha^{p'} \left|\{\, x \;:\; |g_1(x)| > \alpha \,\}\right| \frac{d\alpha}{\alpha}\right)^{\frac{1}{p'}} \\ &= \left(\int_0^R \alpha^{p'} \left|\{\, x \;:\; |g_1(x)| > \alpha \,\}\right| \frac{d\alpha}{\alpha}\right)^{\frac{1}{p'}}\end{aligned}$$

where we have used the layer-cake representation for the $L^{p'}$ norm and (see Chapter 3) and the fact that $|g_1| \le R$.

Continuing,

$$\begin{aligned}\|f * g_1\|_{L^\infty} &\lesssim \left(\int_0^R \sup_{\alpha>0} \left[\alpha^q \left|\{\, x \;:\; |g_1(x)| > \alpha \,\}\right|\right] \alpha^{p'-q} \frac{d\alpha}{\alpha}\right)^{\frac{1}{p'}} \\ &= \left(\sup_{\alpha>0} \left[\alpha^q \left|\{\, x \;:\; |g_1(x)| > \alpha \,\}\right|\right]\right)^{\frac{1}{p'}} \left(\int_0^R \alpha^{p'-q} \frac{d\alpha}{\alpha}\right)^{\frac{1}{p'}}.\end{aligned}$$

The quantity on the left is precisely $(\|g\|^*_{L^{q,\infty}})^{\frac{1}{p'}}$, which we are assuming is 1. The integral on the right is integrable if $p' - q > 0$. By assumption,

$$1 - \frac{1}{p} = \frac{1}{q} - \frac{1}{r} \qquad \Rightarrow \qquad \frac{1}{p'} < \frac{1}{q} \qquad \Rightarrow \qquad p' > q.$$

So the integral is finite, and in particular, $\left(\int_0^R \alpha^{p'-q} \frac{d\alpha}{\alpha}\right)^{\frac{1}{p'}} \lesssim R^{\frac{p'-q}{p'}}$. Putting this all together gives

$$\|f * g_1\|_{L^\infty} \lesssim R^{\frac{p'-q}{p'}}.$$

Recall that we are estimating $\left|\left\{ x : |f * g_1|(x) > \frac{\lambda}{2} \right\}\right|$. Thus, if we choose R small enough (dependent on λ, say $R = c\lambda$ for some sufficiently small constant c, then $\|f * g_1\|_{L^\infty} \le \lambda/2$, and so

$$\left|\left\{ x : |f * g_1|(x) > \frac{\lambda}{2} \right\}\right| = 0.$$

Thus, it remains to estimate $\left|\left\{ x : |f * g_2|(x) > \frac{\lambda}{2} \right\}\right|$. By Chebychev's inequality,

$$\begin{aligned}\left|\left\{ x : |f * g_2|(x) > \frac{\lambda}{2} \right\}\right| &\lesssim \lambda^{-p} \|f * g_2\|_{L^p}^p \lesssim \lambda^{-p} \left(\|f\|_{L^p} \|g_2\|_{L^1}\right)^p \\ &= \lambda^{-p} \left(\int_0^\infty \alpha \left|\{ x : |g_2(x)| > \alpha \}\right| \frac{d\alpha}{\alpha} \right)^p.\end{aligned}$$

We consider the integral separately.

$$\begin{aligned}&\int_0^\infty \alpha \left|\{ x : |g_2(x)| > \alpha \}\right| \frac{d\alpha}{\alpha} \\ &\quad = \int_0^R \alpha \left|\{ x : |g_2(x)| > \alpha \}\right| \frac{d\alpha}{\alpha} + \int_R^\infty \alpha \left|\{ x : |g_2(x)| > \alpha \}\right| \frac{d\alpha}{\alpha} \\ &\quad \le R \cdot \left|\{ x : |g(x)| > R \}\right| + \sup_{\alpha > 0} \alpha^q \left|\{ x : |g(x)| > \alpha \}\right| \int_R^\infty \alpha^{1-q} \frac{d\alpha}{\alpha}.\end{aligned}$$

As before, $\sup_{\alpha>0} \alpha^q \left|\{ x : |g(x)| > \alpha \}\right| = \|g\|_{L^{q,\infty}}^* = 1$. Furthermore, since $q > 1$, we have $\int_R^\infty \alpha^{1-q} \frac{d\alpha}{\alpha} \lesssim R^{1-q}$. So

$$\int_0^\infty \alpha \left|\{ x : |g_2(x)| > \alpha \}\right| \frac{d\alpha}{\alpha} \lesssim R^q \cdot \left|\{ x : |g(x)| > R \}\right| R^{1-q} + R^{1-q}.$$

But since $R^q \cdot \left|\{ x : |g(x)| > R \}\right| \le \|g\|_{L^{q,\infty}}^* = 1$, we then have

$$\int_0^\infty \alpha \left|\{ x : |g_2(x)| > \alpha \}\right| \frac{d\alpha}{\alpha} \lesssim R^{1-q}.$$

Plugging this into our original estimate and using $R^{1-\frac{q}{p'}} = c\lambda$,

$$\left|\{ x : |f * g_2| > \lambda/2 \}\right| \lesssim \lambda^{-p} R^{p(1-q)} \lesssim \lambda^{-p} R^{p(1-q)} \lesssim \lambda^{-p} \lambda^{\left(\frac{p'}{p'-q}\right) p(1-q)} \lesssim \lambda^{-r}.$$

□

5.2 The Sobolev Embedding Theorem

As a consequence of the Hardy-Littlewood Sobolev inequalities, we prove a version of the Sobolev embedding theorem.

Towards this goal, we begin with a computation. We wish to (formally) take the Fourier transform of functions of the form $\frac{1}{|x|^{d-\alpha}}$. As this function is not in $L^p(\mathbb{R}^d)$, we consider the Fourier transform in the sense of distributions. Recall that if $T \in \mathcal{S}'(\mathbb{R}^d)$ is a tempered distribution, its Fourier transform is defined by $\hat{T}(f) := T(\hat{f})$ for $f \in \mathcal{S}(\mathbb{R}^d)$.

Proposition 5.3. *Let $0 < \alpha < d$. Then $\widehat{\frac{1}{|x|^{d-\alpha}}} \sim \frac{1}{|x|^\alpha}$, in the sense of tempered distributions. Explicitly,*

$$\left(\pi^{-\frac{d-\alpha}{2}} \Gamma\left(\frac{d-\alpha}{2}\right) \frac{1}{|x|^{d-\alpha}} \right)^{\wedge} = \pi^{-\frac{\alpha}{2}} \Gamma\left(\frac{\alpha}{2}\right) \frac{1}{|x|^\alpha}. \tag{5.1}$$

Proof. First, observe that

$$\int_0^\infty e^{-\pi t|x|^2} t^{\frac{d-\alpha}{2}} \frac{dt}{t} = \int_0^\infty e^{-u} \left(\frac{u}{\pi|x|^2}\right)^{\frac{d-\alpha}{2}} \frac{du}{u} = \pi^{-\frac{d-\alpha}{2}} \Gamma\left(\frac{d-\alpha}{2}\right) \frac{1}{|x|^{d-\alpha}}.$$

So for $f \in \mathcal{S}(\mathbb{R}^d)$,

$$\begin{aligned}\left(\pi^{-\frac{d-\alpha}{2}} \Gamma\left(\frac{d-\alpha}{2}\right) \frac{1}{|x|^{d-\alpha}}\right)\hat{}\,(f) &= \left(\pi^{-\frac{d-\alpha}{2}} \Gamma\left(\frac{d-\alpha}{2}\right) \frac{1}{|x|^{d-\alpha}}\right)(\hat{f}) \\ &= \int_{\mathbb{R}^d} \pi^{-\frac{d-\alpha}{2}} \Gamma\left(\frac{d-\alpha}{2}\right) \frac{1}{|x|^{d-\alpha}} \hat{f}(x)\, dx.\end{aligned}$$

Our first observation, together with the definition of the Fourier transform, gives

$$\begin{aligned}&= \int_{\mathbb{R}^d} \int_0^\infty \int_{\mathbb{R}^d} e^{-\pi t|x|^2} t^{\frac{d-\alpha}{2}} \frac{dt}{t} \int_{\mathbb{R}^d} e^{-2\pi i x\cdot y} f(y)\, dy\, dx \\ &= \int_{\mathbb{R}^d} \int_0^\infty \left(\int_{\mathbb{R}^d} e^{-\pi t|x|^2} e^{-2\pi i x\cdot y}\, dx\right) t^{-\frac{d-\alpha}{2}} \frac{dt}{t} f(y)\, dy.\end{aligned}$$

The quantity in the parentheses is precisely the Fourier transform of $e^{-\pi|x|^2}$. Recall that

$$\left(e^{-x\cdot Ax}\right)\hat{} = \pi^{\frac{d}{2}} (\det A)^{-\frac{1}{2}} e^{-\pi^2 \xi\cdot A^{-1}\xi}.$$

Thus,

$$\int_{\mathbb{R}^d} e^{-\pi t|x|^2} e^{-2\pi i x\cdot y}\, dx = \pi^{\frac{d}{2}} (\pi t)^{-\frac{d}{2}} e^{-\pi^2 y\cdot \frac{1}{\pi t} y} = t^{\frac{d}{2}} e^{-\frac{\pi|y|^2}{t}}.$$

Continuing the computation,

$$\begin{aligned}&= \int_{\mathbb{R}^d} \int_0^\infty e^{-\frac{\pi|y|^2}{t}} t^{-\frac{\alpha}{2}} \frac{dt}{t} f(y)\, dy = \int_{\mathbb{R}^d} \int_0^\infty t^{\frac{d}{2}} e^{-\frac{\pi|y|^2}{t}} t^{-\frac{d-\alpha}{2}} \frac{dt}{t} f(y)\, dy \\ &= \int_{\mathbb{R}^d} \int_0^\infty \left(\frac{\pi|y|^2}{u}\right)^{-\frac{\alpha}{2}} e^{-u} \frac{du}{u} f(y)\, dy \\ &= \int_{\mathbb{R}^d} \int_0^\infty \pi^{-\frac{\alpha}{2}} |y|^{-\alpha} u^{\frac{\alpha}{2}} e^{-u} \frac{du}{u} f(y)\, dy \\ &= \int_{\mathbb{R}^d} \pi^{-\frac{\alpha}{2}} \Gamma\left(\frac{\alpha}{2}\right) |y|^{-\alpha} f(y)\, dy \\ &= \left(\pi^{-\frac{\alpha}{2}} \Gamma\left(\frac{\alpha}{2}\right) |y|^{-\alpha}\right)(f).\end{aligned}$$

This proves (5.1) in the sense of tempered distributions. □

The are various forms of the Sobolev embedding theorem. The one we consider here involves a fractional derivative operator $|\nabla|^s$, which generalizes the usual notion of derivative. The operator $|\nabla|^s$ is defined via its action in the Fourier domain.

Definition 5.4. Fix $s > -d$. For $f \in \mathcal{S}(\mathbb{R}^d)$, define $|\nabla|^s$ via

$$(|\nabla|^s f)\hat{}\,(\xi) := (2\pi|\xi|)^s \hat{f}(\xi),$$

where this equality is understood in the sense of tempered distributions.

We require $s > -d$ so that the quantity $(2\pi|\xi|)^s$ makes sense as a distribution on $\mathbb{R}^d$. As $(D^\alpha f)\hat{}\,(\xi) = (2\pi i\xi)^\alpha \hat{f}(\xi)$, it is clear why the operator $|\nabla|^s$ generalizes the usual notion of a derivative. It is natural to wonder how estimates on $|\nabla|^s$ compare to estimates on the usual derivative operator ∇. It turns out to be a fact that $\||\nabla| f\|_{L^p} \lesssim \|\nabla f\|_{L^p}$. We will prove this in the future when we discuss Calderon-Zygmund convolution kernels.

Theorem 5.5 (Sobolev embedding). *Fix $1 < p < \infty$, $s > 0$. Let $f \in \mathcal{S}(\mathbb{R}^d)$ such that $|\nabla|^s f \in L^p(\mathbb{R}^d)$.*[1] *Then*

$$\|f\|_{L^q} \lesssim \||\nabla|^s f\|_{L^p}$$

provided $\frac{1}{p} = \frac{1}{q} + \frac{s}{d}$.

Proof. Recall that $\mathcal{S}(\mathbb{R}^d)$ is dense in $L^{q'}$. Thus, by duality of L^p-norms and Plancharel's theorem,

$$\begin{aligned}\|f\|_{L^q} &= \sup_{g \in \mathcal{S}(\mathbb{R}^d) : \|g\|_{L^{q'}} = 1} \langle f, g \rangle = \sup_{g \in \mathcal{S} : \|g\|_{L^{q'}} = 1} \left\langle \hat{f}, \hat{g} \right\rangle \\ &= \sup_{g \in \mathcal{S} : \|g\|_{L^{q'}} = 1} \left\langle (2\pi|\xi|)^s \hat{f}(\xi), (2\pi|\xi|)^{-s} \hat{g}(\xi) \right\rangle.\end{aligned}$$

By our previous comments, $(2\pi|\xi|)^s \hat{f}(\xi) \in \mathcal{S}'(\mathbb{R}^d)$. We would like for $(2\pi|\xi|)^{-s}\hat{g}(\xi)$ to be a Schwartz function but if $\tilde{g}$ does not vanish near the origin then the singularity at $\xi = 0$ may prevent this from being true. Thus, we consider a slightly smaller space of functions.

We claim that the family

$$\mathcal{F} := \left\{ g \in \mathcal{S}(\mathbb{R}^d) \ : \ \hat{g} \text{ vanishes on a neighborhood of } \xi = 0 \right\}$$

is dense in L^p for $1 < p < \infty$. Obviously it suffices to show that $\mathcal{F}$ is dense in $\mathcal{S}$. Fix $g \in \mathcal{S}(\mathbb{R}^d)$. Let φ be a smooth bump function on $\mathbb{R}^d$ such that $\varphi(\xi) = 1$ for $|\xi| \leq 1$ and $\varphi(\xi) = 0$ for $|\xi| \geq 2$. Define

$$\hat{g}_\varepsilon(\xi) := \hat{g}(\xi) \left(1 - \varphi\left(\frac{|\xi|}{\varepsilon}\right)\right).$$

Note that

$$\hat{g}(\xi)\left(1 - \varphi\left(\frac{|\xi|}{\varepsilon}\right)\right) = \hat{g}(\xi) - \hat{g}(\xi)\varphi\left(\frac{|\xi|}{\varepsilon}\right) = \left(g(x) - g * \left(\varphi\left(\frac{|\xi|}{\varepsilon}\right)\right)^{\check{}}(x)\right)^{\hat{}}.$$

Hence

$$\|g - g_\varepsilon\|_{L^p} = \left\| g * \left(\varphi\left(\frac{|\xi|}{\varepsilon}\right)\right)^{\check{}} \right\|_{L^p} = \left\| g * \varepsilon^d \check{\varphi}(\varepsilon|x|) \right\|_{L^p} \lesssim \varepsilon^d \|g\|_{L^1} \varepsilon^{-\frac{d}{p}} \|\check{\varphi}\|_{L^p}.$$

Since g and φ are fixed and because $p > 1$, as $\varepsilon \to 0$, the above quantity $\to 0$.

This estimate clearly fails for $p = 1$. Intuitively, $\mathcal{F}$ is not dense in L^1 because $\hat{g}(0)$ is the total integral over $\mathbb{R}^d$ of g. Functions with mean zero are certainly not dense in L^1.

For $1 < p < \infty$, we have a new dense subset $\mathcal{F}$. So

$$\|f\|_{L^q} = \sup_{g \in \mathcal{F} : \|g\|_{L^{q'}} = 1} \left\langle (2\pi|\xi|)^s \hat{f}(\xi), (2\pi|\xi|)^{-s} \hat{g}(\xi) \right\rangle.$$

The functions $(2\pi|\xi|)^{-s}\hat{g}(\xi)$ are certainly Schwartz, since $\hat{g}$ is a Schwartz function which vanishes in a neighborhood of 0.

Applying Plancharel again gives

$$\|f\|_{L^q} = \sup_{g \in \mathcal{F} : \|g\|_{L^{q'}} = 1} \left\langle |\nabla|^s f, |\nabla|^{-s} g \right\rangle \lesssim \sup_{g \in \mathcal{F} : \|g\|_{L^{q'}} = 1} \||\nabla|^s f\|_{L^p} \left\||\nabla|^{-s} g\right\|_{L^{p'}}.$$

We then have

$$\left(|\nabla|^{-s} g\right)(x) = \left((2\pi|\xi|)^{-s} \hat{g}(\xi)\right)^{\check{}}(x).$$

[1] In the sense of distributions, i.e., $|\nabla|^s f$ is given by integration against a function in L^p.

Now, recall that if T is a tempered distribution and g is a Schwartz function, then $(T * g)\hat{} = \hat{T} \cdot \hat{g}$. Since $(2\pi|\xi|)^{-s}$ is a tempered distribution and $\hat{g}$ is Schwartz, it follows that

$$\left(|\nabla|^{-s} g\right)(x) = \left(\left(((2\pi|\xi|)^{-s})\check{} * g\right)\hat{}\right)\check{}(x).$$

Our computation from (5.1) tells us that $(2\pi|\xi|)^{-s})\check{} \sim_{d,s} |x|^{s-d}$. Therefore,

$$|\nabla|^{-s} g \sim_{d,s} |x|^{s-d} * g.$$

By the first Hardy-Littlewood-Sobolev inequality,

$$\left\| g * \frac{1}{|x|^{d-s}} \right\|_{L^{p'}} \lesssim \|g\|_{L^{q'}}$$

provided $1 + \frac{1}{p'} = \frac{1}{q'} + \frac{d-s}{d}$. But this is true precisely when

$$1 + 1 - \frac{1}{p} = 1 - \frac{1}{q} + 1 - \frac{s}{d} \quad \iff \quad -\frac{1}{p} = -\frac{1}{q} - \frac{s}{d} \quad \iff \quad \frac{1}{p} = \frac{1}{q} + \frac{s}{d}$$

which holds by assumption.

Therefore,

$$\|f\|_{L^q} \lesssim \sup_{g \in \mathcal{F} : \|g\|_{L^{q'}} = 1} \||\nabla|^s f\|_{L^p} \left\||\nabla|^{-s} g\right\|_{L^{p'}} \lesssim \sup_{g \in \mathcal{F} : \|g\|_{L^{q'}} = 1} \||\nabla|^s f\|_{L^p} \|g\|_{L^{q'}} = \||\nabla|^s f\|_{L^p}$$

as desired. □

It will be useful to record the special case of Sobolev embedding when $p = 2$ and $s = 1$.

Corollary 5.6 (Sobolev embedding, $p = 2$)**.** *Fix $d \geq 3$. Suppose that $f \in \mathcal{S}(\mathbb{R}^d)$ with $|\nabla| f \in L^2(\mathbb{R}^d)$. Then*

$$\|f\|_{L^{\frac{2d}{d-2}}} \lesssim \||\nabla| f\|_{L^2}.$$

Proof. By the Sobolev embedding theorem, $\|f\|_{L^q} \lesssim \||\nabla| f\|_{L^2}$ provided $\frac{1}{2} + \frac{1}{q} + \frac{1}{d}$. This means

$$\frac{1}{q} = \frac{d-2}{2d}$$

and so $q = \frac{2d}{d-2}$ as desired. □

Using the yet-unproven fact that $\||\nabla| f\|_{L^2} \lesssim \|\nabla f\|_{L^2}$, this special case of the Sobolev embedding theorem gives the inclusion $\dot{H}^1(\mathbb{R}^d) \subseteq L^{\frac{2d}{d-2}}(\mathbb{R}^d)$.

5.3 The Gagliardo-Nirenberg Inequality

In this section, we use the Sobolev embedding theorem to prove a version of the Gagliardo-Nirenberg inequality. As remarked at the beginning of this chapter, this theorem will be restated and revisited in different contexts after developing more sophisticated tools.

The following estimate is stated for $d \geq 3$, but in fact it holds for $d = 1$ and $d = 2$, where $0 < p < \infty$. The case $d = 1$ is an easy calculus exercise, while the $d = 2$ case is more subtle.

Theorem 5.7 (Gagliardo-Nirenberg)**.** *Fix $d \geq 3$ and $f \in \mathcal{S}(\mathbb{R}^d)$. Then for all $0 < p < \frac{4}{d-2}$,*

$$\|f\|_{p+2}^{p+2} \lesssim \|f\|_2^{p+2-\frac{pd}{2}} \||\nabla| f\|_2^{\frac{pd}{2}}.$$

Proof. For convenience, we recall the log-convexity of L^p norms: for any $0 < \theta < 1$ and $\frac{1}{r} = \frac{1-\theta}{p} + \frac{\theta}{q}$, we have $\|f\|_r \leq \|f\|_p^{1-\theta} \|f\|_q^{\theta}$.

Let $0 < p < \frac{4}{d-2}$ and $f \in \mathcal{S}(\mathbb{R}^d)$. We need to show that

$$\|f\|_{p+2} \leq \|f\|_2^{1-\frac{pd}{2(p+2)}} \||\nabla| f\|_2^{\frac{pd}{2(p+2)}}.$$

Let $\theta = \frac{pd}{2(p+2)}$. Because $p > 0$, $\theta > 0$. Because $0 < p < \frac{4}{d-2}$, $p(d-2) < 4$, hence $pd < 4 + 2p$. Thus,

$$\theta = \frac{pd}{2(p+2)} = \frac{pd}{4+2p} < 1$$

and so $0 < \theta < 1$. Observe that

$$1 - \theta = 1 - \frac{pd}{2(p+2)} = \frac{4+2p-pd}{2(p+2)}.$$

Thus, for any $A, B > 0$,

$$\frac{1-\theta}{A} + \frac{\theta}{B} = \frac{4+2p-pd}{2A(p+2)} + \frac{pd}{2B(p+2)} = \frac{1}{p+2}\left(\frac{4+2p-pd}{2A} + \frac{pd}{2B}\right).$$

With log-convexity of L^p norms in mind, we want to choose A and B so that $\frac{4+2p-pd}{2A} + \frac{pd}{2B} = 1$. Towards our end goal, set $A = 2$. Then

$$\frac{4+2p-pd}{4} + \frac{pd}{2B} = 1 \quad \Rightarrow \quad 1 + \frac{p}{2} - \frac{pd}{4} + \frac{pd}{2B} = 1 \quad \Rightarrow \quad \frac{1}{2} - \frac{d}{4} + \frac{d}{2B} = 0.$$

Solving for B yields

$$\frac{d}{2B} = \frac{d}{4} - \frac{1}{2} = \frac{d-2}{4} \quad \Rightarrow \quad \frac{2B}{d} = \frac{4}{d-2} \quad \Rightarrow \quad B = \frac{2d}{d-2}.$$

These computations demonstrate that

$$\frac{1}{p+2} = \frac{1-\theta}{2} + \frac{\theta}{\frac{2d}{d-2}}.$$

Thus,

$$\|f\|_{p+2} \leq \|f\|_2^{1-\theta} \|f\|_{\frac{2d}{d-2}}^{\theta}.$$

Next, by the $p = 2$ Sobolev embedding theorem, $\|f\|_{\frac{2d}{d-2}} \lesssim \||\nabla| f\|_2$. So

$$\|f\|_{p+2} \lesssim \|f\|_2^{1-\theta} \||\nabla| f\|_2^{\theta} = \|f\|_2^{1-\frac{pd}{2(p+2)}} \||\nabla| f\|_2^{\frac{pd}{2(p+2)}}$$

as desired.

□

Chapter 6

Fourier Multipliers

In Chapter 5, we introduced the operator $|\nabla|^s$. This was our first experience with a nontrivial *Fourier multiplier operator*, i.e., an operator defined by inverting a multiplication operator in the Fourier domain. The operator $|\nabla|^s$ is a Fourier multiplier with *symbol* $(2\pi|\xi|)^s$. We saw other simple examples of when first defining the Fourier transform. For example, the regular partial derivative operator ∂_{x_k} is a Fourier multiplier with symbol $2\pi i\xi_k$; the translation operator $f(\cdot) \mapsto f(\cdot - y)$ is a Fourier multiplier with symbol $e^{2\pi i y\cdot\xi}$, etc. More generally, the Fourier multiplier with symbol $m(\xi)$ is the operator

$$f(x) \mapsto \Big(m(\xi)\hat{f}(\xi)\Big)\check{}\,(x) = \int e^{2\pi i x\cdot\xi} m(\xi)\hat{f}(\xi)\,d\xi.$$

In this chapter, we study the L^p-boundedness of various Fourier multipliers.

6.1 Calderon-Zygmund Convolution Kernels

When the operator $|\nabla|^s$ was introduced in Chapter 5, we remarked that the estimate $\||\nabla|f\|_{L^p} \lesssim \|\nabla f\|_{L^p}$ holds for Schwartz functions f. The usefulness of this estimate is apparent from the fact that $|\nabla|^s$ is a *nonlocal* operator, in that it was defined by multiplication in the Fourier domain and hence by convolution in the physical domain. On the other hand, ∇ is a purely local operator. Recall that $\widehat{|\nabla|f}(\xi) = 2\pi|\xi|\hat{f}(\xi)$. Write

$$2\pi|\xi| = 2\pi\frac{|\xi|^2}{|\xi|} = 2\pi\sum\frac{\xi_j^2}{|\xi|} = \sum -i\frac{\xi_j}{|\xi|}\cdot 2\pi i\xi_j = \sum m_j(\xi)2\pi i\xi_j$$

where $m_j(\xi) := -i\frac{\xi_j}{|\xi|}$ are called **Riesz multipliers**. To show the desired estimate $\||\nabla|f\|_{L^p} \lesssim \|\nabla f\|_{L^p}$, it is enough to prove that the Riesz multipliers are bounded on L^p, that is, the operators R_j defined via $\widehat{R_j f}(\xi) := m_j(\xi)\hat{f}(\xi)$, or equivalently $R_j f = \check{m}_j * f$, are bounded on L^p. Indeed,

$$\widehat{|\nabla|f}(\xi) = 2\pi|\xi|\hat{f}(\xi) = \sum m_j(\xi)2\pi i\xi_j\hat{f}(\xi)$$

so that $|\nabla|f = \sum R_j(\partial_j f)$. By the triangle inequality and the equivalence of norms on finite dimensional spaces,

$$\||\nabla|f\|_{L^p} \le \sum\|R_j(\partial_j f)\|_{L^p} \lesssim \sum\|\partial_j f\|_{L^p} \lesssim \|\nabla f\|_{L^p}.$$

To prove that this is the case, we will prove L^p-boundedness for a more general class of objects. In particular, by passing to distributions we can view any Fourier multiplier as a *convolution kernel*, i.e., an operator defined by $f \mapsto K * f$ for some distribution K. This follows immediately from the fact that the (inverse) Fourier transform turns multiplication into convolution. In this section, we consider a special kind of convolution kernel.

Definition 6.1. A function $K : \mathbb{R}^d \setminus \{0\} \to \mathbb{C}$ is a **Calderon-Zygmund convolution kernel** if it satisfies:

1. $|K(x)| \lesssim \frac{1}{|x|^d}$ uniformly in $x \neq 0$;
2. $\int_{R_1 \leq |x| \leq R_2} K(x)\, dx = 0$ for all $0 < R_1 < R_2 < \infty$;
3. $\int_{|x| \geq 2|y|} |K(x) - K(x+y)|\, dx \lesssim 1$ uniformly in $y \in \mathbb{R}^d$.

The first condition says that K is allowed to have a singularity at 0, as long as the singularity is not too bad. The second condition says that K satisfies a kind of cancellation property; note that if K is an odd function, this condition is satisfied. The third condition is a smoothness requirement.

As a first remark, we claim that if $|\nabla K|(x) \lesssim \frac{1}{|x|^{1+d}}$, then K satisfies condition 3. Indeed, by the fundamental theorem of calculus,

$$K(x) - K(x+y) = -\int_0^1 y \cdot \nabla K(x + \theta y)\, d\theta.$$

So

$$\int_{|x| \geq 2|y|} |K(x) - K(x+y)|\, dx \leq \int_{|x| \geq 2|y|} \int_0^1 |y| \cdot \frac{1}{|x + \theta y|^{1+d}}\, d\theta\, dx.$$

Since $|x| \geq 2|y|$, $|x + \theta y| \gtrsim |x|$. Thus,

$$\begin{aligned}\int_{|x| \geq 2|y|} |K(x) - K(x+y)|\, dx &\lesssim \int_{|x| \geq 2|y|} \int_0^1 |y| \cdot \frac{1}{|x|^{1+d}}\, d\theta\, dx = \int_{|x| \geq 2|y|} |y| \cdot \frac{1}{|x|^{1+d}}\, dx \\ &\lesssim |y| \cdot \frac{1}{|y|} = 1.\end{aligned}$$

Therefore, K satisfies the third as claimed.

Relevant to our motivating discussion, we claim that the kernel of the Riesz multiplier defined above is a Calderon-Zygmund convolution kernels. To see this, first note that

$$K_j := \check{m}_j = \left(-i \frac{\xi_j}{|\xi|}\right)^{\vee} = \left(-\frac{1}{2\pi} \cdot \frac{2\pi i \xi_j}{|\xi|}\right)^{\vee}.$$

In the Sobolev embedding theorem section, we computed the inverse Fourier transform of functions of the form $\frac{1}{|\xi|^s}$. Recalling this computation (see Equation (5.1)) we have

$$K_j(x) \sim_d -\frac{1}{2\pi} \partial_j \left(\frac{1}{|x|^{d-1}}\right) \sim_d \frac{x_j}{|x|^{d+1}}$$

since $\partial_j |x| = \frac{x_j}{|x|}$. Clearly, $|K_j(x)| \lesssim_d \frac{1}{|x|^d}$ and $\int_{R_1 \leq |x| \leq R_2} K_j(x)\, dx = 0$ (since $x \mapsto x_j$ is odd). Finally, we wish to show $|\nabla K_j| \lesssim \frac{1}{|x|^{1+d}}$ to verify the third condition. Note that

$$\begin{aligned}\partial_k K_j(x) \sim \partial_k x_j |x|^{-(d+1)} &= \delta_{jk} |x|^{-(d+1)} - (d+1)|x|^{-(d+1)-1} x_k |x|^{-1} \\ &= \frac{\delta_{jk}}{|x|^{d+1}} - (d-1)\frac{x_k}{|x|^{d+3}},\end{aligned}$$

so certainly $|\nabla K_j| \lesssim \frac{1}{|x|^{1+d}}$. Thus, K_j is a Calderon-Zygmund convolution kernel.

The first result on Calderon-Zygmund convolution kernels is an $L^2 \to L^2$ type estimate.

Theorem 6.2. *Suppose K is a Calderon-Zygmund convolution kernel. For $\varepsilon > 0$, let $K_\varepsilon := K \cdot \chi_{\varepsilon \leq |x| \leq 1/\varepsilon}$. Then $\|K_\varepsilon * f\|_{L^2} \lesssim \|f\|_{L^2}$ uniformly in $\varepsilon > 0$. Moreover, $K * f = \lim_{\varepsilon \to 0} K_\varepsilon * f$ and the operator $K * f$ extends as a bounded operator from $\mathcal{S}(\mathbb{R}^d)$ to $L^2(\mathbb{R}^d)$.*

Proof. First, we will show that K_ε is a Calderon-Zygmund convolution kernel. Because K_ε is just a restriction of K, conditions 1 and 2 are immediate. We only need to verify that $\int_{|x|\geq 2|y|} K_\varepsilon(x) - K_\varepsilon(x+y)|\, dy \leq 1$ uniformly in $y \in \mathbb{R}^d$ and $\varepsilon > 0$.

To do this, consider the following three subregions of $|x| \geq 2|y|$:

$$A = \{|x| \geq 2|y|,\, \varepsilon \leq |x|, |x+y| \leq 1/\varepsilon\}$$
$$B = \{|x| \geq 2|y|,\, \varepsilon \leq |x| \leq 1/\varepsilon,\, |x+y| < \varepsilon \text{ or } |x+y| > 1/\varepsilon\}$$
$$C = \{|x| \geq 2|y|,\, \varepsilon \leq |x+y| \leq 1/\varepsilon,\, |x| < \varepsilon \text{ or } |x| > 1/\varepsilon\}.$$

Then

$$\int_{|x|\geq 2|y|} |K_\varepsilon(x) - K_\varepsilon(x+y)|\, dy \leq \int_A |K(x) - K(x+y)|\, dy + \int_B |K(x)|\, dy + \int_C |K(x+y)|\, dy.$$

The integral over A is $\lesssim 1$ because K is a Calderon-Zygmund convolution kernel.

Consider the integral over B. Assume first that $|x+y| < \varepsilon$. Then $|x| \leq |x+y| + |y| < \varepsilon + \frac{|x|}{2}$, hence $|x| < 2\varepsilon$. The contribution of this part of the integral is bounded by

$$\int_{\varepsilon \leq |x| \leq 2\varepsilon} |K(x)|\, dx \leq \int_{\varepsilon \leq |x| \leq 2\varepsilon} \frac{1}{|x|^d}\, dx \lesssim 1.$$

If $|x+y| > 1/\varepsilon$, then $|x| \geq |x+y| - |y| > 1/\varepsilon - |x|/2$, hence $|x| \geq \frac{2}{3\varepsilon}$. The contribution of this part of the integral is bounded by

$$\int_{\frac{2}{3\varepsilon} \leq |x| \leq \frac{1}{\varepsilon}} |K(x)|\, dx \leq \int_{\frac{2}{3\varepsilon} \leq |x| \leq \frac{1}{\varepsilon}} \frac{1}{|x|^d}\, dx \lesssim 1.$$

Thus, $\int_B |K(x)|\, dy \lesssim 1$.

Bounding the integral over C is essentially the same argument, paired with a change of variables. Thus, $\int_C |K(x+y)|\, dy \lesssim 1$. This verifies the third condition, hence K_ε is a Calderon-Zygmund convolution kernel.

Next, we want to show $\|K_\varepsilon * f\|_{L^2} \lesssim \|f\|_{L^2}$ uniformly in $\varepsilon > 0$. By Plancharel, we have

$$\|K_\varepsilon * f\|_{L^2} = \left\|\widehat{K_\varepsilon * f}\right\|_{L^2} = \left\|\hat{K_\varepsilon} \cdot \hat{f}\right\|_{L^2} \leq \left\|\hat{K_\varepsilon}\right\|_{L^\infty} \left\|\hat{f}\right\|_{L^2} \left\|\hat{K_\varepsilon}\right\|_{L^\infty} \|f\|_{L^2}.$$

Thus, it suffices to prove that $\left\|\hat{K_\varepsilon}\right\|_{L^\infty} \lesssim 1$ uniformly in ε.

We will decompose the Fourier transform of K_ε into an integral without any oscillation and an integral with oscillation. Explicitly,

$$\hat{K_\varepsilon}(\xi) = \int e^{-2\pi i x\cdot\xi} K_\varepsilon(x)\, dx = \int_{|x|\leq 1/|\xi|} e^{-2\pi i x\cdot\xi} K_\varepsilon(x)\, dx + \int_{|x|\geq 1/|\xi|} e^{-2\pi i x\cdot\xi} K_\varepsilon(x)\, dx.$$

In the first integral, $|x\cdot\xi| \leq |x||\xi| \leq 1$, thus the term $e^{-2\pi i x\cdot\xi}$ is restricted in its oscillation. Using the cancellation property of Calderon-Zygmund convolution kernels, we have $\int_{|x|\leq 1/|\xi|} K_\varepsilon(x)\, dx = \int_{\varepsilon\leq|x|\leq 1/|\xi|} K(x)\, dx = 0$. Thus,

$$\left|\int_{|x|\leq 1/|\xi|} e^{-2\pi i x\cdot\xi} K_\varepsilon(x)\, dx\right| = \left|\int_{|x|\leq 1/|\xi|} \left(e^{-2\pi i x\cdot\xi} - 1\right) K_\varepsilon(x)\, dx\right| \leq \int_{|x|\leq 1/|\xi|} |e^{-2\pi i x\cdot\xi} - 1||K_\varepsilon(x)|\, dx.$$

Because K_ε is a Calderon-Zygmund convolution kernel, $|K_\varepsilon(x)| \lesssim |x|^{-d}$. The quantity $|e^{-2\pi x\cdot\xi} - 1|$ is the distance from the point $e^{-2\pi i x\cdot\xi}$ of the unit circle to 1. This is bounded by the arc-length of that segment on the unit circle corresponds to the phase, which is $|x\cdot\xi| \le |x||\xi|$. Therefore,

$$\left|\int_{|x|\le 1/|\xi|} e^{-2\pi i x\cdot\xi} K_\varepsilon(x)\,dx\right| \le \int_{\varepsilon\le|x|\le 1/|\xi|} |x||\xi||x|^{-d}\,dx \lesssim |\xi|\cdot\frac{1}{|\xi|} = 1.$$

Next, consider the second integral. Here we invoke the smoothness condition of Calderon-Zygmund convolution kernels. Write

$$1 = \frac{1-e^{i\pi}}{2} = \frac{1-e^{2\pi i\frac{\xi\cdot\xi}{2|\xi|^2}}}{2}.$$

Then

$$\begin{aligned}
\int_{|x|>1/|\xi|} e^{-2\pi i x\cdot\xi} K_\varepsilon(x)\,dx &= \int_{|x|>1/|\xi|} e^{-2\pi i x\cdot\xi}\left(\frac{1-e^{2\pi i\frac{\xi\cdot\xi}{2|\xi|^2}}}{2}\right) K_\varepsilon(x)\,dx \\
&= \frac{1}{2}\int_{|x|>1/|\xi|}\left(e^{-2\pi i x\cdot\xi} - e^{-2\pi i\xi\cdot\left(x-\frac{\xi}{2|\xi|^2}\right)}\right) K_\varepsilon(x)\,dx \\
&= \frac{1}{2}\int_{|x|>1/|\xi|} e^{-2\pi i x\cdot\xi} K_\varepsilon(x)\,dx \\
&\qquad - \frac{1}{2}\int_{\left|x+\frac{\xi}{2|\xi|^2}\right|>1/|\xi|} e^{-2\pi i x\cdot\xi} K_\varepsilon\left(x+\frac{\xi}{2|\xi|^2}\right)\,dx
\end{aligned}$$

We want to combine these integrals, but have to account for the shifted circles which define the domains; it helps to draw a picture here. We have:

$$\begin{aligned}
&= \frac{1}{2}\int_{|x|>1/|\xi|} e^{-2\pi i x\cdot\xi}\left(K_\varepsilon(x) - K_\varepsilon\left(x+\frac{\xi}{2|\xi|^2}\right)\right)\,dx \\
&\quad + \frac{1}{2}\int_A e^{-2\pi i x\cdot\xi} K_\varepsilon\left(x+\frac{\xi}{2|\xi|^2}\right)\,dx - \frac{1}{2}\int_B e^{-2\pi i x\cdot\xi} K_\varepsilon\left(x+\frac{\xi}{2|\xi|^2}\right)\,dx
\end{aligned}$$

where

$$\begin{aligned}
A &= \left\{x \,:\, |x| \ge \frac{1}{|\xi|} \ge \left|x+\frac{\xi}{2|\xi|^2}\right|\right\}, \\
B &= \left\{x \,:\, |x| \le \frac{1}{|\xi|} \le \left|x+\frac{\xi}{2|\xi|^2}\right|\right\}.
\end{aligned}$$

Estimate the first of these three integrals by employing the smoothness condition of the Calderon-Zygmund convolution kernels:

$$\begin{aligned}
&\left|\frac{1}{2}\int_{|x|>1/|\xi|} e^{-2\pi i x\cdot\xi}\left(K_\varepsilon(x) - K_\varepsilon\left(x+\frac{\xi}{2|\xi|^2}\right)\right)\,dx\right| \\
&\qquad\qquad \lesssim \int_{|x|>2\left|\frac{\xi}{2|\xi|^2}\right|}\left|K_\varepsilon(x) - K_\varepsilon\left(x+\frac{\xi}{2|\xi|^2}\right)\right|\,dx \lesssim 1.
\end{aligned}$$

Next, consider the integral over A. By the reverse triangle inequality,

$$|x| \ge \frac{1}{|\xi|} \ge \left|x+\frac{\xi}{2|\xi|^2}\right| \qquad \Rightarrow \qquad \left|x+\frac{\xi}{2|\xi|^2}\right| \ge |x| - \frac{1}{2|\xi|} \ge \frac{1}{|\xi|}.$$

Thus,

$$\begin{aligned}
\left|\frac{1}{2}\int_A e^{-2\pi i x\cdot\xi} K_\varepsilon\left(x+\frac{\xi}{2|\xi|^2}\right)dx\right| &\lesssim \int_{\frac{1}{2|\xi|}\le\left|x+\frac{\xi}{2|\xi|^2}\right|\le\frac{1}{|\xi|}} \left|K_\varepsilon\left(x+\frac{\xi}{2|\xi|^2}\right)\right| dx \\
&\lesssim \int_{\frac{1}{2|\xi|}\le|x|\le\frac{1}{|\xi|}} \frac{1}{|x|^d}\,dx \\
&\lesssim 1
\end{aligned}$$

where we have made the obvious change of variables and invoked the first property of Calderon-Zygmund convolution kernels. Use a similar trick for bounding the integral over B, noting that

$$|x| \le \frac{1}{|\xi|} \le \left|x+\frac{\xi}{2|\xi|^2}\right| \qquad \Rightarrow \qquad \left|x+\frac{\xi}{2|\xi|^2}\right| \le |x| + \frac{1}{2|\xi|} \le \frac{3}{2|\xi|}.$$

Combining all of this gives $\left|\int_{|x|>1/|\xi|} e^{-2\pi i x\cdot\xi} K_\varepsilon(x)\,dx\right| \lesssim 1$, and hence $|\hat{K}_\varepsilon(\xi)| \lesssim 1$ uniformly in ε and ξ, thus, $\left\|\hat{K}_\varepsilon(\xi)\right\|_{L^\infty} \lesssim 1$ uniformly in ε. Therefore, $\|K_\varepsilon * f\|_{L^2} \lesssim \|f\|_{L^2}$.

Next, we wish to show that $K_\varepsilon * f$ converges in L^2 as $\varepsilon \to 0$. It suffices to show that $\{K_\varepsilon * f\}$ is Cauchy. So fix $0 < \varepsilon_2 < \varepsilon_1$ and fix $f \in \mathcal{S}(\mathbb{R}^d)$. Then

$$\begin{aligned}
(K_{\varepsilon_1} * f - K_{\varepsilon_2} * f)(x) &= \int_{\varepsilon_1\le|y|\le1/\varepsilon_1} K(y)f(x-y)\,dy - \int_{\varepsilon_2\le|y|\le1/\varepsilon_2} K(y)f(x-y)\,dy \\
&= -\int_{\varepsilon_2\le|y|\le\varepsilon_1} K(y)f(x-y)\,dy - \int_{1/\varepsilon_1\le|y|\le1/\varepsilon_2} K(y)f(x-y)\,dy.
\end{aligned}$$

Consider the second integral.

$$\begin{aligned}
\left\|\int_{1/\varepsilon_1\le|y|\le1/\varepsilon_2} K(y)f(x-y)\,dy\right\|_{L^2} &\lesssim \left\|K\cdot\chi_{1/\varepsilon_1\le|y|\le1/\varepsilon_2}\right\|_{L^2}\|f\|_{L^1} \\
&\lesssim_f \left(\int_{1/\varepsilon_1\le|y|\le1/\varepsilon_2} \frac{1}{|x|^{2d}}\,dx\right)^{\frac{1}{2}} \\
&\lesssim \varepsilon_1^{\frac{d}{2}}
\end{aligned}$$

which $\to 0$ as $\varepsilon_1 \to 0$. For the first integral, the cancellation property of K gives

$$\begin{aligned}
\int_{\varepsilon_2\le|y|\le\varepsilon_1} K(y)f(x-y)\,dy &= \int_{\varepsilon_2\le|y|\le\varepsilon_1} K(y)(f(x-y)-f(x))\,dy \\
&= \int_{\varepsilon_2\le|y|\le\varepsilon_1} K(y)\,y\int_0^1 \nabla f(x-\theta y)\,d\theta\,dy.
\end{aligned}$$

Thus,

$$\begin{aligned}
\left\|\int_{\varepsilon_2\le|y|\le\varepsilon_1} K(y)f(x-y)\,dy\right\|_{L^2_x} &\lesssim \int_0^1\int_{\varepsilon_2\le|y|\le\varepsilon_1} \frac{1}{|y|^d}|y|\,\|\nabla f(x-\theta y)\|_{L^2_x}\,dy\,d\theta \\
&\lesssim_f \int_{\varepsilon_2\le|y|\le\varepsilon_1} \frac{1}{|y|^{d-1}}\,dy \le \varepsilon_1 \\
&\to 0
\end{aligned}$$

as $\varepsilon_1 \to 0$.

This shows that for a Schwartz function f, $\{K_\varepsilon * f\}$ is Cauchy. By density, $\{K_\varepsilon * f\}$ is Cauchy for any $f \in L^2$. To see this argument explicitly, fix $\delta > 0$ and let $g \in \mathcal{S}(\mathbb{R}^d)$ be such that $\|f - g\|_{L^2} < \delta$. Then by the usual triangle inequality argument,

$$\|K_{\varepsilon_1} * f - K_{\varepsilon_2} * f\|_{L^2} \le \|K_{\varepsilon_1} * g - K_{\varepsilon_2} * g\|_{L^2} + \|K_{\varepsilon_1} * (f-g)\|_{L^2} + \|K_{\varepsilon_2} * (f-g)\|_{L^2}.$$

The first quantity $\to 0$ as $\varepsilon_1, \varepsilon_2 \to 0$ since g is Schwartz. We have already seen that Calderon-Zygmund convolution kernels are of type $(2,2)$, thus,

$$\|K_{\varepsilon_1} * (f-g)\|_{L^2} \lesssim \|f-g\|_{L^2} \le \delta$$

and likewise for K_{ε_2}. Therefore, taking $\varepsilon_1, \varepsilon_2 \to 0$ and then taking $\delta \to 0$ gives the desired result.

Therefore, we can define $K * f := L^2 - \lim_{\varepsilon \to 0} K_\varepsilon * f$. This operator extends to L^2 by uniform boundedness of the K_ε operators. $\square$

This theorem says that Calderon-Zygmund convolution kernels are operators of type $(2,2)$. Using this fact together with interpolation gives a wider range of estimates.

Theorem 6.3. *Suppose that K is a Calderon-Zygmund convolution kernel. For $\varepsilon > 0$, let $K_\varepsilon := K \cdot \chi_{\varepsilon \le |x| \le 1/\varepsilon}$. Then*

1. *K is of weak-type $(1,1)$, i.e.,*

$$\left|\left\{ x \in \mathbb{R}^d \ : \ |K_\varepsilon * f| > \lambda \right\}\right| \lesssim \frac{1}{\lambda} \|f\|_{L^1(\mathbb{R}^d)}$$

uniformly in $\lambda > 0$ and $\varepsilon > 0$.

2. *K_ε is of (strong) type (p,p) for $1 < p < \infty$, i.e.,*

$$\|K_\varepsilon * f\|_{L^p(\mathbb{R}^d)} \lesssim \|f\|_{L^p(\mathbb{R}^d)}$$

uniformly in $\varepsilon > 0$.

Moreover, for $1 < p < \infty$, $K * f := L^p - \lim_{\varepsilon \to 0} K_\varepsilon * f$ extends as a bounded map on $\mathcal{S}(\mathbb{R}^d)$ to a bounded map on $L^p(\mathbb{R}^d)$.

Proof. Assume first that we have proven 1, that K_ε is of weak-type $(1,1)$. Since K_ε is of type $(2,2)$ by the previous theorem, it follows from the Marcinkiewicz interpolation theorem that K_ε is of type (p,p) for $1 < p < 2$. We then use a duality argument to achieve the estimate for $2 < p < \infty$. Explicitly, for p in this range,

$$\begin{aligned}
\|K_\varepsilon * f\|_{L^p} &= \sup_{\|g\|_{L^{p'}}=1} \langle K_\varepsilon * f, g\rangle = \sup_{\|g\|_{L^{p'}}=1} \int\int K_\varepsilon(x-y) f(y)\, \overline{g(x)}\, dy\, dx \\
&= \sup_{\|g\|_{L^{p'}}=1} \int\int \overline{\overline{(K_\varepsilon)_R(y-x)} g(x)}\, dx\, f(y)\, dy \\
&= \sup_{\|g\|_{L^{p'}}=1} \left\langle f, \overline{(K_\varepsilon)_R} * g \right\rangle
\end{aligned}$$

where $(K_\varepsilon)_R(x) = K_\varepsilon(-x)$ is the reflection of K_ε. By Holder's inequality,

$$\|K_\varepsilon * f\|_{L^p} \le \sup_{\|g\|_{L^{p'}}=1} \|f\|_{L^p} \left\|\overline{(K_\varepsilon)_R} * g\right\|_{L^{p'}}.$$

Note that $\overline{(K_\varepsilon)_R}$ is also a Calderon-Zygmund convolution kernel. Since $1 < p' < 2$, we then have $\left\|\overline{(K_\varepsilon)_R} * g\right\|_{L^{p'}} \lesssim \|g\|_{L^{p'}} = 1$, so that $\|K_\varepsilon * f\|_{L^p} \lesssim \|f\|_{L^p}$ as desired.

For the "moreover" statement, we wish to show that $\|K_{\varepsilon_1} * f - K_{\varepsilon_2} * f\|_{L^p} \to 0$ as $\varepsilon_1, \varepsilon_2 \to 0$ as in the previous theorem. Though the same proof technique as before will work, we present an alternative argument here. For $1 < p < 2$, choose $1 < q < p$ and compute, using the log-concavity of L^p-norms:

$$\|K_{\varepsilon_1} * f - K_{\varepsilon_2} * f\|_{L^p} \lesssim \|K_{\varepsilon_1} * f - K_{\varepsilon_2} * f\|_{L^q}^{\theta} \|K_{\varepsilon_1} * f - K_{\varepsilon_2} * f\|_{L^2}^{1-\theta}$$

where $\frac{1}{p} = \frac{\theta}{q} + \frac{1-\theta}{2}$. We have $\|K_{\varepsilon_1} * f - K_{\varepsilon_2} * f\|_{L^q}^{\theta} \lesssim \|f\|_{L^q}^{\theta}$ and $\|K_{\varepsilon_1} * f - K_{\varepsilon_2} * f\|_{L^2}^{1-\theta} \to 0$ from previous results. For $2 < p < \infty$, choose $p < q < \infty$ and argue similarly.

Thus, to prove the entire theorem it only remains to show 1. Let $f \in L^1(\mathbb{R}^d)$. We perform a modified version of the Calderon-Zygmund decomposition introduced earlier. For a fixed $\lambda > 0$, write $f = g + b$ where:

- $|g| \lesssim \lambda$ a.e.;
- b is supported on a union of cubes $\{Q_k\}$ whose interiors are disjoint and $|\bigcup Q_k| \lesssim \frac{1}{\lambda} \int |f(y)|\, dy$;
- $b\,|_{Q_k} = f\,|_{Q_k} - \frac{1}{|Q_k|} \int_{Q_k} f(y)\, dy$, so that consequently $\int_{Q_k} b(y)\, dy = 0$;
- $\frac{1}{|Q_k|} \int_{Q_k} |b(y)|\, dy \lesssim \lambda$.

Then

$$|\{\, x \,:\, |K_\varepsilon * f| > \lambda \,\}| \leq \left|\left\{\, x \,:\, |K_\varepsilon * g| > \frac{\lambda}{2} \,\right\}\right| + \left|\left\{\, x \,:\, |K_\varepsilon * b| > \frac{\lambda}{2} \,\right\}\right|.$$

We estimate the first quantity on the right using the $(2,2)$ estimate for K_ε. Using Chebychev's inequality and the uniform bound on g, we have:

$$\begin{aligned} \left|\left\{\, x \,:\, |K_\varepsilon * g| > \frac{\lambda}{2} \,\right\}\right| &\lesssim \lambda^{-2} \|K_\varepsilon * g\|_{L^2}^2 \\ &\lesssim \lambda^{-2} \|g\|_{L^2}^2 = \lambda^{-2} \left\||g|^2\right\|_{L^1} \lesssim \lambda^{-1} \|g\|_{L^1} \leq \frac{1}{\lambda} \|f\|_{L^1}. \end{aligned}$$

To estimate the second quantity, we first remove a set of comparable size to $\frac{1}{\lambda}\|f\|_{L^1}$. Let Q_k^* be the cube with same center, call it x_k, as Q_k but with side lengths $2\sqrt{d}\ell(Q_k)$. (To draw a picture, draw Q_k and circumscribe a circle; draw a circle with twice the radius, then circumscribe a cube around this circle.) Then

$$\left|\bigcup Q_k^*\right| \leq \sum |Q_k| \leq (2\sqrt{d})^d \sum |Q_k| \lesssim_d \frac{1}{\lambda} \|f\|_{L^1}.$$

Thus, it suffices to estimate the quantity $\left|\left\{\, x \notin \bigcup Q_k^* \,:\, |K_\varepsilon * b| > \frac{\lambda}{2} \,\right\}\right|$. By Chebychev, we have

$$\left|\left\{\, x \notin \bigcup Q_k^* \,:\, |K_\varepsilon * b| > \frac{\lambda}{2} \,\right\}\right| \lesssim \frac{1}{\lambda} \|K_\varepsilon * b\|_{L^1((\bigcup Q_k^*)^C)}.$$

Writing out the definition of the convolution above gives

$$\begin{aligned} (K_\varepsilon * b)(x) &= \int K_\varepsilon(x-y) b(y)\, dy = \sum \int_{Q_k} K_\varepsilon(x-y) b(y)\, dy \\ &= \sum \int_{Q_k} \left(K_\varepsilon(x-y) - K_\varepsilon(x-x_k)\right) b(y)\, dy \end{aligned}$$

where the last equality follows from the fact that b has mean 0. So

$$\begin{aligned}\|K_\varepsilon * b\|_{L^1((\bigcup Q_k^*)^C)} &= \int_{(\bigcup Q_k^*)^C} |K_\varepsilon * b|(x)\,dx \\ &\leq \sum \int_{(Q_k^*)^C} \int_{Q_k} |K_\varepsilon(x-y) - K_\varepsilon(x-x_k)|\ |b(y)|\,dy\,dx \\ &= \sum \int_{Q_k} |b(y)| \int_{(Q_k^*)^C} |K_\varepsilon(x-y) - K_\varepsilon(x-x_k)|\ dx\,dy.\end{aligned}$$

Making the change of variables $x - x_k \mapsto x$ yields

$$\|K_\varepsilon * b\|_{L^1((\bigcup Q_k^*)^C)} = \sum \int_{Q_k} |b(y)| \int_{(Q_k^*)^C - x_k} |K_\varepsilon(x + x_k - y) - K_\varepsilon(x)|\ dx\,dy$$

where $(Q_k^*)^C - x_k$ denotes the translation of $(Q_k^*)^C$ by the point x_k. Note that if $x \in (Q_k^*)^C - x_k$, then $|x| > 2\sqrt{d}\ell(Q_k)$. Since $\sqrt{d}\ell(Q_k)$ is the diameter of Q_k and $y \in Q_k$, this implies that $|x| \geq 2|y - x_k|$. By the the smoothness property of Calderon-Zygmund convolution kernels,

$$\int_{(Q_k^*)^C - x_k} |K_\varepsilon(x + x_k - y) - K_\varepsilon(x)|\ dx \lesssim 1.$$

Thus,

$$\|K_\varepsilon * b\|_{L^1((\bigcup Q_k^*)^C)} \lesssim \sum \int_{Q_k} |b(y)|\,dy \lesssim \int |f(y)|\,dy$$

and we are done. □

In the proof of this theorem, *having already shown* $(2, 2)$ *boundedness of* K, the only property of Calderon-Zygmund convolution kernels that we explicitly used was the smoothness condition. Interpolation was enough to get boundedness for other p values. This gives us a more general fact.

Proposition 6.4. *If K is any convolution kernel (not necessarily Calderon-Zygmund) which is of type $(2, 2)$ and satisfies $\int_{|x|\geq 2|y|} |K(x+y) - K(x)|\,dx \lesssim 1$ uniformly in y, then K extends to a bounded map on $L^p(\mathbb{R}^d)$ for $1 < p < \infty$.*

6.2 The Hilbert Transform

As a first application of L^p-boundedness of Calderon-Zygmund convolution kernels, we discuss the Hilbert transform. The Hilbert transform is the marquee example of a singular integral operator, and is important in fields such as signal processing.

Definition 6.5. The **Hilbert transform** is the convolution operator corresponding to the kernel $K : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ given by $K(x) = \frac{1}{\pi x}$. Explicitly, the Hilbert transform of a function $f : \mathbb{R} \to \mathbb{R}$ is:

$$(Hf)(x) := \lim_{\varepsilon \to 0} \int_{\varepsilon \leq |y| \leq \frac{1}{\varepsilon}} \frac{1}{\pi y} f(x-y)\,dy.$$

A priori, it is not clear that the Hilbert transform is well-defined for even nice functions f. However, we claim that K is a Calderon-Zygmund convolution kernel. The estimate $|K(x)| \lesssim \frac{1}{|x|}$ is immediate from the definition, and the fact that K satisfies the cancellation property over annuli in $\mathbb{R}$ follows from the fact that $\frac{1}{x}$ is an odd function. Checking the smoothness condition, we have

$$\int_{|x|\geq 2|y|} |K(x+y) - K(x)|\,dx = \frac{1}{\pi} \int_{|x|\geq 2|y|} \left| \frac{1}{x+y} - \frac{1}{x} \right| dx = \frac{1}{\pi} \int_{|x|\geq 2|y|} \frac{|y|}{|x+y||x|}\,dx.$$

Since $|x| \geq 2|y|$, $|x+y| \geq |x| - |y| \gtrsim |x|$, so that

$$\int_{|x|\geq 2|y|} |K(x+y) - K(x)|\, dx \lesssim \int_{|x|\geq |y|} \frac{|y|}{|x|^2}\, dx \lesssim 1.$$

Thus, K is a Calderon-Zygmund convolution kernel. By the results from the previous section, this proves the following.

Proposition 6.6. *The Hilbert transform is a bounded operator from $L^p(\mathbb{R}) \to L^p(\mathbb{R})$ for $1 < p < \infty$.*

It is worth nothing that the strong-type estimates do indeed fail for $p = 1$ and $p = \infty$. For example, let $f = \chi_{[a,b]}$. Then $f \in L^1(\mathbb{R})$ and $f \in L^\infty(\mathbb{R})$, and

$$(Hf)(x) = \frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{\varepsilon \leq |y| \leq \frac{1}{\varepsilon}} \frac{1}{y} \chi_{[a,b]}(x-y)\, dy.$$

Consider the case where $x - a, x + b > 0$. Then for ε sufficiently small,

$$(Hf)(x) = \frac{1}{\pi} \int_{x-a}^{x-b} \frac{1}{y}\, dy = \frac{1}{\pi} \log \left| \frac{x-a}{x-b} \right|$$

for almost all x, which is neither in $L^1(\mathbb{R})$ nor $L^\infty(\mathbb{R})$. The cases $x - a, x - b < 0$ and $x - a < 0 < x - b$ are handled similarly.

It can be shown that the Hilbert transform is a Fourier multiplier with symbol $-i \operatorname{sgn}(\xi)$. That is, for $f \in \mathcal{S}(\mathbb{R})$,

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

As such, the Riesz multipliers $\widehat{R_j f}(\xi) = -i \frac{\xi_j}{|\xi|} \hat{f}(\xi)$ defined earlier in this chapter can be viewed as a higher dimensional analogue of the Hilbert transform.

6.3 The Mikhlin Multiplier Theorem

In this section we prove a multiplier theorem which gives L^p-boundedness of a large class of multipliers. This particular theorem, and its proof technique, offers a segue into the Littlewood-Paley theory developed in the next chapter.

The main result is the following.

Theorem 6.7 (Mikhlin multiplier theorem)**.** *Let $m : \mathbb{R}^d \setminus \{0\} \to \mathbb{C}$ satisfy*

$$\left| D_\xi^\alpha m(\xi) \right| \lesssim \frac{1}{|\xi|^{|\alpha|}} \tag{6.1}$$

uniformly in ξ for all $0 \leq |\alpha| \leq \lceil \frac{d+1}{2} \rceil$. Then $f \mapsto (m \cdot \hat{f})\check{} = \check{m} * f$ is bounded on $L^p(\mathbb{R}^d)$ for $1 < p < \infty$.

To prove this theorem, it will useful to localize at different frequencies in the Fourier domain. With this in mind, we introduce a particular family of bump functions. Let $\varphi \in C^\infty(\mathbb{R}^d)$ be a smooth bump function satisfying

$$\varphi(x) = \begin{cases} 1 & |x| \leq 1 \\ 0 & |x| \geq \frac{11}{10} \end{cases}.$$

Let $\psi(x) = \varphi(x) - \varphi(2x)$. Then

$$\psi(x) = \begin{cases} 0 & |x| \leq \frac{1}{2} \\ 1 & \frac{11}{20} \leq |x| \leq 1 \\ 0 & |x| \geq \frac{11}{10} \end{cases}.$$

For $N \in 2^{\mathbb{Z}}$, define $\psi_N(x) = \psi\left(\frac{x}{N}\right)$. These annular bump functions which are ~ 1 on a scale of N will be used in the proof of the Theorem 6.7, and are central to Littlewood-Paley theory.

Proof of Theorem 6.7. For $p = 2$, the desired conclusion follows from Plancharel. Indeed,

$$\|\check{m} * f\|_{L^2_x} = \left\|m \cdot \hat{f}\right\|_{L^2_\xi} \leq \|m\|_{L^\infty_\xi} \left\|\hat{f}\right\|_{L^2_\xi}.$$

By choosing $\alpha = 0$ in (6.1) it follows that m is uniformly bounded, so $\|m\|_{L^\infty_\xi} < \infty$. Thus, $\|\check{m} * f\|_{L^2_x} \lesssim \left\|\hat{f}\right\|_{L^2_\xi}$.

With this $(2,2)$ estimate, by Proposition 6.4, to prove boundedness on L^p for $1 < p < \infty$ it suffices to prove that $K = \check{m}$ satisfies the smoothness condition of Calderon-Zygmund convolution kernels. Before doing this with the hypotheses of the theorem, we prove a slightly easier case. In particular, suppose that (6.1) holds for all $|\alpha| \leq d+2$ rather than all $|\alpha| \leq \left\lceil \frac{d+1}{2} \right\rceil$. In this case, we will show that $|\nabla K(x)| \lesssim \frac{1}{|x|^{d+1}}$. By previous remarks, this implies that K satisfies the desired smoothness condition. Towards this goal, note that

$$\sum_{N \in 2^{\mathbb{Z}}} \psi_N(x) = 1$$

for almost all $x \in \mathbb{R}^d$. Write

$$m(\xi) = \sum_{N \in 2^{\mathbb{Z}}} m(\xi)\psi_N(\xi) =: \sum_{N \in 2^{\mathbb{Z}}} m_N(\xi).$$

Then $|\nabla K(x)| \leq \sum_{N \in 2^{\mathbb{Z}}} |\nabla \check{m}_N(x)|$. By the properties of the Fourier transform,

$$\|x^\alpha \nabla \check{m}_N(x)\|_{L^\infty_x} \lesssim \left\|D^\alpha_\xi \xi m_N(\xi)\right\|_{L^1_\xi}.$$

By the product rule,

$$D^\alpha_\xi \left(\xi m_N(\xi)\right) = D^\alpha_\xi \left(\xi m(\xi)\psi_N(\xi)\right) = \sum_{\alpha_1+\alpha_2=\alpha} C_{\alpha_1,\alpha_2}\, D^{\alpha_1}_\xi(\xi m(\xi))\, D^{\alpha_2}_\xi(\psi_N(\xi)).$$

By assumption, $|D^{\alpha_1}_\xi m(\xi)| \lesssim |\xi|^{-|\alpha_1|}$. Therefore, $|D^{\alpha_1}_\xi(\xi m(\xi))| \lesssim |\xi|^{1-|\alpha_1|}$, again by the product rule. Next, by the chain rule,

$$|D^{\alpha_2}_\xi \psi_N(\xi)| = \left|D^{\alpha_2}_\xi\left(\psi\left(\frac{\xi}{N}\right)\right)\right| = N^{-|\alpha_2|}\left|(D^{\alpha_2}_\xi \psi)\left(\frac{\xi}{N}\right)\right|.$$

Thus,

$$|D^\alpha_\xi \xi m_N(\xi)| \lesssim \sum_{\alpha_1+\alpha_2=\alpha} |\xi|^{1-|\alpha_1|} N^{-|\alpha_2|}\left|(D^{\alpha_2}_\xi \psi)\left(\frac{\xi}{N}\right)\right|.$$

Since ψ_N and all of its derivatives are supported on an annulus of radius comparable to N,

$$\begin{aligned}\|x^\alpha \nabla \check{m}_N(x)\|_{L^\infty_x} \lesssim \left\|D^\alpha_\xi \xi m_N(\xi)\right\|_{L^1_\xi} &\lesssim \sum_{\alpha_1+\alpha_2=\alpha} \int_{|\xi|\sim N} |\xi|^{1-|\alpha_1|} N^{-|\alpha_2|}\, d\xi \\ &\lesssim \sum_{\alpha_1+\alpha_2=\alpha} N^{1+d-|\alpha_1|} N^{-|\alpha_2|} \\ &\lesssim N^{1+d-|\alpha|}.\end{aligned}$$

By choosing $\alpha = 0$ and $|\alpha| = d + 2$, it then follows that

$$|\nabla \check{m}_N(x)| \lesssim \min\left\{N^{d+1}, \frac{1}{N|x|^{d+2}}\right\}$$

so that

$$|\nabla K(x)| \leq \sum_{N\in 2^{\mathbb{Z}}} |\nabla \check{m}_N(x)| \lesssim \sum_{N\in 2^{\mathbb{Z}}} \min\left\{N^{d+1}, \frac{1}{N|x|^{d+2}}\right\}.$$

Splitting this sum over small and large frequencies, chosen appropriately,

$$|\nabla K(x)| \lesssim \sum_{N\leq \frac{1}{|x|}} N^{d+1} + \sum_{N>\frac{1}{|x|}} \frac{1}{N|x|^{d+2}}.$$

Since both of these sums are over dyadic numbers, the first sum is bounded (up to a constant) by its largest term and the second sum is bounded (up to a constant) by its smallest term. Thus,

$$|\nabla K(x)| \lesssim \frac{1}{|x|^{d+1}} + \frac{1}{\frac{1}{|x|}|x|^{d+2}} \lesssim \frac{1}{|x|^{d+1}}.$$

By our initial remark, we are done.

Next, we return to the original assumption that $|\alpha| \leq \lceil \frac{d+1}{2} \rceil$. The proof is similar, except that we perform the computation in $L^2(\mathbb{R}^d)$ instead of $L^\infty(\mathbb{R}^d)$. Using Plancharel and then proceeding like before, we have:

$$\begin{aligned}
\|(-2\pi i x)^\alpha \check{m}_N(x)\|_{L^2_x} = \|D^\alpha_\xi m_N(\xi)\|_{L^2_\xi} &\lesssim \sum_{\alpha_1+\alpha_2=\alpha} \left\|D^{\alpha_1}_\xi m(\xi)\cdot D^{\alpha_2}_\xi\left(\psi\left(\frac{\xi}{N}\right)\right)\right\|_{L^2_\xi} \\
&\lesssim \sum_{\alpha_1+\alpha_2=\alpha} \left\|\frac{1}{|\xi|^{|\alpha_1|}}\cdot N^{-|\alpha_2|}\right\|_{L^2_\xi(|\xi|\sim N)} \\
&= \sum_{\alpha_1+\alpha_2=\alpha} N^{-|\alpha_2|}\left(\int_{|\xi|\sim N} |\xi|^{-2|\alpha_1|}\,d\xi\right)^{\frac{1}{2}} \\
&\lesssim \sum_{\alpha_1+\alpha_2=\alpha} N^{-|\alpha_2|}\left(N^{-2|\alpha_1|+d}\right)^{\frac{1}{2}} \\
&= N^{-|\alpha|+\frac{d}{2}}.
\end{aligned}$$

In particular, by choosing $\alpha = 0$ we have $\|\check{m}_N(x)\|_{L^2_x} \lesssim N^{\frac{d}{2}}$. Thus, for a fixed $A > 0$, Holder's inequality gives

$$\int_{|x|\leq A} |\check{m}_N(x)|\,dx \lesssim \|\check{m}_N(x)\|_{L^2} \left\|\chi_{|x|\leq A}\right\|_{L^2} \lesssim N^{\frac{d}{2}} A^{\frac{d}{2}}$$

and

$$\int_{|x|>A} |\check{m}_N(x)|\,dx \lesssim \left\||x|^{|\alpha|}\check{m}_N(x)\right\|_{L^2} \left\||x|^{-|\alpha|}\chi_{|x|>A}\right\|_{L^2}.$$

By the above computation, $\left\||x|^{|\alpha|}\check{m}_N(x)\right\|_{L^2} \lesssim N^{-|\alpha|+\frac{d}{2}}$. Computing the other norm,

$$\left\||x|^{-|\alpha|}\chi_{|x|>A}\right\|_{L^2} = \left(\int_{|x|>A} |x|^{-2|\alpha|}\,dx\right)^{\frac{1}{2}} \lesssim A^{-|\alpha|+\frac{d}{2}}$$

provided $|\alpha| \leq \lceil \frac{d+1}{2} \rceil$. Combining these two estimates gives

$$\int_{|x|>A} |\check{m}_N(x)|\,dx \lesssim (NA)^{-\lceil \frac{d+1}{2} \rceil + \frac{d}{2}}.$$

In particular, choosing $A = 1/N$, we have

$$\int |\check{m_N}(x)|\, dx \lesssim 1$$

uniformly in $N \in 2^{\mathbb{Z}}$. Essentially the same computation gives

$$\int |\nabla \check{m_N}(x)|\, dx \lesssim N$$

where the implicit constant is independent of N. The only difference in the calculation is that we begin by estimating $\left\|(-2\pi i x)^{\alpha} \frac{\partial}{\partial x} \check{m_N}(x)\right\|_{L^2_x} \sim \left\| D^{\alpha}_{\xi} \xi m_N(\xi)\right\|_{L^2_{\xi}}$ and consequently pick up an extra power of N throughout. Thus,

$$\begin{aligned}
\int_{|x|\geq 2|y|} |K(x+y) - K(x)|\, dx &\leq \sum_{N\in 2^{\mathbb{Z}}} \int_{|x|\geq 2|y|} |\check{m_N}(x+y) - \check{m_N}(x)|\, dx \\
&= \sum_{N \leq \frac{1}{|y|}} \int_{|x|\geq 2|y|} |\check{m_N}(x+y) - \check{m_N}(x)|\, dx \\
&\quad + \sum_{N > \frac{1}{|y|}} \int_{|x|\geq 2|y|} |\check{m_N}(x+y) - \check{m_N}(x)|\, dx.
\end{aligned}$$

In the first sum, over low frequencies, we can use the fundamental theorem of calculus to get

$$\sum_{N \leq \frac{1}{|y|}} \int_{|x|\geq 2|y|} |\check{m_N}(x+y) - \check{m_N}(x)|\, dx \leq \sum_{N \leq \frac{1}{|y|}} \int_{|x|\geq 2|y|} |y| \int_0^1 |\nabla \check{m_N}(x+\theta y)|\, d\theta\ dx.$$

Using Fubini's theorem and integrating over $\mathbb{R}^d$ instead of $|x| \geq 2|y|$ gives

$$\sum_{N \leq \frac{1}{|y|}} \int_{|x|\geq 2|y|} |\check{m_N}(x+y) - \check{m_N}(x)|\, dx \leq \sum_{N \leq \frac{1}{|y|}} |y|\, \|\nabla \check{m_N}\|_{L^1} \lesssim \sum_{N \leq \frac{1}{|y|}} |y| N \lesssim 1$$

where the last inequality follows from the usual dyadic sum argument. For the second sum, we have the crude estimate

$$\begin{aligned}
\sum_{N > \frac{1}{|y|}} \int_{|x|\geq 2|y|} |\check{m_N}(x+y) - \check{m_N}(x)|\, dx &\leq \sum_{N > \frac{1}{|y|}} \int_{|x|\geq 2|y|} |\check{m_N}(x+y)| + |\check{m_N}(x)|\, dx \\
&\lesssim \sum_{N > \frac{1}{|y|}} \int_{|x|\geq |y|} |\check{m_N}(x)|\, dx \\
&\lesssim \sum_{N > \frac{1}{|y|}} (N|y|)^{-\lceil \frac{d+1}{2} \rceil + \frac{d}{2}}.
\end{aligned}$$

Here, we have used our previous estimate with $A = |y|$. Since $-\left\lceil \frac{d+1}{2} \right\rceil + \frac{d}{2} < 0$ and the sum ranges over dyadic numbers greater than $\frac{1}{|y|}$, we have

$$\sum_{N > \frac{1}{|y|}} (N|y|)^{-\lceil \frac{d+1}{2} \rceil + \frac{d}{2}} \lesssim \left(\frac{1}{|y|}|y|\right)^{-\lceil \frac{d+1}{2} \rceil + \frac{d}{2}} = 1.$$

Therefore,

$$\int_{|x|\geq 2|y|} |K(x+y) - K(x)|\, dx \lesssim 1 + 1 \lesssim 1$$

and we are done. $\square$

Chapter 7

Littlewood-Paley Theory

In this chapter, we develop the theory of Littlewood-Paley projections and discuss a few of their immediate applications. More applications will become apparent in various contexts throughout the rest of this text.

The philosophy behind Littlewood-Paley theory is to decompose functions by localizing at different frequencies in the Fourier domain. As we shall see, this turns out to be a powerful idea. One source of motivation is the desire for a notion of orthogonality in L^p spaces for $p \neq 2$. Indeed, suppose that we can write $f \in L^2$ as $f = \sum_j f_j$, where the functions $\hat{f}$ have disjoint support in the frequency domain. Then by Plancharel,

$$\|f\|_{L^2}^2 = \left\|\sum_j f_j\right\|_{L^2}^2 = \left\|\sum_j \hat{f}_j\right\|_{L^2}^2 = \sum_j \left\|\hat{f}_j\right\|_{L^2}^2 = \sum_j \|f_j\|_{L^2}^2 .$$

The calculation shows that the Fourier transform has some nice orthogonality properties in L^2. Unfortunately, this is not strictly true in other L^p spaces. Littlewood-Paley theory gives a way of understanding the orthogonality properties of the Fourier transform in such spaces.

Some references for the following material are [2] and [3].

7.1 Littlewood-Paley Projections

We begin by defining the Littlewood-Paley projection operators. Throughout the rest of this chapter, we will use the following bump functions: let φ be a smooth function (which we view on the frequency domain) such that

$$\varphi(\xi) = \begin{cases} 1 & |\xi| \leq 1 \\ 0 & |\xi| \geq \frac{11}{10} \end{cases} .$$

Let $\psi(\xi) = \varphi(\xi) - \varphi(2\xi)$. Then

$$\psi(\xi) = \begin{cases} 0 & |\xi| \leq \frac{1}{2} \\ 1 & \frac{11}{20} \leq |\xi| \leq 1 \\ 0 & |\xi| \geq \frac{11}{10} \end{cases} .$$

For $N \in 2^{\mathbb{Z}}$, define $\psi_N(\xi) := \psi\left(\frac{\xi}{N}\right)$.

Definition 7.1. The **Littlewood-Paley projection onto frequencies** $|\xi| \sim N$, denoted P_N, is defined in the Fourier domain for Schwartz functions f via

$$\widehat{P_N f}(\xi) = \psi_N(\xi)\hat{f}(\xi).$$

Equivalently,

$$P_N f = f * N^d \check{\psi}(N\cdot).$$

In words, the operator P_N localizes the function f to frequencies of order N. Note that P_N is not a true projection operator, despite its name, since $P_N^2 \neq P_N$.

Littlewood-Paley projections can naturally be defined on wider ranges of frequencies.

Definition 7.2. The **Littlewood-Paley projection onto low frequencies**, denoted $P_{\leq N}$, is defined in the Fourier domain for Schwartz functions f via

$$\widehat{P_{\leq N} f}(\xi) = \varphi\left(\frac{\xi}{N}\right)\hat{f}(\xi).$$

Equivalently,

$$P_{\leq N} f = f * N^d \check{\varphi}(N\cdot).$$

The **Littlewood-Paley projection onto high frequencies** is $P_{>N} := I - P_{\leq N}$, and the **Littlewood-Paley projection onto medium frequencies** is $P_{M\leq \cdot \leq N} := \sum_{M\leq K\leq N} P_K$.

We often use the shorthand notation f_N, $f_{\leq N}$, $f_{>N}$, and $f_{M\leq\cdot\leq N}$ for $P_N f$, $P_{\leq N} f$, $P_{>N} f$, and $P_{M\leq\cdot\leq N} f$, respectively.

The following proposition contains basic properties of the Littlewood-Paley projections that will be used frequently.

Proposition 7.3. *Fix $N \in 2^{\mathbb{Z}}$.*

1. *The operators P_N and $P_{\leq N}$ are bounded on L^p for $1 \leq p \leq \infty$, i.e., $\|f_N\|_{L^p} + \|f_{\leq N}\|_{L^p} \lesssim \|f\|_{L^p}$.*
2. *We have the pointwise estimate $|f_N(x)| + |f_{\leq N}(x)| \lesssim (Mf)(x)$.*
3. *For $f \in L^p$ with $1 < p < \infty$, $\sum_{N\in 2^{\mathbb{Z}}} P_N f \xrightarrow{L^p} f$.*
4. *(Bernstein inequality I) $\|f_N\|_{L^q} \lesssim N^{\frac{d}{p}-\frac{d}{q}} \|f_N\|_{L^p}$ for $1 \leq p \leq q \leq \infty$.*
5. *(Bernstein inequality II) $\||\nabla|^s f_N\|_{L^p} \sim N^s \|f_N\|_{L^p}$ for $1 \leq p \leq \infty$ and $s \in \mathbb{R}$.*

Proof.

1. We compute:

$$\|f_N\|_{L^p} = \left\|f * N^d\check{\psi}(N\cdot)\right\|_{L^p} \lesssim \|f\|_{L^p} \left\|N^d\check{\psi}(N\cdot)\right\|_{L^1}.$$

By construction, $N^d\check{\psi}(N\cdot)$ is L^1-normalized, so that $\|N^d\check{\psi}(N\cdot)\|_{L^1} = \|\check{\psi}\|_{L^1}$. Thus,

$$\|f_N\|_{L^p} \lesssim \|f\|_{L^p} \|\check{\psi}\|_{L^1}.$$

The same computation works for $\|f_{\leq N}\|_{L^p}$.

2. We show the estimate for $|f_N(x)|$, and as before, the same proof works for $|f_{\leq N}(x)|$. We have

$$f_N(x) = \left(f * N^d\check{\psi}(N\cdot)\right)(x) = N^d \int f(x-y)\check{\psi}(Ny)\, dy.$$

Because ψ is Schwartz, $\check{\psi}$ is Schwartz, hence is bounded and decays up to any order. Thus,

$$|f_N(x)| \lesssim N^d \int \frac{|f(x-y)|}{\langle N|y|\rangle^{100d}}\, dy$$

where we use the *Japanese angle bracket* notation $\langle x\rangle := (1+|x|^2)^{\frac{1}{2}}$. When $|y|$ is small, we exploit the boundedness of the angle bracket, and when $|y|$ is large we exploit its decay. Explicitly,

$$|f_N(x)| \lesssim N^d \int_{|y|\leq \frac{1}{N}} |f(x-y)|\, dy + N^d \int_{|y|>\frac{1}{N}} \frac{|f(x-y)|}{(N|y|)^{100d}}\, dy.$$

Because $|B(0,1/N)| \sim N^{-d}$ so that $1/|B(0,1/N)| \sim N^d$, we have by definition of the maximal function

$$N^d \int_{|y|\leq \frac{1}{N}} |f(x-y)|\, dy \lesssim (Mf)(x).$$

To estimate the second integral, we sum over dyadic annuli.

$$\begin{aligned} N^d \int_{|y|>\frac{1}{N}} \frac{|f(x-y)|}{(N|y|)^{100d}}\, dy &\lesssim N^d \sum_{M\in 2^{\mathbb{Z}};\, M>\frac{1}{N}} \int_{M\leq |y|\leq 2M} \frac{|f(x-y)|}{(NM)^{100d}}\, dy \\ &\lesssim N^d \sum_{M\in 2^{\mathbb{Z}};\, M>\frac{1}{N}} (NM)^{-100d} \frac{M^d}{|B(0,2M)|} \int_{B(0,2M)} |f(x-y)|\, dy \\ &\lesssim \sum_{M\in 2^{\mathbb{Z}};\, M>\frac{1}{N}} (NM)^{-99d} (Mf)(x). \end{aligned}$$

By the usual dyadic sum argument,

$$\sum_{M\in 2^{\mathbb{Z}};\, M>\frac{1}{N}} (NM)^{-99d}(Mf)(x) \lesssim \left(\frac{1}{M}M\right)^{-99d} (Mf)(x) = (Mf)(x)$$

and we are done.

3. By the density of $\mathcal{S}(\mathbb{R}^d) \subseteq L^p(\mathbb{R}^d)$ and by property (1), it suffices by the usual approximation argument to prove the claim for Schwartz functions. Fix $f \in \mathcal{S}(\mathbb{R}^d)$.

   For $p=2$, by Plancharel we have

$$\|f - f_{N\leq\leq M}\|_{L^2} = \left\| \hat{f}\left(1 - \sum_{N\leq K\leq M} \psi_K\right)\right\|_{L^2} \leq \left\|\hat{f}\sum_{K<N}\psi_K\right\|_{L^2} + \left\|\hat{f}\sum_{K>M}\psi_K\right\|_{L^2}.$$

   Since $\sum_{K<N}\psi_K \leq \chi_{B(0,N)}$ and $\sum_{K>M}\psi_K \leq \chi_{|\xi|\geq M}$, by the dominated convergence theorem, the left hands side tends to 0 as $N\to 0$, and the right hand side tends to 0 as $M\to\infty$.

   For $1<p<2$, we use interpolation. Choose θ so that $\frac{1}{p} = \frac{\theta}{1} + \frac{1-\theta}{2}$. Then

$$\|f - f_{N\leq\leq M}\|_{L^p} \leq \|f - f_{N\leq\leq M}\|_{L^1}^{\theta} \|f - f_{N\leq\leq M}\|_{L^2}^{1-\theta}.$$

   By (1),

$$\|f - f_{N\leq\leq M}\|_{L^1}^{\theta} \leq \left(\|f\|_{L^1} + \|f_{N\leq\leq M}\|_{L^1}\right)^{\theta} \lesssim \|f\|_{L^1}^{\theta}.$$

   Our previous computation shows that $\|f - f_{N\leq\leq M}\|_{L^2}^{1-\theta} \to 0$ and thus

$$\|f - f_{N\leq\leq M}\|_{L^p} \to 0.$$

   For $2<p<\infty$, the same trick works:

$$\|f - f_{N\leq\leq M}\|_{L^p} \leq \|f - f_{N\leq\leq M}\|_{L^1}^{1-\frac{2}{p}} \|f - f_{N\leq\leq M}\|_{L^2}^{\frac{2}{p}}.$$

   The left hand side is bounded by (1), and the right hand side $\to 0$ by the previous calculation.

4. By Young's convolution inequality, for r satisfying $\frac{1}{p}+\frac{1}{r}=\frac{1}{q}+1$,

$$\begin{aligned}\|f_N\|_{L^q} &= \left\|f * N^d\check{\psi}(N\cdot)\right\|_{L^q} \lesssim \|f\|_{L^p}\left\|N^d\check{\psi}(N\cdot)\right\|_{L^r} = \|f\|_{L^p} N^{d-\frac{d}{r}}\left\|\check{\psi}\right\|_{L^r}\\ &\lesssim N^{\frac{d}{p}-\frac{d}{q}}\|f\|_{L^p}.\end{aligned}$$

This is great but this isn't the estimate we need. To recover f_N on the right, we use *fattened* Littlewood-Paley projections $\tilde{P_N} := P_{\frac{N}{2}\leq\leq 2N}$. Note that $\tilde{P_N}P_N = P_N$. Also, we have

$$\widehat{\tilde{P_N}f}(\xi) = \left(\sum_{\frac{N}{2}\leq K\leq 2N}\psi_K\right)(\xi)\hat{f}(\xi)$$

so that

$$\begin{aligned}\tilde{P_N}f &= f * \left(\sum_{\frac{N}{2}\leq K\leq 2N}\psi_K\right)^{\vee} = f * \sum_{\frac{N}{2}\leq K\leq 2N} K^d\check{\psi}(K\cdot) \sim f * N^d \sum_{\frac{N}{2}\leq K\leq 2N}\check{\psi}(K\cdot)\\ &= f * N^d\check{\tilde{\psi}}(N\cdot)\end{aligned}$$

for an appropriately defined $\tilde{\psi}$. [1] Thus,

$$\begin{aligned}\|f_N\|_{L^q} &= \left\|\tilde{P_N}f_N\right\|_{L^q} \sim \left\|f_N * N^d\check{\tilde{\psi}}(N\cdot)\right\|_{L^q} \lesssim \|f_N\|_{L^p}\left\|N^d\check{\tilde{\psi}}(N\cdot)\right\|_{L^r}\\ &= \|f_N\|_{L^p} N^{d-\frac{d}{r}}\left\|\check{\tilde{\psi}}\right\|_{L^r}\\ &\lesssim N^{\frac{d}{p}-\frac{d}{q}}\|f_N\|_{L^p}.\end{aligned}$$

5. Fix $s\in\mathbb{R}$. By definition, we have:

$$\widehat{|\nabla|^s f_N}(\xi) \sim |\xi|^s\psi_N(\xi)\hat{f}(\xi) = N^s\left[\left(\frac{|\xi|}{N}\right)^s\psi\left(\frac{\xi}{N}\right)\right]\hat{f}(\xi).$$

Since the support of ψ is localized away from the origin, $\rho(\xi) := |\xi|^s\psi(\xi)\in C_c^\infty(\mathbb{R}^d\setminus\{0\})$ for any value of s. So $\widehat{|\nabla|^s f_N}(\xi)\sim N^s\rho\left(\frac{\xi}{N}\right)\hat{f}(\xi)$, hence

$$|\nabla|^s f_N = N^s\left[f * N^d\check{\rho}(N\cdot)\right].$$

Thus,

$$\||\nabla|^s f_N\|_{L^p} \lesssim N^s\|f\|_{L^p}\left\|N^d\check{\rho}(N\cdot)\right\|_{L^1} = N^s\|f\|_{L^p}\|\check{\rho}\|_{L^1}\lesssim N^s\|f\|_{L^p}.$$

Via the same fattened Littlewood-Paley projections technique, we get the estimate $\||\nabla|^s f_N\|_{L^p}\lesssim N^s\|f_N\|_{L^p}$.

To get the reverse inequality, observe that

$$\hat{f_N}(\xi) = |\xi|^{-s}|\xi|^s\hat{f_N}(\xi)\sim(|\nabla|^{-s}|\nabla|^s f_N)\hat{}(\xi).$$

Thus, applying the inequality we already have to $|\nabla|^{-s}|\nabla|^s f_N$ gives

$$\|f_N\|_{L^p}\sim\left\||\nabla|^{-s}|\nabla|^s f_N\right\|_{L^p}\lesssim N^{-s}\||\nabla|^s f_N\|_{L^p}$$

so that $N^s\|f_N\|_{L^p}\lesssim\||\nabla|^s f_N\|_{L^p}$ as desired.

[1] Note that, since we are summing over dyadic numbers, $\sum_{\frac{N}{2}\leq K\leq 2N}\check{\psi}(K\cdot) = \check{\psi}\left(\frac{N}{2}\cdot\right)+\check{\psi}(N\cdot)+\check{\psi}(2N\cdot)$.

$\square$

We remark that (3) does not hold for $p = 1$ or $p = \infty$. Note that $\int f_N(x)\,dx = \hat{f_N}(0) = 0$ for any $N \in 2^{\mathbb{Z}}$, so $\int f_{N\leq\leq M}(x)\,dx = 0$. But $\sum P_N f \xrightarrow{L^1} f$ implies that their means converge. The claim fails for $p = 1$ by choosing an $f \in L^1$ with $\int f(x)\,dx \neq 0$. To see why the claim fails for $p = \infty$, note that $f_{N\leq\leq M} \in C^\infty$. Since convergence in L^∞ is uniform convergence, $\lim \sum P_N f$ is continuous.

We also remark that the same proof technique in (4) gives the estimate $\|f_{\leq N}\|_{L^q} \lesssim N^{\frac{d}{p}-\frac{d}{q}} \|f_{\leq N}\|_{L^p}$.

Though (3) does not hold for $p = 1$, we do have the following result.

Proposition 7.4. *Let $f \in L^1(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} f(x)\,dx = 0$. Then $\sum P_N f \xrightarrow{L^1} f$.*

Proof. First, we claim that f can be approximated by a smooth function with compact support and mean 0. Indeed, fix $\varepsilon > 0$. Then since $f \in L^1$, there exists an $R > 0$ such that $\int_{|x|>R} |f(x)|\,dx < \varepsilon$. Then

$$\int f(x)\chi_{B(0,R)}(x) + O(\varepsilon)\chi_{R\leq|x|\leq 2R}(x)\,dx = 0$$

for some appropriately chosen constant $O(\varepsilon)$. Convolving with a smooth approximation to the identity the desired approximation.

Thus, we can assume that $f \in C_c^\infty(\mathbb{R}^d)$ and $\int f(x)\,dx = 0$. By the triangle inequality,

$$\left\| f - \sum_{N\leq K\leq M} P_N f \right\|_{L^1} \leq \left\| \sum_{K<N} P_K f \right\|_{L^1} + \sum_{K>M} \|P_K f\|_{L^1}.$$

First, consider the high frequencies. By Bernstein's inequality,

$$\sum_{K>M} \|P_K f\|_{L^1} \lesssim \sum_{K>M} K^{-1} \||\nabla| f_K\|_{L^1}.$$

Since $f \in C_c^\infty(\mathbb{R}^d)$ and the sum is a geometric sum over dyadic numbers,

$$\sum_{K>M} \|P_K f\|_{L^1} \lesssim_f M^{-1}$$

which $\to 0$ as $M \to \infty$. Next, consider the low frequencies. We exploit the fact that convolving with a mean 0 function is like taking a derivative, in the following sense: we have

$$(P_{\leq} f)(x) = \int f(y) N^d \check{\varphi}(N(x-y))\,dy = N^d \int f(y)\left[\check{\varphi}(N(x-y)) - \check{\varphi}(Nx)\right]\,dy$$

since f has mean 0. By the fundamental theorem of calculus,

$$\check{\varphi}(N(x-y)) - \check{\varphi}(Nx) = Ny \int_0^1 \nabla\check{\varphi}(Nx - \theta Ny)\,d\theta.$$

Using the fact that the support of f is contained in $B(0,R)$, along with the fact that $\check{\varphi}$ is Schwartz and hence decays as fast as we need,

$$\begin{aligned} \|P_{\leq N} f\|_{L^1} &\lesssim N^{d+1} \int\int |y||f(y)| \int_0^1 |\nabla\check{\varphi}(Nx - \theta Ny)|\,d\theta\,dy\,dx \\ &\lesssim N^{d+1} R \int\int |f(y)| \frac{1}{\langle Nx\rangle^{100d}}\,dy\,dx. \end{aligned}$$

Here we have also used the fact that $NR << 1$ for $N << 1$, so we don't need to include it in the Japanese angle bracket. So

$$\|P_{\leq N} f\|_{L^1} \lesssim N^{d+1} R \|f\|_{L^1} \int \frac{1}{\langle Nx \rangle^{100d}} \, dx \lesssim N^{d+1} R \|f\|_{L^1} N^{-d}$$

which $\to 0$ as $N \to 0$.

$\square$

Broadly speaking, low frequencies of a function contribute to smoothness, whereas high frequencies contribute to irregularity. This is made somewhat precise by the following proposition.

Proposition 7.5. *Let $f \in L^\infty(\mathbb{R}^d)$ and fix $0 < \alpha < 1$. Then f is α-Holder continuous if and only if $\|P_{\geq N} f\|_{L^\infty} \lesssim N^{-\alpha}$ for all $N \geq 1$.*

Proof. Suppose first that f is α-Holder continuous. Then for all $x, y \in \mathbb{R}^d$ with $y \neq 0$, $|f(x-y) - f(x)| \lesssim |y|^\alpha$. Fix $N \in 2^{\mathbb{Z}}$ such that $N \geq 1$. By definition of the Littlewood-Paley projections,

$$\begin{aligned}
|P_{\geq N} f(x)| &= \left| \sum_{K \geq N} P_K f(x) \right| \leq \sum_{K \geq N} |P_K f(x)| = \sum_{K \geq N} \left| (f * K^d \check{\psi}(K \cdot)(x) \right| \\
&= \sum_{K \geq N} \left| K^d \int_{\mathbb{R}^d} f(x-y) \check{\psi}(Ky) \, dy \right|.
\end{aligned}$$

Observe that $\int_{\mathbb{R}^d} \check{\psi}(y) \, dy = \hat{\check{\psi}}(0) = \psi(0) = 0$. Thus, $\check{\psi}$ has mean zero, hence

$$K^d \int_{\mathbb{R}^d} f(x) \check{\psi}(Ny) \, dy = 0.$$

Inserting this into the above expression gives

$$\begin{aligned}
|P_{\geq N} f(x)| &\leq \sum_{K \geq N} \left| K^d \int_{\mathbb{R}^d} f(x-y) \check{\psi}(Ky) \, dy - K^d \int_{\mathbb{R}^d} f(x) \check{\psi}(Ky) \, dy \right| \\
&= \sum_{K \geq N} \left| K^d \int_{\mathbb{R}^d} [f(x-y) - f(x)] \, \check{\psi}(Ky) \, dy \right| \\
&\leq \sum_{K \geq N} K^d \int_{\mathbb{R}^d} |f(x-y) - f(x)| |\check{\psi}(Ky)| \, dy.
\end{aligned}$$

Since f is α-Holder continuous, we have

$$|P_{\geq N} f(x)| \lesssim \sum_{K \geq N} K^d \int_{\mathbb{R}^d} |y|^\alpha |\check{\psi}(Ky)| \, dy.$$

Making the change of variables $z = Ky$ gives $dy = K^{-d} \, dz$, hence

$$|P_{\geq N} f(x)| \lesssim \sum_{K \geq N} K^d \int_{\mathbb{R}^d} \left| \frac{z}{K} \right|^\alpha |\check{\psi}(z)| \, K^{-d} \, dz = \sum_{K \geq N} K^{-\alpha} \int_{\mathbb{R}^d} |z|^\alpha \, |\check{\psi}(z)| \, dz.$$

Since ψ is a smooth function with compact support on an annulus of radius ~ 1, ψ is Schwartz, and hence $\check{\psi}$ is Schwartz. Thus, $\check{\psi}$ decays up to any polynomial order, and $|z|^\alpha |\check{\psi}(z)|$ is bounded. Therefore,

$$\int_{\mathbb{R}^d} |z|^\alpha \, |\check{\psi}(z)| \, dz \lesssim \int_{|z| \leq 1} |z|^\alpha \, |\check{\psi}(z)| \, dz + \int_{|z| > 1} |z|^\alpha \frac{1}{|z|^{100d}} \, dz \lesssim 1 + 1 = 1.$$

This gives
$$|P_{\geq N}f(x)| \lesssim \sum_{K\geq N} K^{-\alpha} \lesssim N^{-\alpha}$$
This bound is uniform in x. Thus, $\|P_{\geq N}f\|_{L^\infty} \lesssim N^{-\alpha}$ for all dyadic $N \geq 1$.

Next, we show the converse statement.

Suppose that $\|P_{\geq N}f\|_{L^\infty} \lesssim N^{-\alpha}$ for all dyadic $N \geq 1$. We need to show that $|f(x-y) - f(x)| \lesssim |y|^\alpha$ for all $x \in \mathbb{R}^d$ and $y \neq 0 \in \mathbb{R}^d$. Because $f \in L^\infty(\mathbb{R}^d)$ and hence $|f(x-y) - f(x)| \leq 2\|f\|_{L^\infty} \lesssim 1$, it suffices to consider $0 < |y| < 1$.

So fix $x, y \in \mathbb{R}^d$ with $0 < |y| < 1$. Decompose $|f(x-y) - f(x)|$ as
$$|f(x-y) - f(x)| \leq |f_{<1}(x-y) - f_{<1}(x)| + |f_{\geq 1}(x-y) - f_{\geq 1}(x)|.$$
We estimate these terms separately. First, consider the projection of f onto frequencies < 1. By definition of the Littlewood-Paley projection $P_{<1}$,
$$\begin{aligned}|f_{<1}(x-y) - f_{<1}(x)| &= |(f * \check{\varphi})(x-y) - (f * \check{\varphi})(x)| \\ &= \left| \int f(x-y-z)\check{\varphi}(z)\,dz - \int f(x-z)\check{\varphi}(z)\,dz \right|.\end{aligned}$$
Changing variables in the first integral via $y + z \mapsto z$ gives
$$\begin{aligned}|f_{<1}(x-y) - f_{<1}(x)| &= \left| \int f(x-z)\check{\varphi}(z-y)\,dz - \int f(x-z)\check{\varphi}(z)\,dz \right| \\ &\leq \int |f(x-z)|\,|\check{\varphi}(z-y) - \check{\varphi}(z)|\,dz.\end{aligned}$$
By the fundamental theorem of calculus,
$$\begin{aligned}|f_{<1}(x-y) - f_{<1}(x)| &\leq \int |f(x-z)||y| \int_0^1 |\nabla\check{\varphi}(z-\theta y)|\,d\theta\,dz \\ &\leq \|f\|_\infty |y| \int\int_0^1 |\nabla\check{\varphi}(z-\theta y)|\,d\theta\,dz.\end{aligned}$$
Because φ is Schwartz, $\nabla\check{\varphi}$ is Schwartz, hence is bounded and decays up to any order. Consequently,
$$\int\int_0^1 |\nabla\check{\varphi}(z-\theta y)|\,d\theta\,dz \lesssim \int \frac{1}{\langle z\rangle^{100d}}\,dz \lesssim 1.$$
Thus, $|f_{<1}(x-y) - f_{<1}(x)| \lesssim |y|$. Since $|y| < 1$, $|y|^\alpha \geq |y|$, so $|f_{<1}(x-y) - f_{<1}(x)| \lesssim |y|^\alpha$.

Next we consider $|f_{\geq 1}(x-y) - f_{\geq 1}(x)|$. For any $N \geq 1$, by assumption,
$$|f_{\geq N}(x-y) - f_{\geq N}(x)| \leq |f_{\geq N}(x-y)| + |f_{\geq N}(x)| \leq 2\|P_{\geq N}f\|_{L^\infty} \lesssim N^{-\alpha}. \tag{7.1}$$
On the other hand, for any $K \geq 1$ we compute as above with fattened Littlewood-Paley projections to get
$$\begin{aligned}|f_K(x-y) - f_K(x)| &= |\tilde{P_K}f_K(x-y) - \tilde{P_K}f_K(x)| \\ &= \left|(f_K * K^d\check{\tilde{\psi}}(K\cdot))(x-y) - (f_K * K^d\check{\tilde{\psi}}(K\cdot))(x)\right| \\ &= \left|K^d \int_{\mathbb{R}^d} f_K(x-y-z)\check{\tilde{\psi}}(Kz)\,dz - K^d\int_{\mathbb{R}^d} f_K(x-z)\check{\tilde{\psi}}(Kz)\,dz\right| \\ &= \left|K^d \int_{\mathbb{R}^d} f_K(x-z)\check{\tilde{\psi}}(K(z-y))\,dz - K^d\int_{\mathbb{R}^d} f_K(x-z)\check{\tilde{\psi}}(Kz)\,dz\right| \\ &= \left|K^d \int_{\mathbb{R}^d} \left[\check{\tilde{\psi}}(K(z-y)) - \check{\tilde{\psi}}(Kz)\right] f_K(x-z)\,dz\right| \\ &\leq K^d \|f_K\|_{L^\infty} \int_{\mathbb{R}^d} \left|\check{\tilde{\psi}}(Kz-Ky) - \check{\tilde{\psi}}(Kz)\right|\,dz.\end{aligned}$$

Changing variables via $Kz \mapsto z$ and hence $dz \mapsto K^{-d}\,dz$ gives

$$|f_K(x-y) - f_K(x)| \le \|f_K\|_{L^\infty} \int_{\mathbb{R}^d} \left| \check{\check{\psi}}(z - Ky) - \check{\check{\psi}}(z) \right| dz.$$

Since $f_K = f_{\ge K} - f_{\ge 2K}$,

$$\|f_K\|_{L^\infty} \le \|f_{\ge K}\|_{L^\infty} + \|f_{\ge 2K}\|_{L^\infty} \lesssim K^{-\alpha} + (2K)^{-\alpha} \lesssim K^{-\alpha}.$$

Then by this fact and by the fundamental theorem of calculus,

$$|f_K(x-y) - f_K(x)| \le K^{-\alpha} K|y| \int \int_0^1 \left| \nabla \check{\check{\psi}}(z - \theta N y \right| d\theta\, dz.$$

Since ψ is Schwartz, the same argument as above gives $\int \int_0^1 \left| \nabla \check{\check{\psi}}(z - \theta N y \right| d\theta\, dz \lesssim 1$. Thus, for any $K \ge 1$ we have the estimate

$$|f_K(x-y) - f_K(x)| \lesssim K^{1-\alpha}|y|. \tag{7.2}$$

So for any dyadic $N \ge 1$, by (7.1) and (7.2),

$$\begin{aligned} |f_{\ge 1}(x-y) - f_{\ge 1}(x)| &\le \sum_{1 \le K < N} |f_K(x-y) - f_K(x)| \,+\, |f_{\ge N}(x-y) - f_{\ge N}(x)| \\ &\lesssim \sum_{1 \le K < N} K^{1-\alpha}|y| \,+\, N^{-\alpha}. \end{aligned}$$

Since the sum is over dyadic K, $\sum_{1 \le K < N} K^{1-\alpha}|y| \lesssim N^{1-\alpha}|y|$. So for any dyadic $N \ge 1$,

$$|f_{\ge 1}(x-y) - f_{\ge 1}(x)| \lesssim N^{1-\alpha}|y| + N^{-\alpha}.$$

Choose $N \ge 1$ such that

$$\frac{1}{|y|} \le N \le 2\frac{1}{|y|}.$$

Since $|y| < 1$, this is certainly possible. Then $N \sim |y|^{-1}$, and thus

$$\begin{aligned} |f_{\ge 1}(x-y) - f_{\ge 1}(x)| &\lesssim (|y|^{-1})^{1-\alpha}|y| + (|y|^{-1})^{-\alpha} = |y|^{\alpha-1}|y| + |y|^{\alpha} \\ &\lesssim |y|^{\alpha}. \end{aligned}$$

Putting the two estimates together yields

$$\begin{aligned} |f(x-y) - f(x)| &\le |f_{<1}(x-y) - f_{<1}(x)| + |f_{\ge 1}(x-y) - f_{\ge 1}(x)| \lesssim |y|^{\alpha} + |y|^{\alpha} \\ &\lesssim |y|^{\alpha} \end{aligned}$$

so that f is α-Holder continuous as desired. $\square$

7.2 The Littlewood-Paley Square Function

In this section, we introduce the Littlewood-Paley square function, and the key fact that its L^p-norm is comparable to the L^p-norm of the input function.

Definition 7.6. For $f \in \mathcal{S}(\mathbb{R}^d)$, the **Littlewood-Paley square function of** f is given by

$$(Sf)(x) = \left(\sum_{N \in 2^{\mathbb{Z}}} |f_N(x)|^2 \right)^{\frac{1}{2}}.$$

To better estimate the size of the square function, we borrow an inequality from probability theory.

Lemma 7.7 (Kinchin's Inequality)**.** *Let X_n be independent identically distributed random variables such that $X_n = \pm 1$ with equal probability. Then for any $0 < p < \infty$,*

$$\left(\mathbb{E}\left|\sum c_n X_N\right|^p\right)^{\frac{1}{p}} \sim \sqrt{\sum |c_n|^2}.$$

Proof. By considering real and imaginary parts, it suffices to assume that $c_n \in \mathbb{R}$. Then

$$\mathbb{E}\left|\sum c_n X_N\right|^p = p \int_0^\infty \lambda^p \, \mathbb{P}\left(\left|\sum c_n X_N\right| > \lambda\right) \frac{d\lambda}{\lambda}.$$

We have

$$\mathbb{P}\left(\left|\sum c_n X_N\right| > \lambda\right) \leq \mathbb{P}\left(\sum c_n X_N > \lambda\right) + \mathbb{P}\left(\sum c_n X_N < -\lambda\right) = 2\,\mathbb{P}\left(\sum c_n X_N > \lambda\right)$$

because the X_n are i.i.d with $X_n = \pm 1$ with equal probability. Next, recall the exponential Chebychev's inequality: for a random variable X and $t > 0$,

$$\mathbb{P}\left(X \geq \lambda\right) \leq e^{-t\lambda}\,\mathbb{E}\left(e^{tX}\right).$$

Thus, for $t > 0$,

$$\mathbb{P}\left(\left|\sum c_n X_N\right| > \lambda\right) \leq 2e^{-\lambda t}\,\mathbb{E}\left(e^{t\sum c_n X_n}\right).$$

Because the X_n are independent, this gives

$$\begin{aligned}\mathbb{P}\left(\left|\sum c_n X_N\right| > \lambda\right) &\leq 2e^{-\lambda t}\prod \mathbb{E}\left(e^{tc_n X_n}\right) = 2e^{-\lambda t}\prod\left(\frac{1}{2}e^{tc_n} + \frac{1}{2}e^{-tc_n}\right)\\ &= 2e^{-\lambda t}\prod \cosh(tc_n).\end{aligned}$$

Recall that $\cosh(x) \leq e^{\frac{x^2}{2}}$; one quick way to see this is by comparing Taylor series. So

$$\mathbb{P}\left(\left|\sum c_n X_N\right| > \lambda\right) \leq 2e^{-\lambda t}\prod e^{\frac{(tc_n)^2}{2}} = 2e^{-\lambda t}e^{\frac{t^2}{2}\sum c_n^2}.$$

Choosing t such that $\lambda t = t^2 \sum c_n^2$, hence $t = \frac{\lambda}{\sum c_n^2}$, gives

$$\mathbb{P}\left(\left|\sum c_n X_N\right| > \lambda\right) \leq 2e^{-\lambda t}e^{\frac{\lambda t}{2}} = 2e^{-\frac{\lambda t}{2}} = 2e^{-\frac{\lambda^2}{2\sum c_n^2}}.$$

So

$$\mathbb{E}\left|\sum c_n X_N\right|^p \lesssim_p \int_0^\infty \lambda^p e^{-\frac{\lambda^2}{2\sum c_n^2}}\frac{d\lambda}{\lambda}.$$

Making the changes of variables $z = \frac{\lambda}{\sqrt{\sum c_n^2}}$ yields

$$\mathbb{E}\left|\sum c_n X_N\right|^p \lesssim \left(\sum c_n^2\right)^{\frac{p}{2}} \int_0^\infty z^p e^{-\frac{z^2}{2}}\frac{dz}{z} \lesssim \left(\sum c_n^2\right)^{\frac{p}{2}}.$$

This gives the $\lesssim$ direction of the statement.

Next, we need the $\gtrsim$ direction. We first consider the case $1 \leq p < \infty$. Note that $\sum |c_n|^2 = \mathbb{E}\left|\sum c_n X_n\right|^2$. This is because the X_n are independent, and so

$$\mathbb{E}(X_n X_m) = \mathbb{E}(X_n)\,\mathbb{E}(X_m) = 0.$$

Then by Holder, and by the above inequality,

$$\sum |c_n|^2 \leq \left(\mathbb{E}\left|\sum c_n X_n\right|^p\right)^{\frac{1}{p}} \left(\mathbb{E}\left|\sum c_n X_n\right|^{p'}\right)^{\frac{1}{p'}} \lesssim \left(\mathbb{E}\left|\sum c_n X_n\right|^p\right)^{\frac{1}{p}} \left(\sum c_n^2\right)^{\frac{1}{2}}.$$

This gives

$$\left(\sum c_n^2\right)^{\frac{p}{2}} \lesssim \left(\mathbb{E}\left|\sum c_n X_n\right|^p\right)^{\frac{1}{p}}$$

as desired.

For $0 < p < 1$, we use Cauchy-Schwarz:

$$\begin{aligned}\sum |c_n|^2 = \mathbb{E}\left|\sum c_n X_n\right|^2 &= \mathbb{E}\left(\left|\sum c_n X_n\right|^{\frac{p}{2}} \left|\sum c_n X_n\right|^{2-\frac{p}{2}}\right) \\ &\leq \left(\mathbb{E}\left|\sum c_n X_n\right|^p\right)^{\frac{1}{2}} \left(\mathbb{E}\left|\sum c_n X_n\right|^{4-p}\right)^{\frac{1}{2}}.\end{aligned}$$

Again by the first demonstrated inequality we have

$$\sum |c_n|^2 \lesssim \left(\mathbb{E}\left|\sum c_n X_n\right|^p\right)^{\frac{1}{2}} \left(\sum |c_n|^2\right)^{\frac{4-p}{4}} \quad \Rightarrow \quad \left(\sum |c_n|^2\right)^{\frac{p}{4}} \lesssim \left(\mathbb{E}\left|\sum c_n X_n\right|^p\right)^{\frac{1}{2}}.$$

Raising both sides to the power $\frac{2}{p}$ gives the desired result. □

The main result of this section is the following theorem.

Theorem 7.8 (Square function estimate). *For $1 < p < \infty$, we have $\|Sf\|_{L^p} \sim \|f\|_{L^p}$.*

Proof. We will first show $\|Sf\|_{L^p} \lesssim \|f\|_{L^p}$. As a remark, the proof of this direction does not rely on the specific multiplier ψ defining P_1, so this inequality holds in more generality. In particular, ψ can be replaced by any element of $C_c^\infty(\mathbb{R}^d \setminus \{0\})$.

Let $m(\xi) := \sum_{N \in 2^{\mathbb{Z}}} \psi_N(\xi) X_N$, where the X_N are i.i.d random variables such $X_N = \pm 1$ with equal probability. Then $\|m\|_{L^\infty} \lesssim 1$, because for any given ξ, only finitely many of the summands give any nonzero contribution due to the compact supports of the ϕ_N. Also,

$$|D_\xi^\alpha m(\xi)| \leq \sum_{N \in 2^{\mathbb{Z}}} N^{-|\alpha|} \left| D_\xi^\alpha \psi\left(\frac{\xi}{N}\right)\right|.$$

Because ψ_N is smooth and has support $|\xi| \sim N$, and because only finitely summands contribute,

$$|D_\xi^\alpha m(\xi)| \lesssim |\xi|^{-|\alpha|}.$$

As this holds for any $\alpha \in \mathbb{N}^d$, the Mikhlin multiplier theorem gives $\|\check{m} * f\|_{L^p} \lesssim \|f\|_{L^p}$ for all $1 < p < \infty$. Note that $\check{m} * f = \sum_{N \in 2^{\mathbb{Z}}} f_N X_N$. So by Kinchin's inequality,

$$\begin{aligned}(Sf)(x) = \left(\sum_{N \in 2^{\mathbb{Z}}} |f_N(x)|^2\right)^{\frac{1}{2}} &\sim \left(\mathbb{E}\left|\sum_{N \in 2^{\mathbb{Z}}} f_N(x) X_N\right|^p\right)^{\frac{1}{p}} \\ &= \left(\mathbb{E}\,|\check{m} * f|^p\right)^{\frac{1}{p}}.\end{aligned}$$

Using Fubini's theorem,

$$\|Sf\|_{L^p}^p \sim \int \mathbb{E}\,|\check{m} * f|^p\,(x)\,dx \lesssim \mathbb{E}\,\|\check{m} * f\|_{L^p}^p$$

Since the expectation of a constant is itself,

$$\|Sf\|_{L^p}^p \lesssim \mathbb{E}\,\|\check{m} * f\|_{L^p}^p \lesssim \mathbb{E}\,\|f\|_{L^p}^p = \|f\|_{L^p}^p.$$

Next, we show $\|f\|_{L^p} \lesssim \|Sf\|_{L^p}$ using duality and the fattened Littlewood-Paley projections. Recall that $\tilde{P_N} P_N = P_N$. We have

$$\begin{aligned}\|f\|_{L^p} &= \sup_{\|g\|_{L^{p'}=1}} \langle f, g\rangle = \sup_{\|g\|_{L^{p'}=1}} \left\langle \sum P_N f, g\right\rangle = \sup_{\|g\|_{L^{p'}=1}} \left\langle \sum \tilde{P_N} P_N f, g\right\rangle \\ &= \sup_{\|g\|_{L^{p'}=1}} \sum \left\langle P_N f, \tilde{P_N} g\right\rangle\end{aligned}$$

since the operators P_N are self-adjoint. By Cauchy-Schwarz and then Holder,

$$\|f\|_{L^p} \leq \sup_{\|g\|_{L^{p'}=1}} \left\langle \left(\sum |P_N f|^2\right)^{\frac{1}{2}}, \left(\sum |\tilde{P_N} g|^2\right)^{\frac{1}{2}}\right\rangle \leq \sup_{\|g\|_{L^{p'}=1}} \|Sf\|_{L^p} \left\| \left(\sum |\tilde{P_N} g|^2\right)^{\frac{1}{2}}\right\|_{L^{p'}}.$$

By the remark at the beginning of the proof, $\left\| \left(\sum |\tilde{P_N} g|^2\right)^{\frac{1}{2}}\right\|_{L^{p'}} \lesssim \|g\|_{L^{p'}}$. Thus,

$$\|f\|_{L^p} \lesssim \|Sf\|_{L^p}.$$

□

7.3 Applications to Fractional Derivatives

One of the first properties of Littlewood-Paley projections that we proved was Bernstein's inequality, which says $\||\nabla|^s f_N\|_{L^p} \sim N^s \|f_N\|_{L^p}$. This should not be surprising, as differentiation of a function amount to multiplication by ξ in the frequency domain, and the function f_N is localized to frequencies $|\xi| \sim N$. Because the Littlewood-Paley projections interact with differentiation so nicely, they naturally have applications in partial differential equations. In this section, we prove a product rule and chain rule for the fractional derivative operator $|\nabla|^s$.

First, we provide a square function estimate involving the fractional derivative. This captures the notion that the derivative $|\nabla|^s$ can be expressed as a linear combination of Littlewood-Paley projection operators.

Proposition 7.9. *Let $1 < p < \infty$ and suppose that $\hat{f} \in C_c^\infty(\mathbb{R}^d \setminus \{0\})$. Then*

1. *For $s \in \mathbb{R}$,*
$$\||\nabla|^s f\|_{L^p} \sim \left\| \left(\sum_{N\in 2^{\mathbb{Z}}} |N^s f_N|^2\right)^{\frac{1}{2}}\right\|_{L^p}.$$

2. *For $s > 0$,*
$$\||\nabla|^s f\|_{L^p} \sim \left\| \left(\sum_{N\in 2^{\mathbb{Z}}} |N^s f_{\geq N}|^2\right)^{\frac{1}{2}}\right\|_{L^p}.$$

Proof.

1. First, consider the $\gtrsim$ inequality. We have
$$\left\| \left(\sum_{N\in 2^{\mathbb{Z}}} |N^s f_N|^2\right)^{\frac{1}{2}}\right\|_{L^p} = \left\| \left(\sum_{N\in 2^{\mathbb{Z}}} |N^s |\nabla|^{-s} |\nabla|^s f_N|^2\right)^{\frac{1}{2}}\right\|_{L^p}.$$

Also, $|\nabla|^s f_N = |\nabla|^s P_N f = P_N(|\nabla|^s f)$, as these operators are given by multiplication on the Fourier side. Thus,

$$\left\| \left(\sum_{N \in 2^{\mathbb{Z}}} |N^s f_N|^2 \right)^{\frac{1}{2}} \right\|_{L^p} = \left\| \left(\sum_{N \in 2^{\mathbb{Z}}} |N^s |\nabla|^{-s} P_N(|\nabla|^s f)|^2 \right)^{\frac{1}{2}} \right\|_{L^p} .$$

Recall that in the proof of the Littlewood-Paley square function estimate from, we proved one direction in a greater generality; we showed that $\|Sf\|_{L^p} \lesssim \|f\|_{L^p}$ where S is defined via any $\psi \in C_c^\infty(\mathbb{R}^d \setminus \{0\})$. In particular, this holds for $\tilde{\psi}(\xi) := |\xi|^{-s}\psi(\xi)$, where ψ is the usual Littlewood-Paley ψ. Because $\psi \in C_c^\infty(\mathbb{R}^d \setminus \{0\})$, $\tilde{\psi} \in C_c^\infty(\mathbb{R}^d \setminus \{0\})$. Defining $\tilde{\psi_N}(\xi) := N^s|\xi|^{-s}\psi(\xi/N)$, the general square function estimate gives

$$\left\| \left(\sum_{N \in 2^{\mathbb{Z}}} |N^s |\nabla|^{-s} P_N(|\nabla|^s f)|^2 \right)^{\frac{1}{2}} \right\|_{L^p} \lesssim \||\nabla|^s f\|_{L^p} .$$

Next, we show $\lesssim$, using duality and the fattened Littlewood-Paley projections.

$$\begin{aligned}
\||\nabla|^s f\|_{L^p} &= \sup_{\|g\|_{L^{p'}}=1} \langle |\nabla|^s f, g\rangle = \sup_{\|g\|_{L^{p'}}=1} \left\langle \sum P_N(|\nabla|^s f), g \right\rangle \\
&= \sup_{\|g\|_{L^{p'}}=1} \left\langle \sum |\nabla|^s \tilde{P_N} P_N f, g \right\rangle \\
&= \sup_{\|g\|_{L^{p'}}=1} \sum \left\langle P_N f, |\nabla|^s \tilde{P_N} g \right\rangle \\
&= \sup_{\|g\|_{L^{p'}}=1} \sum \left\langle N^s P_N f, N^{-s} |\nabla|^s \tilde{P_N} g \right\rangle .
\end{aligned}$$

Applying Cauchy-Schwarz and then Holder,

$$\begin{aligned}
\||\nabla|^s f\|_{L^p} &\le \sup_{\|g\|_{L^{p'}}=1} \left\langle \left(\sum |N^2 f_N|^2\right)^{\frac{1}{2}}, \left(\sum |N^{-s}|\nabla|^s \tilde{P_N} g|^2\right)^{\frac{1}{2}} \right\rangle \\
&\le \sup_{\|g\|_{L^{p'}}=1} \left\| \left(\sum |N^s f_N|^2\right)^{\frac{1}{2}} \right\|_{L^p} \left\| \left(\sum |N^{-s}|\nabla|^s \tilde{P_N} g|^2\right)^{\frac{1}{2}} \right\|_{L^{p'}} .
\end{aligned}$$

By the same remark from above,

$$\left\| \left(\sum |N^{-s}|\nabla|^s \tilde{P_N} g|^2\right)^{\frac{1}{2}} \right\|_{L^{p'}} \lesssim \|g\|_{L^{p'}} .$$

Thus, $\||\nabla|^s f\|_{L^p} \lesssim \left\| \left(\sum |N^s f_N|^2\right)^{\frac{1}{2}} \right\|_{L^p}$.

2. Note that $P_N = P_{\ge N} - P_{\ge 2N}$. Thus,

$$\begin{aligned}
\||\nabla|^s f\|_{L^p} &\sim \left\| \left(\sum |N^s f_N|^2\right)^{\frac{1}{2}} \right\|_{L^p} \\
&\le \left\| \left(\sum |N^s f_{\ge N}|^2\right)^{\frac{1}{2}} \right\|_{L^p} + \left\| \left(\sum |N^s f_{\ge 2N}|^2\right)^{\frac{1}{2}} \right\|_{L^p}
\end{aligned}$$

where we have invoked the ℓ^2 triangle inequality followed by the L^p triangle inequality. But up to a factor of 2,

$$\left\| \left(\sum |N^s f_{\ge 2N}|^2\right)^{\frac{1}{2}} \right\|_{L^p} \lesssim \left\| \left(\sum |N^s f_{\ge N}|^2\right)^{\frac{1}{2}} \right\|_{L^p} .$$

So

$$\| |\nabla|^s f \|_{L^p} \lesssim \left\| \left(\sum |N^s f_{\geq N}|^2 \right)^{\frac{1}{2}} \right\|_{L^p}.$$

Next, we show $\gtrsim$. We have

$$\begin{aligned} \sum |N^s f_{\geq N}|^2 = \sum N^{2s} |f_{\geq N}|^2 &\leq \sum_{N \in 2^{\mathbb{Z}}} N^{2s} \sum_{K \geq N} |f_K|^2 \\ &\leq \sum_{N \in 2^{\mathbb{Z}}} N^{2s} \left(\sum_{N_1 \geq N} |f_{N_1}| \right) \left(\sum_{N_2 \geq N} |f_{N_2}| \right). \end{aligned}$$

Multiplying the latter sums out, rearranging, and picking up a factor of 2 from symmetry gives

$$\sum |N^s f_{\geq N}|^2 \leq 2 \sum_{N \in 2^{\mathbb{Z}}} N^{2s} \sum_{N \leq N_1 \leq N_2} |f_{N_1}| |f_{N_2}|.$$

Cleverly inserting constants yields

$$\begin{aligned} \sum |N^s f_{\geq N}|^2 &\leq 2 \sum_{N \in 2^{\mathbb{Z}}} N^{2s} \sum_{N \leq N_1 \leq N_2} \frac{1}{N_1^s N_2^s} |N_1^s f_{N_1}| |N_2^s f_{N_2}| \\ &= 2 \sum_{N, N_1, N_2;\, N \leq N_1 \leq N_2} \frac{N^{2s}}{N_1^s N_2^s} |N_1^s f_{N_1}| |N_2^s f_{N_2}|. \end{aligned}$$

Freeze N_1 and N_2 and consider the sum over N. This is a dyadic sum, and by the usual argument we can bound this by plugging in the largest term, which is N_1. Thus,

$$\sum |N^s f_{\geq N}|^2 \lesssim \sum_{N_1 \leq N_2} \left(\frac{N_1}{N_2} \right)^s |N_1^s f_{N_1}| |N_2^s f_{N_2}|.$$

This is summable in both N_1 and N_2, since $s > 0$. By Schur's test, it follows that

$$\sum |N^s f_{\geq N}|^2 \lesssim \left(\sum_{N_2} |N_2^s f_{N_2}|^2 \right)^{\frac{1}{2}} \left(\sum_{N_1} |N_1^s f_{N_1}|^2 \right)^{\frac{1}{2}} = \sum |N^s f_N|^2.$$

Taking the square root of both sides and then the L^p norm gives

$$\left\| \left(\sum |N^s f_{\geq N}|^2 \right)^{\frac{1}{2}} \right\|_{L^p} \lesssim \left\| \left(\sum |N^s f_N|^2 \right)^{\frac{1}{2}} \right\|_{L^p}.$$

Applying the result from the previous part gives the desired result.

□

Using these estimates, we now prove the product rule and chain rule for the operator $|\nabla|^s$, both results due to Christ and Weinstein [1]. As $|\nabla|^s$ is a non-local operator, these rules are given as L^p norm estimates, rather than as pointwise facts.

Theorem 7.10 (Fractional product rule)**.** *Let* $1 < p < \infty$ *and* $s > 0$*. Then*

$$\| |\nabla|^s (fg) \|_{L^p} \lesssim \| |\nabla|^s f \|_{L^{p_1}} \|g\|_{L^{p_2}} + \| |\nabla|^s g \|_{L^{q_1}} \|f\|_{L^{q_2}}$$

where $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{q_1} + \frac{1}{q_2}$.

Proof. The previous proposition gives $\||\nabla|^s(fg)\|_{L^p} \sim \left\|\left(\sum N^{2s}|P_N(fg)|^2\right)^{\frac{1}{2}}\right\|_{L^p}$. We perform a *paraproduct decomposition* on fg:

$$fg = f_{\geq\frac{N}{8}}g + f_{<\frac{N}{8}}g = f_{\geq\frac{N}{8}}g + f_{<\frac{N}{8}}g_{\geq\frac{N}{8}} + f_{<\frac{N}{8}}g_{<\frac{N}{8}}.$$

Then

$$P_N(fg) = P_N(f_{\geq\frac{N}{8}}g) + P_N(f_{<\frac{N}{8}}g_{\geq\frac{N}{8}}) + P_N(f_{<\frac{N}{8}}g_{<\frac{N}{8}}).$$

Note that $f_{<\frac{N}{8}}$ and $g_{<\frac{N}{8}}$ have frequency supports bounded by $2\frac{N}{8} = \frac{N}{4}$. Thus, the maximum frequency attained by $f_{<\frac{N}{8}}g_{<\frac{N}{8}}$ is bounded by $\frac{N}{4} + \frac{N}{4} = \frac{N}{2}$. Thus, $P_N(f_{<\frac{N}{8}}g_{<\frac{N}{8}}) = 0$. Thus, it remains to consider the first two terms.

Using the fact that the Littlewood-Paley projections are bounded by the maximal function,

$$|P_N(fg)| \lesssim M(f_{\geq\frac{N}{8}}g) + M(f_{<\frac{N}{8}}g_{\geq\frac{N}{8}}) \lesssim M(f_{\geq\frac{N}{8}}g) + M\left[(Mf)\, g_{\geq\frac{N}{8}}\right].$$

Therefore, the ℓ^2 and L^p triangle inequalities yield

$$\begin{aligned}&\left\|\left(\sum N^{2s}|P_N(fg)|^2\right)^{\frac{1}{2}}\right\|_{L^p}\\ &\qquad\lesssim \left\|\left(\sum \left|M(N^s f_{\geq\frac{N}{8}}g)\right|^2\right)^{\frac{1}{2}}\right\|_{L^p} + \left\|\left(\sum \left|M\left[(Mf)\, N^s g_{\geq\frac{N}{8}}\right]\right|^2\right)^{\frac{1}{2}}\right\|_{L^p}.\end{aligned}$$

Next, apply the vector-valued Hardy-Littlewood maximal inequality to both norms:

$$\begin{aligned}\left\|\left(\sum N^{2s}|P_N(fg)|^2\right)^{\frac{1}{2}}\right\|_{L^p} &\lesssim \left\|\left(\sum \left|N^s f_{\geq\frac{N}{8}}g\right|^2\right)^{\frac{1}{2}}\right\|_{L^p} + \left\|\left(\sum \left|(Mf)\, N^s g_{\geq\frac{N}{8}}\right|^2\right)^{\frac{1}{2}}\right\|_{L^p}\\ &= \left\||g|\left(\sum \left|N^s f_{\geq\frac{N}{8}}\right|^2\right)^{\frac{1}{2}}\right\|_{L^p} + \left\||Mf|\left(\sum \left|N^s g_{\geq\frac{N}{8}}\right|^2\right)^{\frac{1}{2}}\right\|_{L^p}.\end{aligned}$$

Applying Holder's inequality,

$$\begin{aligned}&\left\|\left(\sum N^{2s}|P_N(fg)|^2\right)^{\frac{1}{2}}\right\|_{L^p}\\ &\qquad\lesssim \|g\|_{L^{p_2}}\left\|\left(\sum \left|N^s f_{\geq\frac{N}{8}}\right|^2\right)^{\frac{1}{2}}\right\|_{L^{p_1}} + \|Mf\|_{L^{q_2}}\left\|\left(\sum \left|N^s g_{\geq\frac{N}{8}}\right|^2\right)^{\frac{1}{2}}\right\|_{L^{q_1}}.\end{aligned}$$

Since the maximal function is of type (q_2, q_2), $\|Mf\|_{L^{q_2}} \lesssim \|f\|_{L^{q_2}}$. Applying part 2 of the previous proposition to the p_1 and q_1 norms yields

$$\||\nabla|^s(fg)\|_{L^p} \sim \left\|\left(\sum N^{2s}|P_N(fg)|^2\right)^{\frac{1}{2}}\right\|_{L^p} \lesssim \|g\|_{L^{p_2}}\, \||\nabla|^s f\|_{L^{p_1}} + \|f\|_{L^{q_2}}\, \||\nabla|^s g\|_{L^{q_1}}.$$

□

Theorem 7.11 (Fractional chain rule). *Suppose $F : \mathbb{C} \to \mathbb{C}$ satisfies*

$$|F(u) - F(v)| \lesssim |u - v| \cdot |G(u) - G(v)|$$

for all functions $u, v : \mathbb{R}^d \to \mathbb{C}$ and for some function $G : \mathbb{C} \to [0, \infty)$. Then for $1 < p < \infty$ and $0 < s < 1$,

$$\||\nabla|^s F(u)\|_{L^p} \lesssim \||\nabla|^s u\|_{L^{p_1}}\, \|G(u)\|_{L^{p_2}}$$

where $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$ and $1 < p_2 \leq \infty$.

Proof. As before, the previous proposition yields

$$\||\nabla|^s F(u)\|_{L^p} \sim \left\| \left(\sum N^{2s} |P_N F(u)|^2\right)^{\frac{1}{2}} \right\|_{L^p}.$$

Observe that the function $\check{\psi}$ is a mean zero function. Indeed,

$$\int_{\mathbb{R}^d} \check{\psi}(x)\, dx = \hat{\check{\psi}}(0) = \psi(0) = 0.$$

We use this and the fact that convolving with a mean zero function is like differentiation as follows:

$$\begin{aligned} P_N F(u)(x) &= (F(u) * N^d \check{\psi}(N\cdot))(x) = N^d \int \check{N}y F(u(x-y))\, dy \\ &= N^d \int \check{\psi}(Ny) \left[F(u(x-y)) - F(u(x))\right]\, dy \end{aligned}$$

and so

$$|P_N F(u)(x)| \lesssim N^d \int |\check{\psi}(Ny)|\, |u(x-y) - u(x)|\, |G(u(x-y)) - G(u(x)|\, dy.$$

Decomposing $|u(x-y) - u(x)|$ based on where we expect cancellation, we have

$$|u(x-y) - u(x)| \le |u_{>N}(x-y)| + |u_{>N}(x)| + \sum_{K \le N} |u_K(x-y) - u_K(x)|.$$

We proceed by considering each of these terms separately. First, we claim that

$$|u_K(x-y) - u_K(x)| \lesssim K|y| \left|(Mu_K)(x-y) + (Mu_K)(x)\right|.$$

Indeed, if $K|y| \gtrsim 1$, then

$$\begin{aligned} |u_K(x-y) - u_K(x)| \le |u_K(x-y)| + |u_K(x)| &\lesssim |(Mu_K)(x-y) + (Mu_K)(x)| \\ &\lesssim K|y| \left|(Mu_K)(x-y) + (Mu_K)(x)\right|. \end{aligned}$$

If $K|y| << 1$, then

$$\begin{aligned} u_K(x-y) - u_K(x) &= \tilde{P_K}(u_K(x-y) - u_K(x)) \\ &= K^d \int \check{\tilde{\psi}}(Ky) \left[u_K(x-y-z) - u_K(x-z)\right]\, dz \\ &= \int K^d \left[\check{\tilde{\psi}}(K(z-y)) - \check{\tilde{\psi}}(Kz)\right] u_K(x-z)\, dz \end{aligned}$$

via the change of variables $y + z \mapsto z$. By the fundamental theorem of calculus,

$$|u_K(x-y) - u_K(x)| \le \int K^d \int_0^1 K|y| \left|\nabla\check{\psi}(Kz - \theta K y)\right|\, d\theta\, |u_K(x-z)|\, dz.$$

Because $\check{\tilde{\psi}}$ is Schwartz and thus is bounded and decays as quickly as we need, we have

$$|u_K(x-y) - u_K(x)| \lesssim K|y| \int K^d \frac{1}{\langle Kz \rangle^{100d}} |u_K(x-z)|\, dz.$$

In previous arguments, by considering $z < 1/K$ and $z > 1/K$ separately, we have seen that

$$\int K^d \frac{1}{\langle Kz \rangle^{100d}} |u_K(x-z)|\, dz \lesssim (Mu_K)(x).$$

Our claim then follows from this fact.

Next, we consider the contribution of $|u_{>N}(x-y)|$ to $|P_N F(u)(x)|$. The contribution is bounded by

$$\int N^d \,|\check{\psi}(Ny)|\,|u_{>N}(x-y)|\,[G(u)(x-y)+G(u)(x)]\; dy.$$

By the same trick as above, use the quickly decaying and bounded nature of $\check{\psi}$ to estimate this integral by

$$M\left(u_{>N}\cdot G(u)\right)(x)+G(u)(x)\cdot M(u_{>N})(x) \lesssim M\left(u_{>N}\cdot G(u)\right)(x)+M(G(u))(x)\cdot M(u_{>N})(x).$$

Similarly, the contribution of $|u_{>N}(x)|$ to $|P_N F(u)(x)|$ is bounded by

$$\begin{aligned}\int N^d \,|\check{\psi}(Ny)|\,|u_{>N}(x)|\,&[G(u)(x-y)+G(u)(x)]\; dy\\ &\lesssim |u_{>N}(x)|\cdot M(G(u))(x)+|u_{>N}(x)|\cdot G(u)(x)\\ &\qquad\qquad\lesssim M(u_{>N})(x)\cdot M(G(u))(x).\end{aligned}$$

Thus, the contribution of $|u_{>N}(x-y)|+|u_{>N}(x)|$ to $|P_N F(u)(x)|$ is bounded by

$$M\left(u_{>N}\cdot G(u)\right)(x)+M(G(u))(x)\cdot M(u_{>N})(x).$$

Estimate the contribution of $M\left(u_{>N}\cdot G(u)\right)(x)$ to $\left\| |\nabla|^s F(u)\right\|_{L^p}$ as follows:

$$\begin{aligned}\left\|\left(\sum_{N\in 2^{\mathbb{Z}}} N^{2s}\left|M(u_{>N}G(u))\right|^2\right)^{\frac{1}{2}}\right\|_{L^p} &= \left\|\left(\sum_{N\in 2^{\mathbb{Z}}} \left|M(N^s u_{>N}G(u))\right|^2\right)^{\frac{1}{2}}\right\|_{L^p}\\ &\lesssim \left\|\left(\sum_{N\in 2^{\mathbb{Z}}} \left|N^s u_{>N}G(u)\right|^2\right)^{\frac{1}{2}}\right\|_{L^p}\end{aligned}$$

using the vector-valued Hardy-Littlewood maximal inequality. Next, pull the $G(u)$ out of the sum and apply Holder's inequality followed by our derivative estimate from earlier:

$$\|G(u)\|_{L^{p_2}}\left\|\left(\sum_{N\in 2^{\mathbb{Z}}} \left|N^s u_{>N}\right|^2\right)^{\frac{1}{2}}\right\|_{L^{p_1}} \lesssim \|G(u)\|_{L^{p_2}}\left\||\nabla|^s u\right\|_{L^{p_1}}.$$

The contribution of $M(G(u))(x)\cdot M(u_{>N})(x)$ to $\left\||\nabla|^s F(u)\right\|_{L^p}$ is almost the exact same calculation.

Thus, it only remains to bound the contribution of the $\sum_{K\leq N}|u_K(x-y)-u_K(x)|$ term.

The contribution to $|P_N F(u)|$ is

$$
\begin{aligned}
\sum_{K\le N} &\int N^d |\check{\psi}(Ny)||u_K(x-y)-u_K(x)||G(u)(x-y)-G(u)(x)|\,dy \\
&\lesssim \sum_{K\le N} \int N^d |\check{\psi}(Ny)|K|y|\,|(Mu_K)(x-y)+(Mu_K)(x)|\,|G(u)(x-y)-G(u)(x)|\,dy \\
&= \sum_{K\le N} \frac{K}{N} \int N^d |\check{\psi}(Ny)|N|y|\,|(Mu_K)(x-y)+(Mu_K)(x)|\,|G(u)(x-y)-G(u)(x)|\,dy \\
&\lesssim \sum_{K\le N} \frac{K}{N} [M(M(u_K)G(u))(x) + M(M(u_K))(x)\,G(u)(x) \\
&\qquad\qquad + (M(u_K)M(G(u)))(x) + (M(u_K)G(u))(x)] \\
&\lesssim \sum_{K\le N} \frac{K}{N} [M(M(u_K)G(u))(x) + M(M(u_K))(x)\,G(u)(x)] \\
&=: \sum_{K\le N} \frac{K}{N} C_K.
\end{aligned}
$$

Towards the goal of estimating the contribution of *this* term to $\||\nabla|^s F(u)\|_{L^p}$, we first compute:

$$
\begin{aligned}
\sum_N N^{2s} \left| \sum_{K\le N} \frac{K}{N} C_K \right|^2 &= \sum_N \left(\sum_{K\le N} \frac{K}{N} C_K \right) \left(\sum_{L\le N} \frac{L}{N} C_L \right) \\
&\le 2 \sum_N N^{2s} \sum_{K\le L\le N} \frac{KL}{N^2} C_K C_L
\end{aligned}
$$

where the factor of 2 comes from the symmetry of the summations. Inserting K^s and L^s terms and then summing in N gives

$$
\begin{aligned}
&= 2 \sum_{N,K,L;\, K\le L\le N} N^{2(s-1)} K^{1-s} L^{1-s} (K^s C_K)(L^s C_L) \\
&\lesssim \sum_{K\le L} L^{2(s-1)} K^{1-s} L^{1-s} (K^s C_K)(L^s C_L) \\
&= \sum_{K\le L} \left(\frac{K}{L} \right)^{1-s} (K^s C_K)(L^s C_L).
\end{aligned}
$$

This is summable in both K and L. Hence, by Schur's test,

$$
\sum_N N^{2s} \left| \sum_{K\le N} \frac{K}{N} C_K \right|^2 \lesssim \left(\sum_K (K^s C_K)^2 \right)^{\frac{1}{2}} \left(\sum_L (L^s C_L)^2 \right)^{\frac{1}{2}} = \sum_K K^{2s} C_K^2.
$$

Using this, we can estimate the contribution to $\||\nabla|^s F(u)\|_{L^p}$ by

$$
\left\| \left(\sum_K K^{2s} \, |M(M(u_K)\,G(u))|^2 \right)^{\frac{1}{2}} \right\|_{L^p} + \left\| \left(\sum_K K^{2s} \, |M(M(u_K))\,M(G(u))|^2 \right)^{\frac{1}{2}} \right\|_{L^p}.
$$

The rest of the computation uses the vector valued maximal inequality in exactly the same manner as above. To avoid unnecessary repetition, the details are left to the reader. □

An example of such a function F which appears in applications in partial differential equations is $F(u) = |u|^p u$. Indeed, by the fundamental theorem of calculus,

$$|F(u) - F(v)| \leq |u - v| \cdot ||u|^p + |v|^p| \, .$$

Nonlinearities in equations such as the Schrodinger equation often take this form.

Also, this fractional product rule is useful for the case $s > 1$, as we can typically write derivative operators as an integer part plus a fractional part $0 < s < 1$.

Chapter 8

Oscillatory Integrals

In this chapter, we develop the basic tools to study oscillatory integrals. The fundamental example of an oscillatory integral is the Fourier transform:

$$\int_{\mathbb{R}^d} e^{-2\pi i x\cdot\xi} f(x)\, dx.$$

The exponential term in the above integral has unit modulus, but nontrivial phase oscillation. The most naive way to estimate the size of a particular Fourier transform is to use the integral triangle inequality:

$$\left|\int_{\mathbb{R}^d} e^{-2\pi i x\cdot\xi} f(x)\, dx\right| \le \int_{\mathbb{R}^d} |e^{-2\pi i x\cdot\xi} f(x)|\, dx = \int_{\mathbb{R}^d} |f(x)|\, dx.$$

Though useful in some contexts, this estimate is crude because it ignores any oscillation coming from the phase $e^{-2\pi i x\cdot\xi}$. For example, if the function f is not integrable, estimating in this way is useless. Oftentimes, taking advantage of phase oscillation and cancellation is a more effective way to bound the size of an integral like the Fourier transform.

More generally, we define two different notions of oscillatory integrals.

Definition 8.1. An **oscillatory integral of the first kind** is an integral of the form

$$I(\lambda) = \int_{\mathbb{R}^d} e^{i\lambda\phi(x)}\, \psi(x)\, dx$$

where $\phi : \mathbb{R}^d \to \mathbb{R}$, $\psi : \mathbb{R}^d \to \mathbb{C}$, and $\lambda > 0$.

The main goal of this chapter is to understand the asymptotic behavior of these integrals as $\lambda \to \infty$. We will eventually see applications of this to various PDEs.

Oscillatory integrals of the first kind will be our only focus, but we provide the following definition for the sake of completeness.

Definition 8.2. An **oscillatory integral of the second kind** is an operator of the form

$$(T_\lambda f)(x) = \int_{\mathbb{R}^d} e^{i\lambda\phi(x,y)}\, K(x,y)\, f(y)\, dy$$

where $\phi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, $K : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$, $f : \mathbb{R}^d \to \mathbb{C}$, and $\lambda > 0$.

In this case, it is desirable to understand the asymptotic behavior of $\|T_\lambda\|_{op}$ as $\lambda \to \infty$. We refer to [10] for more details on oscillatory integrals of the second kind, as well as the other material in this chapter.

8.1 Oscillatory Integrals of the First Kind, $d = 1$

We begin by studying oscillatory integrals of the first kind in one dimension. In this setting, the theory is completely understood and admits a simpler presentation. There are three main techniques presented in this section: nonstationary phase, the van der Corput lemma, and stationary phase.

We begin with nonstationary phase. Roughly, the principle of nonstationary phase says that if the phase $e^{i\lambda\phi(x)}$ is nonstationary in the sense that $\phi'(x) \neq 0$, then the oscillatory integral $I(\lambda)$ is rapidly decaying. In other words, integrating a wildly varying phase leads to cancellation, and hence decay. This should make sense from the perspective of adding sinusoidal waves: if the waves have similar phases, their sum amplifies, whereas summing waves with different phases leads to cancellation in amplitude.

More directly, the principle of nonstationary phase uses the compact support of ψ together with integration by parts to achieve decay up to any order.

Proposition 8.3 (Nonstationary phase)**.** *Let $\phi : \mathbb{R} \to \mathbb{R}$ and $\psi : \mathbb{R} \to \mathbb{C}$ be smooth with ψ compactly supported in (a, b). Assume $\phi'(x) \neq 0$ for all $x \in [a, b]$. Then*

$$\left| \int_a^b e^{i\lambda\phi(x)}\psi(x)\, dx \right| \lesssim_N \lambda^{-N} \tag{8.1}$$

for any $N \geq 0$. Here the implicit constant does not depend on a and b.

Proof. As remarked, the proof of this proposition is simply repeated applications of integration by parts. The details are as follows.

Because $\phi'(x) \neq 0$, we have $e^{i\lambda\phi(x)} = \frac{1}{i\lambda\phi'(x)}\frac{d}{dx}e^{i\lambda\phi(x)}$. Let $D := \frac{1}{i\lambda\phi'(x)}\frac{d}{dx}$. Then $e^{i\lambda\phi(x)} = D^N\left(e^{i\lambda\phi(x)}\right)$ for any $N \geq 0$. To integrate by parts with D, we need to compute its adjoint. For any smooth f, normal integration by parts gives

$$\int_{\mathbb{R}} Df(x)\,\psi(x)\,dx = \int_{\mathbb{R}} \frac{1}{i\lambda\phi'(x)}\frac{df}{dx}(x)\,\psi(x)\,dx = -\int_{\mathbb{R}} f(x)\frac{d}{dx}\left(\frac{1}{i\lambda\phi'(x)}\,\psi(x)\right)\,dx.$$

Here, the boundary terms vanish because of the compact support of ψ.

Thus, the adjoint of D is given by

$${}^tDf(x) := -\frac{d}{dx}\left[\frac{1}{i\lambda\phi'(x)}f(x)\right].$$

We have

$$\begin{aligned}\int_a^b e^{i\lambda\phi(x)}\psi(x)\,dx &= \int_a^b D^N(e^{i\lambda\phi(x)})\psi(x)\,dx = \int_a^b e^{i\lambda\phi(x)}\,({}^tD)^N\psi(x)\,dx \\ &= \int_a^b e^{i\lambda\phi(x)}\left[-\frac{d}{dx}\frac{1}{i\lambda\phi'(x)}\right]^N(\psi)(x)\,dx.\end{aligned}$$

Therefore,

$$\left|\int_a^b e^{i\lambda\phi(x)}\psi(x)\,dx\right| \leq \lambda^{-N}\int_a^b \left|\left[\frac{d}{dx}\frac{1}{\phi'(x)}\right]^N(\psi)(x)\right|\,dx.$$

Inside this integral, the derivative can hit either the $\phi'(x)$ in the denominator or the $\psi(x)$. So

$$\int_a^b \left|\left[\frac{d}{dx}\frac{1}{\phi'(x)}\right]^N(\psi)(x)\right|\,dx$$

is some finite constant which depends on $N + 1$ derivatives of $\phi(x)$ and N derivatives of $\psi(x)$. Thus,

$$\left|\int_a^b e^{i\lambda\phi(x)}\psi(x)\,dx\right| \lesssim_N \lambda^{-N}$$

for any $N \geq 0$, where the implicit constant depends on $N+1$ derivatives of $\phi(x)$ and N derivatives of $\psi(x)$. $\square$

If ψ does not have compact support, the best decay in λ we can expect is λ^{-1}. For example, when $\psi(x) = 1$ and $\phi(x) = x$,

$$\left| \int_a^b e^{i\lambda x}\, dx \right| = \left| \frac{e^{i\lambda b} - e^{i\lambda a}}{i\lambda} \right| \lesssim \lambda^{-1}.$$

Summarizing our analysis so far, we have the following informal heuristics:

$$\left\{ \text{ control of } \phi' \right\} \qquad \Rightarrow \qquad I(\lambda) \text{ decays like } \lambda^{-1}$$

whereas

$$\left\{ \begin{array}{c} \text{control of } \phi' \\ \& \\ \text{compact support of } \psi \end{array} \right\} \qquad \Rightarrow \qquad I(\lambda) \text{ decays to any order.}$$

The first heuristic is suggested by the preceding example, and the second heuristic is the principle of nonstationary phase. Here, "control" means bounding *away* from 0. As ϕ determines the phase oscillation of $I(\lambda)$, these ideas suggest that a lot of oscillation leads to a lot of cancellation, hence decay.

The next principle is a generalization of the first heuristic, which analyses the decay of $I(\lambda)$ if we only have control of the k-th derivative of ϕ, rather than ϕ'.

Lemma 8.4 (van der Corput). *Let $\phi : \mathbb{R} \to \mathbb{R}$ be smooth. Fix $k \geq 1$, and assume that $|\phi^{(k)}(x)| \geq 1$. If $k = 1$, assume further that ϕ' is monotone. Then*

$$\left| \int_a^b e^{i\lambda\phi(x)}\, dx \right| \lesssim_k \lambda^{-\frac{1}{k}} \tag{8.2}$$

where the implicit constant does not depend on a or b.

Proof. We proceed by induction on k.

Consider the $k = 1$ case. Because $\phi' \neq 0$,

$$\int_a^b e^{i\lambda\phi(x)}\, dx = \int_a^b \frac{1}{i\lambda\phi(x)} \frac{d}{dx} \left[e^{i\lambda\phi'(x)} \right] dx.$$

Integrating by parts (with boundary terms!) gives

$$\begin{aligned}
\int_a^b e^{i\lambda\phi(x)}\, dx &= -\int_a^b \frac{d}{dx}\left[\frac{1}{i\lambda\phi'(x)}\right] e^{i\lambda\phi(x)}\, dx \;+\; \frac{e^{i\lambda\phi(b)}}{i\lambda\phi'(b)} - \frac{e^{i\lambda\phi(a)}}{i\lambda\phi'(a)} \\
&= \frac{1}{\lambda}\left[\frac{e^{i\phi(b)}}{i\lambda\phi'(b)} - \frac{e^{i\lambda\phi(a)}}{i\phi'(a)}\right] - \frac{1}{i\lambda}\int_a^b e^{i\lambda\phi(x)} \frac{d}{dx}\left[\frac{1}{\phi'(x)}\right] dx.
\end{aligned}$$

Thus,

$$\left| \int_a^b e^{i\lambda\phi(x)}\, dx \right| \lesssim \frac{1}{\lambda} + \frac{1}{\lambda}\int_a^b \left| \frac{d}{dx}\frac{1}{\phi'(x)} \right| dx.$$

Because ϕ' is monotone, $\frac{d}{dx}\frac{1}{\phi'(x)} = \pm\left|\frac{d}{dx}\frac{1}{\phi'(x)}\right|$. So we can estimate further by

$$\begin{aligned}
\left| \int_a^b e^{i\lambda\phi(x)}\, dx \right| &\lesssim \frac{1}{\lambda} + \frac{1}{\lambda}\int_a^b \left| \frac{d}{dx}\frac{1}{\phi'(x)} \right| dx \\
&= \frac{1}{\lambda} + \frac{1}{\lambda}\left| \int_a^b \frac{d}{dx}\frac{1}{\phi'(x)}\, dx \right| \\
&= \frac{1}{\lambda} + \frac{1}{\lambda}\left| \frac{1}{\phi'(b)} - \frac{1}{\phi'(a)} \right| \\
&\lesssim \frac{1}{\lambda}.
\end{aligned}$$

This establishes the base case $k = 1$.

Next, assume by induction that the desired estimate holds for some $k \geq 1$. Suppose that $|\phi^{(k+1)}(x)| \geq 1$. By replacing ϕ with $-\phi$ if necessary, we may assume without loss of generality that $\phi^{(k+1)}(x) \geq 1$. Then $\phi^{(k)}$ is strictly increasing, and consequently there is at most one point $c \in [a, b]$ such that $\phi^{(k)}(c) = 0$.

Case 1. First, consider the case where there is such a c. Since $\phi^{(k+1)} \geq 1$, $|\phi^{(k)}(x)| \geq \delta$ for all x of distance at least δ from c, i.e., for all $x \in [a, b] \setminus (c - \delta, c + \delta)$. Close to c, we have the trivial estimate

$$\left| \int_{[a,b]\cap(c-\delta,c+\delta)} e^{i\lambda\phi(x)}\, dx \right| \leq 2\delta.$$

Away from c, we estimate as follows. Because $|\phi^{(k)}(x)| \geq \delta$ on the complement of $(c - \delta, c + \delta)$, we can rescale ϕ to use the inductive hypothesis. Let $\phi_\delta(x) := \phi(\delta^{-\frac{1}{k}} x)$. Then $\phi_\delta^{(k)}(x) = \delta^{-1}\phi^{(k)}(\delta^{-\frac{1}{k}} x)$ and hence $|\phi_\delta^{(k)}(x)| \geq 1$ outside of $(c - \delta, c + \delta)$. Thus,[1]

$$\begin{aligned} \left| \int_a^{c-\delta} e^{i\lambda\phi(x)}\, dx \right| &= \left| \delta^{-\frac{1}{k}} \int_{\delta^{-\frac{1}{k}} a}^{\delta^{-\frac{1}{k}}(c-\delta)} e^{i\lambda\phi(\delta^{-\frac{1}{k}} y)}\, dy \right| = \left| \delta^{-\frac{1}{k}} \int_{\delta^{-\frac{1}{k}} a}^{\delta^{-\frac{1}{k}}(c-\delta)} e^{i\lambda\phi_\delta(y)}\, dy \right| \\ &\lesssim \delta^{-\frac{1}{k}} \lambda^{-\frac{1}{k}}. \end{aligned}$$

The same computation gives $\left| \int_{c+\delta}^b e^{i\lambda\phi(x)}\, dx \right| \lesssim \delta^{-\frac{1}{k}} \lambda^{-\frac{1}{k}}$. Combining these estimates with the first trivial estimate yields

$$\left| \int_a^b e^{i\lambda\phi(x)}\, dx \right| \lesssim \delta + \delta^{-\frac{1}{k}} \lambda^{-\frac{1}{k}}$$

for any $\delta > 0$. We optimize the choice of δ by requiring $\delta = \delta^{-\frac{1}{k}} \lambda^{-\frac{1}{k}}$, and thus $\delta = \lambda^{-\frac{1}{k+1}}$. This gives

$$\left| \int_a^b e^{i\lambda\phi(x)}\, dx \right| \lesssim_k \lambda^{-\frac{1}{k+1}}$$

as desired.

Case 2. Next, suppose that no such c exists, i.e., $\phi^{(k)}(x) \neq 0$ for all $x \in [a, b]$. It is either the case that $\phi^{(k)}(a) > 0$ or $\phi^{(k)}(a) < 0$. Suppose first that $\phi^{(k)}(a) > 0$. Arguing as before, because $\phi^{(k+1)}(x) \geq 1$ we know that $\phi^{(k)}(x) \geq \delta$ for all $x \in (a + \delta, b]$. Then

$$\left| \int_a^b e^{i\lambda\phi(x)}\, dx \right| \leq \left| \int_a^{a+\delta} e^{i\lambda\phi(x)}\, dx \right| + \left| \int_{a+\delta}^b e^{i\lambda\phi(x)}\, dx \right|.$$

The first quantity is trivially bounded by δ. In the second integral, we rescale by δ and use the inductive hypothesis exactly as before to get $\left| \int_{a+\delta}^b e^{i\lambda\phi(x)}\, dx \right| \lesssim \delta^{-\frac{1}{k}} \lambda^{-\frac{1}{k}}$. Optimizing in δ by choosing $\delta = \lambda^{-\frac{1}{k+1}}$ gives

$$\left| \int_a^b e^{i\lambda\phi(x)}\, dx \right| \lesssim_k \lambda^{-\frac{1}{k+1}}$$

as desired.

Suppose now that $\phi^{(k)}(a) < 0$. Then $\phi^{(k)}(b) < 0$. The argument from here is analogous, this time decomposing $[a, b]$ into $[a, b - \delta]$ and $(b - \delta, b]$.

$\square$

[1] The following integral is understood to be 0 if $c - \delta \leq a$.

If $k = 1$, then the monotonicity assumption on ϕ' is in fact needed. Note that

$$\left| \int_a^b e^{i\phi(x)}\, dx \right| = \left| \int_a^b \cos(\phi(x)) + i \sin(\phi(x))\, dx \right| \geq \left| \int_a^b \sin(\phi(x))\, dx \right|.$$

It is possible to construct a function ϕ such that $\phi'(x)$ is large when $\sin(\phi(x))$ is small, and $\phi'(x)$ is small when $\sin(\phi(x))$ is large. Consequently,

$$|\{ x \ :\ \sin\phi(x) \}| << |\{ x \ :\ \sin\phi(x) > 0 \}|$$

and

$$\left| \int_a^b \sin(\phi(x))\, dx \right| \to \infty$$

as $|a|, |b| \to \infty$. We leave the details to the reader.

Finally, we come to the principle of stationary phase. By the principle of *non*stationary phase, $I(\lambda)$ decays as rapidly as we like when $\phi' \neq 0$. Thus, it remains to consider what happens near critical points of ϕ, i.e., where the phase is not oscillating rapidly.

The following proposition gives a complete description of the asymptotic behavior of $I(\lambda)$ near a well-behaved critical point.

Proposition 8.5 (Stationary phase)**.** *Let $\phi : \mathbb{R} \to \mathbb{R}$ be smooth, and assume that ϕ has a nondegenerate critical point*[2] *at x_0. If ψ is smooth and supported in a sufficiently small*[3] *neighborhood neighborhood of x_0, then*

$$I(\lambda) = \int_{\mathbb{R}} e^{i\lambda\phi(x)} \psi(x)\, dx$$

satisfies

$$I(\lambda) = \left[\sqrt{2\pi i}\, |\phi''(x_0)|^{-\frac{1}{2}}\, e^{i\frac{\pi}{4}\operatorname{sgn}\phi''(x_0)} e^{i\lambda\phi(x_0)} \psi(x_0) \right] \lambda^{-\frac{1}{2}} + O\left(\lambda^{-\frac{3}{2}}\right) \tag{8.3}$$

as $\lambda \to \infty$.

This statement of the principle of stationary phase gives an explicit constant in (8.3) that will prove to be useful. Before providing the proof of this fact, we can easily derive the general behavior if we ignore the explicit constant.

In particular, let $a \in C_c^\infty(\mathbb{R})$ be a smooth bump function satisfying $a(x) = 1$ for $|x| \leq 1$ and $a(x) = 0$ for $|x| \geq 2$. We use this bump function to localize close to and far away from x_0 as follows:

$$|I(\lambda)| \leq \left| \int_{\mathbb{R}} e^{i\lambda\phi(x)} \psi(x)\, a\left(\lambda^{\frac{1}{2}}(x - x_0)\right)\, dx \right| + \left| \int_{\mathbb{R}} e^{i\lambda\phi(x)} \psi(x) \left[1 - a\left(\lambda^{\frac{1}{2}}(x - x_0)\right)\right]\, dx \right|.$$

Consider the first integral. The support of the rescaled a function is $\lambda^{-\frac{1}{2}}$, so ignoring any oscillation and using the triangle inequality gives

$$\left| \int_{\mathbb{R}} e^{i\lambda\phi(x)} \psi(x)\, a\left(\lambda^{\frac{1}{2}}(x - x_0)\right)\, dx \right| \lesssim \lambda^{-\frac{1}{2}}.$$

Consider the second integral. As $\psi(x)$ has compact support and $1 - a\left(\lambda^{\frac{1}{2}}(x - x_0)\right)$ has support away from x_0, $\phi' \neq 0$ on the support of $\psi(x)\left[1 - a\left(\lambda^{\frac{1}{2}}(x - x_0)\right)\right]$. We would like to use the principle of nonstationary phase to estimate this term, but because the function

[2] That is, $\phi'(x_0) = 0$ and $\phi''(x_0) \neq 0$.

[3] The precise nature of *sufficiently small* will be clear from the proof. Primarily, the support of ψ must not contain any other critical points of ϕ.

$\psi(x)\left[1-a\left(\lambda^{\frac{1}{2}}(x-x_0)\right)\right]$ has support which scales with λ, the implicit constant in (8.1) depends on λ. However, as a is rescaled by a factor of $\lambda^{\frac{1}{2}}$, every iteration of integration by parts in the proof of nonstationary phase picks up a a factor of $\lambda^{-1}\cdot\lambda^{\frac{1}{2}}=\lambda^{-\frac{1}{2}}$. Thus,

$$\left|\int_{\mathbb{R}} e^{i\lambda\phi(x)}\psi(x)\,\left[1-a\left(\lambda^{\frac{1}{2}}(x-x_0)\right)\right]\,dx\right|\lesssim\lambda^{-\frac{N}{2}}$$

for any order N. We can certainly choose N large enough so that this term is $O\left(\lambda^{-\frac{3}{2}}\right)$. Thus,

$$|I(\lambda)|\lesssim\lambda^{-\frac{1}{2}}+O\left(\lambda^{-\frac{3}{2}}\right)$$

as desired.

Now we provide the full proof, keeping track of the explicit constants.

Proof of Proposition 8.5. Because $\phi'(x_0)=0$ and $\phi''(x_0)=0$, ϕ is well-approximated by its order-2 Taylor expansion near x_0. Expanding gives us

$$\begin{aligned}\phi(x)&=\phi(x_0)+(x-x_0)\phi'(x_0)+\frac{1}{2}(x-x_0)^2\phi''(x_0)+O(|x-x_0|^3)\\&=\phi(x_0)+\phi''(x_0)(x-x_0)^2\left(1+\eta(x)\right)\end{aligned}$$

where $\eta:\mathbb{R}\to\mathbb{R}$ is smooth and $\eta(x)=O(|x-x_0|)$.

Let U be an open neighborhood of x_0, chosen sufficiently small enough so that $|\eta(x)|<1$ for all $x\in U$, and $\phi(x)\neq 0$ for all $x\neq x_0\in U$. Suppose that the support of ψ is contained in U.

Change variables via $y=(x-x_0)\sqrt{1+\eta(x)}$. This is a diffeomorphism from U to a small neighborhood of $y=0$. Then

$$\begin{aligned}I(\lambda)&=\int_{\mathbb{R}} e^{i\lambda\left[\phi(x_0)+\frac{1}{2}\phi''(x_0)(x-x_0)^2(1+\eta(x))\right]}\psi(x)\,dx\\&=e^{i\lambda\phi(x_0)}\int_{\mathbb{R}} e^{i\frac{1}{2}\phi''(x_0)\lambda y^2}\tilde{\psi}(y)\,dy\end{aligned}$$

where $\tilde{\psi}\in C_c^\infty(\mathbb{R})$ is the appropriately transformed function under the indicated diffeomorphism. Note that $\tilde{\psi}$ has support contained the the small neighborhood surrounding $y=0$. Let $\tilde{\lambda}:=\frac{1}{2}\phi''(x_0)\lambda$ and let $\gamma\in C_c^\infty(\mathbb{R})$ be a function which is identically 1 on the support of $\tilde{\psi}$. Then

$$\begin{aligned}I(\lambda)&=e^{i\lambda\phi(x_0)}\int_{\mathbb{R}} e^{i\tilde{\lambda}y^2}\tilde{\psi}(y)\,dy\\&=e^{i\lambda\phi(x_0)}\int_{\mathbb{R}} e^{i\tilde{\lambda}y^2}e^{-y^2}\left[e^{y^2}\tilde{\psi}(y)\right]\gamma(y)\,dy.\end{aligned}$$

Taylor expand $e^{y^2}\tilde{\psi}(y)$ by

$$e^{y^2}\tilde{\psi}(y)=\sum_{j=0}^{N}a_jy^j\;+\;y^{N+1}R_N(y)=:P(N)$$

where R_N is smooth and bounded. With this, we decompose $I(\lambda)$ as follows:

$$\begin{aligned}I(\lambda)=e^{i\lambda\phi(x_0)}\sum_{j=0}^{N}a_j\int_{\mathbb{R}} e^{i\tilde{\lambda}y^2}e^{-y^2}y^j\,dy&\\+\,e^{i\lambda\phi(x_0)}\int_{\mathbb{R}} e^{i\tilde{\lambda}y^2}e^{-y^2}y^{N+1}R_N(y)\gamma(y)\,dy&\\+\,e^{i\lambda\phi(x_0)}\int_{\mathbb{R}} e^{i\tilde{\lambda}y^2}e^{-y^2}P(y)\left[\gamma(y)-1\right]\,dy.&\end{aligned}$$

Let (1), (2), and (3) denote the first, second, and third lines of this decomposition, respectively.

Consider (3) first. Note that $e^{-y^2}P(y)\,[\gamma(y)-1]$ is supported outside of a neighborhood of $y=0$. As $(y^2)'=2y\neq 0$ away from 0, by the principle of nonstationary phase we have $|(3)|\lesssim|\tilde{\lambda}|^{-m}$ for any order m.

Next, we estimate (2). Let $a\in C_c^\infty(\mathbb{R})$ be a smooth bump function satisfying $a(x)=1$ for $|x|\leq 1$ and $a(x)=0$ for $|x|\geq 2$. For any $\varepsilon>0$, we have a decomposition

$$(2)=e^{i\lambda\phi(x_0)}\int_{\mathbb{R}}e^{i\tilde{\lambda}y^2}e^{-y^2}y^{N+1}R_N(y)\gamma(y)\varphi(y/\varepsilon)\,dy \\ +e^{i\lambda\phi(x_0)}\int_{\mathbb{R}}e^{i\tilde{\lambda}y^2}e^{-y^2}y^{N+1}R_N(y)\gamma(y)\,[1-\varphi(y/\varepsilon)]\,dy.$$

Label the first line of this decomposition (2i), and the second line (2ii). We will estimate each term separately and then optimize in ε. Because the support of $\varphi(y/\varepsilon)$ is small (of size $\sim\varepsilon$, when ε is small we do not expect much oscillation from the phase in this integral. So to estimate (2i) we simply use the triangle inequality, boundedness of $R_N(y)$ and e^{-y^2}, and the support of $\varphi(y/\varepsilon)$ to get

$$|(2\text{i})|\lesssim\int_{\mathbb{R}}\left|y^{N+1}\varphi(y/\varepsilon)\,dy\right|\lesssim\varepsilon^{N+1}\varepsilon=\varepsilon^{N+2}.$$

To estimate (2ii) we use the nonstationary oscillation of the phase $e^{i\tilde{\lambda}y^2}$. Because we need to keep track of the ε's, we will integrate by parts directly instead of applying our previous nonstationary phase result. To simplify notation, let $X(y):=e^{-y^2}R_N(y)\gamma(y)$. Define the differential operator D by

$$D:=\frac{1}{2i\tilde{\lambda}y}\frac{d}{dy}$$

so that $De^{i\tilde{\lambda}y^2}=e^{i\tilde{\lambda}y^2}$. Then the adjoint is given by

$$({}^tDf)(y)=-\frac{d}{dy}\left(\frac{1}{2i\tilde{\lambda}y}f(y)\right).$$

Thus,

$$\begin{aligned}(2\text{ii})&=e^{i\lambda\phi(x_0)}\int_{\mathbb{R}}e^{i\tilde{\lambda}y^2}e^{-y^2}y^{N+1}R_N(y)\gamma(y)\,[1-\varphi(y/\varepsilon)]\,dy\\&=e^{i\lambda\phi(x_0)}\int_{\mathbb{R}}D^m[e^{i\tilde{\lambda}y^2}]\,y^{N+1}X(y)\,[1-\varphi(y/\varepsilon)]\,dy\\&=e^{i\lambda\phi(x_0)}\int_{\mathbb{R}}e^{i\tilde{\lambda}y^2}\,({}^tD)^m\left\{y^{N+1}X(y)\,[1-\varphi(y/\varepsilon)]\right\}\,dy.\end{aligned}$$

We carefully analyze the $({}^tD)^m\left\{y^{N+1}X(y)\,[1-\varphi(y/\varepsilon)]\right\}$ term. For clarity, consider a single application of ${}^tD$:

$${}^tD\left\{y^{N+1}X(y)\,[1-\varphi(y/\varepsilon)]\right\}=-\frac{d}{dy}\left[\frac{1}{2i\tilde{\lambda}y}\left\{y^{N+1}X(y)\,[1-\varphi(y/\varepsilon)]\right\}\right].$$

By applying ${}^tD$ to $\left\{y^{N+1}X(y)\,[1-\varphi(y/\varepsilon)]\right\}$ m times, we pick up m copies of y in the denominator. Moreover, for each application of ${}^tD$, the derivative can either hit the y in the denominator, the y^{N+1} term, $X(y)$, or $1-\varphi(y/\varepsilon)$. Let k denote the number of times the derivative hits the y denominator, let α_1 denote the number of times the derivative hits the y^{N+1}, and let α_2 the number of times the derivative hits $1-\varphi(y/\varepsilon)$. As

$X(y) = e^{-y^2}R_N(y)\gamma(y)$, all of its derivatives are bounded and hence uninteresting. Applying the integral triangle inequality to (2ii) and summing over all possible combinations of derivations described above, we have

$$|(2\text{ii})| \lesssim \frac{1}{|\tilde{\lambda}|^m} \sum_{k=0}^{m} \sum_{\alpha_1+\alpha_2 \le m-k} \int_{|y|\ge\varepsilon} \frac{|y|^{N+1-\alpha_1}\varepsilon^{-\alpha_2}}{|y|^{m+k}}\, dy.$$

The $\le$ sign in the second sum is to account for the fact that we are ignoring the derivatives of $X(y)$. The region of integration is due to the support of $1-\varphi(y/\varepsilon)$. Continuing the estimate,

$$|(2\text{ii})| \lesssim |\tilde{\lambda}|^{-m} \sum_{k=0}^{m} \sum_{\alpha_1+\alpha_2 \le m-k} \varepsilon^{-\alpha_2} \int_{|y|\ge\varepsilon} |y|^{N+1-\alpha_1-(m+k)}\, dy.$$

The most singular integral in this sum corresponds to $\alpha_1 = 0$ and $k = 0$. This particular integral is finite provided $N+1-m < -1$, i.e., $N+2 < m$, and is bounded by $\varepsilon^{N+2-(\alpha_1+\alpha_2)-(m+k)}$. So, as long as m is chosen so that $N+2 < m$, we have

$$|(2\text{ii})| \lesssim |\tilde{\lambda}|^{-m} \sum_{k=0}^{m} \sum_{\alpha_1+\alpha_2 \le m-k} \varepsilon^{N+2-(\alpha_1+\alpha_2)-(m+k)}.$$

Because ε is small, the summand is majorized when $\alpha_1+\alpha_2 = m-k$. Thus,

$$\begin{aligned} |(2\text{ii})| &\lesssim |\tilde{\lambda}|^{-m} \sum_{k=0}^{m} \sum_{\alpha_1+\alpha_2 \le m-k} \varepsilon^{N+2-(m-k)-(m+k)} = |\tilde{\lambda}|^{-m} \sum_{k=0}^{m} \sum_{\alpha_1+\alpha_2 \le m-k} \varepsilon^{N+2-2m} \\ &\lesssim |\tilde{\lambda}|^{-m}\varepsilon^{N+2-2m}. \end{aligned}$$

We now optimize in ε. Choose ε so that the (2i) and (2ii) estimates are equal:

$$\varepsilon^{N+2} = |\tilde{\lambda}|^{-m}\varepsilon^{N+2-2m} \quad \Rightarrow \quad \varepsilon = |\tilde{\lambda}|^{-\frac{1}{2}}.$$

Then

$$|(2)| \lesssim |\tilde{\lambda}|^{-\frac{N+2}{2}}.$$

This estimate is acceptable for $N \ge 1$.

It remains to estimate (1):

$$(1) = e^{i\lambda\phi(x_0)} \sum_{j=0}^{N} a_j \int_{\mathbb{R}} e^{i\tilde{\lambda}y^2} e^{-y^2} y^j\, dy.$$

Note that $a_0 = \tilde{\psi}(0) = \psi(x_0)$. For $j \ge 0$,

$$\int_{\mathbb{R}} e^{i\tilde{\lambda}y^2} e^{-y^2} y^j\, dy = \int_{\mathbb{R}} e^{-(1-i\tilde{\lambda})y^2} y^j\, dy.$$

Let $z = (1-i\tilde{\lambda})^{\frac{1}{2}}$. Then

$$\int_{\mathbb{R}} e^{-(1-i\tilde{\lambda})y^2} y^j\, dy = (1-i\tilde{\lambda})^{-\frac{j}{2}-\frac{1}{2}} \int_{(1-i\tilde{\lambda})^{\frac{1}{2}}\mathbb{R}} e^{-z^2} z^j\, dz$$

where we have chosen a branch of the function $z^{-\frac{j}{2}-\frac{1}{2}}$ with cut along the negative real axis. Using the decay of e^{-z^2} to shift the contour, we have

$$\int_{\mathbb{R}} e^{i\tilde{\lambda}y^2} e^{-y^2} y^j\, dy = (1-i\tilde{\lambda})^{-\frac{j}{2}-\frac{1}{2}} \int_{\mathbb{R}} e^{-y^2} y^j\, dy$$

When j is odd, the function $e^{-y^2}y^j$ is odd, and consequently the above integral is 0. When $j \geq 2$ is even, the asymptotic contribution of the above integral is $|\tilde{\lambda}|^{-\frac{j+1}{2}}$. As $j \geq 2$, this quantity is $O(\lambda^{-\frac{3}{2}})$ and hence is acceptable. It remains to consider the contribution of the $j = 0$ term. Write $(1 - i\tilde{\lambda}) = re^{i\theta}$. Then $r = \sqrt{1 + \tilde{\lambda}^2}$ and $\tan\theta = \frac{-\tilde{\lambda}}{1} = -\tilde{\lambda}$. Recall that $\int_{\mathbb{R}} e^{-y^2}\,dy = \sqrt{\pi}$. Thus, the contribution of the $j = 0$ term to (1) is therefore

$$e^{i\lambda\phi(x_0)}a_0(1 - i\tilde{\lambda})^{-\frac{1}{2}}\sqrt{\pi} = e^{i\lambda\phi(x_0)}\psi(x_0)r^{-\frac{1}{2}}e^{-i\frac{\theta}{2}}.$$

As $\tilde{\lambda} \to \infty$, $\theta \to -\frac{\pi}{2}$, and as $\tilde{\lambda} \to -\infty$, $\theta \to \frac{\pi}{2}$. Recall that $\tilde{\lambda} = \frac{1}{2}\phi''(x_0)\lambda$. Thus, as $\lambda \to \infty$ we have $e^{-i\frac{\theta}{2}} \to e^{i\frac{\pi}{4}\operatorname{sgn}\phi''(x_0)}$. Thus, the contribution to (1) is

$$e^{i\lambda\phi(x_0)}\psi(x_0)\sqrt{\pi}\left[\frac{2}{\lambda|\phi''(x_0)|}\right]^{\frac{1}{2}} e^{i\frac{\pi}{4}\operatorname{sgn}\phi''(x_0)}.$$

Careful inspection shows that this is precisely the first term of (8.3), which completes the proof. □

8.2 The Morse Lemma

The remainder of this chapter is dedicated to proving the principles of nonstationary and stationary phase in higher dimensions. In the proof of stationary phase, it will be useful to change variables in a particularly nice way, as described by the following lemma.

Lemma 8.6 (Morse)**.** *Suppose that $\phi : \mathbb{R}^d \to \mathbb{R}$ is smooth with a nondegenerate critical point at $x_0 \in \mathbb{R}^d$. There is a change of coordinates $x \mapsto y(x)$ such that $y(x_0) = 0$, $\frac{\partial y}{\partial x}(x_0) = I_d$, and*

$$\phi(x) - \phi(x_0) = \frac{1}{2}\lambda_1 y_1^2 + \cdots \frac{1}{2}\lambda_d y_d^2,$$

where $\lambda_1, \ldots, \lambda_d$ are the eigenvalues of $D^2\phi(x_0)$.

We have dedicated an entire section to the Morse lemma, as its proof is technical and does not strictly involve any harmonic analysis or techniques from oscillatory integrals. The higher dimensional analogues of nonstationary and stationary phase are contained in the next section, and the reader may be inclined to skip ahead.

Proof. Because $D^2\phi(x_0)$ is symmetric, we may change coordinates if necessary and assume without loss of generality that $D^2\phi(x_0) = \operatorname{diag}(\lambda_1, \ldots, \lambda_d)$. By Taylor expansion and integration by parts, we may write

$$\begin{aligned}\phi(x) &= \phi(x_0) + (x - x_0)\nabla\phi(x_0) + \int_0^1 (1-t)\frac{d^2}{dt^2}\left[\phi(x_0 + t(x - x_0))\right]\,dt \\ &= \phi(x_0) + \int_0^1 (1-t)\sum_{i,j=1}^d (x - x_0)_i(x - x_0)_j \frac{\partial^2\phi}{\partial x_i\,\partial x_j}(x_0 + t(x - x_0))\,dt.\end{aligned}$$

So $\phi(x) - \phi(x_0) = \sum_{i,j=1}^d (x - x_0)_i(x - x_0)_j\mu_{ij}(x)$, where

$$\mu_{ij}(x) = \int_0^1 (1-t)\frac{\partial^2\phi}{\partial x_i\,\partial x_j}(x_0 + t(x - x_0))\,dt.$$

Since ϕ is smooth, μ_{ij} is smooth. Also, $\mu_{ij} = \mu_{ji}$ and

$$\mu_{ij}(x_0) = \frac{1}{2}\frac{\partial^2\phi}{\partial x_i\,\partial x_j}(x_0).$$

From here, we argue by induction. Suppose that we have found a change of variables $x \mapsto y(x)$ such that $y(x_0) = 0$, $\frac{\partial y}{\partial x}(x_0) = I$, and

$$\phi(x) - \phi(x_0) = \frac{1}{2}\lambda_1 y_1^2 + \cdots \frac{1}{2}\lambda_{r-1} y_{r-1}^2 \; + \sum_{i,j \geq r}^{d} y_i y_j \tilde{\mu}_{ij}(y).$$

For the base case, we take $r = 1$, $y = x - x_0$, and $\tilde{\mu}_{ij}(y) = \mu_{ij}(x)$.

Next, we perform some computations assuming the above decomposition. First,

$$\frac{\partial^2}{\partial x_i\, \partial x_j}\left(\frac{1}{2}\lambda_k y_k^2\right)\Bigg|_{x=x_0} = \lambda_k \frac{\partial y_k}{\partial x_i}\Bigg|_{x_0} \cdot \frac{\partial y_k}{\partial x_j}\Bigg|_{x_0} \; + \; \lambda_k \frac{\partial^2 y_k}{\partial x_i\, \partial x_j}\Bigg|_{x_0} \cdot y_k\Bigg|_{x_0}.$$

The second summand is 0 because $y_k(x_0) = 0$. Because $\frac{\partial y}{\partial x}(x_0) = I$ by assumption, we are left with

$$\frac{\partial^2}{\partial x_i\, \partial x_j}\left(\frac{1}{2}\lambda_k y_k^2\right)\Bigg|_{x=x_0} = \lambda_k \delta_{ij}\delta_{ik}.$$

As $D^2\phi(x_0) = \operatorname{diag}(\lambda_1, \ldots, \lambda_d)$, we get

$$D^2\left(\sum_{i,j \geq r}^{d} y_i y_j \tilde{\mu}_{ij}(y)\right)\Bigg|_{x=x_0} = \operatorname{diag}(\lambda_r, \ldots, \lambda_d).$$

Note that

$$\begin{aligned}
\frac{\partial^2}{\partial x_l\, \partial x_k}\left(\sum_{i,j \geq r}^{d} y_i y_j \tilde{\mu}_{ij}(y)\right)\Bigg|_{x=x_0} &= \sum_{i,j \geq r}\left(\frac{\partial y_i}{\partial x_l}\Bigg|_{x_0} \cdot \frac{\partial y_j}{\partial x_k}\Bigg|_{x_0} + \frac{\partial y_i}{\partial x_k}\Bigg|_{x_0} \cdot \frac{\partial y_j}{\partial x_l}\Bigg|_{x_0}\right)\tilde{\mu}_{ij}(x = x_0) \\
&= \sum_{i,j \geq r} (\delta_{il}\delta_{jk} + \delta_{ik}\delta_{jl})\tilde{\mu}_{ij}(0) \\
&= 2\tilde{m}_{lk}(0).
\end{aligned}$$

This implies that

$$\tilde{\mu}_{ij}(0) = \operatorname{diag}(\lambda_r, \ldots, \lambda_d).$$

Next, perform the change of variables $y \mapsto y'(y)$ defined by

$$\begin{cases} y'_j := y_j \text{ for } j \neq r; \\ y'_r := \sqrt{\frac{\tilde{\mu}_{rr}(y)}{\lambda_r/2}}\left(y_r + \sum_{j \geq r+1} \frac{\tilde{\mu}_{jr}(y)}{\tilde{\mu}_{rr}(y)} y_j\right) \end{cases}.$$

We need to show that this diffeomorphism satisfies $y'(x_0) = 0$, $\frac{\partial y'}{\partial x}(x = x_0) = I$, and

$$\phi(x) - \phi(x_0) = \frac{1}{2}\lambda_1 (y'_1)^2 + \cdots \frac{1}{2}\lambda_r (y'_r)^2 \; + \sum_{i,j \geq r+1}^{d} y'_i y'_j \tilde{\mu}'_{ij}(y).$$

for some functions $\tilde{\mu}'_{ij}(y)$. It is clear that $y'(x_0) = 0$, since $y_k(x_0) = 0$ for all k. next, we have

$$\frac{\partial y'_r}{\partial x_i}\Bigg|_{x_0} = \sqrt{\frac{\tilde{\mu}_{rr}(y(x_0))}{\lambda_r/2}}\left(\delta_{ir} + \sum_{j \geq r+1} \delta_{ij} \cdot 0\right) = \delta_{ir}$$

and so $\frac{\partial y'}{\partial x}(x = x_0) = I$. Finally,

$$\sum_{i,j\geq r}^{d} y_i y_j \tilde{\mu}_{ij}(y) - \frac{1}{2}\lambda_r (y_r')^2$$

$$= \sum_{i,j\geq r}^{d} y_i y_j \tilde{\mu}_{ij}(y) - \tilde{m}_{rr}(y)\left[y_r^2 + 2y_r \sum_{j\geq r+1} \frac{\tilde{\mu}_{jr}(y)}{\tilde{\mu}_{rr}(y)} y_j + \left(\sum_{j\geq r+1} \frac{\tilde{\mu}_{jr}(y)}{\tilde{\mu}_{rr}(y)} y_j\right)^2\right]$$

$$= \sum_{i,j\geq j+1} y_i y_j \left[\tilde{\mu}_{ij}(y) - \frac{\tilde{\mu}_{ir}(y)\tilde{\mu}_{jr}(y)}{\tilde{\mu}_{rr}(y)}\right].$$

Setting $\tilde{\mu}'_{ij} = \tilde{\mu}_{ij}(y) - \frac{\tilde{\mu}_{ir}(y)\tilde{\mu}_{jr}(y)}{\tilde{\mu}_{rr}(y)}$ gives the desired conclusion. $\square$

8.3 Oscillatory Integrals of the First Kind, $d > 1$

In this section we consider the asymptotics of oscillatory integrals on $\mathbb{R}^d$. We begin with the principle of nonstationary phase. The statement and proof is essentially the same as in the one-dimensional case.

Proposition 8.7 (Nonstationary phase). *Let $\phi : \mathbb{R}^d \to \mathbb{R}$ and $\psi : \mathbb{R}^d \to \mathbb{C}$ be smooth. Assume that ψ has compact support, and that $\nabla\phi \neq 0$ on the support of ψ. Then*

$$\left|\int_{\mathbb{R}^d} e^{i\lambda\phi(x)}\psi(x)\,dx\right| \lesssim \lambda^{-N}$$

for any $N \geq 0$.

Proof. As in the one-dimensional case, we simply use integration by parts to exploit the phase oscillation.

Because $\nabla\phi \neq 0$ on the support of ψ, we have

$$e^{i\lambda\phi(x)} = \frac{\nabla\phi(x)}{i\lambda|\nabla\phi(x)|^2}\,\nabla e^{i\lambda\phi(x)}.$$

Then

$$I(\lambda) = \int_{\mathbb{R}^d} e^{i\lambda\phi(x)}\psi(x)\,dx = \int_{\mathbb{R}^d} \frac{\nabla\phi(x)}{i\lambda|\nabla\phi(x)|^2}\,\nabla e^{i\lambda\phi(x)}\psi(x)\,dx$$
$$= -\frac{1}{i\lambda}\int_{\mathbb{R}^d} e^{i\lambda\phi(x)}\,\nabla\left[\frac{\nabla\phi(x)}{i\lambda|\nabla\phi(x)|^2}\psi(x)\right]\,dx.$$

Thus, $|I(\lambda)| \lesssim \lambda^{-1}$, where the implicit constant depends on two derivatives of ϕ and one derivative of ψ. Iterating integration by parts as before gives decay of $|I(\lambda)|$ to any order. $\square$

Next, we consider the principle of stationary phase. The one-dimensional principle of stationary phase described how $I(\lambda)$ decays like $\lambda^{-\frac{1}{2}}$ near a nondegenerate critical point of the phase. In the higher-dimensional case, the decay is like becomes $\lambda^{-\frac{d}{2}}$.

The proof is structurally the same as in the one-dimensional case, with some slightly more involved computations.

Proposition 8.8 (Stationary phase). *Let $\phi : \mathbb{R}^d \to \mathbb{R}$ and $\psi : \mathbb{R}^d \to \mathbb{C}$ be smooth. Assume that ϕ has a nondegenerate critical point at $x_0 \in \mathbb{R}^d$, and that ψ is supported in a sufficiently small neighborhood of x_0. Then $I(\lambda) = \int_{\mathbb{R}^d} e^{i\lambda\phi(x)}\psi(x)\,dx$ satisfies*

$$I(\lambda) = \left[e^{i\lambda\phi(x_0)}\psi(x_0)(2\pi i)^{\frac{d}{2}}\prod_{j=1}^{d}\mu_j^{-\frac{1}{2}}\right]\lambda^{-\frac{d}{2}} + O\left(\lambda^{-\frac{d}{2}-\frac{1}{2}}\right) \tag{8.4}$$

as $\lambda \to \infty$, where $\mu_1, \ldots, \mu_d$ are the eigenvalues of $D^2\phi(x_0)$.

Proof. By the Morse lemma, there is a change of variables $x \mapsto y(x)$ such that $y(x_0) = 0$, $\frac{\partial y}{\partial x}(x_0) = I_d$, and

$$\phi(x) - \phi(x_0) = \frac{1}{2}\mu_1 y_1^2 + \cdots + \frac{1}{2}\mu_d y_d^2.$$

With this change of variables, we have

$$I(\lambda) = e^{i\lambda\phi(x_0)} \int e^{i\lambda \sum_{j=1}^d \frac{1}{2}\mu_j y_j^2} \tilde{\psi}(y)\, dy$$

where $\tilde{\psi}(y)$ absorbs the Jacobian from the change of variables. As in the proof of the one-dimensional case, let $\gamma \in C_c^\infty(\mathbb{R}^d)$ be identically 1 on the support of $\tilde{\psi}$. Then

$$I(\lambda) = e^{i\lambda\phi(x_0)} \int e^{i\lambda \sum_{j=1}^d \frac{1}{2}\mu_j y_j^2} e^{-|y|^2} \left[e^{|y|^2}\tilde{\psi}(y)\right] \gamma(y)\, dy.$$

By Taylor expansion, we can write

$$e^{|y|^2}\tilde{\psi}(y) = \sum_{|\alpha| \le N} c_\alpha y^\alpha + \sum_{|\beta| = N+1} y^\beta R_\beta(y) =: P(y) + \sum_{|\beta|=N+1} y^\beta R_\beta(y).$$

As in the one-dimensional case, we then have a decomposition

$$\begin{aligned} I(\lambda) = e^{i\lambda\phi(x_0)} &\sum_{|\alpha|\le N} c_\alpha \int e^{i\lambda \sum_{j=1}^d \frac{1}{2}\mu_j y_j^2} e^{-|y|^2} y^\alpha\, dy \\ &+ e^{i\lambda\phi(x_0)} \sum_{|\beta|=N+1} \int e^{i\lambda \sum_{j=1}^d \frac{1}{2}\mu_j y_j^2} e^{-|y|^2} y^\beta R_\beta(y)\gamma(y)\, dy \\ &\qquad + e^{i\lambda\phi(x_0)} \int e^{i\lambda \sum_{j=1}^d \frac{1}{2}\mu_j y_j^2} e^{-|y|^2} P(y)\left(\gamma(y) - 1\right)\, dy. \end{aligned}$$

Let (1), (2), and (3) denote the first, second, and third lines of this decomposition, respectively.

Estimating (1).

For $|\alpha| \le N$, we have

$$\begin{aligned} \int e^{i\lambda \sum_{j=1}^d \frac{1}{2}\mu_j y_j^2} e^{-|y|^2} y^\alpha\, dy &= \prod_{j=1}^d \int e^{\frac{i\lambda\mu_j^2}{2} y_j^2} e^{-y_j^2} y_j^{\alpha_j}\, dy_j \\ &= \prod_{j=1}^d \left(1 - \frac{i\lambda\mu_j^2}{2}\right)^{-\frac{1}{2} - \frac{\alpha_j}{2}} \int e^{-y_j^2} y_j^{\alpha_j}\, dy_j. \end{aligned}$$

For $|\alpha| \ge 2$, this computation gives

$$\left| \int e^{i\lambda \sum_{j=1}^d \frac{1}{2}\mu_j y_j^2} e^{-|y|^2} y^\alpha\, dy \right| = O\left(|\lambda|^{-\frac{d}{2} - \frac{|\alpha|}{2}}\right) = O\left(\lambda^{-\frac{d}{2}-1}\right)$$

as $\lambda \to \infty$. When $|\alpha| = 1$, the contribution is 0 because $\int e^{-y_j^2} y_j^{\alpha_j}\, dy_j = 0$ when α_j is odd. When $\alpha = 0$, write $1 - \frac{i\lambda\mu_j^2}{2} = r_j e^{i\theta_j}$ where $r_j = \sqrt{1 + \frac{\lambda^2\mu_j^2}{4}}$ and $\tan\theta_j = -\frac{\lambda\mu_j}{2}$. As $\lambda \to \infty$, $\theta_j \to -\frac{\pi}{2}\operatorname{sgn}\mu_j$. Thus,

$$\prod_{j=1}^d r_j^{-\frac{1}{2}} e^{-i\frac{\theta_j}{2}} \to \left(\frac{\lambda}{2}\right)^{-\frac{d}{2}} \prod_{j=1}^d e^{i\frac{\pi}{4}\operatorname{sgn}\mu_j} \left(\pi|\mu_j|\right)^{-\frac{1}{2}}$$

as $\lambda \to \infty$. Finally, observe that $c_0 = \psi(x_0)$. Rewriting the above constant and combining the estimates thus far gives

$$(1) = \left[e^{i\lambda\phi(x_0)}\psi(x_0)(2\pi i)^{\frac{d}{2}} \prod_{j=1}^{d} \mu_j^{-\frac{1}{2}} \right] \lambda^{-\frac{d}{2}} + O\left(\lambda^{-\frac{d}{2}-\frac{1}{2}}\right).$$

Estimating (2).

Let $a \in C_c^\infty(\mathbb{R}^d)$ satisfy $a(x) = 1$ for $|x| \leq 1$ and $a(x) = 0$ for $|x| \geq 2$. Let $X_\beta(y) = e^{-|y|^2} R_\beta(y)\gamma(y)$. For $\varepsilon > 0$ (to be determined later), we have

$$(2) = e^{i\lambda\phi(x_0)} \sum_{|\beta|=N+1} \int e^{i\lambda \sum_{j=1}^{d} \frac{1}{2}\mu_j y_j^2} y^\beta X_\beta(y) a(y/\varepsilon)\, dy \\ + e^{i\lambda\phi(x_0)} \sum_{|\beta|=N+1} \int e^{i\lambda \sum_{j=1}^{d} \frac{1}{2}\mu_j y_j^2} y^\beta X_\beta(y) \left[1 - a(y/\varepsilon)\right]\, dy.$$

Let (2i) and (2ii) denote the first and second lines of this decomposition, respectively. As in the one-dimensional case, we do not expect any cancellation in (2i) because of the small support of $a(\cdot/\varepsilon)$, thus we estimate (2i) crudely by

$$|(2\text{i})| \lesssim \int_{|y|\leq 2\varepsilon} |y|^{N+1}\, dy \lesssim \varepsilon^{N+1+d}.$$

Next we consider (2ii). We will effectively use nonstationary phase on this integral, but we have to be explicit with the integration by parts because we are optimizing our choice of ε. Let

$$\Gamma_j = \left\{ x \in \mathbb{R}^d \ :\ x_j > \frac{|x|^2}{d} \right\}.$$

Then $\mathbb{R}^d = \bigcup_{j=1}^{d} \Gamma_j$. Fatten the regions as follows:

$$\tilde{\Gamma}_j = \left\{ x \in \mathbb{R}^d \ :\ x_j > \frac{|x|^2}{2d} \right\}.$$

We will define a particular partition of unity subordinate to this collection of sets. Let $\tilde{\varphi}_j : S^{d-1} \subseteq \mathbb{R}^d \to \mathbb{R}$ be functions with support in $\tilde{\Gamma}_j \cap S^{d-1}$ satisfying $\tilde{\varphi}_j(x) = 1$ for $x \in \Gamma_j \cap S^{d-1}$. Set $\varphi_j = \frac{\tilde{\varphi}_j}{\sum_{j=1}^{d} \tilde{\varphi}_j}$. Finally, extend φ_j to all of $\mathbb{R}^d$ by setting $\varphi_j(x) = \varphi_j(x/|x|)$.

Next, let $Q(y) = \sum_{j=1}^{d} \frac{1}{2}\mu_j y_j^2$. Then $\nabla Q(y) = (\mu_1 y_1, \ldots, \mu_d y_d)$, and hence Q has one critical point at $y = 0$. We have

$$(2\text{ii}) = e^{i\lambda\phi(x_0)} \sum_{j=1}^{d} \sum_{|\beta|=N+1} \int e^{i\lambda Q(y)} y^\beta X_\beta(y) \left[1 - a(y/\varepsilon)\right] \varphi_j(y)\, dy.$$

Note that

$$e^{i\lambda Q(y)} = \frac{1}{i\lambda\mu_j y_j} \frac{\partial}{\partial y_j} e^{i\lambda Q(y)}.$$

The computation from here is similar to the one-dimensional case. Integrating by parts m times via $\frac{\partial}{\partial y_j}$ in each set $\tilde{\Gamma}_j$, and using the fact that $y_j \geq \frac{|y|^2}{d}$ in $\tilde{\Gamma}_j$, we compute as follows. Let k denote the number of times the adjoint derivative hits the y_i in the denominator of the above fraction. Then

$$|(2\text{ii})| \lesssim \frac{1}{\lambda^m} \sum_{k=0}^{m} \sum_{\alpha_1+\alpha_2 \leq m-k} \int_{|y|\geq\varepsilon} \frac{|y|^{N+1-|\alpha_1|}}{|y|^{m+k}} \varepsilon^{-|\alpha_2|}\, dy.$$

This integral is finite provided $m + k - N - 1 + |\alpha_1| > d$, in which case

$$|(2ii)| \lesssim \frac{1}{\lambda^m} \sum_{k=0}^{m} \sum_{\alpha_1+\alpha_2 \le m-k} \varepsilon^{N+1+d-(|\alpha_1|+|\alpha_2|)-(m-k)}$$
$$\lesssim \frac{1}{\lambda^m} \varepsilon^{N+1+d-2m}.$$

To optimize our choice in $\varepsilon > 0$ between (2i) and (2ii), we require

$$\varepsilon^{N+1+d} = \frac{1}{\lambda^m} \varepsilon^{N+1+d-2m}$$

and thus $\varepsilon = \lambda^{-\frac{1}{2}}$. With this choice,

$$|(2)| \lesssim \lambda^{-\frac{d}{2}-\frac{N+1}{2}},$$

which is acceptable for $N \ge 1$.

Estimating (3).

Finally, we estimate (3). As $\gamma(y) - 1$ is supported away from the origin, and because we don't need to worry about optimizing between different estimates, we can blindly apply the principle of nonstationary phase to get $|(3)| \lesssim \lambda^{-m}$ for any order m, which is obviously acceptable. This completes the proof. $\square$

8.4 The Fourier Transform of a Surface Measure

We end this chapter by applying the theory of oscillatory integrals to the surface measure on the unit sphere $S^{d-1} \subseteq \mathbb{R}^d$. This will have applications when we study dispersive PDEs in Chapter 9, and also when we study restriction theory in Chapter 10.

Let $d\sigma$ denote the surface measure supported on $S^{d-1} \subseteq \mathbb{R}^d$. Define the inverse Fourier transform of the measure $d\sigma$ to be the function on $\mathbb{R}^d$ given by

$$\check{d\sigma}(x) := \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{S^{d-1}} e^{ix\cdot\xi}\, d\sigma(\xi).$$

This function satisfies the following decay property.

Lemma 8.9. *If $d\sigma$ is the surface measure on $S^{d-1} \subseteq \mathbb{R}^d$, then $|\check{d\sigma}(x)| \lesssim \langle x\rangle^{-\frac{d-1}{2}}$.*

Proof. First, observe that

$$\check{d\sigma}(0) = \int_{S^{d-1}} d\sigma(\xi) \sim 1.$$

Thus, to prove the lemma it suffices to prove $|\check{d\sigma}(x)| \lesssim |x|^{-\frac{d-1}{2}}$.

As $d\sigma$ is rotation invariant, $\check{d\sigma}$ is also rotation invariant. In particular,

$$\check{d\sigma}(x) = \check{d\sigma}((0, \dots, 0, |x|))$$
$$\sim \int_{S^{d-1}} e^{i|x|\xi_d}\, d\sigma(\xi).$$

Let θ denote the angle between ξ and the north pole $(0, \dots, 0, 1)$. Changing variables via spherical coordinates gives

$$\check{d\sigma}(x) \sim \int_0^\pi e^{i|x|\cos\theta} \sin^{d-2}(\theta)\, d\theta.$$

Note that $\frac{d}{d\theta}e^{i|x|\cos\theta} = e^{i|x|\cos\theta}(-i|x|\sin\theta)$. Thus, if $d \geq 4$ we can integrate by parts to get

$$\begin{aligned}
\check{d\sigma}(x) &\sim -\int_0^\pi \frac{d}{d\theta}\left[e^{i|x|\cos\theta}\right]\frac{1}{i|x|\sin\theta}\sin^{d-2}(\theta)\,d\theta \\
&= -\frac{1}{i|x|}\int_0^\pi \frac{d}{d\theta}\left[e^{i|x|\cos\theta}\right]\sin^{d-3}(\theta)\,d\theta \\
&= -\frac{1}{i|x|}e^{i|x|\cos\theta}\sin^{d-3}(\theta)\Big|_0^\pi + \frac{d-3}{i|x|}\int_0^\pi e^{i|x|\cos\theta}\sin^{d-4}(\theta)\cos\theta\,d\theta.
\end{aligned}$$

The boundary terms are 0, since $\sin 0 = \sin\pi = 0$. So if $d \geq 4$, one iteration of integration by parts gives

$$\check{d\sigma}(x) \sim |x|^{-1}\int_0^\pi e^{i|x|\cos\theta}\sin^{d-4}(\theta)\cos\theta\,d\theta.$$

In words, one iteration of integration by parts picked up a power of $|x|^{-1}$, decreased the power of $\sin\theta$ in the integral by 2, and picked up a power of $\cos\theta$.

If d is even, we can integrate by parts $\frac{d-2}{2}$ times to eliminate the original $d-2$ copies of $\sin\theta$. Along the way we pick up more copies of $\sin\theta$ and $\cos\theta$ from differentiating products of $\sin\theta$ and $\cos\theta$, but the elimination of the original $d-2$ copies of $\sin\theta$ ensures

$$\check{d\sigma}(x) \sim |x|^{-\frac{d-2}{2}}\int_0^\pi e^{i|x|\cos\theta}P(\theta)\,d\theta$$

where $P(\theta)$ is a trigonometric polynomial containing at least one term without a $\sin\theta$.

Similarly, if d is odd we can integrate by parts $\frac{d-3}{2}$ times to get

$$\check{d\sigma}(x) \sim |x|^{-\frac{d-3}{2}}\int_0^\pi e^{i|x|\cos\theta}\sin\theta Q(\theta)\,d\theta$$

for some trigonometric polynomial $Q(\theta)$. Integrating by parts one more time gives the desired estimate:

$$\begin{aligned}
\check{d\sigma}(x) &\sim |x|^{-\frac{d-3}{2}}|x|^{-1}\int_0^\pi \frac{d}{d\theta}\left[e^{i|x|\cos\theta}\right]Q(\theta)\,d\theta \\
&\sim |x|^{-\frac{d-1}{2}}.
\end{aligned}$$

Thus, it remains to consider when d is even. Let $a_1 \in C_c^\infty(-\varepsilon,\varepsilon)$, $a_2 \in C_c^\infty(\varepsilon/2, \pi-\varepsilon/2)$, and $a_3 \in C_c^\infty(\pi-\varepsilon,\pi+\varepsilon)$ satisfy $a_1+a_2+a_3 = 1$ on $(0,\pi)$. In the support of a_2, the phase $\phi(\theta) = \cos\theta$ has no critical points, so nonstationary phase gives

$$\left|\int_0^\pi e^{i|x|\cos\theta}P(\theta)a_2(\theta)\,d\theta\right| \lesssim |x|^{-1}.$$

On the support of a_1, we have

$$\int_0^\pi e^{i|x|\cos\theta}P(\theta)a_2(\theta)\,d\theta = \int_0^\varepsilon e^{i|x|\cos\theta}P(\theta)a_2(\theta)\,d\theta.$$

The phase $\cos\theta$ has one nondegenerate critical point at 0. The principle of stationary phase gives a contribution of $O(|x|^{-\frac{1}{2}})$, and similarly on the support of a_3. Thus,

$$\check{d\sigma}(x) \lesssim |x|^{-\frac{d-2}{2}}|x|^{-\frac{1}{2}} = |x|^{-\frac{d-1}{2}}$$

as desired.

The lower dimensional cases are straightforward exercises.

□

Chapter 9

Dispersive Partial Differential Equations

ADD INTRODUCTION

The most important examples of dispersive partial differential equations are the *nonlinear Schrödinger equation*

$$i\partial_t u + \Delta u = F(u, Du) \tag{9.1}$$

and the *Korteweg-de Vries (KdV) equation*

$$u_t + u_{xxx} + \gamma u u_x = 0. \tag{9.2}$$

In (9.1), $u : \mathbb{R}^d_x \times \mathbb{R}_t \to \mathbb{C}$ or $u : \mathbb{T}^d_x \times \mathbb{R}_t \to \mathbb{C}$, and in (9.2), $u : \mathbb{R}_x \times \mathbb{R}_t \to \mathbb{R}$ or $u : \mathbb{T}_x \times \mathbb{R}_t \to \mathbb{R}$. These equations arise in a number of physical contexts and are an active area of research.

FINISH

9.1 Dispersion

What does it mean for a partial differential equation to be *dispersive*? Informally, an equation is dispersive if its solutions are waves whose spatial profiles spread out over time, if no boundary conditions are imposed. This heuristic is nice, but is not mathematically rigorous in any way. The goal of this section is to give a (mostly) formal definition of dispersion.

A slightly more tractable way of identifying a dispersive equation is via the frequency support of its fundamental solution. Consider the *general linear evolution equation*

$$\partial_t u = iP(D)u. \tag{9.3}$$

Here, $u : \mathbb{R}^d_x \times \mathbb{R}_t \to \mathbb{C}$ is a complex-valued function of space and time, and $P(D)$ is a linear differential operator with Fourier symbol $2\pi\varphi(\xi)$. Taking the space-time Fourier transformation of (9.3) gives $2\pi i\tau\hat{u}(\xi,\tau) = 2\pi i\varphi(\xi)\hat{u}(\xi,\tau)$. Rearranging and dividing by $2\pi i$ gives

$$[\tau - \varphi(\xi)]\,\hat{u}(\xi,\tau) = 0. \tag{9.4}$$

This equation tells us that the space-time Fourier transform of a solution of (9.3) is supported on the surface $\Sigma := \{\,(\xi,\tau) \;:\; \tau = \varphi(\xi)\,\}$. We may identify an evolution equation of the form (9.3) as *dispersive* if the surface Σ is curved.[1] Indeed, the surface corresponding to the Schrödinger equation (9.1) is a paraboloid, and the surface corresponding to the KdV equation (9.2) is a cubic curve. An example of an equation which is *not* dispersive in this sense is the *transport equation*

$$u_t - cu_x = 0. \tag{9.5}$$

[1] Here we are being intentionally vague about curvature to avoid any unnecessary excursions into differential geometry.

The corresponding surface is the hyperplane $\{\tau = c\xi\}$, which is decidedly not curved.

To be more formal, consider the evolution equation

$$\begin{cases} \partial_t u = Lu \\ u(x,0) = u_0(x) \end{cases}. \tag{9.6}$$

Here, $u : \mathbb{R}^d_x \times \mathbb{R}_t \to V$, where V is either $\mathbb{C}^d$ or $\mathbb{R}^d$, and L is a *skew-adjoint constant-coefficient linear differential operator*, that is,

$$\int_{\mathbb{R}^d} \langle Lu(x), v(x)\rangle_V \, dx = -\int_{\mathbb{R}^d} \langle u(x), Lv(x)\rangle_V \, dx$$

for all functions u and v. With $D = \frac{1}{i}\nabla_x$, we may write

$$Lu = \sum_{|\alpha|\le k} c_\alpha \partial_x^\alpha u := ih(D)u,$$

where $h : \mathbb{R}^d \to \operatorname{End} V$ is the polynomial

$$h(\xi) \sim_\pi \sum_{|\alpha|\le k} i^{|\alpha|-1} c_\alpha \xi^\alpha.$$

We refer to the symbol h as the *dispersion relation* of (9.6). For the linear Schrödinger equation, the dispersion relation is $h(\xi) = -|\xi|^2$, and for the Airy equation (the linear part of the KdV equation) the dispersion relation is $h(\xi) = \xi^3$. For a fixed $v \in \mathbb{R}^d$, the general transport equation

$$\partial_t u = -v \cdot \nabla_x u \tag{9.7}$$

has dispersion relation $h(\xi) = -v \cdot \xi$.

By considering these dispersion relations, we are irrevocably led to the following definition.

Definition 9.1. An evolution equation of the form (9.6) is **dispersive** if ∇h is not constant.

9.2 The Linear Schrödinger Equation

A necessary prerequisite for understanding the NLS equation (9.1) or the KdV equation (9.2) is to develop a solid understanding of their linear counterparts. Thus, we begin with the **linear Schrödinger equation**, given by

$$\begin{cases} iu_t + \frac{\Delta}{2} u \\ u(x,0) = u_0(x) \end{cases}. \tag{9.8}$$

Here, $u : \mathbb{R}^d_x \times \mathbb{R}_t \to \mathbb{C}$ and $u_0 : \mathbb{R}^d_x \to \mathbb{C}$. For now, we will assume that $u_0 \in \mathcal{S}(\mathbb{R}^d)$.

9.2.1 The Fundamental Solution

Our first order of business is to compute a concrete representation of a solution to the linear Schrödinger equation. To solve (9.8), we first apply the spatial Fourier transform.[2] As $\hat{\Delta} = (i\xi)\cdot(i\xi) = -|\xi|^2$, this gives

$$\begin{cases} i\hat{u}_t(\xi) = \frac{|\xi|^2}{2}\hat{u}(\xi) \\ \hat{u}(0,\xi) = \hat{u_0}(\xi) \end{cases}$$

[2]In this chapter we will adopt the convention that $\hat{f}(\xi) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x)\, dx$.

This is an ODE with respect to t, and can be solved by separation of variables. Indeed, the solution is given on the Fourier side by

$$\hat{u}(t,\xi) = e^{-i\frac{|\xi|^2}{2}t}\,\hat{u_0}(\xi). \tag{9.9}$$

Inverting the Fourier transform, we have

$$\begin{aligned} u(t,x) &= \frac{1}{(2\pi)^{\frac{d}{2}}}\int_{\mathbb{R}^d} e^{ix\cdot\xi - i\frac{|\xi|^2}{2}t}\,\hat{u_0}(\xi)\,d\xi \\ &= \frac{1}{(2\pi)^{\frac{d}{2}}}\int_{\mathbb{R}^d} e^{it\left(\frac{x}{t}\cdot\xi - \frac{|\xi|^2}{2}\right)}\,\hat{u_0}(\xi)\,d\xi. \end{aligned}$$

9.2.2 Dispersive Estimates

9.2.3 Strichartz Estimates

Chapter 10

Restriction Theory

In this chapter, we study one of the most important unsolved problems in harmonic analysis: the restriction problem. Vaguely, the restriction problem asks when one can meaningfully restrict the Fourier transform of a function to a subspace of $\mathbb{R}^d$. While this is an interesting question in harmonic analysis on its own, we will see that the restriction problem has close ties with dispersive partial differential equations. There are also connections to other major problems in harmonic analysis, such as the Kakeya conjecture, though we will only focus on the dispersive PDE connection in this chapter. For more information on restriction theory and its connections to other parts of harmonic analysis, refer to [10] and [11].

We begin by formalizing the statement of the restriction problem.

10.1 The Restriction Conjecture

A simple motivator for the restriction problem stems from the following observation. Fix $d \geq 2$. In Chapter 2 we defined the Fourier transform for functions in $L^p(\mathbb{R}^d)$ for $1 \leq p \leq 2$. If $f \in L^1(\mathbb{R}^d)$, then by the Riemann-Lebesgue lemma we know that $\hat{f} \in C_0(\mathbb{R}^d)$. In particular, we can meaningfully restrict $\hat{f}$ to a set of a measure 0 in $\mathbb{R}^d$, as $\hat{f}$ is continuous. On the other hand, if $f \in L^2(\mathbb{R}^d)$, we only know that $\hat{f} \in L^2(\mathbb{R}^d)$. As an arbitrary L^2 function can be redefined on sets of measure zero, we cannot restrict $\hat{f}$ to a sets of measure 0.

What about functions $f \in L^p(\mathbb{R}^d)$ for $1 < p < 2$? By the above discussion, restriction to measure zero sets is completely well-defined for $p = 1$ and entirely not well-defined for $p = 2$. Ideally, we would like a critical value between 1 and 2 for which restriction is well-defined for p smaller than this critical value, but this is unfortunately not the case. It turns out that there are functions that belong to $L^p(\mathbb{R}^d)$ *for all* $p > 1$ but for which $\hat{f}$ blows up on a hyperplane.

Example 10.1. Let $\psi : \mathbb{R}^{d-1} \to \mathbb{C}$ be a smooth bump function. Define $f : \mathbb{R}^d \to \mathbb{C}$ by

$$f(x_1, \ldots, x_d) = \frac{\psi(x_2, \ldots, x_d)}{1 + |x_1|}.$$

Observe that

$$\int_{\mathbb{R}^d} |f(x)|^p \, dx = \int_{\mathbb{R}^{d-1}} \psi(x_2, \ldots, x_d) \, dx_2 \ldots dx_d \cdot \int \frac{1}{(1 + |x_1|)^p} \, dx_1 < \infty$$

for any $p > 1$, hence $f \in L^p(\mathbb{R}^d)$ for all $p > 1$. The Fourier transform of f is

$$\begin{aligned}\hat{f}(\xi) &= (2\pi)^{-\frac{d}{2}} \int e^{-ix\cdot\xi} \frac{\psi(x_2, \ldots, x_d)}{1 + |x_1|}\, dx \\ &= (2\pi)^{-\frac{1}{2}} \hat{\psi}(\xi_2, \ldots, \xi_d) \int \frac{e^{-ix_1\xi_1}}{1 + |x_1|}\, dx_1.\end{aligned}$$

On the hyperplane $\{\xi_1 = 0\}$, the integral $\int \frac{e^{-ix_1\xi_1}}{1+|x_1|}\, dx_1 = \int \frac{1}{1+|x_1|}\, dx_1$ is divergent. Thus, the restriction of $\hat{f}$ to this set is not well-defined.

Stein discovered (see ??) that if a surface S of measure 0 has "sufficient curvature," then one may restrict $\hat{f}$ to S when $f \in L^p(\mathbb{R}^d)$ for certain values of p. As such, we can informally state the question of restriction of the Fourier transform as follows: for which zero-measure sets S and which $1 < p < 2$ can one meaningfully restrict $\hat{f}$ to S if $f \in L^p(\mathbb{R}^d)$?

There are three surfaces which have received a lot of attention:

- The sphere, $S_{sphere} = \left\{ \xi \in \mathbb{R}^d \;:\; |\xi| = 1 \right\}$.
- The paraboloid,[1] $S_{parab} = \left\{ \xi \in \mathbb{R}^d \;:\; \xi_d = \frac{1}{2}|\bar{\xi}|^2 \right\}$.
- The cone, $S_{cone} = \left\{ \xi \in \mathbb{R}^d \;:\; \xi_d = |\bar{\xi}| \right\}$.

These three surfaces are endowed with the following canonical measures:

- The sphere is given the usual surface measure $d\sigma$.
- The paraboloid is given the measure $d\sigma$ defined via

$$\int_{S_{parab}} f(\xi)\, d\sigma(\xi) := \int_{\mathbb{R}^{d-1}} f\left(\bar{\xi}, \frac{1}{2}|\bar{\xi}|^2\right) d\bar{\xi}.$$

- The cone is given the measure $d\sigma$ defined via

$$\int_{S_{cone}} f(\xi)\, d\sigma(\xi) := \int_{\mathbb{R}^{d-1}} f\left(\bar{\xi}, |\bar{\xi}|\right) \frac{d\bar{\xi}}{|\bar{\xi}|}.$$

With these surfaces and measures, we can give a quantitative formulation of the restriction problem. Explicitly, we seek an estimate of the form

$$\left\| \hat{f}\; |_S \right\|_{L^q(S, d\sigma)} \lesssim_{p,q,S} \|f\|_{L^p(\mathbb{R}^d)} \tag{10.1}$$

We shall refer to an estimate of the form (10.1) by $R(p \to q)$. For now, our focus will be on the case $S = S_{sphere}$.

First, a few remarks.

1. If $p = 1$, then $R(p \to q)$ holds for all $1 \le q \le \infty$. Indeed, by Holder's inequality,

$$\left\| \hat{f}\; |_S \right\|_{L^q(S, d\sigma)} \lesssim \sigma(S)^{\frac{1}{q}} \left\| \hat{f} \right\|_{L^\infty(S, d\sigma)} \lesssim \|f\|_{L^1(\mathbb{R}^d)}$$

   because S has finite measure.

2. If $p = 2$, then by our initial discussion $R(p \to q)$ fails for every $1 \le q \le \infty$.

[1] Here we are using the notation $\bar{\xi} = (x_1, \ldots, x_{d-1})$.

3. If $R(p \to q)$ holds, then $R(\tilde{p} \to \tilde{q})$ holds for all $\tilde{p} \leq p$ and $\tilde{q} \leq q$. Indeed, by Holder's inequality,

$$\left\| \hat{f} \mid_S \right\|_{L^{\tilde{q}}(S,d\sigma)} \lesssim \sigma(S)^{\frac{1}{\tilde{q}} - \frac{1}{q}} \left\| \hat{f} \mid_S \right\|_{L^q(S,d\sigma)}.$$

From here, let φ be a function such that $\hat{\varphi}$ is compactly supported and $\hat{\varphi} = 1$ on a ball containing S. Then

$$\left\| \hat{f} \mid_S \right\|_{L^{\tilde{q}}(S,d\sigma)} \lesssim \left\| \hat{f} \mid_S \cdot \hat{\varphi} \right\|_{L^q(S,d\sigma)} \lesssim \left\| \widehat{f * \varphi} \right\|_{L^q(S,d\sigma)} \lesssim \| f * \varphi \|_{L^p(\mathbb{R}^d)}$$

where the last inequality holds because $R(p \to q)$ holds. Next, we apply Young's inequality with r chosen so that $1 + \frac{1}{p} = \frac{1}{\tilde{p}} + \frac{1}{r}$. Note that, for such an r, we have $\frac{1}{r} = 1 + \frac{\tilde{p} - p}{\tilde{p}p} \leq 1$ so that this choice is possible. Then

$$\left\| \hat{f} \mid_S \right\|_{L^{\tilde{q}}(S,d\sigma)} \lesssim \| f \|_{L^{\tilde{p}}(\mathbb{R}^d)} \| \varphi \|_{L^r(\mathbb{R}^d)} \lesssim \| f \|_{L^{\tilde{p}}(\mathbb{R}^d)}$$

as desired.

Because of this, when studying restriction problems we adopt a philosophy of maximizing p and q for which $R(p \to q)$ holds.

In modern papers dealing with the restriction problem, the formulation is often presented using the *adjoint* of the expression $R(p \to q)$. In particular, let $R : L^p(\mathbb{R}^d) \to L^q(S, d\sigma)$ be the restriction operator given by $Rf = \hat{f} \mid_S$. Consider the adjoint operator $R^* : L^{q'}(S, d\sigma) \to L^{p'}(\mathbb{R}^d)$. For a Schwartz function f, we have

$$\begin{aligned}
\langle Rf, g \rangle_{L^2(S,d\sigma)} &= \int_S (Rf)(\xi) \overline{g(\xi)} \, d\sigma(\xi) = (2\pi)^{-\frac{d}{2}} \int_S \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x) \, dx \, \overline{g(\xi)} \, d\sigma(\xi) \\
&= \int_{\mathbb{R}^d} f(x) (2\pi)^{-\frac{d}{2}} \int_S \overline{e^{ix\cdot\xi} g(\xi) \, d\sigma(\xi)} \, dx \\
&= \int_{\mathbb{R}^d} f(x) \, \overline{(g \, d\sigma)\check{}} \, dx \\
&= \left\langle f, (g \, d\sigma)\check{} \right\rangle_{L^2(S,d\sigma)}.
\end{aligned}$$

This implies that $R^* g = (g \, d\sigma)\check{}$. Thus, $R(p \to q)$ holds if and only if $R^*(q' \to p')$ holds, i.e., if there is an estimate of the form

$$\left\| (g \, d\sigma)\check{} \right\|_{L^{p'}(\mathbb{R}^d)} \lesssim_{p,q,S} \| g \|_{L^{q'}(S,d\sigma)}. \tag{10.2}$$

We begin by considering necessary conditions for $R^*(q' \to p')$ to hold. It is natural to first analyze what happens when $g \equiv 1$. By Lemma 8.9, $|(d\sigma)\check{}| \lesssim \langle x \rangle^{-\frac{d-1}{2}}$. Thus, $(d\sigma)\check{} \in L^{p'}$ precisely when $\frac{d-1}{2} p' > d$, which is equivalent to $p < \frac{2d}{d+1}$. So for the function $g \equiv 1$, $R^*(q' \to p')$ holds when

$$p < \frac{2d}{d+1}.$$

This provides one necessary condition. Another condition comes from the following example.

Example 10.2 (Knapp). Let $\kappa \subseteq S$ be a cap on the sphere centered at the point $\xi_0 = (0, \ldots, 0, -1)$ of horizontal radius $1/R$ for some $R >> 1$. Near ξ_0, we can parametrize the cap κ and Taylor expand via

$$\xi_d = -\sqrt{1 - |\bar{\xi}|^2} = -\left(1 - \frac{|\bar{\xi}|^2}{2} + O(|\bar{\xi}|^4) \right) = -1 + O(R^{-2})$$

because $|\bar{\xi}| \leq 1/R$ on κ. This means that the ξ_d-height of the cap is on the order of R^{-2}, i.e., the cap κ is contained in a cylinder D of radius $1/R$ and thickness $\sim 1/R^2$. Let T denote the "dual" cylinder to D, i.e., the cylinder centered at 0 with radius R and height $\sim R^2$.

Let $g = \chi_\kappa$. Then

$$\|g\|_{L^{q'}(S,d\sigma)} = \sigma(\kappa)^{\frac{1}{q'}} \lesssim (R^{-1})^{\frac{d-1}{q'}}.$$

Also,

$$(g\, d\sigma)\check{}(x) = (2\pi)^{-\frac{d}{2}} \int_\kappa e^{ix\cdot\xi}\, d\sigma(\xi) = (2\pi)^{-\frac{d}{2}} e^{-ix_d} \int_\kappa e^{ix(\xi-\xi_0)}\, d\sigma(\xi).$$

Note that, for $\xi \in \kappa$,

$$|(\xi - \xi_0)_i| \lesssim \begin{cases} 1/R & 1 \leq i \leq d-1 \\ 1/R^2 & i = d \end{cases}$$

and for $x \in T$,

$$|x_i| \lesssim \begin{cases} R & 1 \leq i \leq d-1 \\ R^2 & i = d \end{cases}.$$

Thus, for most $x \in T$, we have

$$|(g\, d\sigma)\check{}(x)| \gtrsim \sigma(\kappa) \sim R^{-(d-1)}.$$

So

$$\left\|(g\, d\sigma)\check{}\right\|_{L^{p'}(\mathbb{R}^d)} \gtrsim \left\|(g\, d\sigma)\check{}\right\|_{L^{p'}(T)} \gtrsim R^{-(d-1)}|T|^{\frac{1}{p'}} \sim R^{-(d-1)} R^{\frac{d-1+2}{p'}} = R^{-(d-1)} R^{\frac{d+1}{p'}}.$$

Therefore, if $R^*(q' \to p')$ holds, then $R^{-(d-1)} R^{\frac{d+1}{p'}} \lesssim R^{-\frac{d-1}{q'}}$. Letting $R \to \infty$ gives the condition

$$\frac{d+1}{p'} \leq \frac{d-1}{q}.$$

From these two examples, two necessary scaling conditions for $R^*(q' \to p')$ to hold are $1 \leq \frac{2d}{d+1}$ and $\frac{d+1}{p'} \leq \frac{d-1}{q}$. One might wonder whether such conditions are also sufficient. Formally, we pose the following conjecture.

Conjecture 10.3 (Restriction conjecture for the sphere)**.** *Let $d \geq 2$, and let $S = S^{d-1} \subseteq \mathbb{R}^d$ be the unit sphere endowed with the canonical surface measure. Then*

$$\left\|\hat{f}\,|_S\right\|_{L^q(S,d\sigma)} \lesssim_{p,q} \|f\|_{L^p(\mathbb{R}^d)},$$

or equivalently

$$\left\|(g\, d\sigma)\check{}\right\|_{L^{p'}(\mathbb{R}^d)} \lesssim_{p,q} \|g\|_{L^{q'}(S,d\sigma)},$$

whenever $1 \leq p \leq \frac{2d}{d+1}$ and $1 \leq q \leq \frac{d-1}{d+1}p'$.

Zygmund proved this conjecture for $d = 2$ (see ??); for all other d, this remains an open problem.

10.2 The Tomas-Stein Inequality

Though the restriction conjecture for the sphere is open for $d > 2$, we have the following result, due to Tomas and Stein.

Theorem 10.4 (Tomas-Stein)**.** *Let $d \geq 2$, and let $S = S^{d-1} \subseteq \mathbb{R}^d$ be the unit sphere endowed with the canonical surface measure. Then*

$$\left\| \hat{f} \mid_S \right\|_{L^2(S,d\sigma)} \lesssim_p \|f\|_{L^p(\mathbb{R}^d)}$$

whenever $1 \leq p \leq \frac{2(d+1)}{d+3}$.

When $q = 2$, the scaling conditions in Conjecture 10.3 become $1 \leq p \leq \frac{2d}{d+1}$ and $p' \geq \frac{2(d+1)}{d-1}$, which is equivalent to $p \leq \frac{2(d+1)}{d-3}$. Moreover, for $d > 1$, $\frac{2(d+1)}{d+3} < \frac{2d}{d+1}$. Thus, the Tomas-Stein inequality is indeed the full restriction conjecture for the sphere in the special case $q = 2$.

Proof. For the sake of instruction, we begin with a proof attempt which will not prove the full statement of Theorem 10.4.

Attempt 1: The first approach relies on using the decay condition $|\check{\sigma}(x)| \lesssim \langle x \rangle^{-\frac{d-1}{2}}$. We wish to show that the operator $R : L^p(\mathbb{R}^d) \to L^2(S, d\sigma)$ is bounded, which is equivalent to showing boundedness of $R^* : L^2(S, d\sigma) \to L^{p'}(\mathbb{R}^d)$, which is equivalent to showing boundedness of $R^*R : L^p(\mathbb{R}^d) \to L^{p'}(\mathbb{R}^d)$.

Let f be a Schwartz function. By definition,

$$\begin{aligned}
(R^*Rf)(x) = (Rf\, d\sigma)\check{}(x) &= (2\pi)^{-\frac{d}{2}} \int_S e^{ix\cdot\xi}(Rf)(\xi)\, d\sigma(\xi) = (2\pi)^{-\frac{d}{2}} \int_S e^{ix\cdot\xi} \hat{f}(\xi)\, d\sigma(\xi) \\
&= (2\pi)^{-\frac{d}{2}} (2\pi)^{-\frac{d}{2}} \int_S e^{ix\cdot\xi} \int_{\mathbb{R}^d} e^{-iy\cdot\xi} f(y)\, dy\, d\sigma(\xi) \\
&= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} f(y) \check{\sigma}(x-y)\, dy.
\end{aligned}$$

Hence, $R^*Rf = f * \check{\sigma}$. Next, because $|\check{\sigma}(x)| \lesssim \langle x \rangle^{-\frac{d-1}{2}}$, we have $\check{\sigma} \in L^{\frac{2d}{d-1},\infty}(\mathbb{R}^d)$. By the Hardy-Littlewood-Sobolev inequality,

$$\|R^*Rf\|_{L^{p'}(\mathbb{R}^d)} \lesssim \|f\|_{L^p(\mathbb{R}^d)} \|\check{\sigma}\|_{L^{\frac{2d}{d-1},\infty}(\mathbb{R}^d)} \lesssim \|f\|_{L^p}$$

provided that $1 + \frac{1}{p'} = \frac{1}{p} + \frac{d-1}{2d}$. This is equivalent to $\frac{2}{p} = \frac{d-1}{2d}$, hence $p' = \frac{4d}{d-1}$, hence $p = \frac{4d}{3d+1}$. So $R(p \to 2)$ holds for all $1 \leq p \leq \frac{4d}{3d+1}$. But a quick calculation shows that $\frac{4d}{3d+1} < \frac{2(d+1)}{d+3}$, so we do not have the full scaling range in the statement of the theorem.

Attempt 2: To remedy this, we use both the decay of $|\check{\sigma}(x)|$ and the oscillation. We have

$$\varphi(x) + \sum_{N>1} \psi_N(x) = 1$$

where φ and ψ_N are the usual Littlewood projections, and N is a dyadic number. So we have a decomposition

$$f * \check{\sigma} = f * (\varphi\check{\sigma}) + \sum_{N>1} f * (\psi_N \check{\sigma}).$$

By the triangle inequality,

$$\|R^*Rf\|_{L^{p'}(\mathbb{R}^d)} \leq \|f * (\varphi\check{\sigma})\|_{L^{p'}(\mathbb{R}^d)} + \sum_{N>1} \|f * (\psi_N\check{\sigma})\|_{L^{p'}(\mathbb{R}^d)}.$$

Consider the first term. Note that $1 + \frac{1}{p'} = \frac{1}{p} + \frac{1}{r}$ when $r = \frac{p'}{2}$. Because $p < 2$, $p' > 2$, hence $r = \frac{p'}{2}$ is a sensible choice. So by Young's convolution inequality, we have

$$\|f * (\varphi\check{\sigma})\|_{L^{p'}(\mathbb{R}^d)} \lesssim \|f\|_{L^p(\mathbb{R}^d)} \|\varphi\check{\sigma}\|_{L^{\frac{p'}{2}}(\mathbb{R}^d)}.$$

As φ is a bounded function with compact support, $\|\varphi\check{\sigma}\|_{L^{\frac{p'}{2}}(\mathbb{R}^d)} < \infty$. Thus,

$$\|f * (\varphi\check{\sigma})\|_{L^{p'}(\mathbb{R}^d)} \lesssim \|f\|_{L^p(\mathbb{R}^d)}.$$

Next, we consider the sum. In order to retain summability, we seek an estimate of the form

$$\|f * (\psi_N\check{\sigma})\|_{L^{p'}(\mathbb{R}^d)} \lesssim N^{-\varepsilon(p)} \|f\|_{L^p(\mathbb{R}^d)}$$

for some $\varepsilon(p) > 0$. The strategy is to use real interpolation between a $L^1 \to L^\infty$ and $L^2 \to L^2$ type estimate.

First, because ψ_N is supported on an annulus of radius $\sim N$, the decay of $|\check{\sigma}|$ gives

$$\|f * (\psi_N\check{\sigma})\|_{L^\infty(\mathbb{R}^d)} \lesssim \|f\|_{L^1(\mathbb{R}^d)} \|\psi_N\check{\sigma}\|_{L^\infty(\mathbb{R}^d)} \lesssim N^{-\frac{d-1}{2}} \|f\|_{L^1(\mathbb{R}^d)}.$$

For the $L^2 \to L^2$ type estimate, we use Plancharel to get

$$\begin{aligned}\|f * (\psi_N\check{\sigma})\|_{L^2(\mathbb{R}^d)} &= \left\|\hat{f} \cdot (\psi_N\check{\sigma})\hat{}\right\|_{L^2(\mathbb{R}^d)} \\ &\leq \left\|\hat{f}\right\|_{L^2(\mathbb{R}^d)} \left\|\hat{\psi_N} * d\sigma\right\|_{L^\infty(\mathbb{R}^d)} = \|f\|_{L^2(\mathbb{R}^d)} \left\|\hat{\psi_N} * d\sigma\right\|_{L^\infty(\mathbb{R}^d)}.\end{aligned}$$

Expanding the convolution,

$$(\hat{\psi_N} * d\sigma)(x) = \int_S \hat{\psi_N}(x-y)d\sigma(y) = N^d \int_S \hat{\psi}(N(x-y))\, d\sigma(y).$$

So for any m, since ψ is Schwartz,

$$\begin{aligned}|(\hat{\psi_N} * d\sigma)(x)| &\lesssim N^d \int_S \frac{1}{\langle N(x-y)\rangle^m}\, d\sigma(y) \\ &= N^d \int_{|x-y|\leq 1/N;\, y\in S} d\sigma(y) + N^d \sum_{M\leq N} \int_{|x-y|\sim 1/M;\, y\in S} \left(\frac{M}{N}\right)^m d\sigma(y).\end{aligned}$$

The first integral $\int_{|x-y|\leq 1/N;\, y\in S} d\sigma(y)$ is the measure of a cap on the sphere of radius $\lesssim 1/N$, hence $\int_{|x-y|\leq 1/N;\, y\in S} d\sigma(y) \lesssim (1/N)^{d-1}$. Similarly, $\int_{|x-y|\sim 1/M;\, y\in S} d\sigma(y)$ is the measure of two caps on the sphere with radius $\lesssim 1/M$, so in total we have

$$|(\hat{\psi_N} * d\sigma)(x)| \lesssim N^d \left(\frac{1}{N}\right)^{d-1} + N^d \sum_{M\leq N} \left(\frac{M}{N}\right)^m \left(\frac{1}{M}\right)^{d-1} = N + N \sum_{M\leq N} \left(\frac{M}{N}\right)^{m-(d-1)}.$$

The sum in question is finite if $m > d-1$. For such a choice of m, we then have $|(\hat{\psi_N} * d\sigma)(x)| \lesssim N$, hence $\left\|\hat{\psi_N} * d\sigma\right\|_{L^\infty(\mathbb{R}^d)} \lesssim N$. Therefore, $\|f * (\psi_N\check{\sigma})\|_{L^2(\mathbb{R}^d)} \lesssim N \|f\|_{L^2(\mathbb{R}^d)}$.

Now we interpolate between the estimates $\|f * (\psi_N\check{\sigma})\|_{L^\infty(\mathbb{R}^d)} \lesssim N^{-\frac{d-1}{2}} \|f\|_{L^1(\mathbb{R}^d)}$ and $\|f * (\psi_N\check{\sigma})\|_{L^2(\mathbb{R}^d)} \lesssim N \|f\|_{L^2(\mathbb{R}^d)}$. Note that

$$\frac{1}{p'} = \frac{2/p'}{2} + \frac{1-2/p'}{\infty}.$$

Thus,

$$\begin{aligned}\|f * (\psi_N\check{\sigma})\|_{L^{p'}(\mathbb{R}^d)} &\lesssim \|f * (\psi_N\check{\sigma})\|_{L^2(\mathbb{R}^d)}^{\frac{2}{p'}} \|f * (\psi_N\check{\sigma})\|_{L^\infty(\mathbb{R}^d)}^{1-\frac{2}{p'}} \\ &\lesssim N^{\frac{2}{p'}} N^{-\frac{d-1}{2}\left(1-\frac{2}{p'}\right)} \|f\|_{L^2(\mathbb{R}^d)}^{\frac{2}{p'}} \|f\|_{L^1(\mathbb{R}^d)}^{1-\frac{2}{p'}} \\ &\lesssim N^{-\frac{d-1}{2}+\frac{d+1}{p'}} \|f\|_{L^p(\mathbb{R}^d)}.\end{aligned}$$

Set $\varepsilon(p) = -\frac{d-1}{2} + \frac{d+1}{p'}$. Then $\varepsilon(p) > 0$ precisely when $p' > \frac{2(d+1)}{d-1}$, or equivalently $p < \frac{2(d+1)}{d+3}$. This gives the desired estimate in all cases, save for the endpoint $p = \frac{2(d+1)}{d+3}$.

Attempt 3: The above argument, which proves the theorem except for the endpoint case, is due to Tomas. The loss of the endpoint is essentially due to the use of the triangle inequality in the dyadic sum. To recover the endpoint case, Stein used complex interpolation with estimates of the form

$$\left\| \sum_{N>1} N^{\frac{d-1}{2}+it} f * (\psi_N \check{\sigma}) \right\|_{L^\infty(\mathbb{R}^d)} \lesssim \|f\|_{L^1(\mathbb{R}^d)}$$

$$\left\| \sum_{N>1} N^{-1+it} f * (\psi_N \check{\sigma}) \right\|_{L^2(\mathbb{R}^d)} \lesssim \|f\|_{L^2(\mathbb{R}^d)} .$$

We will not consider the details of the argument here; for reference, see [10].

□

10.3 Restriction Theory and Strichartz Estimates

As alluded to in the introduction of this chapter, the restriction of the Fourier transform is deeply connected to dispersive partial differential equations. In this section, we will describe this connection by studying an analogue of Theorem 10.4 for the paraboloid, and showing its equivalence to a symmetric Strichartz inequality.

In this setting, because Strichartz estimates involve space and time, it will be convenient to increase the dimension of our restriction estimates from d to $d+1$. As such, we now seek an estimate of the form

$$\left\| (g\, d\sigma)\check{} \right\|_{L^{\frac{2(d+2)}{d}}(\mathbb{R}^{d+1})} \lesssim \|g\|_{L^2(S_{parab}, d\sigma)} \tag{10.3}$$

where $d\sigma$ is the canonical measure on the paraboloid. This is dual formulation of the endpoint case for the Tomas-Stein inequality on the paraboloid.

For the reader's convenience, we recall some estimates from Chapter 9. Recall that the free propagator $e^{it\frac{\Delta}{2}}$ is defined for Schwartz functions by

$$e^{it\frac{\Delta}{2}} f = \mathcal{F}_\xi^{-1}\left(e^{-it\frac{|\xi|^2}{2}} f(\xi) \right)$$

where $u_0(x) = u(0,x)$. By interpolating the estimates

$$\left\| e^{it\frac{\Delta}{2}} f \right\|_{L^2_x} = \|f\|_{L^2_x} \quad \text{and} \quad \left\| e^{it\frac{\Delta}{2}} f \right\|_{L^\infty_x} \lesssim |t|^{-\frac{d}{2}} \|f\|_{L^1_x}$$

we obtained (see Theorem ??) the dispersive estimate

$$\left\| e^{it\frac{\Delta}{2}} f \right\|_{L^p_x} \lesssim |t|^{-d\left(\frac{1}{2}-\frac{1}{p}\right)} \|f\|_{L^{p'}_x} \tag{10.4}$$

for $2 \le p \le \infty$. From this, we derived (see Theorem ??) the following inequality.

Theorem 10.5 (Strichartz inequality). *Let $f \in \mathcal{S}(\mathbb{R}^d_x)$. Then*

$$\left\| e^{it\frac{\Delta}{2}} f \right\|_{L^q_t L^r_x} \lesssim \|f\|_{L^2_x}$$

for $2 \le q, r \le \infty$ satisfying $\frac{2}{q} + \frac{d}{r} = \frac{d}{2}$ and $(d,q,r) \neq (2,2,\infty)$.

When $q = r$, the above inequality is known as a *symmetric Strichartz inequality.* In this case, the scaling relation gives

$$\frac{2}{q} + \frac{d}{q} = \frac{d}{2} \quad \Rightarrow \quad q = \frac{2(d+2)}{d}.$$

As $\frac{2(d+2)}{d} > 2$, this proves:

Theorem 10.6 (Symmetric Strichartz inequality)**.** *Let $f \in \mathcal{S}(\mathbb{R}^d_x)$. Then*

$$\left\| e^{it\frac{\Delta}{2}} f \right\|_{L_{t,x}^{\frac{2(d+2)}{d}}} \lesssim \|f\|_{L^2_x}.$$

We will now prove the following result.

Proposition 10.7. *The Tomas-Stein inequality for the paraboloid is equivalent to the symmetric Strichartz inequality; that is, the estimate*

$$\left\| (g\, d\sigma)\check{} \right\|_{L^{\frac{2(d+2)}{d}}(\mathbb{R}^{d+1})} \lesssim \|g\|_{L^2(S_{parab}, d\sigma)}$$

is equivalent to the estimate

$$\left\| e^{it\frac{\Delta}{2}} f \right\|_{L_{t,x}^{\frac{2(d+2)}{d}}} \lesssim \|f\|_{L^2_x}.$$

Proof. Let $u_0 \in \mathcal{S}(\mathbb{R}^d)$ and $\varphi \in \mathcal{S}(\mathbb{R} \times \mathbb{R}^d)$. We adopt the notation $\tilde{\varphi} = \mathcal{F}_{t,x}\varphi$ to denote the space-time Fourier transform of φ. Let $\psi = \mathcal{F}_x^{-1}\tilde{\varphi}$.

We wish to calculate the space-time Fourier transform of $e^{it\frac{\Delta}{2}}u_0$, in the distributional sense. We have

$$\begin{aligned}
\left\langle \widetilde{e^{it\frac{\Delta}{2}}u_0}, \varphi \right\rangle_{t,x} &= \left\langle e^{it\frac{\Delta}{2}}u_0, \tilde{\varphi} \right\rangle_{t,x} = (2\pi)^{-\frac{d}{2}} \left\langle e^{it\frac{\Delta}{2}}u_0, \int_{\mathbb{R}^d} e^{-ix\cdot\xi}\psi(t,\xi)\, d\xi \right\rangle_{t,x} \\
&= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}}\int_{\mathbb{R}^d} \left(e^{it\frac{\Delta}{2}}u_0\right)(x) e^{-ix\cdot\xi}\psi(t,\xi)\, dx\, dt\, d\xi \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}} e^{-it\frac{|\xi|^2}{2}}\hat{u_0}(\xi)\psi(t,\xi)\, dt\, d\xi \\
&= (2\pi)^{\frac{1}{2}} \int_{\mathbb{R}^d} \hat{u_0}(\xi)\ [\mathcal{F}_t\psi(\xi)] \left(\frac{|\xi|^2}{2}\right) d\xi.
\end{aligned}$$

Observe that $[\mathcal{F}_t\psi(\xi)]\left(\frac{|\xi|^2}{2}\right) = \varphi\left(\frac{|\xi|^2}{2}, \xi\right)$. Thus,

$$\left\langle \widetilde{e^{it\frac{\Delta}{2}}u_0}, \varphi \right\rangle_{t,x} = (2\pi)^{\frac{1}{2}} \left\langle \hat{u_0}\, d\sigma_{parab}, \varphi \right\rangle_{t,x}.$$

So, in the distributional sense, and up to a factor of $(2\pi)^{\frac{1}{2}}$,

$$\widetilde{e^{it\frac{\Delta}{2}}u_0} = \hat{u_0}\, d\sigma_{parab} \quad \Rightarrow \quad e^{it\frac{\Delta}{2}}u_0 = \mathcal{F}_{t,x}\left(\hat{u_0}\, d\sigma_{parab}\right).$$

Let $g\left(\xi, \frac{|\xi|^2}{2}\right) = \hat{u_0}(\xi)$. Then

$$\|g\|_{L^2(S_{parab}, d\sigma)} = \|\hat{u_0}\|_{L^2} = \|u_0\|_{L^2}$$

and $(g\, d\sigma_{parab})\check{} = e^{it\frac{\Delta}{2}}u_0$. Thus, the estimate $\left\| e^{it\frac{\Delta}{2}}u_0 \right\|_{L_{t,x}^{\frac{2(d+2)}{d}}} \lesssim \|u_0\|_{L^2}$ yields the Tomas-Stein estimate

$$\left\| (g\, d\sigma)\check{} \right\|_{L^{\frac{2(d+2)}{d}}(\mathbb{R}^{d+1})} \lesssim \|g\|_{L^2(S_{parab}, d\sigma)}.$$

The converse implication is essentially the same. $\square$

This proves that the endpoint Tomas-Stein inequality holds for the paraboloid by way of the symmetric Strichartz estimate, which by our remark earlier in this chapter proves the full Tomas-Stein inequality for the paraboloid. However, it is of interest to prove that the Tomas-Stein inequality on the sphere also implies the Tomas-Stein estimate on the paraboloid, without appealing directly to Strichartz estimates.

Proposition 10.8. *The Tomas-Stein estimate on the sphere implies the Tomas-Stein estimate on the paraboloid.*

Before proving this proposition, we explicitly compute the surface area measure on the sphere.

Lemma 10.9. *The canonical surface measure on the unit sphere $S^d \subseteq \mathbb{R}^{d+1}$ is given explicitly by*

$$d\sigma_{S^d} = \frac{1}{\sqrt{1-|\xi|^2}}\, d\xi.$$

Proof. Let $\phi : \mathbb{R}^d \to \mathbb{R}^{d+1}$ be a coordinate parametrization of the sphere given by $\phi(\xi) = (\xi, \sqrt{1-|\xi|^2})$. Then

$$d\sigma_{S^d} = \sqrt{\det\left[(\nabla\phi)^T(\nabla\phi)\right]}\, d\xi.$$

It remains to compute $\sqrt{\det\left[(\nabla\phi)^T(\nabla\phi)\right]}$.

We have

$$\nabla\phi(\xi) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \\ \frac{-\xi_1}{\sqrt{1-|\xi|^2}} & \frac{-\xi_2}{\sqrt{1-|\xi|^2}} & \cdots & \frac{-\xi_d}{\sqrt{1-|\xi|^2}} \end{pmatrix}.$$

Thus,

$$(\nabla\phi(\xi))^T(\nabla\phi(\xi)) = \left[\delta_{ij} + \frac{\xi_i\xi_j}{1-|\xi|^2}\right]_{1\le i,j\le d} = I + \frac{|\xi|^2}{1-|\xi|^2}P_\xi$$

where $P_\xi = \left[\frac{\xi_i\xi_j}{|\xi|^2}\right]$ is the projection matrix onto the ξ direction.

Consequently,

$$(\nabla\phi(\xi))^T(\nabla\phi(\xi))\xi = \left(I + \frac{|\xi|^2}{1-|\xi|^2}P_\xi\right)\xi = \xi + \frac{|\xi|^2}{1-|\xi|^2}\xi = \left(\frac{1}{1-|\xi|^2}\right)\xi$$

so that $(\nabla\phi)^T(\nabla\phi)$ has an eigenvector ξ with eigenvalue $\frac{1}{1-|\xi|^2}$. Next, observe that if x is a vector orthogonal to ξ, then

$$\left(I + \frac{|\xi|^2}{1-|\xi|^2}P_\xi\right)x = x$$

so that $(\nabla\phi)^T(\nabla\phi)$ has a $d-1$-dimensional eigenspace corresponding to the eigenvalue 1. Thus,

$$\sqrt{\det\left[(\nabla\phi)^T(\nabla\phi)\right]} = \frac{1}{\sqrt{1-|\xi|^2}}$$

as desired.

□

With this, we can prove Proposition 10.8.

Proof. Assume the Tomas-Stein estimate on the sphere, that is,

$$\left\|(g\, d\sigma_{S^d})\check{}\right\|_{L^{\frac{2(d+2)}{d}}(\mathbb{R}^{d+1})} \lesssim \|g\|_{L^2(S^d, d\sigma_{S^d})}.$$

Because Tomas-Stein on the paraboloid is equivalent to the symmetric Strichartz estimate, our goal is to derive an estimate of the form

$$\left\|e^{it\frac{\Delta}{2}}u_0\right\|_{L_{t,x}^{\frac{2(d+2)}{d}}(\mathbb{R}\times\mathbb{R}^d)} \lesssim \|u_0\|_{L^2_x}$$

where we can take $\hat{u_0} \in C_c^\infty(\mathbb{R}^d)$ by density.

Fix $\lambda >> 1$. Define a function g_λ on the sphere S^d as follows:

$$\begin{cases} g_\lambda(\xi, \sqrt{1-|\xi|^2}) = \lambda^{\frac{d}{2}}\hat{u_0}(\lambda\xi)\sqrt{1-|\xi|^2} \\ g_\lambda(\xi, -\sqrt{1-|\xi|^2}) = 0 \end{cases}.$$

Because $\hat{u_0} \in C_c^\infty(\mathbb{R}^d)$, for λ large, g_λ is supported on a small ξ-region, i.e., a small cap around the north pole. Using the surface area lemma, we have

$$\begin{aligned}
\|g_\lambda\|^2_{L^2(S^d, d\sigma_{S^d})} &= \lambda^d \int |\hat{u_0}(\lambda\xi)|^2(1-|\xi|^2)\frac{1}{\sqrt{1-|\xi|^2}}\, d\xi \\
&= \lambda^d \int |\hat{u_0}(\lambda\xi)|^2\sqrt{1-|\xi|^2}\, d\xi \\
&= \int |\hat{u_0}(\xi)|^2\sqrt{1-\frac{|\xi|^2}{\lambda}}\, d\xi.
\end{aligned}$$

So for very large λ we have $\|g_\lambda\|^2_{L^2(S^d, d\sigma_{S^d})} \sim \|\hat{u_0}\|_{L^2_\xi} = \|u_0\|_{L^2_x}$.

Next, we compute $(g_\lambda\, d\sigma_{S^d})\check{}$:

$$\begin{aligned}
(g_\lambda\, d\sigma_{S^d})\check{}(t,x) &= (2\pi)^{-\frac{d+1}{2}} \int e^{it\omega + ix\cdot\xi}\lambda^{\frac{d}{2}}\hat{u_0}(\lambda\xi)\sqrt{1-|\xi|^2}\frac{1}{\sqrt{1-|\xi|^2}}\, d\xi\, \delta(\omega = \sqrt{1-|\xi|^2}) \\
&= (2\pi)^{-\frac{d+1}{2}} \int e^{it\sqrt{1-|\xi|^2}+ix\cdot\xi}\lambda^{\frac{d}{2}}\hat{u_0}(\lambda\xi)\, d\xi \\
&= (2\pi)^{-\frac{d+1}{2}}\lambda^{-\frac{d}{2}} \int e^{it\sqrt{1-\frac{|\xi|^2}{\lambda^2}}+i\frac{x}{\lambda}\cdot\xi}\hat{u_0}(\xi)\, d\xi.
\end{aligned}$$

By Taylor expansion, we have $\sqrt{1-\frac{|\xi|^2}{\lambda^2}} = 1 - \frac{|\xi|^2}{2\lambda^2} + O\left(\frac{|\xi|^4}{\lambda^4}\right)$. Also note that

$$(2\pi)^{-\frac{d+1}{2}}\lambda^{-\frac{d}{2}} \int e^{it-\frac{it|\xi|^2}{2\lambda^2}+i\frac{x}{\lambda}\cdot\xi}\hat{u_0}(\xi)\, d\xi = (2\pi)^{-\frac{d+1}{2}}\lambda^{-\frac{d}{2}}e^{it}\cdot\left[e^{i\frac{t}{\lambda^2}\frac{\Delta}{2}}u_0\right]\left(\frac{x}{\lambda}\right).$$

Thus, we can write

$$\begin{aligned}
(g_\lambda\, d\sigma_{S^d})\check{}(t,x) &= (2\pi)^{-\frac{d+1}{2}}\lambda^{-\frac{d}{2}}e^{it}\cdot\left[e^{i\frac{t}{\lambda^2}\frac{\Delta}{2}}u_0\right]\left(\frac{x}{\lambda}\right) \\
&\quad + (2\pi)^{-\frac{d+1}{2}}\lambda^{-\frac{d}{2}} \int e^{it-\frac{it|\xi|^2}{2\lambda^2}+i\frac{x}{\lambda}\cdot\xi}\hat{u_0}(\xi)\left[e^{it\left(\sqrt{1-\frac{|\xi|^2}{\lambda^2}}-1+\frac{|\xi|^2}{2\lambda^2}\right)} - 1\right] d\xi.
\end{aligned}$$

By the Taylor expansion comment and by the usual arc-length type estimate, for λ large we have

$$\left|e^{it\left(\sqrt{1-\frac{|\xi|^2}{\lambda^2}}-1+\frac{|\xi|^2}{2\lambda^2}\right)} - 1\right| \lesssim |t|\frac{|\xi|^4}{\lambda^4}.$$

Consequently,

$$\left|(2\pi)^{-\frac{d+1}{2}}\lambda^{-\frac{d}{2}}\int e^{it-\frac{it|\xi|^2}{2\lambda^2}+i\frac{x}{\lambda}\cdot\xi}\hat{u_0}(\xi)\left[e^{it\left(\sqrt{1-\frac{|\xi|^2}{\lambda^2}}-1+\frac{|\xi|^2}{2\lambda^2}\right)}-1\right]d\xi\right|$$
$$\lesssim \lambda^{-\frac{d}{2}-4}|t|\int|\hat{u_0}(\xi)||\xi|^4\,d\xi \lesssim \lambda^{-\frac{d}{2}-4}|t|$$

where the last inequality follows from the fact that $\hat{u_0}\in C_c^\infty(\mathbb{R}^d)$. Thus,

$$\begin{aligned}\left\|(g_\lambda\,d\sigma_{S^d})\check{}\right\|_{L^{\frac{2(d+2)}{d}}(\mathbb{R}^{d+1})} &\gtrsim \left\|(g_\lambda\,d\sigma_{S^d})\check{}\right\|_{L^{\frac{2(d+2)}{d}}\left(|t|\le\lambda^{2(1+\varepsilon)},|x|\le\lambda^{1+\varepsilon}\right)}\\ &\gtrsim \lambda^{-\frac{d}{2}}\left\|\left[e^{i\frac{t}{\lambda^2}\frac{\Delta}{2}}u_0\right]\left(\frac{x}{\lambda}\right)\right\|_{L^{\frac{2(d+2)}{d}}\left(|t|\le\lambda^{2(1+\varepsilon)},|x|\le\lambda^{1+\varepsilon}\right)}\\ &\quad-\lambda^{-\frac{d}{2}-4}\lambda^{2(1+\varepsilon)}\lambda^{(2(1+\varepsilon)+d(1+\varepsilon))\frac{d}{2(d+2)}}\end{aligned}$$

where the final power of λ comes from considering the volume of the set. Changing variables and simplifying the power of λ in the error term then gives

$$\left\|(g_\lambda\,d\sigma_{S^d})\check{}\right\|_{L^{\frac{2(d+2)}{d}}(\mathbb{R}^{d+1})}\gtrsim\lambda^{-\frac{d}{2}}\lambda^{(2+d)\frac{d}{2(d+2)}}\left\|e^{it\frac{\Delta}{2}}u_0\right\|_{L^{\frac{2(d+2)}{d}}(|t|\le\lambda^{2\varepsilon},|x|\le|\lambda|^{\varepsilon})}-\lambda^{-2+2\varepsilon+\frac{d\varepsilon}{2}}.$$

Thus,

$$\left\|e^{it\frac{\Delta}{2}}u_0\right\|_{L_{t,x}^{\frac{2(d+2)}{d}}(|t|\le\lambda^{2\varepsilon},|x|\le|\lambda|^{\varepsilon})}+O\left(\lambda^{-2+\frac{d+4}{2}\varepsilon}\right)\lesssim\|u_0\|_{L_x^2}.$$

As $\lambda\to\infty$, for $\varepsilon << 1$ we then get

$$\left\|e^{it\frac{\Delta}{2}}u_0\right\|_{L_{t,x}^{\frac{2(d+2)}{d}}\left(\mathbb{R}^{d+1}\right)}\lesssim\|u_0\|_{L_x^2}$$

as desired. □

In the end, via either method, we have proven the Tomas-Stein inequality on the paraboloid.

Theorem 10.10 (Tomas-Stein)**.** *Let $d\ge d$, and let $S_{parab}\subseteq\mathbb{R}^{d+1}$ be the paraboloid endowed with the canonical surface measure. Then*

$$\left\|(g\,d\sigma)\check{}\right\|_{L^{p'}(\mathbb{R}^{d+1})}\lesssim\|g\|_{L^2(S_{parab},d\sigma)}$$

whenever $p'\ge\frac{2(d+2)}{d}$.

Chapter 11

Rearrangement Theory

Our next topic of study is the theory of rearrangements. Rearrangement inequalities have applications to problems involving the minimization of energy functionals and other variational problems, and can be used in showing the existence and uniqueness of minimizers. Later in this chapter, we will see an explicit application of this type to the energy functional of the hydrogen atom. There are also connections to classical problems such as the isoperimetric inequality.

A good reference for the following material is Chapter 3 of [6].

11.1 Definitions and Basic Estimates

We begin by recalling the *layer-cake* decomposition of a Borel measurable function $f : \mathbb{R}^d \to [0, \infty)$; namely, the identity

$$f(x) = \int_0^\infty \chi_{\{f>\lambda\}}(x)\, d\lambda.$$

This equality follows from the observation that

$$\chi_{\{f>\lambda\}}(x) = \begin{cases} 1 & f(x) > \lambda \\ 0 & f(x) \le \lambda \end{cases} = \begin{cases} 1 & \lambda < f(x) \\ 0 & \lambda \ge f(x) \end{cases} = \chi_{[0,f(x))}(\lambda)$$

and so

$$\int_0^\infty \chi_{\{f>\lambda\}}(x)\, d\lambda = \int_0^\infty \chi_{[0,f(x))}(\lambda)\, d\lambda = \int_0^{f(x)} d\lambda = f(x).$$

This decomposition will frequently be used in the proofs of the inequalities to come. With this in mind, we begin discussing rearrangements.

Definition 11.1. If $A \subseteq \mathbb{R}^d$ is Borel measurable, its **symmetric rearrangement** is the set

$$A^* = \Big\{ x \in \mathbb{R}^d \ : \ |x| < r \Big\}$$

where r is chosen so that $|A^*| = |A|$.

Definition 11.2. Let $f : \mathbb{R}^d \to \mathbb{C}$ be a Borel measurable function which **vanishes at infinity**, that is, $|\{|f| > \lambda\}| < \infty$ for all $\lambda > 0$. The **symmetric decreasing rearrangement** of f is given by

$$f^*(x) = \int_0^\infty \chi_{\{|f|>\lambda\}^*}(x)\, d\lambda.$$

The intuition and geometry behind the rearrangement of a set A is clear: A^* is simply the open ball centered at the origin with the same measure as A. To give intuition for the symmetric rearrangement of a function, we compute a simple example. Suppose that $f : \mathbb{R} \to \mathbb{C}$ is given by $f(x) = \chi_{[-2,-1]}(x) + \chi_{[1,2]}(x)$. Fix $\lambda > 0$. Then

$$\{|f| > \lambda\}^* = \begin{cases} ([-2,-1] \cup [1,2])^* & 0 < \lambda < 1 \\ \emptyset^* & \lambda \geq 1 \end{cases} = \begin{cases} (-1,1) & 0 < \lambda < 1 \\ \emptyset & \lambda \geq 1 \end{cases}.$$

Therefore,

$$f^*(x) = \int_0^1 \chi_{\{|f|>\lambda\}^*}(x)\, d\lambda = \int_0^1 \chi_{(-1,1)}(x)\, d\lambda = \begin{cases} 1 & x \in (-1,1) \\ 0 & x \notin (-1,1) \end{cases} = \chi_{(-1,1)}(x).$$

For a slightly more interesting example, suppose that $g : \mathbb{R} \to \mathbb{C}$ is given by $g(x) = \chi_{[-2,-1]}(x) + 2\chi_{[1,2]}(x)$. Fix $\lambda > 0$. Then

$$\{|g| > \lambda\}^* = \begin{cases} ([-2,-1] \cup [1,2])^* & 0 < \lambda < 1 \\ ([1,2])^* & 1 \leq \lambda < 2 \\ \emptyset^* & \lambda \geq 2 \end{cases} = \begin{cases} (-1,1) & 0 < \lambda < 1 \\ \left(-\frac{1}{2}, \frac{1}{2}\right) & 1 \leq \lambda < 2 \\ \emptyset & \lambda \geq 2 \end{cases}.$$

Therefore,

$$g^*(x) = \int_0^1 \chi_{(-1,1)}(x)\, d\lambda + \int_1^2 \chi_{\left(-\frac{1}{2},\frac{1}{2}\right)}(x)\, d\lambda = \chi_{(-1,1)}(x) + \chi_{\left(-\frac{1}{2},\frac{1}{2}\right)}(x).$$

Roughly, the symmetric decreasing rearrangement of a function takes the mass under the graph of the function and rearranges it to return a function which is radially symmetric and decreasing with the same total mass.

These examples suggest the following basic properties of symmetric decreasing rearrangements.

Proposition 11.3. *If $f : \mathbb{R}^d \to \mathbb{C}$ is a Borel measurable function which vanishes at infinity, then f^* is nonnegative, radially symmetric, radially non-increasing, and lower-semicontinuous.*

Proof. Nonnegativity is clear. Radial symmetry of f^* follows from the fact that the set $\{|f| > \lambda\}^*$ is radially symmetric. The fact that f^* is radially decreasing follows from the observation that if $\lambda > \mu$, then $\{|f| > \lambda\}^* \subseteq \{|f| > \mu\}^*$.

To prove lower-semicontinuity, recall that f^* is lower-semicontinuous if and only if $\{f^* > \lambda\}$ is open for all λ. We have

$$\left\{ x \in \mathbb{R}^d \ : \ f^*(x) > \lambda \right\} = \left\{ x \in \mathbb{R}^d \ : \ \int_0^\infty \chi_{\{|f|>t\}^*}(x)\, dt > \lambda \right\} = \bigcup_{t>\lambda} \{|f| > t\}^*.$$

Thus, $\{f^* > \lambda\}$ is open. ☐

Next, observe that the rearrangement function satisfies the following monotonicity property.

Proposition 11.4. *If $f, g : \mathbb{R}^d \to \mathbb{C}$ are Borel measurable functions which vanishes at infinity, and if $|f| \leq |g|$, then $f^* \leq g^*$.*

Proof. If $|f| \leq |g|$, then $\{|f| > \lambda\} \subseteq \{|g| > \lambda\}$. Thus, $\{|f| > \lambda\}^* \subseteq \{|g| > \lambda\}^*$, and the claim follows. ☐

Next, we make an observation that will be used frequently. By definition,

$$f^*(x) = \int_0^\infty \chi_{\{|f|>\lambda\}^*}(x)\, d\lambda.$$

On the other hand, we have the layer-cake decomposition of f^*, which gives

$$f^*(x) = \int_0^\infty \chi_{\{f^*>\lambda\}}(x)\, d\lambda.$$

This suggests the equality $\{|f| > \lambda\}^* = \{f^* > \lambda\}$. Indeed, this is the case.

Proposition 11.5. *If $f : \mathbb{R}^d \to \mathbb{C}$ is a Borel measurable function which vanishes at infinity, then*

$$\{|f| > \lambda\}^* = \{f^* > \lambda\}.$$

Proof. Note that

$$|\{|f| > t\}^*| = |\{|f| > t\}| = \int_{\mathbb{R}^d} \chi_{\{|f|>t\}}(x)\, dx.$$

As $t \to \lambda$ from above, by the dominated convergence theorem, it then follows that

$$|\{|f| > t\}^*| \to \int_{\mathbb{R}^d} \chi_{\{|f|>\lambda\}}(x)\, dx = |\{|f| > \lambda\}| = |\{|f| > \lambda\}^*|.$$

Because $\{f^* > \lambda\} = \bigcup_{t>\lambda}\{|f| > t\}^*$, we have thus shown $|\{f^* > \lambda\}| = |\{|f| > \lambda\}^*|$. Because f^* is radially symmetric and lower-semicontinuous, $\{f^* > \lambda\}$ is an open ball centered at the origin. By definition of a set rearrangement, $\{|f| > \lambda\}^*$ is also an open ball centered at the origin. As two open balls of equal measure with same center, $\{|f| > \lambda\}^* = \{f^* > \lambda\}$. □

Corollary 11.6. *If $f : \mathbb{R}^d \to \mathbb{C}$ is a Borel measurable function which vanishes at infinity, then*

$$\|f\|_{L^p(\mathbb{R}^d)} = \|f^*\|_{L^p(\mathbb{R}^d)}$$

for $1 \leq p \leq \infty$.

Proof. This follows from

$$\|f\|^p_{L^p(\mathbb{R}^d)} = p\int_0^\infty \lambda^{p-1}|\{|f| > \lambda\}|\, d\lambda$$

when $1 \leq p < \infty$. The case $p = \infty$ is similar. □

In preparation for our first major result, we make a few remarks. We say that a function $f : \mathbb{R}^d \to [0, \infty)$ is *strictly symmetric decreasing* if f is radial and if $|x| > |y|$ implies that $f(x) < f(y)$. Note that this implies $f(x) > 0$ for all $x \in \mathbb{R}^d$, and furthermore $f^* = f$.

Theorem 11.7 (Hardy-Littlewood). *Let $f, g : \mathbb{R}^d \to [0, \infty)$ be vanishing at infinity. Then*

$$\int f(x)g(x)\, dx \leq \int f^*(x)g^*(x)\, dx.$$

Moreover, if f is strictly symmetric decreasing, then equality holds if and only if $g = g^*$ almost everywhere.

Proof. First, we demonstrate the inequality. We compute, using a layer-cake decomposition:

$$\begin{aligned}\int f(x)g(x)\, dx &= \int_{\mathbb{R}^d}\int_0^\infty \chi_{\{f>\lambda\}}(x)\, d\lambda \int_0^\infty \chi_{\{g>\mu\}}\, d\mu\, dx \\ &= \int_0^\infty \int_0^\infty |\{f > \lambda\} \cap \{g > \mu\}|\, d\lambda\, d\mu.\end{aligned}$$

Suppose without loss of generality that $|\{f > \lambda\}| \leq |\{g > \mu\}|$. Then $\{f > \lambda\}^* \subseteq \{g > \mu\}^*$. So

$$\begin{aligned}|\{f > \lambda\} \cap \{g > \mu\}| &\leq |\{f > \lambda\}| = |\{f > \lambda\}^*| = |\{f > \lambda\}^* \cap \{g > \mu\}^*| \\ &= \int \chi_{\{f>\lambda\}^*}(x)\chi_{\{g>\mu\}^*}(x)\, dx\end{aligned}$$

and hence

$$\int f(x)g(x)\,dx \leq \int_{\mathbb{R}^d} \int_0^\infty \chi_{\{f>\lambda\}^*}(x)\,d\lambda \int_0^\infty \chi_{\{g>\mu\}^*}(x)\,d\mu\,dx = \int f^*(x)g^*(x)\,dx.$$

Next, we consider the case of equality. Suppose that f is strictly symmetric decreasing and that $\int f(x)g(x)\,dx = \int f^*(x)g^*(x)\,dx$. Then $\int f(x)g(x)\,dx = \int f(x)g^*(x)\,dx$, hence

$$\int_0^\infty \int_{\mathbb{R}^d} f(x)\chi_{\{g>\mu\}}(x)\,dx\,d\mu = \int_0^\infty \int_{\mathbb{R}^d} f(x)\chi_{\{g>\mu\}^*}(x)\,dx\,d\mu.$$

Because

$$\int_{\mathbb{R}^d} f(x)\chi_{\{g>\mu\}}(x)\,dx \leq \int_{\mathbb{R}^d} f(x)\chi_{\{g>\mu\}^*}(x)\,dx$$

by the demonstrated inequality, it follows that

$$\int_{\mathbb{R}^d} f(x)\chi_{\{g>\mu\}}(x)\,dx = \int_{\mathbb{R}^d} f(x)\chi_{\{g>\mu\}^*}(x)\,dx$$

for almost every μ.

We claim that $|\{g > \mu\}\Delta\{g > \mu\}^*| = 0$. To see this, first observe that all open balls centered at 0 occur as super-level sets of f. That is, $\{f > \lambda\} = B(0, r(\lambda))$ where $\lambda \mapsto r(\lambda)$ is non-increasing and $r(\lambda) \to \infty$ as $\lambda \to 0$, since f is *strictly* symmetric decreasing. Note that $r(\lambda)$ is continuous. If it were not, then by monotonicity there would be a jump discontinuity at some point λ_0, so that

$$\lim_{\varepsilon\to 0} |\{f > \lambda_0 + \varepsilon\}| - |\{f > \lambda_0 - \varepsilon\}| < 0$$

and thus

$$\lim_{\varepsilon\to 0} |\{\lambda_0 - \varepsilon \leq f < \lambda_0 + \varepsilon\}| > 0.$$

This implies that $|\{f = \lambda_0\}| > 0$, which is a contradiction, as f is *strictly* symmetric decreasing. Now, by the above equality, we have

$$\int_0^\infty \int \chi_{\{f>\lambda\}}(x)\chi_{\{g>\mu\}}(x)\,dx\,d\lambda = \int_0^\infty \int \chi_{\{f>\lambda\}}(x)\chi_{\{g^*>\mu\}}(x)\,dx\,d\lambda.$$

If $A \subseteq \mathbb{R}^d$ is a Borel measurable function, then the function

$$F_A(\lambda) := \int \chi_{\{f>\lambda\}}(x)\chi_A(x)\,dx = |A \cap B(0, r(\lambda))|$$

is continuous. By our demonstrated inequality,

$$F_{\{g>\mu\}}(\lambda) \leq \int \chi_{\{f>\lambda\}}(x)\chi_{\{g^*>\mu\}}(x)\,dx = F_{\{g^*>\mu\}}(\lambda).$$

The assumed equality implies

$$\int_0^\infty F_{\{g>\mu\}}(\lambda)\,d\lambda = \int_0^\infty F_{\{g^*>\mu\}}(\lambda)\,d\lambda$$

and thus $F_{\{g>\mu\}}(\lambda) = F_{\{g^*>\mu\}}(\lambda)$ for almost every λ. Continuity of these functions gives equality for all $\lambda > 0$. That is,

$$|B(0, r(\lambda)) \cap \{g > \mu\}| = |B(0, r(\lambda)) \cap \{g^* > \mu\}|$$

for all $\lambda > 0$. Because $r(\lambda)$ is continuous and $r(\lambda) \to \infty$ as $\lambda \to 0$,

$$|B(0, r) \cap \{g > \mu\}| = |B(0, r) \cap \{g^* > \mu\}|$$

for all $r > 0$. From this, it follows that $|\{g > \mu\}\Delta\{g > \mu\}^*| = 0$ for almost every μ as desired. From this, we have $g = g^*$ almost everywhere. $\square$

A quick corollary of this inequality is that the symmetric decreasing rearrangement compresses the L^2 distance between functions.

Corollary 11.8. *Let $f, g \in L^2(\mathbb{R}^d \to [0,\infty))$. Then $\|f^* - g^*\|_{L^2} \leq \|f - g\|_{L^2}$.*

Proof. We have

$$\begin{aligned} \|f^* - g^*\|_{L^2}^2 = \int (f^*(x) - g^*(x))^2 \, dx &= \|f^*\|_{L^2}^2 + \|g^*\|_{L^2}^2 - 2 \int f^*(x) g^*(x) \, dx \\ &\leq \|f\|_{L^2}^2 + \|g\|_{L^2}^2 - 2 \int f(x) g(x) \, dx \\ &= \|f - g\|_{L^2}^2 \, . \end{aligned}$$

□

This estimate holds for more general L^p spaces. In fact, we have the following more general theorem.

Theorem 11.9. *Let $\phi : \mathbb{R} \to [0,\infty)$ be convex with $\phi(0) = 0$, and suppose that $f, g : \mathbb{R}^d \to [0,\infty)$ are vanishing at infinity. Then*

$$\int \phi(f^* - g^*)(x) \, dx \leq \int \phi(f - g)(x) \, dx.$$

Moreover, if ϕ is strictly convex and f is strictly symmetric decreasing, then equality holds if and only if $g = g^*$ almost everywhere.

Proof. Write $\phi = \phi_+ + \phi_-$, where

$$\phi_+(t) = \begin{cases} \phi(t) & t \geq 0 \\ 0 & t < 0 \end{cases} ; \qquad \phi_-(t) = \begin{cases} 0 & t \geq 0 \\ \phi(t) & t < 0 \end{cases} .$$

Both ϕ_+ and ϕ_- are convex, and it suffices to prove the inequality for only ϕ_+.

Recall that convex functions are Lipschitz on compact domains, and that if $c < a < x < y < b < d$, we have

$$\frac{\phi_+(a) - \phi_+(c)}{a - c} \leq \frac{\phi_+(y) - \phi_+(x)}{y - x} \leq \frac{\phi_+(d) - \phi_+(b)}{d - b}.$$

Thus, convex functions are absolutely continuous, hence differentiable almost everywhere, and so $\phi_+(t) = \int_0^t \phi_+'(s) \, ds$. Moreover, ϕ_+' is a nondecreasing function, and strictly increasing if ϕ is strictly convex. Thus, ϕ_+' has countably many (jump) discontinuities, and so we may modify ϕ_+' on a countable set to ensure that ϕ_+' is left-continuous.

So,

$$\begin{aligned} \phi_+(f(x) - g(x)) &= \int_0^{f(x)-g(x)} \phi_+'(t) \, dt = \int_{g(x)}^{f(x)} \phi_+'(f(x) - \mu) \, d\mu \\ &= \int_0^\infty \phi_+'(f(x) - \mu) \chi_{\{g \leq \mu\}}(x) \, d\mu \end{aligned}$$

where the last equality follows from the fact that $\phi_+'(t) = 0$ for $t < 0$. Thus,

$$\int \phi_+(f(x) - g(x)) \, dx = \int_0^\infty \int \phi_+'(f(x) - \mu) \chi_{\{g \leq \mu\}}(x) \, dx \, d\mu.$$

We now make two claims, which are left as exercises to the reader.

1. We have the inequality

$$\int f(x)\chi_{\{g\le\mu\}}(x)\,dx \ge \int f^*(x)\chi_{\{g^*\le\mu\}}(x)\,dx.$$

2. If F is nondecreasing, left continuous, and satisfies $F(0)=0$, then $(F\circ f)^* = F(f^*)$.

Using these two facts, we have

$$\begin{aligned}\int \phi_+(f(x)-g(x))\,dx &\ge \int_0^\infty \int [\phi_+'(f(x)-\mu)]^*\chi_{\{g^*\le\mu\}}(x)\,dx\,d\mu\\ &= \int_0^\infty \int \phi_+'(f^*(x)-\mu)\chi_{\{g^*\le\mu\}}(x)\,dx\,d\mu\end{aligned}$$

where we have adopted the obvious convention that $(f-\mu)^* := f^*-\mu$ to account for the fact that $f-\mu$ does not vanish at ∞. Unraveling the above integral, it follows that

$$\int \phi_+(f(x)-g(x))\,dx \ge \int \phi_+(f^*(x)-g^*(x))\,dx,$$

which is the desired inequality.

Next, we consider the equality case. Assume that ϕ is strictly convex, f is strictly symmetric decreasing, and that $\int \phi_+(f(x)-g(x))\,dx = \int \phi_+(f^*(x)-g^*(x))\,dx$. Then for almost every μ,

$$\int \phi_+'(f(x)-\mu)\chi_{\{g\le\mu\}}(x)\,dx = \int \phi_+'(f(x)-\mu)\chi_{\{g^*\le\mu\}}(x)\,dx$$

because $f=f^*$. Let $h(x) := \phi_+'(f(x)-\mu)$. Because f is strictly symmetric decreasing and ϕ_+' is strictly increasing, h is strictly symmetric decreasing on the set $\{f(x)>\mu\}$. Rewriting the above equality with a layer-cake decomposition then gives

$$\int_0^\infty \int \chi_{\{h>\lambda\}}(x)\chi_{\{g\le\mu\}}(x)\,dx\,d\lambda = \int_0^\infty \int \chi_{\{h>\lambda\}}(x)\chi_{\{g^*\le\mu\}}(x)\,dx\,d\lambda.$$

Define

$$\begin{aligned}F_{\{g\le\mu\}}(\lambda) &:= \int \chi_{\{h>\lambda\}}(x)\chi_{\{g\le\mu\}}(x)\,dx;\\ F_{\{g^*\le\mu\}}(\lambda) &:= \int \chi_{\{h>\lambda\}}(x)\chi_{\{g^*\le\mu\}}(x)\,dx.\end{aligned}$$

By claim 1, $F_{\{g\le\mu\}}(\lambda) \ge F_{\{g^*\le\mu\}}(\lambda)$. Since $\int_0^\infty F_{\{g\le\mu\}}(\lambda)\,d\lambda = \int_0^\infty F_{\{g^*\le\mu\}}(\lambda)\,d\lambda$, we then have $F_{\{g\le\mu\}}(\lambda) = F_{\{g^*\le\mu\}}(\lambda)$ for almost all $\lambda>0$. Since these functions are continuous in λ, equality in fact holds for all $\lambda>0$. So for all $\lambda>0$,

$$|\{h>\lambda\}\cap\{g\le\mu\}| = |\{h>\lambda\}\cap\{g^*\le\mu\}|$$

and so

$$|\{f>r\}\cap\{g\le\mu\}| = |\{f>r\}\cap\{g^*\le\mu\}| \quad\Rightarrow\quad |\{f>r\}\cap\{g>\mu\}| = |\{f>r\}\cap\{g^*>\mu\}|$$

for all $r>\mu$, since h is strictly symmetric decreasing. Working through the exact same argument with ϕ_- in place of ϕ_+ gives the above equality for all $r<\mu$. As before, it follows that

$$|\{g>\mu\}\Delta\{g^*>\mu\}| = 0$$

and consequently $g=g^*$ almost everywhere.

□

11.2 The Riesz Rearrangement Inequality, $d = 1$

Next, we consider a classical rearrangement inequality due to Riesz. The motivation for studying inequalities of the kind to follow originates with Poincare and his work on the shapes of fluid bodies in equilibrium. The resulting variational problem deals with integrals of the form

$$\iint f(x)h(y)|x-y|^{-1}\,dx\,dy.$$

More generally, variational problems involving symmetric decreasing functions $h(x-y)$ such as the *heat kernel* $(4\pi t)^{\frac{1}{1}}e^{-\frac{|x-y|^2}{t}}$ are of interest.

The following inequality holds more generally in Euclidean space of any dimension, but because of the technical nature of the proof, we begin by considering the one dimensional case.

Theorem 11.10 (Riesz rearrangement inequality, $d = 1$)**.** *Let $f, g, h : \mathbb{R} \to [0, \infty)$ be vanishing at infinity. Then*

$$I(f,g,h) := \iint f(x)g(x-y)h(y)\,dx\,dy \le \iint f^*(x)g^*(x-y)h^*(y)\,dx\,dy = I(f^*,g^*,h^*).$$

Moreover, if g is strictly symmetric decreasing, then equality holds if and only if there exists $x_0 \in \mathbb{R}$ such that $f(x) = f^*(x-x_0)$ and $h(x) = h^*(x-x_0)$ for almost all $x \in \mathbb{R}$.

Proof. Using a layer-cake decomposition on f, g and h, we have

$$I(f,g,h) = \int_0^\infty \int_0^\infty \int_0^\infty \iint \chi_{\{f>\lambda\}}(x)\chi_{\{g>\mu\}}(x-y)\chi_{\{h>\nu\}}(y)\,dx\,dy\,d\lambda\,d\mu\,d\nu.$$

Thus, it suffices to prove the inequality for $f = \chi_A$, $g = \chi_B$, and $h = \chi_C$ for finite-measure sets A, B, C. Approximating these sets from above by open sets A_n, B_n, and C_n, we have $I(\chi_{A_n}, \chi_{B_n}, \chi_{C_n}) \to I(\chi_A, \chi_B, \chi_C)$ and $I(\chi_{A_n^*}, \chi_{B_n^*}, \chi_{C_n^*}) \to I(\chi_{A^*}, \chi_{B^*}, \chi_{C^*})$ by the dominated convergence theorem, hence we may reduce further and assume without loss of generality that A, B, and C are open finite-measure sets.

Because A is an open subset of $\mathbb{R}$, we can write $A = \bigcup_{j\ge1} I_j$ where I_j are open, disjoint intervals such that $|I_j| \ge |I_{j+1}|$. Let $A_m := \bigcup_{j=1}^m I_j$. Let B_m and C_m denote the analogous decompositions for B and C. By the monotone convergence theorem, $I(\chi_{A_m}, \chi_{B_m}, \chi_{C_m}) \to I(\chi_A, \chi_B, \chi_C)$ and $I(\chi_{A_m^*}, \chi_{B_m^*}, \chi_{C_m^*}) \to I(\chi_{A^*}, \chi_{B^*}, \chi_{C^*})$. Consequently, we may assume without loss of generality that A, B, C are each finite unions of disjoint open intervals.

Write

$$\chi_A(x) = \sum_{j=1}^J f_j(x-a_j)$$

$$\chi_B(x) = \sum_{k=1}^K g_k(x-b_k)$$

$$\chi_C(x) = \sum_{l=1}^L h_l(x-c_l)$$

where f_j, g_k, h_l are characteristic functions of open intervals which are centered at 0. Then we have

$$I(\chi_A, \chi_B, \chi_C) = \sum_{j=1}^J \sum_{k=1}^K \sum_{l=1}^L \iint f_j(x-a_j)g_k(x-y-b_k)h_l(y-c_l)\,dx\,dy.$$

Observe that $(f_j(x-a_j))^* = f_j^*$. To demonstrate the desired inequality, our goal is to take each of the characteristic functions in the integral and shift them towards 0. For $t \in [0,1]$ we define

$$I_{jkl}(t) := \iint f_j(x-ta_j)g_k(x-y-tb_k)h_l(y-tc_l)\,dx\,dy.$$

We claim that this function is non-increasing with respect to t. To see this, change variables via $u = x - ta_j$ and $u - v = x - y - tb_k$, so that $y = v + t(a_j - b_k)$. Then

$$I_{jkl}(t) = \iint f_j(u)g_k(u-v)h_l(v + t(a_j - b_k - c_l))\,du\,dv.$$

Let $w_{jk}(v) = \int f_j(u)g_k(u-v)\,du$. Because f_j and g_k are both characteristic functions of intervals centered at 0, this is a symmetric decreasing function in v. Thus

$$I_{jkl}(t) = \int w_{jk}(v)h_l(v + t(a_j - b_k - c_l))\,dv$$

is non-increasing in t by the same reasoning.

Next, define

$$I_t(\chi_A, \chi_B, \chi_C) := \sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{l=1}^{L} \iint f_j(x-ta_j)g_k(x-y-tb_k)h_l(y-tc_l)\,dx\,dy.$$

As $t \to 0$, the intervals corresponding to the functions $f_j(x-ta_j)$, $g_k(x-y-tb_k)$, and $h_l(y-tc_l)$ shift towards the origin. As t descends from 1 to 0, when two intervals corresponding to the same set A, B, or C, touch, redefine the f_j's, g_k's, and h_l's accordingly. Carrying this out for finitely many times (because there are finitely many intervals comprising A, B, and C), we end up with A^*,B^*, and C^* intervals centered at 0. Moreover, since the functions $I_{jkl}(t)$ are non-increasing,

$$I(\chi_A, \chi_B, \chi_C) = I_1(\chi_A, \chi_B, \chi_C) \le I_0(\chi_A, \chi_B, \chi_C).$$

By construction, $I_0(\chi_A, \chi_B, \chi_C) = I(\chi_{A^*}, \chi_{B^*}, \chi_{C^*})$. Thus, we have demonstrated the inequality

$$I(\chi_A, \chi_B, \chi_C) \le I(\chi_{A^*}, \chi_{B^*}, \chi_{C^*})$$

as desired.

Next, we consider the case of equality. Assume that g is strictly symmetric decreasing and that $I(f,g,h) = I(f^*, g^*, h^*)$. Applying a layer-cake decomposition to f and h yields the equality

$$\int_0^\infty \int_0^\infty \iint \chi_{\{f>\lambda\}}(x)g(x-y)\chi_{\{h>\mu\}}(y)dxdy\,d\lambda\,d\mu$$
$$= \int_0^\infty \int_0^\infty \iint \chi_{\{f^*>\lambda\}}(x)g(x-y)\chi_{\{h^*>\mu\}}(y)dxdy\,d\lambda\,d\mu$$

since $g = g^*$. Consider the inner $dx\,dy$ integrals. By the demonstrated inequality,

$$\iint \chi_{\{f>\lambda\}}(x)g(x-y)\chi_{\{h>\mu\}}(y)\,dx\,dy \le \iint \chi_{\{f^*>\lambda\}}(x)g(x-y)\chi_{\{h^*>\mu\}}(y)\,dx\,dy.$$

Hence, equality of the above integrals implies that

$$\iint \chi_{\{f>\lambda\}}(x)g(x-y)\chi_{\{h>\mu\}}(y)\,dx\,dy = \iint \chi_{\{f^*>\lambda\}}(x)g(x-y)\chi_{\{h^*>\mu\}}(y)\,dx\,dy$$

for almost all $\lambda, \mu > 0$. Fix λ and μ for which this holds.

Set $A = \{f > \lambda\}$ and $B = \{h > \mu\}$. We wish to show that there exists $x_0 \in \mathbb{R}$ such that $f(x) = f^*(x - x_0)$ and $h(x) = h^*(x - x_0)$ for almost all $x \in \mathbb{R}$. That is, we wish to show that A and B are intervals (up to a zero-measure set) centered at the same point, and that this central point does not depend on λ or μ. Fixing λ and letting μ vary (and vice versa) shows that the central point will not depend on λ or μ.

Performing a layer-cake decomposition on g in the above equality yields

$$\int_0^\infty \iint \chi_{\{f>\lambda\}}(x)\chi_{\{g>\sigma\}}(x-y)\chi_{\{h>\mu\}}(y)\,dx\,dy\,d\sigma$$
$$= \int_0^\infty \iint \chi_{\{f^*>\lambda\}}(x)\chi_{\{g>\sigma\}}(x-y)\chi_{\{h^*>\mu\}}(y)\,dx\,dy\,d\sigma.$$

Thus, applying our demonstrated inequality to the $dx\,dy$ integrals as before, it follows that

$$\iint \chi_{\{f>\lambda\}}(x)\chi_{\{g>\sigma\}}(x-y)\chi_{\{h>\mu\}}(y)\,dx\,dy$$
$$= \iint \chi_{\{f^*>\lambda\}}(x)\chi_{\{g>\sigma\}}(x-y)\chi_{\{h^*>\mu\}}(y)\,dx\,dy$$

for almost all $\sigma > 0$. But

$$\sigma \mapsto \iint \chi_{\{f>\lambda\}}(x)\chi_{\{g>\sigma\}}(x-y)\chi_{\{h>\mu\}}(y)\,dx\,dy;$$
$$\sigma \mapsto \iint \chi_{\{f^*>\lambda\}}(x)\chi_{\{g>\sigma\}}(x-y)\chi_{\{h^*>\mu\}}(y)\,dx\,dy$$

are continuous functions in σ, since g is strictly symmetric decreasing. Therefore, equality holds for all σ. The fact that g is strictly symmetric decreasing then gives the equality

$$\iint \chi_A(x)\chi_{(-\frac{r}{2},\frac{r}{2})}(x-y)\chi_B(y)\,dx\,dy = \iint \chi_{A^*}(x)\chi_{(-\frac{r}{2},\frac{r}{2})}(x-y)\chi_{B^*}(y)\,dx\,dy$$

for all $r > 0$, with A and B as defined above.

We claim that A and B must be intervals. To see this, pick $r > |A|+|B|$. With this choice of r,

$$\iint \chi_{A^*}(x)\chi_{(-\frac{r}{2},\frac{r}{2})}(x-y)\chi_{B^*}(y)\,dx\,dy = |A| \cdot |B|.$$

Then for all $r > |A| + |B|$, we must have

$$|A| \cdot |B| = \iint \chi_A(x)\chi_{(-\frac{r}{2},\frac{r}{2})}(x-y)\chi_B(y)\,dx\,dy$$
$$= \int \chi_{(-\frac{r}{2},\frac{r}{2})}(x) \int \chi_A(x+y)\chi_B(y)\,dy\,dx.$$

By this equality, the continuous function $w(x) := \int \chi_A(x+y)\chi_B(y)\,dy$ must have support in the interval $\left(-\frac{r}{2}, \frac{r}{2}\right)$. Let I_A be the smallest interval such that $|I_A \cap A| = |A|$, let I_B be the smallest interval such that $|I_B \cap B| = |B|$, and let I_{AB} be the smallest interval containing the support of w.

We leave it as an exercise to the reader to show that $|I_{AB}| = |I_A| + |I_B|$. With this claim, we then have $|I_A| + |I_B| < r$ for all $r > |A| + |B|$. Taking the infimum over such r gives $|I_A| + |I_B| \leq |A| + |B|$. Because I_A and I_B were chosen to be intervals that essentially contain A and B respectively, it follows that $I_A = A$ and $I_B = B$, up to sets of measure zero, so that A and B are intervals.

Thus

$$\iint \chi_A(x)\chi_{(-\frac{r}{2},\frac{r}{2})}(x-y)\chi_B(y)\,dx\,dy = \iint \chi_{A^*}(x)\chi_{(-\frac{r}{2},\frac{r}{2})}(x-y)\chi_{B^*}(y)\,dx\,dy$$

for all $r > 0$, where A and B are intervals. Letting $r \to 0$ yields

$$\int \chi_A(x)\chi_B(x)\,dx = \int \chi_{A^*}(x)\chi_{B^*}(x)\,dx.$$

From here we consider two cases. If $|A| = |B|$, then $\int \chi_{A^*}(x)\chi_{B^*}(x)\,dx = |A|$. But if A and B are not centered at the same point, then $\int \chi_A(x)\chi_B(x)\,dx < |A|$, which is a contradiction. Thus, A and B are intervals centered at the same point, as desired. If $|A| > |B|$, let $r = |A| - |B|$. Then

$$\iint \chi_{A^*}(x)\chi_{(-\frac{r}{2},\frac{r}{2})}(x-y)\chi_{B^*}(y)\,dx\,dy = \int_{B^*}\int_{A^*\cap\left(y-\frac{r}{2},y+\frac{r}{2}\right)} dx\,dy = r|B|.$$

On the other hand, if A and B are not centered at the same point, then

$$\iint \chi_{A}(x)\chi_{(-\frac{r}{2},\frac{r}{2})}(x-y)\chi_{B}(y)\,dx\,dy = \int_{B}\int_{A\cap\left(y-\frac{r}{2},y+\frac{r}{2}\right)} dx\,dy < r|B|.$$

This is a contradiction, thus A and B are centered at the same point. This completes the proof. □

11.3 The Riesz Rearrangement Inequality, $d > 1$

In this section, we prove a generalization of the previous theorem for Euclidean space $\mathbb{R}^d$. In order to do this, it is necessary another notion of symmetrizing sets and functions.

Definition 11.11. Let $A \subseteq \mathbb{R}^d$ be a Borel measurable set of finite measure. Let $e \in S^{d-1} \subseteq \mathbb{R}^d$. The **Steiner symmetrization of** A **along** e is the set A^{*e} which is obtained from A by symmetrizing along lines parallel to e.

Definition 11.12. Let $f : \mathbb{R}^d \to [0,\infty)$ be Borel measurable and vanishing at infinity. For $e \in S^{d-1} \subseteq \mathbb{R}^d$, let $\rho : \mathbb{R}^d \to \mathbb{R}^d$ be the rotation taking e to e_1. Write $\rho f(x) := f(\rho^{-1}(x))$, and let $(\rho f)^{*1}$ be the one-dimensional symmetric decreasing rearrangement of ρf along x_1, keeping $x_2, \ldots, x_d$ fixed. The **Steiner symmetrization of** f **along** e is $f^{*e} := \rho^{-1}(\rho f)^{*1}$.

To clarify the definition of the Steiner symmetrization of a function, observe that f^{*1} can be written as follows:

$$f^{*1}(x_1, x_2, \ldots, x_d) = (x_1 \mapsto f(x_1, x_2, \ldots, x_d))^*.$$

Also note that these quantities are well-defined by the usual considerations of product measures and measurability of slices.

Theorem 11.13 (Riesz rearrangement inequality)**.** *Let $f, g, h : \mathbb{R}^d \to [0,\infty)$ be vanishing at infinity. Then*

$$I(f,g,h) := \iint f(x)g(x-y)h(y)\,dx\,dy \le \iint f^*(x)g^*(x-y)h^*(y)\,dx\,dy = I(f^*,g^*,h^*).$$

Moreover, if g is strictly symmetric decreasing, then equality holds if and only if there exists $x_0 \in \mathbb{R}^d$ such that $f(x) = f^*(x-x_0)$ and $g(x) = g^*(x-x_0)$ for almost all $x \in \mathbb{R}^d$.

Proof. The proof proceeds as follows. We consider the inequality when $d = 2$ separately, then prove the general inequality by induction. Finally, we consider the case of equality.

The case $d = 2$.

Suppose that $d = 2$. By the usual layer-cake decomposition argument that we have given multiple times, it suffices to prove the inequality when $f = \chi_A$, $g = \chi_B$, and $h = \chi_C$ for finite measure sets $A, B, C \subseteq \mathbb{R}^2$. To simplify notation, we write $I(A, B, C) := I(\chi_A, \chi_B, \chi_C)$. First, observe that $I(A, B, C) \leq I(A^{*e}, B^{*e}, C^{*e})$ for any $e \in S^1$. This follows from Fubini's theorem and the one-dimensional Riesz rearrangement inequality, since in this case Steiner symmetrization is essentially symmetric rearrangement one direction.

Next, we define sets A_k, B_k, and C_k for $k \geq 1$ by more or less iterating Steiner symmetrization in various directions, with the goal of showing $A_k \to A^*, B_k \to B^*$, and $C_k \to C^*$. So let α be an irrational multiple of 2π and let R_α denote the rotation around the origin by α. Let X and Y denote the Steiner symmetrizations along the x and y axes, respectively. Note then that $Xf = f^{*1}$ and $Yf = R_{\pi/2}\left(R_{-\pi/2}f\right)^{*1}$. For $k \geq 1$, define

$$\begin{aligned} A_k &= (YXR_\alpha)^k(A) \\ B_k &= (YXR_\alpha)^k(B) \\ C_k &= (YXR_\alpha)^k(C). \end{aligned}$$

In words, we obtain A_k by rotating A, symmetrizing in the X direction and then symmetrizing in the Y direction, and then iterating this process k times. Because these operations are invariant under measure, $|A_k| = A$. Moreover, A_k has double-reflection symmetry over both the x and y axes, and hence we can write

$$A_k = \left\{\, (x, y) \in \mathbb{R}^2 \;:\; |y| < \omega_k^A(|x|) \,\right\}$$

where $\omega_k^A : [0, \infty) \to [0, \infty)$ is non-increasing, symmetric, and lower-semicontinuous. The analogous statements obviously hold for B and C.

Our goal is to show that $\chi_{A_k} \xrightarrow{L^2} \chi_{A^*}$, and similarly for B and C. To see why this is sufficient, first observe that

$$\begin{aligned} &|I(A^*, B^*, C^*) - I(A_k, B_k, C_k)| \\ &\quad\leq \left|\iint (\chi_{A^*} - \chi_{A_k})(x)\chi_{B^*}(x-y)\chi_{C^*}(y)\,dx\,dy\right| \\ &\qquad + \left|\iint \chi_{A_k}(x)(\chi_{B^*} - \chi_{B_k})(x-y)\chi_{C^*}(y)\,dx\,dy\right| \\ &\qquad\quad + \left|\iint \chi_{A_k}(x)\chi_{B_k}(x-y)(\chi_{C^*} - \chi_{C_k})(y)\,dx\,dy\right| \\ &\quad\leq \|\chi_{A^*} - \chi_{A_k}\|_{L^2}\|\chi_{B^*} * \chi_{C^*}\|_{L^2} + \|\chi_{B^*} - \chi_{B_k}\|_{L^2}\|\chi_{A_k} * \chi_{C^*}\|_{L^2} \\ &\qquad + \|\chi_{C^*} - \chi_{C_k}\|\,\|\chi_{A_k} * \chi_{B_k}\|_{L^2} \\ &\quad\leq \|\chi_{A^*} - \chi_{A_k}\|_{L^2}\|\chi_{B^*}\|_{L^1}\|\chi_{C^*}\|_{L^2} + \|\chi_{B^*} - \chi_{B_k}\|_{L^2}\|\chi_{A_k}\|_{L^2}\|\chi_{C^*}\|_{L^1} \\ &\qquad + \|\chi_{C^*} - \chi_{C_k}\|\,\|\chi_{A_k}\|_{L^1}\|\chi_{B_k}\|_{L^1}. \end{aligned}$$

Because A, B, and C are finite measure sets, if we can show that $\chi_{A_k} \to \chi_{A^*}$ and likewise for B and C, then this shows that $I(A_k, B_k, C_k) \to I(A^*, B^*, C^*)$. Next, observe that

$$\begin{aligned} I(A_{k+1}, B_{k+1}, C_{k+1}) &= I(YXR_\alpha A_k, YXR_\alpha B_k, YXR_\alpha C_k) \\ &\geq I(XR_\alpha A_k, XR_\alpha B_k, XR_\alpha C_k) \\ &\geq I(R_\alpha A_k, R_\alpha B_k, R_\alpha C_k) \\ &= I(A_k, B_k, C_k). \end{aligned}$$

The first and second inequalities follow from the one-dimensional Riesz rearrangement inequality, and the last equality is because the quantity $I(f, g, h)$ is invariant under rotation, via the obvious change of variables. Hence, $I(A_k, B_k, C_k)$ is an increasing sequence which converges to $I(A^*, B^*, C^*)$. In particular,

$$I(A, B, C) = I(A_0, B_0, C_0) \leq I(A^*, B^*, C^*)$$

as desired.

With this goal in mind, we make the following observations. Suppose that F and G are finite measure sets. Then $\|R_\alpha \chi_F - R_\alpha \chi_G\|_{L^2} = \|\chi_F - \chi_G\|_{L^2}$, and by the inequality $\|f^* - g^*\|_{L^2} \leq \|f - g\|_{L^2}$ we also have

$$\begin{aligned}\|X\chi_F - X\chi_G\|_{L^2} &\leq \|\chi_F - \chi_G\|_{L^2} \\ \|Y\chi_F - Y\chi_G\|_{L^2} &\leq \|\chi_F - \chi_G\|_{L^2}.\end{aligned}$$

Consequently, we may assume without loss of generality that A, B, and C are bounded sets. Indeed, for every $\varepsilon > 0$ there exists an $r > 0$ and a set $A' \subseteq B_r(0)$ such that $\|\chi_A - \chi_{A'}\|_{L^2} < \varepsilon$. Let $A'_k = (YXR_\alpha)^k(A')$. Then by the above inequalities we have

$$\left\|\chi_{A_k} - \chi_{A'_k}\right\|_{L^2} \leq \|\chi_A - \chi_{A'}\|_{L^2} < \varepsilon.$$

Thus, it is enough to consider convergence of the finite measure sets A'_k, and similarly for B and C.

From now on we suppose that A, B, and C are bounded. This implies that the functions ω_k^A, ω_k^B, and ω_k^C are uniformly bounded. For now we focus on A, as the arguments for B and C are exactly the same. Using a diagonalization argument and passing to a subsequence $l(k)$, we see that $\{\omega_{l(k)}^A\}$ converges at every single rational number. By monotonicity, $\{\omega_{l(k)}^A\}$ converges everywhere except at countably many points, at most. This implies that $\chi_{A_{l(k)}} \to \chi_{\tilde{A}}$ almost everywhere for some set $\tilde{A}$. Necessarily, $\tilde{A}$ has double reflection symmetry across the x and y axes. To conclude that $\tilde{A} = A^*$, which then concludes the proof of the $d = 2$ inequality, it suffices to show that $|\tilde{A}|$ is a ball, since $|\tilde{A}| = \lim_{k\to\infty} |A_{l(k)}| = |A|$.

We will accomplish this by essentially showing that $R_\alpha \tilde{A} = \tilde{A}$, and using that fact that α is an irrational multiple of 2π and thus $\tilde{A}$ has rotational symmetry on a dense set of angles. To do this, we take advantage of the equality case in the $\|f^* - g^*\|_{L^2} \leq \|f - g\|_{L^2}$, which presumes f to be strictly symmetric decreasing. Let

$$\gamma(x, y) = e^{-|x|^2 - |y|^2}$$

be a Gaussian on $\mathbb{R}^2$, which is indeed strictly symmetric decreasing. Let $a_k = \|\gamma - \chi_{A_k}\|_{L^2}$. Note that $R_\alpha \gamma = \gamma$, $X\gamma = \gamma$, and $Y\gamma = \gamma$. So

$$\begin{aligned}a_{k+1} = \left\|\gamma - \chi_{A_{k+1}}\right\|_{L^2} &= \|YXR_\alpha\gamma - YXR_\alpha\chi_{A_k}\|_{L^2} \leq \|\gamma - \chi_{A_k}\|_{L^2} \\ &= a_k.\end{aligned}$$

Thus, $\{a_k\}$ is a decreasing sequence. Because $a_k \geq 0$, $a := \lim_{k\to\infty} a_k = \lim_{k\to\infty} a_{l(k)}$ exists. Moreover, by the dominated convergence theorem, $a = \left\|\gamma - \chi_{\tilde{A}}\right\|_{L^2}$. Next, note that

$$\begin{aligned}\left|a_{l(k)+1} - \left\|\gamma - YXR_\alpha\chi_{\tilde{A}}\right\|_{L^2}\right| &= \left|\left\|\gamma - \chi_{A_{l(k)+1}}\right\|_{L^2} - \left\|\gamma - YXR_\alpha\chi_{\tilde{A}}\right\|_{L^2}\right| \\ &\leq \left\|\chi_{A_{l(k)+1}} - YXR_\alpha\chi_{\tilde{A}}\right\|_{L^2} \\ &= \left\|YXR_\alpha\chi_{A_{l(k)}} - YXR_\alpha\chi_{\tilde{A}}\right\|_{L^2} \\ &\leq \left\|\chi_{A_{l(k)}} - \chi_{\tilde{A}}\right\|_{L^2}\end{aligned}$$

which $\to 0$ as $k \to \infty$. On the other hand,

$$\left|a_{l(k)+1} - \left\|\gamma - YXR_\alpha\chi_{\tilde{A}}\right\|_{L^2}\right| \to \left|a - \left\|\gamma - YXR_\alpha\chi_{\tilde{A}}\right\|_{L^2}\right|$$

as $k \to \infty$. Thus, $a = \left\|\gamma - YXR_\alpha\chi_{\tilde{A}}\right\|_{L^2}$. But we have

$$\begin{aligned} a = \left\|\gamma - YXR_\alpha\chi_{\tilde{A}}\right\|_{L^2} &= \left\|Y\gamma - YXR_\alpha\chi_{\tilde{A}}\right\|_{L^2} \le \left\|\gamma - XR_\alpha\chi_{\tilde{A}}\right\|_{L^2} \\ &= \left\|X\gamma - XR_\alpha\chi_{\tilde{A}}\right\|_{L^2} \le \left\|\gamma - R_\alpha\chi_{\tilde{A}}\right\|_{L^2} \\ &= \left\|\gamma - \chi_{\tilde{A}}\right\|_{L^2} \\ &= a. \end{aligned}$$

Therefore, we must have $\left\|\gamma - R_\alpha\chi_{\tilde{A}}\right\|_{L^2} = \left\|\gamma - XR_\alpha\chi_{\tilde{A}}\right\|_{L^2}$, i.e.,

$$\iint |\gamma - R_\alpha\chi_{\tilde{A}}|^2\,dx\,dy = \iint |\gamma - XR_\alpha\chi_{\tilde{A}}|^2\,dx\,dy.$$

Because $\int |\gamma - R_\alpha\chi_{\tilde{A}}|^2\,dx \le \int |\gamma - XR_\alpha\chi_{\tilde{A}}|^2\,dx$, it follows that $\int |\gamma - R_\alpha\chi_{\tilde{A}}|^2\,dx = \int |\gamma - XR_\alpha\chi_{\tilde{A}}|^2\,dx$ for almost every y. Thus, for almost every y, $XR_\alpha\chi_{\tilde{A}} = R_\alpha\chi_{\tilde{A}}$ for almost every x. By Fubini's theorem, $XR_\alpha\chi_{\tilde{A}} = R_\alpha\chi_{\tilde{A}}$ almost everywhere in $\mathbb{R}^2$. The same argument shows that $YXR_\alpha\chi_{\tilde{A}} = XR_\alpha\chi_{\tilde{A}}$ almost everywhere in $\mathbb{R}^2$, and hence $YXR_\alpha\chi_{\tilde{A}} = R_\alpha\chi_{\tilde{A}}$ almost everywhere in $\mathbb{R}^2$. Thus, $R_\alpha\chi_{\tilde{A}}$ has double reflection symmetry about both axes. If P denotes reflection across the y-axis, then

$$R_\alpha\chi_{\tilde{A}} = PR_\alpha\chi_{\tilde{A}} = R_{-\alpha}P\chi_{\tilde{A}} = R_{-\alpha}\chi_{\tilde{A}}$$

which implies that $\chi_{\tilde{A}} = R_{2\alpha}\chi_{\tilde{A}}$. Because α, and hence 2α, is an irrational multiple of 2π, the set $\{n(2\alpha) \mod 2\pi\}$ is dense in $[0, 2\pi)$. Thus, the function

$$\mu(\theta) = \left\|\chi_{\tilde{A}} - R_\theta\right\|$$

has dense set of zeroes. To conclude that $\tilde{A}$ is a circle, and thus conclude the proof of the Riesz rearrangement inequality in the $d = 2$ case, it suffices to prove that μ is continuous in θ. By writing and expanding μ^2 as an inner product, it suffices to show continuity of the map

$$\theta \mapsto \iint \chi_{\tilde{A}} \cdot R_\theta\chi_{\tilde{A}}\,dx\,dy.$$

To see this, let $f_n \in C^\infty$ such that $f_n \to \chi_{\tilde{A}}$. Then

$$\left|\iint \chi_{\tilde{A}} \cdot R_\theta\chi_{\tilde{A}}\,dx\,dy - \iint f_n \cdot R_\theta\chi_{\tilde{A}}\,dx\,dy\right| \le \left\|\chi_{\tilde{A}} - f_n\right\|_{L^2} \cdot |A|^{\frac{1}{2}} \to 0$$

uniformly in θ. Because $\theta \mapsto \iint f_n \cdot R_\theta\chi_{\tilde{A}}\,dx\,dy$ is continuous, it then follows that $\theta \mapsto \iint \chi_{\tilde{A}} \cdot R_\theta\chi_{\tilde{A}}\,dx\,dy$ is a uniform limit of continuous functions, hence continuous. Thus, μ is continuous, and so $\tilde{A}$ is a ball and therefore $\tilde{A} = A^*$.

The case $d > 2$.

Let $e \in S^{d-1} \subseteq \mathbb{R}^d$. Recall that the Steiner symmetrization of $f : \mathbb{R}^d \to [0, \infty)$ along e is $f^{*e}(x) = (f \circ \rho^{-1})^{*1}(\rho x)$, where ρ is a rotation defined by $\rho e = e_1$. Define the *Schwartz symmetrization along directions perpendicular to e* by

$$f^{*e^\perp}(x) := (f \circ \rho^{-1})^{*e_1^\perp}(\rho x),$$

where $*e_1^\perp$ indicates symmetrization in the variables $x_2, \ldots, x_d$, keeping x_1 fixed.

As before, to prove the $d > 2$ case it suffices to consider $f = \chi_A, g = \chi_B$, and $h = \chi_C$ for finite and bounded measure sets A, B, C. Let R be a rotation such that $Re_d =$

e_{d-1}. Let Y denote the Steiner symmetrization along e_d, and let X denote the Schwartz symmetrizations perpendicular to e_d. Let

$$\begin{aligned} A_k &= (YXR)^k A \\ B_k &= (YXR)^k B \\ C_k &= (YXR)^k C. \end{aligned}$$

Write $x = (x', x_d)$ where $x' = (x_1, \ldots, x_{d-1})$. By induction, we have

$$\iint f(x', x_d) g(x' - y', x_d - y_d) h(y', y_d)\, dx'\, dy' \\ \leq \iint (Xf)(x', x_d)(Xg)(x' - y', x_d - y_d)(Xh)(y', y_d)\, dx'\, dy'.$$

In particular, $I(A_{k+1}, B_{k+1}, C_{k+1}) \geq I(A_k, B_k, C_k)$, so by the same argument as before it suffices to prove $\chi_{A_k} \xrightarrow{L^2} \chi_{A^*}$, and likewise for B and C.

By construction, we can write

$$A_k = \left\{ x \in \mathbb{R}^d \ : \ |x_d| < \omega_k^A(|x'|) \right\}$$

where ω_k^A is a symmetric and non-increasing function. The ω_k^A's are uniformly bounded by boundedness of A. Exactly as before, we may pass to a subsequence to get $\chi_{A_{l(k)}} \to \chi_{\tilde{A}}$ almost everywhere. As each $A_{l(k)}$ is rotationally symmetric around the x_d axis, the limiting set $\tilde{A}$ is also rotationally symmetric around the x_d axis. Also as before, define the Gaussian γ and set $a_k = \|\gamma - \chi_{A_k}\|_{L^2}$. Running through the same computation gives $YXR\tilde{A} = R\tilde{A}$ almost everywhere. Hence, $R\tilde{A}$ is rotationally symmetric along the x_d axis, which implies that $\tilde{A}$ is rotationally symmetric around the x_d axis *and* the x_{d-1} axis. We will show that this implies that $\tilde{A}$ is a ball. Then because $|\tilde{A}| = |A| = |A^*|$ it will follow that $\tilde{A} = A^*$ almost everywhere, as desired.

Let φ_ε be a radial smooth approximation to the identity and set $\chi_\varepsilon(x) = (\varphi_\varepsilon * \chi_{\tilde{A}})(x)$. Then χ_ε and $\chi_{\tilde{A}}$ have the same symmetry properties. As such, we can write

$$\chi_\varepsilon(x_1, \ldots, x_d) = u(\sqrt{x_1^2 + \cdots + x_{d-1}^2}, x_d) = v(\sqrt{x_1^2 + \cdots + x_{d-2}^2 + x_d^2}, x_{d-1})$$

for smooth functions u and v of two variables. Writing $\rho^2 = x_1^2 + \cdots + x_{d-1}^2$, we have $u(\sqrt{\rho^2 + x_{d-1}^2}, x_d) = v(\sqrt{\rho^2 + x_d^2}, x_{d-1})$ for all $x_{d-1}, x_d \in \mathbb{R}$ and for all $\rho > 0$. In particular, setting $x_d = 0$ we have

$$u(\sqrt{\rho^2 + x_{d-1}^2}, 0) = v(\sqrt{\rho^2}, x_{d-1})$$

for all $x_{d-1} \in \mathbb{R}$ and for all $\rho > 0$. Choosing $\rho^2 = x_1^2 + \cdots + x_{d-2}^2 + x_d^2$ gives $\chi_\varepsilon(x) = u(|x|, 0)$, hence χ_ε is spherically symmetric and so $\tilde{A}$ is a ball.

This completes the proof of the inequality for $d > 2$.

The case of equality.

Finally, we consider the equality statement. Suppose that g is strictly symmetric decreasing and that equality holds. As before, we assume that $f = \chi_A$ and $h = \chi_B$ for finite measure sets A and B. Using the (x', x_d) notation from above, we have

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}^{d-1}} \int_{\mathbb{R}^{d-1}} f(x', x_d) g(x' - y', x_d - y_d) h(y', y_d)\, dx'\, dy'\, dx_d\, dy_d \\ = \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}^{d-1}} \int_{\mathbb{R}^{d-1}} f^*(x', x_d) g^*(x' - y', x_d - y_d) h^*(y', y_d)\, dx'\, dy'\, dx_d\, dy_d.$$

By the inequality in the statement of the theorem,

$$\int_{\mathbb{R}^{d-1}}\int_{\mathbb{R}^{d-1}} f(x',x_d)g(x'-y',x_d-y_d)h(y',y_d)\,dx'\,dy' \\ \leq \int_{\mathbb{R}^{d-1}}\int_{\mathbb{R}^{d-1}} f^*(x',x_d)g^*(x'-y',x_d-y_d)h^*(y',y_d)\,dx'\,dy'.$$

Hence, the assumed equality gives

$$\int_{\mathbb{R}^{d-1}}\int_{\mathbb{R}^{d-1}} f(x',x_d)g(x'-y',x_d-y_d)h(y',y_d)\,dx'\,dy' \\ = \int_{\mathbb{R}^{d-1}}\int_{\mathbb{R}^{d-1}} f^*(x',x_d)g^*(x'-y',x_d-y_d)h^*(y',y_d)\,dx'\,dy'$$

for almost all (x_d, y_d).

By induction we get sets A_{x_d} and B_{y_d} corresponding to the functions $f(\cdot, x_d)$ and $h(\cdot, y_d)$ which are balls in $\mathbb{R}^{d-1}$ centered at the same point, by the equality case of the one dimensional Riesz rearrangement inequality. Also, by the same argument as before, this central point is independent of x_d and y_d. Thus, A and B are rotationally symmetric with respect to a line L_1 which is parallel to the x_d axis. Similarly, A and B must also be rotationally symmetric with respect to another line L_2 parallel to the e_{d-1} axis.

We claim that $L_1 \cap L_2 \neq \emptyset$. In $d = 2$, this is obvious, but the existence of skew-symmetric lines in higher dimensions makes the $d \geq 3$ case more complicated. This can be seen via a *ping pong* argument:[1] if L_1 and L_2 do not intersect, then by rotating around L_2 180 degrees it follows that the sets are rotationally symmetric about L_1 *and* and an affine shift of L_1. By rotating around this shift of L_1, the sets are then also rotationally symmetric around an affine shift of L_2. Continuing this process, we contradict the finite measure of A and B.

Thus, A and B are rotationally symmetric about two axes. As we argued previously, it follows that A and B are balls.

□

11.4 The Polya-Szego inequality

Our first application of the Riesz rearrangement inequality is a special case of the Polya-Szego inequality. We begin with a lemma.

Lemma 11.14. *Let $f : \mathbb{R}^d \to \mathbb{C}$ be such that $f \in L^1_{loc}$ and $\nabla f \in L^1_{loc}$. Then*

$$\nabla|f|(x) = \begin{cases} \frac{(u\nabla u + v\nabla v)(x)}{|f(x)|} & \text{if } f = u+iv \neq 0 \\ 0 & \text{if } f = u+iv = 0 \end{cases}$$

in the sense of distributions. In particular, $|\nabla|f|| \leq |\nabla f|$ almost everywhere.

Proof. First we demonstrate the "in particular" statement, assuming the formula for $\nabla|f|$. For $f(x) \neq 0$ (up to a choice of representative), we have by Cauchy-Schwarz

$$\begin{aligned} |\nabla|f(x)||^2 &= \frac{\left[u^2|\nabla u|^2 + v^2|\nabla v|^2 + 2uv\nabla u\nabla v\right](x)}{|f(x)|^2} \\ &\leq \frac{\left[u^2|\nabla u|^2 + v^2|\nabla v|^2 + u^2|\nabla v|^2 + v^2|\nabla u|^2\right](x)}{|f(x)|^2} \\ &= \frac{(u^2+v^2)(|\nabla u|^2 + |\nabla v|^2)}{|f|^2}(x) \\ &= |\nabla f(x)|^2 \end{aligned}$$

[1]Drawing a picture here is helpful.

as desired.

To prove the formula in the statement, let $\psi \in C_c^\infty$. We wish to show

$$\int \nabla\psi\,|f|\,dx = -\int_{\{f\neq 0\}} \psi\,\frac{u\nabla u + v\nabla v}{|f|}\,dx.$$

The first issue we need to deal with is moving the derivative onto $|f|$. To do this, for $\delta > 0$ consider $\sqrt{\delta^2+|f|^2}$. Note that for small δ, $\sqrt{\delta^2+|f|^2} \leq 1+|f|$, so by the dominated convergence theorem

$$\int \nabla\psi\,\sqrt{\delta^2+|f|^2}\,dx \to \int \nabla\psi\,|f|\,dx$$

as $\delta \to 0$. Next, let φ_ε be an approximation to the identity, $u_\varepsilon = u * \varphi_\varepsilon$, and $v_\varepsilon = v * \varphi_\varepsilon$. Then $u_\varepsilon \to u$ and $v_\varepsilon \to v$ almost everywhere and in L^1_{loc}. Also, $\nabla(u_\varepsilon + iv_\varepsilon) \to \nabla f$ almost everywhere and in L^1_{loc}. Note that

$$\begin{aligned}\int \nabla\psi \left(\sqrt{\delta^2+u^2+v^2} - \sqrt{\delta^2+u_\varepsilon^2+v_\varepsilon^2}\right)\,dx \\ = \int \nabla\psi\,\frac{(u-u_\varepsilon)(u+u_\varepsilon)+(v-v_\varepsilon)(v+v_\varepsilon)}{\sqrt{\delta^2+u^2+v^2}+\sqrt{\delta^2+u_\varepsilon^2+v_\varepsilon^2}}\,dx\end{aligned}$$

so that

$$\begin{aligned}&\left|\int \nabla\psi \left(\sqrt{\delta^2+u^2+v^2} - \sqrt{\delta^2+u_\varepsilon^2+v_\varepsilon^2}\right)\,dx\right| \\ &\qquad\leq \int |\nabla\psi|\,\frac{|u-u_\varepsilon|(|u|+|u_\varepsilon|)+|v-v_\varepsilon|(|v|+|v_\varepsilon|)}{\sqrt{\delta^2+u^2+v^2}+\sqrt{\delta^2+u_\varepsilon^2+v_\varepsilon^2}}\,dx \\ &\qquad\leq 2\int |\nabla\psi|\left(|u-u_\varepsilon|+|v-v_\varepsilon|\right)\,dx.\end{aligned}$$

Because $u_\varepsilon \to u$ and $v_\varepsilon \to v$ in L^1_{loc}, since ψ has compact support, it follows that

$$\int \nabla\psi\sqrt{\delta^2+u_\varepsilon^2+v_\varepsilon^2}\,dx \to \int \nabla\psi\sqrt{\delta^2+u^2+v^2}\,dx$$

as $\varepsilon \to 0$. Integrating by parts, we then have

$$\begin{aligned}\int \nabla\psi\sqrt{\delta^2+u^2+v^2}\,dx &= \lim_{\varepsilon\to 0}\int \nabla\psi\sqrt{\delta^2+u_\varepsilon^2+v_\varepsilon^2}\,dx \\ &= -\lim_{\varepsilon\to 0}\int \psi\,\frac{u_\varepsilon\nabla u_\varepsilon + v_\varepsilon\nabla v_\varepsilon}{\sqrt{\delta^2+u_\varepsilon^2+v_\varepsilon^2}}\,dx \\ &= -\lim_{\varepsilon\to 0}\int \psi\,\frac{u_\varepsilon\nabla u + v_\varepsilon\nabla v}{\sqrt{\delta^2+u_\varepsilon^2+v_\varepsilon^2}}\,dx \\ &\qquad -\lim_{\varepsilon\to 0}\int \psi\,\frac{u_\varepsilon(\nabla u_\varepsilon - \nabla u) + v_\varepsilon(\nabla v_\varepsilon - \nabla v)}{\sqrt{\delta^2+u_\varepsilon^2+v_\varepsilon^2}}\,dx.\end{aligned}$$

Consider the second integral. We have

$$\left|\int \psi\,\frac{u_\varepsilon(\nabla u_\varepsilon - \nabla u) + v_\varepsilon(\nabla v_\varepsilon - \nabla v)}{\sqrt{\delta^2+u_\varepsilon^2+v_\varepsilon^2}}\,dx\right| \leq \int |\psi|\left(|\nabla u_\varepsilon - \nabla u| + |\nabla v_\varepsilon - \nabla v|\right)\,dx$$

Again by L^1_{loc} convergence and compact support of ψ, this integral converges to 0 as $\varepsilon \to 0$. In the first integral, observe that

$$\left|\int \psi\,\frac{u_\varepsilon\nabla u + v_\varepsilon\nabla v}{\sqrt{\delta^2+u_\varepsilon^2+v_\varepsilon^2}}\,dx\right| \leq \int |\psi|\left(|\nabla u| + |\nabla v|\right)\,dx < \infty.$$

Thus, by the dominated convergence theorem,

$$\begin{aligned}\int \nabla\psi\sqrt{\delta^2+u^2+v^2}\,dx &= -\lim_{\varepsilon\to 0}\int \psi\,\frac{u_\varepsilon\nabla u+v_\varepsilon\nabla v}{\sqrt{\delta^2+u_\varepsilon^2+v_\varepsilon^2}}\,dx\\ &= -\int \psi\,\frac{u\nabla u+v\nabla v}{\sqrt{\delta^2+u^2+v^2}}\,dx.\end{aligned}$$

Letting $\delta\to 0$ gives

$$\int \nabla\psi|f|\,dx = -\int \psi\,\frac{u\nabla u+v\nabla v}{|f|}\,dx$$

as desired. □

Theorem 11.15 (Polya-Szego). *Let $f:\mathbb{R}^d\to\mathbb{C}$ be vanishing at infinity such that $f\in L^1_{loc}$ and $\nabla f\in L^2$. Then $\|\nabla f^*\|_{L^2}\le\|\nabla f\|_{L^2}$.*

Proof. First, note that $\nabla f\in L^2$ implies $\nabla f\in L^1_{loc}$, hence by the previous lemma we have $|\nabla|f||\le|\nabla f|$. Consequently, it suffices to consider functions f satisfying $f\ge 0$.

Next, we make a further reduction. For $\varepsilon>0$, define

$$f_\varepsilon := \min\left\{\frac{1}{\varepsilon},[f-\varepsilon]_+\right\} = \begin{cases} 0 & f\le\varepsilon\\ f-\varepsilon & \varepsilon\le f\le\varepsilon+\varepsilon^{-1}\\ \frac{1}{\varepsilon} & f\ge\varepsilon+\varepsilon^{-1}\end{cases} = \varphi_\varepsilon\circ f$$

where

$$\varphi_\varepsilon(x) = \begin{cases} 0 & x\le\varepsilon\\ x-\varepsilon & \varepsilon\le x\le\varepsilon+\varepsilon^{-1}\\ \frac{1}{\varepsilon} & x\ge\varepsilon+\varepsilon^{-1}\end{cases}.$$

Note that

$$\begin{aligned}\int|f_\varepsilon|\,dx &= \frac{1}{\varepsilon}\left|\{f\ge\varepsilon+\varepsilon^{-1}\}\right| + \int_{\{\varepsilon\le f\le\varepsilon+\varepsilon^{-1}\}}|f-\varepsilon|\,dx\\ &\le \frac{1}{\varepsilon}\left|\{f\ge\varepsilon+\varepsilon^{-1}\}\right| + \frac{1}{\varepsilon}\left|\{f\ge\varepsilon\}\right|.\end{aligned}$$

Because f vanishes at infinity, the above super-level sets have finite measure, hence $\int|f_\varepsilon|\,dx$ is finite and so $f_\varepsilon\in L^1$. Similarly, it can be shown that $f_\varepsilon\in L^2$. Also note that

$$\nabla f_\varepsilon = \nabla f\chi_{\{\varepsilon\le f\le\varepsilon+\varepsilon^{-1}\}} \xrightarrow{L^2} \nabla f$$

as $\varepsilon\to 0$ by the monotone convergence theorem.

Next, recall that because φ_ε is symmetric and non-decreasing, $(f_\varepsilon)^* = (\varphi_\varepsilon\circ f)^* = \varphi_\varepsilon\circ f^* = (f^*)_\varepsilon$. Thus, we have $\nabla(f_\varepsilon)^* = \nabla(f^*)_\varepsilon \xrightarrow{L^2} \nabla f^*$. This implies that it suffices to consider $f\in L^1\cap L^2$ and $\nabla f\in L^2$.

By Plancharel and the monotone convergence theorem, we have

$$\int|\nabla f|^2\,dx = \int|\xi|^2|\hat{f}(\xi)|^2\,d\xi = \lim_{t\to 0}\int\frac{1-e^{-t|\xi|^2}}{t}|\hat{f}(\xi)|^2\,d\xi.$$

Because $f\in L^1$ and $f\ge 0$, we can distribute and use Fubini to get

$$\begin{aligned}\int|\nabla f|^2\,dx &= \lim_{t\to 0}\frac{1}{t}\left[\|f\|_{L^2}^2-(2\pi)^{-d}\int e^{-t|\xi|^2}\int e^{-ix\cdot\xi}f(x)\,dx\int e^{iy\cdot\xi}f(y)\,dy\,d\xi\right]\\ &= \lim_{t\to 0}\frac{1}{t}\left[\|f\|_{L^2}^2-(2\pi)^{-d}\iint f(x)f(y)\int e^{-t|\xi|^2-i(x-y)\cdot\xi}\,d\xi\,dx\,dy\right]\\ &= \lim_{t\to 0}\frac{1}{t}\left[\|f\|_{L^2}^2-(2\pi)^{-d}\iint f(x)f(y)\int e^{-t\left(\xi+\frac{i(x-y)}{2t}\right)^2}e^{-\frac{|x-y|^2}{4t}}\,d\xi\,dx\,dy\right]\end{aligned}$$

where we are using the convention that for $z \in \mathbb{C}^d$, $z^2 = \sum_{j=1}^d z_j^2$. Changing variables and integrating the resulting ξ-Gaussian then gives

$$\int |\nabla f|^2 \, dx = \lim_{t \to 0} \frac{1}{t} \left[\|f\|_{L^2}^2 - (2\pi)^{-d} t^{-\frac{d}{2}} \pi^{\frac{d}{2}} \iint f(x)\, e^{-\frac{|x-y|^2}{4t}}\, f(y)\, dx\, dy \right].$$

By the Riesz rearrangement inequality, and by equi-measurability of f^* with f,

$$\int |\nabla f|^2 \, dx \geq \lim_{t \to 0} \frac{1}{t} \left[\|f^*\|_{L^2}^2 - (2\pi)^{-d} t^{-\frac{d}{2}} \pi^{\frac{d}{2}} \iint f^*(x)\, e^{-\frac{|x-y|^2}{4t}}\, f^*(y)\, dx\, dy \right].$$

Beginning with the expression on the right and undoing all of the previous calculations, we recover the inequality $\|\nabla f\|_{L^2}^2 \geq \|\nabla f^*\|_{L^2}^2$. □

We make passing observation that the heat kernel $(4\pi t)^{-\frac{d}{2}} e^{-\frac{|x-y|^2}{4t}}$ appears in the above proof. This use of the heat kernel is possible because the stated estimate is in L^2. A generalization of the Polya-Szego inequality holds for other values of p, but the above method does not give this result.

11.5 The Energy of the Hydrogen Atom

As an application of the Polya-Szego inequality, we consider the energy functional

$$E(f) := \int_{\mathbb{R}^3} |\nabla f(x)|^2 - \frac{|f(x)|^2}{|x|}\, dx$$

for $f \in H^1(\mathbb{R}^3)$. Such a functional can be taken to represent the energy of a hydrogen atom. We will demonstrate that

$$\inf_{\|f\|_{L^2}=1} E(f)$$

is achieved by a symmetric decreasing function. Via a scaling argument, we show that the infimum is negative. Let $\varphi \in C_c^\infty(\mathbb{R}^3)$ with $\|\varphi\|_{L^2} = 1$. For $\lambda > 0$, let $\varphi_\lambda(x) = \lambda^{-\frac{3}{2}} \varphi\left(\frac{x}{\lambda}\right)$. Then $\|\varphi_\lambda\|_{L^2} = \|\varphi\|_2 = 1$, and

$$\begin{aligned} E(\varphi_\lambda) &= \lambda^{-2}\lambda^{-3} \int \left| (\nabla \varphi)\left(\frac{x}{\lambda}\right) \right|^2 dx - \lambda^{-3} \int \frac{\left|\varphi\left(\frac{x}{\lambda}\right)\right|^2}{|x|}\, dx \\ &= \lambda^{-2} \|\nabla \varphi\|_{L^2}^2 - \lambda^{-1} \int \frac{|\varphi(x)|^2}{|x|}\, dx \end{aligned}$$

Because $\|\nabla\varphi\|_{L^2}^2$ and $\frac{|\varphi(x)|^2}{|x|}\, dx$ are fixed numbers, for large λ the above quantity is negative. Hence, $E_{min} := \inf_{\|f\|_2=1} E(f) < 0$. To show that the problem is well-posed, we also must prove that the infimum is not identically $-\infty$. To do this, we first recall Hardy's inequality:

Proposition 11.16 (Hardy's Inequality). *Let $f \in \mathcal{S}(\mathbb{R}^d)$ and $0 \leq s < d$. Then*

$$\left\| \frac{f(x)}{|x|^s} \right\|_{L^p} \lesssim \| |\nabla|^s f \|_{L^p}$$

for all $1 < p < \frac{d}{s}$.

By this inequality, we have

$$\begin{aligned} \left\| \frac{f(x)}{|x|^{\frac{1}{2}}} \right\|_{L^2} \lesssim \left\| |\nabla|^{\frac{1}{2}} f(x) \right\|_{L^2} &= \left\| |\xi|^{\frac{1}{2}} \hat{f}(\xi) \right\|_{L^2} = \left\| |\xi|^{\frac{1}{2}} |\hat{f}(\xi)|^{\frac{1}{2}} |\hat{f}(\xi)|^{\frac{1}{2}} \right\|_{L^2} \\ &\leq \left\| |\xi|^{\frac{1}{2}} |\hat{f}(\xi)|^{\frac{1}{2}} \right\|_{L^4} \left\| |\hat{f}(\xi)|^{\frac{1}{2}} \right\|_{L^4} \\ &= \|\nabla f\|_{L^2}^{\frac{1}{2}} \|f\|_{L^2}^{\frac{1}{2}}. \end{aligned}$$

So for $\|f\|_{L^2} = 1$, $E(f) \geq \|\nabla f\|_{L^2}^2 - C\|\nabla f\|_{L^2}$ where the constant C comes from Hardy's inequality. So $E_{min} \geq \inf_{x>0} x^2 - Cx > -\infty$ as desired.

Next, we show that the infimum is achieved. Let $f_n \in H^1(\mathbb{R}^3)$ be a decreasing minimizing sequence, that is, $\|f_n\|_{L^2} = 1$ for all n and $E(f_n) \to E_{min}$. First, we claim that we can find a minimizing sequence of radially symmetric decreasing functions. Indeed, consider f_n^*. We have $\|f_n^*\|_{L^2} = \|f_n\|_{L^2} = 1$, and by the Polya-Szego inequality, $\|\nabla f_n^*\|_{L^2} \leq \|\nabla f_n\|_{L^2}$. By the Hardy-Littlewood rearrangement estimate,

$$\int \frac{|f_n(x)|^2}{|x|}\,dx = \left\langle \frac{1}{|x|}, |f_n(x)|^2 \right\rangle \leq \left\langle \frac{1}{|x|}, (|f_n(x)|^2)^* \right\rangle = \left\langle \frac{1}{|x|}, |f_n^*(x)|^2 \right\rangle = \int \frac{|f_n^*(x)|^2}{|x|}\,dx.$$

Hence, $E(f_n^*) \leq E(f_n)$, which implies that f_n^* is also a minimizing sequence. Consequently, we may assume from now on that each f_n is radially symmetric and decreasing.

Next, we claim that the sequence $\{f_n\}$ is bounded in $H^1(\mathbb{R}^3)$. If it were not, then because $\|f_n\|_{L^2} = 1$, the quantities $\|\nabla f\|_{L^2}$ are unbounded. But since $E(f) \geq \|\nabla f\|_{L^2}^2 - C\|\nabla f\|_{L^2}$, this implies that $\limsup E(f_n) = \infty$. This contradicts the assumption that $E(f_n)$ decreases to $E_{min} < 0$. Hence, $\{f_n\}$ is indeed bounded in $H^1(\mathbb{R}^3)$. Because $H^1(\mathbb{R}^3)$ can be reformulated on the Fourier side as an L^2-space with a weighted measure, and bounded sets in L^2 spaces are weakly compact, it follows that (passing to a subsequence) $f_n \rightharpoonup f$ weakly in $H^1(\mathbb{R}^3)$.

We need to demonstrate that the weak limit f is an optimizer. First, we will show that

$$E(f) \leq \liminf E(f_n) = E_{min}$$

as desired. By weak lower semi-continuity of the L^2 norm, $\|\nabla f\|_{L^2} \leq \liminf \|\nabla f_n\|_{L^2}$. This is great, but it doesn't help in demonstrating the energy inequality because it tells us nothing about the potential energy term $-\int \frac{|f(x)|^2}{|x|}\,dx$. To remedy this, we invoke the following fact.

Lemma 11.17 (Strauss)**.** *The space $H^1(\mathbb{R}^3)$ embeds compactly into $L^p(\mathbb{R}^3)$ for $2 < p < 6$.*

The proof of this lemma is found in Chapter 12, after discussing some compactness results in L^p spaces. Here we only remark that the upper bound 6 for p comes from the scaling condition of the Sobolev embedding theorem.

In particular, this lemma implies that because $f_n \rightharpoonup f$ weakly in $H^1(\mathbb{R}^3)$, then $f_n \to f$ strongly in $L^p(\mathbb{R}^3)$ for $2 < p < 6$. With this in mind, we estimate as follows:

$$\begin{aligned}\left|\int \frac{|f_n(x)|^2}{|x|}\,dx - \int \frac{|f(x)|^2}{|x|}\,dx\right| &\leq \left|\int \frac{(|f_n| - |f|)(|f_n| + |f|)}{|x|}\,dx\right| \\ &\leq \int \frac{|f_n - f|(|f_n| + |f|)}{|x|}\,dx.\end{aligned}$$

In order to use the fact that $f_n \to f$ strongly in certain L^p spaces, we want to apply Holder's inequality. But since $\frac{1}{|x|}$ is only in *weak* L^p spaces, we split up the integral:

$$\begin{aligned}&\left|\int \frac{|f_n(x)|^2}{|x|}\,dx - \int \frac{|f(x)|^2}{|x|}\,dx\right| \\ &\qquad\leq \int_{|x|\leq 1} \frac{|f_n - f|(|f_n| + |f|)}{|x|}\,dx + \int_{|x|\geq 1} \frac{|f_n - f|(|f_n| + |f|)}{|x|}\,dx.\end{aligned}$$

In the first integral, $\frac{1}{|x|} \in L^p(B(0,1) \subseteq \mathbb{R}^3)$ for $p < 3$. Applying Holder's inequality twice gives

$$\begin{aligned}\int_{|x|\leq 1} \frac{|f_n - f|(|f_n| + |f|)}{|x|}\,dx &\leq \left\|\frac{1}{|x|}\right\|_{L^2(B(0,1))} \||f_n - f|(|f_n| + |f|)\|_{L^2(B(0,1))} \\ &\leq \left\|\frac{1}{|x|}\right\|_{L^2(B(0,1))} \|f_n - f\|_{L^4}\left(\|f_n\|_{L^4} + \|f\|_{L^4}\right).\end{aligned}$$

The first norm is finite by the above comment. We claim that the quantities $\|f_n\|_{L^4} + \|f\|_{L^4}$ are bounded in n. To see this, recall the Gagliardo-Nirenberg inequality:

Proposition 11.18 (Gagliardo-Nirenberg inequality)**.** *Fix $d \geq 1$ and $0 < p < \infty$ for $d = 1, 2$, or $0 < p < \frac{4}{d-2}$ for $d \geq 3$. Then for all $f \in \mathcal{S}(\mathbb{R}^d)$,*

$$\|f\|_{L^{p+2}}^{p+2} \lesssim \|f\|_{L^2}^{p+2-\frac{pd}{2}} \|\nabla f\|_{L^2}^{\frac{pd}{2}}.$$

A proof of this inequality is also contained in Chapter 12. When $p = 2$ and $d = 3$, the inequality above takes the form $\|f\|_{L^4}^4 \lesssim \|f\|_{L^2} \|\nabla f\|_{L^2}^3$. Because $\{f_n\}$ is a bounded sequence in $H^1(\mathbb{R}^3)$, it follows that $\|f_n\|_{L^4} + \|f\|_{L^4}$ is bounded. Then because $\|f_n - f\|_{L^4} \to 0$,

$$\int_{|x|\leq 1} \frac{|f_n - f|(|f_n| + |f|)}{|x|}\, dx \to 0.$$

Similarly,

$$\int_{|x|\geq 1} \frac{|f_n - f|(|f_n| + |f|)}{|x|}\, dx \leq \left\| \frac{1}{|x|} \right\|_{L^4(|x|\geq 1)} \|f_n - f\|_{L^4} \left(\|f_n\|_{L^2} + \|f\|_{L^2}\right) \lesssim \|f_n - f\|_{L^4}.$$

Thus,

$$\int \frac{|f_n - f|(|f_n| + |f|)}{|x|}\, dx \to 0.$$

From this and the fact that $\|\nabla f\|_{L^2} \leq \liminf \|\nabla f_n\|_{L^2}$ it follows that

$$E(f) \leq \liminf E(f_n) = E_{min}$$

as desired.

In fact, a stronger statement is true. *All* of the optimizers of the energy functional $E(f)$ above are radially symmetric; in particular, any optimizer satisfies $f = e^{i\theta} f^*$. We sketch the argument here. If f is an optimizer, then as seen above it follows that f^* is also an optimizer. From $E(f) = E(f^*)$ and Polya-Szego we deduce the equalities $\|\nabla f\|_{L^2} = \|\nabla f^*\|_{L^2}$ and $\left\| \frac{f(x)}{|x|^{\frac{1}{2}}} \right\|_{L^2} = \left\| \frac{f^*(x)}{|x|^{\frac{1}{2}}} \right\|_{L^2}$. By the equality case of our first rearrangement inequality, $|f| = f^*$ almost everywhere. Unfortunately this does not rule out the possibility that f itself is not radial. However, we then have

$$\|\nabla f\|_{L^2} = \|\nabla f^*\|_{L^2} = \|\nabla |f|\|_{L^2}.$$

Having previously shown $|\nabla |f|| \leq |\nabla f|$ almost everywhere, it follows that $|\nabla |f|| = |\nabla f|$ almost everywhere. Then if $f = u + iv$, we have $u \nabla v = v \nabla u$ almost everywhere. Writing $f = re^{i\theta}$ where r and θ are functions of x, we then have $r \nabla \theta = 0$ almost everywhere. Because f is not identically 0, θ is constant as desired.

Chapter 12

Compactness in L^p Spaces

The analysis of the energy of the hydrogen atom in Chapter 11 is a model problem in the calculus of variations. Given a functional on some spaces of functions, we wish to determine its optimimum value, either a supremum or infimum, and moreover show that this optimum is achieved. By definition of the supremum and infimum, we can always find a sequence of functions whose functional values tend towards the optimum. From here, we would ideally invoke some sort of compactness of the function space to extract a limiting function which optimizes the functional. However, as we know from previous mathematical experiences, compactness in infinite dimensional spaces, and in particular function spaces, is far from ideal. For example, it is well-known that a closed and bounded set in an infinite dimensional Banach space is *not* compact in the norm-topology. We can, however, achieve varying degrees of compactness by considering weaker topologies. Also well-known is the the following result, which we have invoked throughout the text already.

Theorem 12.1 (Banach-Alaoglu)**.** *Let X be a Banach space. A closed and bounded set in X^* is compact in the weak-* topology.*

If X is reflexive, it immediately follows that a closed and bounded set in X is *weakly* compact. In the context of L^p-spaces (where we also have separability), this is manifested by the following theorem.

Theorem 12.2 (Weak compactness)**.** *If $\{f_n\}$ is a bounded sequence in $L^p(\mathbb{R}^d)$ for $1 < p < \infty$, then, passing to a subsequence, $f_n \rightharpoonup f$ for some $f \in L^p(\mathbb{R}^d)$.*

Oftentimes, weak compactness is not enough to solve optimization problems. For example, if f_n is a translation of a bump function whose profile marches off to infinity, then $f_n \rightharpoonup 0$. As such, we must seek compactness with more care.

The next four chapters of this text may be taken a single unit; a careful study of compactness in various function spaces and its consequences in harmonic analysis and partial differential equations. In the present chapter, we begin by developing a foundation of standard compactness results, and in Chapters 13, 14, and 15 we study the delicate notion of concentration compactness, which measures the defects of compactness in various contexts.

12.1 The Riesz Compactness Theorem

The first theorem we prove addresses the question of precompactness in the norm topology. For continuous functions, the answer is given by the well-known Arzela-Ascoli theorem. The main compactness tool used in this chapter is an analogous result for L^p spaces.

Theorem 12.3 (Riesz compactness)**.** *Fix $1 \leq p < \infty$. A family $\mathcal{F} \subseteq L^p(\mathbb{R}^d)$ is precompact in $L^p(\mathbb{R}^d)$ if and only if it satisfies the following three properties:*

1. *(Boundedness) There exists $A > 0$ such that $\|f\|_{L^p} \leq A$ for all $f \in \mathcal{F}$.*

2. *(Equicontinuity) For every $\varepsilon > 0$ there exists a $\delta > 0$ such that $\int |f(x+y) - f(x)|^p \, dx < \varepsilon$ for all $|y| < \delta$ and for all $f \in \mathcal{F}$.*

3. *(Tightness) For every $\varepsilon > 0$, there exists an $R > 0$ such that $\int_{|x| \geq R} |f(x)|^p \, dx < \varepsilon$ for all $f \in \mathcal{F}$.*

Proof. The forward direction is fairly straightforward. Suppose that $\mathcal{F}$ is precompact in $L^p(\mathbb{R}^d)$. Then for every $\varepsilon > 0$, there exists $f_1, \ldots, f_n \in \mathcal{F}$ such that $\mathcal{F} \subseteq \bigcup_{j=1}^n B(f_j, \frac{\varepsilon}{10})$. Because any function $f \in \mathcal{F}$ lies in one of these balls, $\|f\|_{L^p} \leq \frac{\varepsilon}{10} + \max_j \|f_j\|_{L^p}$. Thus, $\mathcal{F}$ is bounded. Next, for $|y| < \delta = \delta(\varepsilon) << 1$ sufficiently small, we have

$$\|f(\cdot + y) - f\|_{L^p} \leq 2\frac{\varepsilon}{10} + \max_j \|f_j(\cdot + y) - f_j\|_{L^p} < \varepsilon,$$

so that $\mathcal{F}$ is equicontinuous. Finally,

$$\|f\|_{L^p(|x| \geq R)} \leq \frac{\varepsilon}{10} + \max_j \|f_j\|_{L^p(|x| \geq R)} < \varepsilon$$

for R sufficiently large. Hence, $\mathcal{F}$ is tight.

The other direction is more involved. Suppose that $\mathcal{F}$ is bounded, equicontinuous, and tight. Fix $\varepsilon > 0$. We need to show the existence of $f_1, \ldots, f_n \in \mathcal{F}$ so that $\mathcal{F} \subseteq \bigcup_{j=1}^n B(f_j, \varepsilon)$.

We will approximate $\mathcal{F}$ by a continuous family of functions and then apply Arzela-Ascoli. Let $\varphi \in C_c^\infty$ be a radial bump function satisfying $\int \varphi \, dx = 1$ and

$$\varphi(x) = \begin{cases} c_d & |x| \leq \frac{1}{2} \\ 0 & |x| > 1 \end{cases}$$

for some dimensional constant c_d. For $R > 0$ and $f \in \mathcal{F}$, define

$$f_R(x) = \varphi\left(\frac{x}{R}\right) \int f\left(x - \frac{x}{R}\right) \varphi(y) \, dy = \varphi\left(\frac{x}{R}\right) \int f(y) \, R^d \varphi(R(x-y)) \, dy.$$

The passage from f to f_R "fuzzes" f at a scale $1/R$. Define $\mathcal{F}_R = \{ f_R \,:\, f \in \mathcal{F} \}$. Because f_R is defined as convolution with a smooth function multiplied by a compactly supported function, $\mathcal{F}_R$ is a continuous family of functions supported on $B(0, R)$. Also,

$$\|f - f_R\|_{L^p} = \left\|\left(1 - \varphi\left(\frac{x}{R}\right)\right) f(x)\right\|_{L^p} + \left\|\varphi\left(\frac{x}{R}\right)\left(f(x) - \int f\left(x - \frac{x}{R}\right)\varphi(y)\,dy\right)\right\|_{L^p}.$$

Because $\mathcal{F}$ is tight, we can estimate the first norm by

$$\left\|\left(1 - \varphi\left(\frac{x}{R}\right)\right) f(x)\right\|_{L^p} = \|f\|_{L^p(|x| > R)} < \frac{\varepsilon}{10}$$

for $R = R(\varepsilon) >> 1$ sufficiently large. For the second norm, we use the fact that φ has unit mass. We have

$$\begin{aligned} \left\|\varphi\left(\frac{x}{R}\right)\left(f(x) - \int f\left(x - \frac{x}{R}\right)\varphi(y)\,dy\right)\right\|_{L^p} &\leq \left\|\int \left(f(x) - f\left(\frac{x}{R}\right)\right)\varphi(y)\,dx\right\|_{L^p_x} \\ &\leq \int |\varphi(y)| \left\|f(x) - f\left(\frac{x}{R}\right)\right\|_{L^p_x} dy \\ &\leq \frac{\varepsilon}{10} \end{aligned}$$

for R sufficiently large, by equicontinuity of $\mathcal{F}$. Thus, $\|f - f_R\|_{L^p} < \frac{\varepsilon}{5}$ for R sufficiently large.

Next, we claim that $\mathcal{F}_R$ is uniformly bounded and equicontinuous. To see boundedness, by Young's inequality and a change of variables we have

$$\|f_R\|_{L^\infty} \lesssim \left\| f\left(x - \frac{y}{R}\right)\right\|_{L^p_y} \|\varphi\|_{L^{p'}} \lesssim R^{\frac{d}{p}} \|f\|_{L^p} \lesssim R^{\frac{d}{p}} A.$$

To see equicontinuity, write

$$\begin{aligned}
|f_R(x+y) - f_R(x)| &\leq \left| \left[\varphi\left(\frac{x+y}{R}\right) - \varphi\left(\frac{x}{R}\right)\right] \int f\left(x + y - \frac{z}{R}\right) \varphi(z)\, dz \right| \\
&\quad + \left| \varphi\left(\frac{x}{R}\right) \int \left[f\left(x+y-\frac{z}{R}\right) - f\left(x - \frac{z}{R}\right)\right] \varphi(z)\, dz\right| \\
&\lesssim \left| \varphi\left(\frac{x+y}{R}\right) - \varphi\left(\frac{x}{R}\right)\right| R^{\frac{d}{p}} \|f\|_{L^p} \|\varphi\|_{L^{p'}} \\
&\quad + \left\| f\left(x+y-\frac{z}{R}\right) - f\left(x-\frac{z}{R}\right)\right\|_{L^p_z} \|\varphi\|_{L^{p'}} \\
&\lesssim \left| \varphi\left(\frac{x+y}{R}\right) - \varphi\left(\frac{x}{R}\right)\right| R^{\frac{d}{p}} A + R^{\frac{d}{p}} \|f(\cdot + y) - f(\cdot)\|_{L^p}.
\end{aligned}$$

Because φ is smooth and because $\mathcal{F}$ is tight, for $|y| < \delta = \delta(\varepsilon, R)$ with δ sufficiently small, $|f_R(x+y) - f_R(x)| < \varepsilon$ uniformly in f.

Therefore, as $\mathcal{F}_R$ is a continuous family of uniformly bounded and equicontinuous compactly supported functions, it follows by the Arzela-Ascoli theorem that $\mathcal{F}_R$ is precompact in the L^∞ topology. So for every $\tilde{\varepsilon} > 0$ there exists $f_1, \ldots, f_n \in \mathcal{F}$ such that $\mathcal{F}_R \subseteq \bigcup_{j=1}^n B_{L^\infty}((f_j)_R, \tilde{\varepsilon})$.

Fix $f \in \mathcal{F}$. There exists a j such that $f_R \in B_{L^\infty}((f_j)_R, \tilde{\varepsilon})$. By the triangle inequality, we have

$$\begin{aligned}
\|f - f_j\|_{L^p} &\leq \|f - f_R\|_{L^p} + \|f_j - (f_j)_R\|_{L^p} + \|f_R - (f_j)_R\|_{L^p} \\
&< \frac{\varepsilon}{5} + \frac{\varepsilon}{5} + c_d \|f_j - (f_j)_R\|_{L^\infty} R^{\frac{d}{p}}.
\end{aligned}$$

Because $\|f_j - (f_j)_R\|_{L^\infty} < \tilde{\varepsilon}$, by choosing $\tilde{\varepsilon}$ sufficiently small it follows that $\|f - f_j\|_{L^p} < \varepsilon$, and hence $\mathcal{F} \subseteq \bigcup_{j=1}^n B(f_j, \varepsilon)$ as desired.

$\square$

Our first corollary of the Riesz compactness theorem is an alternative characterization of precompact families when $p = 2$. In particular, in this case we can replace the equicontinuity condition by tightness in the Fourier domain.

Corollary 12.4. *A family $\mathcal{F} \subseteq L^2(\mathbb{R}^d)$ is precompact if and only if it satisfies the following two properties:*

1. *There exists $A > 0$ such that $\|f\|_{L^2} \leq A$ for all $f \in \mathcal{F}$.*
2. *For every $\varepsilon > 0$, there exists $R > 0$ such that*

$$\int_{|x| \geq R} |f(x)|^2\, dx + \int_{|\xi| \geq R} |\hat{f}(\xi)|^2\, d\xi < \varepsilon$$

for all $f \in \mathcal{F}$.

Proof. The forward direction is almost exactly the same as in the proof of the Riesz compactness theorem; tightness in the Fourier domain (i.e. the estimate $\int_{|\xi| \geq R} |\hat{f}(\xi)|^2\, d\xi < \varepsilon$) comes from the fact that the Fourier transform is an isometry on $L^2(\mathbb{R}^d)$.

Conversely, suppose the two conditions hold. By the Reisz compactness theorem, it suffices to verify that $\mathcal{F}$ is equicontinuous.

Fix $\varepsilon > 0$. Using the fact that the Fourier transform is an isometry on $L^2(\mathbb{R}^d)$, for any $R > 0$ we can write

$$\begin{aligned} \|f(x+y) - f(x)\|_{L^2_x} &= \left\| e^{iy\cdot\xi}\hat{f}(\xi) - \hat{f}(\xi) \right\|_{L^2_\xi} \\ &\le \left\| e^{iy\cdot\xi}\hat{f}(\xi) - \hat{f}(\xi) \right\|_{L^2(|\xi|\ge R)} + \left\| \hat{f}(\xi)\left[e^{iy\cdot\xi} - 1\right] \right\|_{L^2(|\xi|\le R)}. \end{aligned}$$

When $|\xi|$ is large, there is a lot of oscillation coming from the $e^{iy\cdot\xi}$ phase. Thus, we simply use the triangle quality and tightness of $\hat{f}$ to estimate the first term:

$$\left\| e^{iy\cdot\xi}\hat{f}(\xi) - \hat{f}(\xi) \right\|_{L^2(|\xi|\ge R)} \le 2\left\|\hat{f}\right\|_{L^2(|\xi|\ge R)} < \frac{\varepsilon}{2}$$

for R sufficiently large. For the second term, we have

$$\left\| \hat{f}(\xi)\left[e^{iy\cdot\xi} - 1\right] \right\|_{L^2(|\xi|\le R)} \le \left\| \hat{f}\cdot|y|\cdot|\xi| \right\|_{L^2(|\xi|\le R)} \le |y|RA.$$

Then $\left\| \hat{f}(\xi)\left[e^{iy\cdot\xi} - 1\right] \right\|_{L^2(|\xi|\le R)} < \varepsilon/2$ for $|y| < \delta$ where $\delta(R,\varepsilon)$ is sufficiently small. For such a choice of δ, $\|f(x+y) - f(x)\|_{L^2_x} < \varepsilon$ uniformly in f and hence $\mathcal{F}$ is equicontinuous. □

12.2 The Rellich-Kondrachov Theorem

In this section, we consider the issue of compactness in the Sobolev space $H^1(\mathbb{R}^d)$ and the *homogeneous* Sobolev space $\dot{H}^1(\mathbb{R}^d)$. For convenience, we recall their definitions here.

Definition 12.5.

1. The Hilbert space $H^1(\mathbb{R}^d)$ is the completion of $\mathcal{S}(\mathbb{R}^d)$ under the norm $\|f\|_{H^1} := \|f\|_{L^2} + \|\nabla f\|_{L^2}$.
2. The Hilbert space $H^1(\mathbb{R}^d)$ is the completion of $\mathcal{S}(\mathbb{R}^d)$ under the norm $\|f\|_{H^1} := \|f\|_{L^2} + \|\nabla f\|_{L^2}$.

In words, functions in $H^1(\mathbb{R}^d)$ are L^2-functions whose (distributional) derivative is also in $L^2(\mathbb{R}^d)$. For a function in $\dot{H}^1(\mathbb{R}^d)$ we only assume that the gradient is in $L^2(\mathbb{R}^d)$; the function itself may not be in $L^2(\mathbb{R}^d)$.

The two most important estimates in the study of the Sobolev spaces $H^1(\mathbb{R}^d)$ and $\dot{H}^1(\mathbb{R}^d)$ are the Gagliardo-Nirenberg inequality and the $p = 2$ Sobolev embedding theorem, respectively. Both results have already been stated and proven in this text, but we recall them here for convenience.

Proposition 12.6 (Gagliardo-Nirenberg)**.** *Fix $d \ge 1$. Let p satisfy $2 \le p < \infty$ if $d = 1$ or $d = 2$, and $2 \le p < \frac{2d}{d-2}$ if $d \ge 3$. Then for $f \in H^1(\mathbb{R}^d)$,*

$$\|f\|_{L^p} \lesssim \|f\|_{L^2}^{\frac{2d-p(d-2)}{2p}} \|\nabla f\|_{L^2}^{\frac{d(p-2)}{2p}}.$$

Proposition 12.7 (Sobolev embedding, $p = 2$)**.** *Fix $d \ge 3$. Suppose that $f \in \mathcal{S}(\mathbb{R}^d)$ with $\nabla f \in L^2(\mathbb{R}^d)$. Then*

$$\|f\|_{L^{\frac{2d}{d-2}}} \lesssim \|\nabla f\|_{L^2}.$$

From Gagliardo-Nirenberg we have the relationship $\|f\|_p \lesssim \|f\|_{H^1}$, and from Sobolev embedding we have $\|f\|_{L^{\frac{2d}{d-2}}} \lesssim \|f\|_{\dot{H}^1}$. As promised in Chapter 5, we provide a slick proof of Proposition 12.6 using Littlewood-Paley theory, which we previously did not have at our disposal.

Proof of Proposition 12.6. It suffices to consider Schwartz functions f.

Because the Littlewood-Paley projections of f converge to f in L^p for the prescribed values of p, we have

$$\|f\|_{L^p} \le \sum_{N\in 2^{\mathbb{Z}}} \|f_N\|_{L^p} .$$

By Bernstein's first inequality,

$$\|f_N\|_{L^p} \lesssim N^{\frac{d}{2}-\frac{d}{p}} \|f_N\|_{L^2} \lesssim N^{\frac{d}{2}-\frac{d}{p}} \|f\|_{L^2} .$$

On the other hand, by Bernstein's second inequality we have

$$\|f_N\|_{L^p} \lesssim N^{\frac{d}{2}-\frac{d}{p}} N^{-1} \|\nabla f_N\|_{L^2} \lesssim N^{\frac{d}{2}-\frac{d}{p}-1} \|\nabla f\|_{L^2} .$$

Together, this implies that

$$\|f\|_{L^p} \lesssim \sum_{N\in 2^{\mathbb{Z}}} \min\left\{ N^{\frac{d}{2}-\frac{d}{p}} \|f\|_{L^2} \, , \, N^{\frac{d}{2}-\frac{d}{p}-1} \|\nabla f\|_{L^2} \right\} .$$

Because the Gagliardo-Nirenberg inequality is trivial when $p = 2$, we may assume that $p > 2$. Consequently, $\frac{d}{2} - \frac{d}{p} > 0$. Also, $\frac{d}{2} - \frac{d}{p} - 1 = \frac{d-2}{2} - \frac{d}{p} < 0$. This latter statement is clear when $d = 1$ or $d = 2$, and when $d \ge 3$ this follows from the fact that $p < \frac{2d}{d-2}$. Using the positivity and negativity of the above exponents, we split the sum into two summable series to get the desired inequality. Explicitly,

$$\begin{aligned}
\|f\|_{L^p} &\lesssim \sum_{N \lesssim \frac{\|\nabla f\|_{L^2}}{\|f\|_{L^2}}} N^{\frac{d}{2}-\frac{d}{p}} \|f\|_{L^2} + \sum_{N \gtrsim \frac{\|\nabla f\|_{L^2}}{\|f\|_{L^2}}} N^{\frac{d}{2}-\frac{d}{p}-1} \|\nabla f\|_{L^2} \\
&\lesssim \|\nabla f\|_{L^2}^{\frac{d}{2}-\frac{d}{p}} \|f\|_{L^2}^{1-\frac{d}{2}+\frac{d}{p}} + \|\nabla f\|_{L^2}^{\frac{d}{2}-\frac{d}{p}} \|f\|_{L^2}^{1+\frac{d}{p}-\frac{d}{2}} \\
&\lesssim \|f\|_{L^2}^{\frac{2d-p(d-2)}{2p}} \|\nabla f\|_{L^2}^{\frac{d(p-2)}{2p}} .
\end{aligned}$$

□

For an H^1-bounded family of functions, the Reisz compactness theorem leads to the following important result.

Theorem 12.8 (Rellich-Kondrachov, $H^1(\mathbb{R}^d)$). *Let χ_R be a cutoff function on $B_R(0) \subseteq \mathbb{R}^d$, smooth or not. Then the family $\mathcal{F} := \{ \chi_R f \ : \ \|f\|_{H^1} \le 1 \}$ is compact in $L^p(\mathbb{R}^d)$ for*

$$\begin{cases} 2 \le p < \infty & d = 1,2 \\ 2 \le p < \frac{2d}{d-2} & d \ge 3 \end{cases} .$$

Proof. Let B_1 denote the unit ball in $H^1(\mathbb{R}^d)$. We begin with the observation that $\mathcal{F}$ is closed. To see this, let $f_n \in B_1$ be a sequence of functions such that $\chi_R f_n$ converges in L^p. Because $f_n \in B_1$, the sequence f_n is bounded in L^2. Passing to a subsequence, we have $f_n \rightharpoonup f$ weakly in L^2. This implies that $\chi_R f_n \rightharpoonup \chi_R f$ weakly in L^2. As weak limits are unique, it follows that $\chi_R f_n \to \chi_r f$ in L^p and hence $\mathcal{F}$ is closed. Thus, to prove the theorem it suffices to prove that $\mathcal{F}$ is precompact.

We will invoke the Reisz compactness theorem. As such, we need to demonstrate boundedness, equicontinuity, and tightness. By Gagliardo-Nirenberg, $\|\chi_R f\|_{L^p} \le \|f\|_{L^p} \lesssim \|f\|_{H^1}$ and hence the family $\mathcal{F}$ is bounded. Tightness of $\mathcal{F}$ is immediate, as each function in $\mathcal{F}$ has compact support.

It remains to show equicontinuity. Fix $\varepsilon > 0$. We have

$$\begin{aligned}
&\|\chi_R f(x+y) - \chi_R f(x)\|_{L^p_x} \\
&\qquad\qquad \le \|[\chi_R(x+y) - \chi_R(x)]f(x+y)\|_{L^p_x} + \|\chi_R(x)[f(x+y) - f(x)]\|_{L^p_x}.
\end{aligned}$$

To estimate the first term, we want to use Holder, the L^p continuity property of translations on χ_R to get smallness, and then Gagliardo-Nirenberg on the f term to get uniformity in f. However, because translation of χ_R is not continuous in the L^∞-norm, we have to be careful. We have

$$\|[\chi_R(x+y) - \chi_R(x)]f(x+y)\|_{L^p_x} \le \|\chi_R(x+y) - \chi_R(x)\|_{L^r_x} \|f(x+y)\|_{L^q_x}$$

for r and q satisfying

$$p < q < \begin{cases} \infty & d = 1,2 \\ \frac{2d}{d-2} & d \ge 3 \end{cases}; \qquad\qquad \frac{1}{p} = \frac{1}{r} + \frac{1}{q}.$$

By Gagliardo-Nirenberg, $\|f(x+y)\|_{L^q_x}$ is uniformly bounded in f. For $|y| < \delta(\varepsilon) << 1$ where δ is sufficiently small, we have $\|\chi_R(x+y) - \chi_R(x)\|_{L^r_x} < \varepsilon$.

To estimate the second term, we simply ignore the χ_R and use Gagliardo-Nirenberg:

$$\begin{aligned}
\|\chi_R(x)[f(x+y) - f(x)]\|_{L^p_x} &\le \|f(x+y) - f(x)\|_{L^p_x} \\
&\lesssim \|f(x+y) - f(x)\|^\theta_{L^2_x} \|\nabla_x[f(x+y) - f(x)]\|^{1-\theta}_{L^2_x}
\end{aligned}$$

where θ is the correct exponent dictated by Gagliardo-Nirenberg. By the triangle inequality,

$$\|\nabla_x[f(x+y) - f(x)]\|^{1-\theta}_{L^2_x} \lesssim \|\nabla f\|^{1-\theta}_{L^2}$$

and hence is uniformly bounded in f. By the fundamental theorem of calculus,

$$\|f(x+y) - f(x)\|^\theta_{L^2_x} \lesssim (|y| \, \|\nabla f\|_{L^2})^\theta.$$

Thus, for $|y|$ sufficiently small, $\|\chi_R(x)[f(x+y) - f(x)]\|_{L^p_x} < \varepsilon$.

These two estimates give equicontinuity of $\mathcal{F}$. By the Riesz compactness theorem, the proof is complete.

□

The same result holds for $\dot{H}^1(\mathbb{R}^d)$. The proof is for the most part the same, with the Sobolev embedding theorem replacing the Gagliardo-Nirenberg inequality as the main estimation tool.

Theorem 12.9 (Rellich-Kondrachov, $\dot{H}^1(\mathbb{R}^d)$)**.** *Fix $d \ge 3$ and $2 \le p < \frac{2d}{d-2}$. Let χ_R be a cutoff function on $B_R(0) \subseteq \mathbb{R}^d$, smooth or not. Then the family $\mathcal{F} := \{ \chi_R f \ : \ \|f\|_{\dot{H}^1} \le 1 \}$ is compact in $L^p(\mathbb{R}^d)$.*

Proof. The proof of closedness of $\mathcal{F}$ is the same, this time using the fact that the sequence is bounded in $L^{\frac{2d}{d-2}}(\mathbb{R}^d)$ (by Sobolev embedding) rather than $L^2(\mathbb{R}^d)$.

To use the Riesz compactness theorem, we need to show boundedness, tightness, and equicontinuity. As before, tightness is immediate by compact support. To see boundedness, we first apply Holder, and then Sobolev embedding:

$$\|\chi_R f\|_{L^p} \le \|\chi_R\|_{L^\alpha} \|f\|_{L^{\frac{2d}{d-2}}} \lesssim \|f\|_{\dot{H}^1} \lesssim 1$$

where $\frac{1}{\alpha} = \frac{1}{p} - \frac{1}{2d/(d-2)}$.

To prove equicontinuity, write

$$\begin{aligned}&\|\chi_R f(x+y) - \chi_R f(x)\|_{L^p_x} \\ &\qquad\qquad \leq \|[\chi_R(x+y) - \chi_R(x)]f(x+y)\|_{L^p_x} + \|\chi_R(x)[f(x+y) - f(x)]\|_{L^p_x}\end{aligned}$$

as before. Estimating the first term, we have

$$\begin{aligned}\|[\chi_R(x+y) - \chi_R(x)]f(x+y)\|_{L^p_x} &\leq \|\chi_R(x+y) - \chi_R(x)\|_{L^\alpha_x} \|f(x+y)\|_{L_x^{\frac{2d}{d-2}}} \\ &\lesssim \|\chi_R(x+y) - \chi_R(x)\|_{L^\alpha_x}.\end{aligned}$$

Continuity of translations in $L^\alpha(\mathbb{R}^d)$ gives equicontinuity in this term. To estimate the second term, write

$$\begin{aligned}\|\chi_R(x)[f(x+y) - f(x)]\|_{L^p_x} &\leq \|f(x+y) - f(x)\|_{L^p_x} \\ &\leq \|f(x+y) - f(x)\|_{L^2_x}^{1-\theta} \|f(x+y) - f(x)\|^\theta_{L_x^{\frac{2d}{d-2}}}\end{aligned}$$

where $\theta = d\left(\frac{1}{2} - \frac{1}{2d/(d-2)}\right)$. By Sobolev embedding, the latter term is uniformly bounded. The first term we estimate by first using the fundamental theorem of calculus:

$$\|f(x+y) - f(x)\|_{L^2_x}^{1-\theta} \lesssim (|y| \, \|\nabla f\|_{L^2})^{1-\theta} \lesssim |y|^{1-\theta}$$

which gives equicontinuity.. □

12.3 The Strauss Lemma

In Chapter 11, we studied the energy functional of the hydrogen atom. A crucial step in our analysis was the Strauss lemma, which we did not prove. In this section, we provide a proof.

First, we have the following theorem, which describes the decay of radial functions in $\dot{H}^1(\mathbb{R}^d)$.

Proposition 12.10 (Radial Sobolev embedding). *Fix $d \geq 2$ and $1 \leq p < \infty$. For a radial function $f \in \dot{H}^1(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$, we have*

$$r^{\frac{2(d-1)}{p+2}} |f(r)| \lesssim \|f\|_{L^p}^{\frac{p}{2+p}} \|\nabla f\|_{L^2}^{\frac{2}{2+p}}$$

for almost all $r > 0$.

Proof. It suffices to prove the inequality for Schwartz functions. Indeed, $\mathcal{S}(\mathbb{R}^d)$ is dense in $\dot{H}^1(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$, and by passing to a subsequence we can ensure almost everywhere convergence. As $|\nabla |f|| \leq |\nabla f|$ almost everywhere, we can further assume that $f \geq 0$.

By the fundamental theorem of calculus we have

$$\begin{aligned}r^{d-1} f(r)^{1+\frac{p}{2}} = \left| r^{d-1} \int_r^\infty f'(\sigma) f(\sigma)^{\frac{p}{2}} \, d\sigma \right| &\lesssim \int_r^\infty \sigma^{d-1} |f'(\sigma)| f(\sigma)^{\frac{p}{2}} \, d\sigma \\ &\lesssim \left(\int_r^\infty \sigma^{d-1} |f'(\sigma)|^2 \, d\sigma \right)^{\frac{1}{2}} \left(\int_r^\infty \sigma^{d-1} |f(\sigma)|^p \, d\sigma \right)^{\frac{1}{2}} \\ &\lesssim \|\nabla f\|_{L^2} \|f\|_{L^p}^{\frac{p}{2}}.\end{aligned}$$

The final inequality here comes from the fact that f is radial and noting that $\sigma^{d-1} \, d\sigma$ describes a polar coordinate transformation. Raising both sides of this inequality to the power $\frac{2}{p+2}$ completes the proof. □

Observe that as p increases in the above inequality, the decay condition weakens. In particular, the case $p = 2$ gives particularly good decay. As $\dot{H}^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d) = H^1(\mathbb{R}^d)$, this proves the following.

Corollary 12.11. *If $f \in H^1_{rad}(\mathbb{R}^d)$ for $d \geq 2$, then*

$$r^{\frac{d-1}{2}} |f(r)| \lesssim \|f\|_{L^2}^{\frac{1}{2}} \|\nabla f\|_{L^2}^{\frac{1}{2}}$$

for almost all $r > 0$.

Finally, we prove the Strauss lemma.

Lemma 12.12 (Strauss). *Fix $d \geq 2$ and $2 < p < \infty$ if $d = 2$ or $2 < p < \frac{2d}{d-2}$ if $d \geq 3$. Then $H^1_{rad}(\mathbb{R}^d)$ embeds compactly into $L^p(\mathbb{R}^d)$.*

Proof. By normalizing, it suffices to prove that the unit ball $B_1 \subseteq H^1_{rad}$ is precompact in $L^p(\mathbb{R}^d)$. Naturally, we will apply the Riesz compactness theorem.

The arguments for boundedness and equicontinuity are similar to those used in the proof of the Rellich-Kondrachov theorem. For $f \in B_1$, Gagliardo-Nirenberg gives

$$\|f\|_{L^p} \lesssim \|f\|_{L^2}^{\theta} \|\nabla f\|_{L^2}^{1-\theta} \lesssim 1$$

for some the appropriate exponent θ. Thus, B_1 is L^p-bounded. Equicontinuity follows from Gagliardo-Nirenberg and the fundamental theorem of calculus:

$$\begin{aligned} \|f(x+y) - f(x)\|_{L^p_y} &\lesssim \|f(x+y) - f(x)\|_{L^2}^{\theta} \|\nabla_x (f(x+y) - f(x))\|_{L^2}^{1-\theta} \\ &\lesssim (|y| \|\nabla f\|_{L^2})^{\theta} \|\nabla f\|_{L^2}^{1-\theta} \\ &\lesssim |y|^{\theta}. \end{aligned}$$

It remains to show tightness. We wish to use Corollary 12.11. If we try to use the theorem directly, we get

$$\begin{aligned} \int_{|x| \geq R} |f(x)|^p \, dx &\lesssim \int_{|x| \geq R} r^{-\frac{d-1}{2}p} \, dx = \int_R^{\infty} r^{-\frac{d-1}{2}p} r^{d-1} \, dr \\ &\lesssim R^{d - \frac{d-1}{2}p} \end{aligned}$$

provided that $d - \frac{d-1}{2}p < 0$, hence $p > \frac{2d}{d-1}$. But the conclusion holds for $p > 2$, so this is not quite strong enough. Instead, we first use the fact that $f \in L^2(\mathbb{R}^d)$. Writing

$$\int_{|x| \geq R} |f(x)|^p \, dx = \int_{|x| \geq R} |f(x)|^2 \, |f(x)|^{p-2} \, dx$$

and using the radial Sobolev embedding theorem on $|f(x)|^{p-2}$ rather than $|f(x)|^p$ gives

$$\int_{|x| \geq R} |f(x)|^p \, dx \lesssim \|f\|_{L^2}^2 R^{-\frac{d-1}{2}(p-2)} \|f\|_{L^2}^2 .$$

This can be made uniformly small for $p > 2$ by choosing R sufficiently large.

By the Riesz compactness theorem, the proof is complete. $\square$

The endpoint cases in the Strauss lemma, namely $p = 2$ and $p = \frac{2d}{d-2}$ when $d \geq 3$, do not hold. That is, $H^1_{rad}(\mathbb{R}^d)$ does not embed compactly into $L^2(\mathbb{R}^d)$ or $L^{\frac{2d}{d-2}}(\mathbb{R}^d)$. To see that $H^1_{rad}(\mathbb{R}^d)$ does not embed compactly into $L^2(\mathbb{R}^d)$, we choose a radial bump function and rescale on a large scale. Explicitly, let $\varphi \in C_c^\infty$ be radial, and define $\varphi_\lambda(x) := \lambda^{-\frac{d}{2}}\varphi(\frac{x}{\lambda})$. Then $\|\varphi_\lambda\|_{L^2} = \|\varphi\|_{L^2}$. Also,

$$\|\nabla\varphi_\lambda\|_{L^2} = \lambda^{-\frac{d}{2}}\lambda^{-1}\lambda^{\frac{d}{2}}\|\nabla\varphi\|_{L^2} = \lambda^{-1}\|\nabla\varphi\|_{L^2}$$

which is $\lesssim 1$ for $\lambda >> 1$, hence $\varphi_\lambda \in H^1(\mathbb{R}^d)$. Note that

$$|\langle\varphi_\lambda, \psi\rangle| \lesssim \lambda^{-\frac{d}{2}}\|\varphi\|_{L^\infty}\|\psi\|_{L^1}$$

and so $\varphi \rightharpoonup 0$as $\lambda \to \infty$. However, as the L^2-norm of φ_λ is constant in λ, the φ_λ's cannot converge strongly to 0 in $L^2(\mathbb{R}^d)$. The argument for showing that $H^1_{rad}(\mathbb{R}^d)$ does not embed compactly into $L^{\frac{2d}{d-2}}(\mathbb{R}^d)$ is similar; here, we rescale to small scales. Let $\varphi_\lambda(x) = \lambda^{-\frac{d-2}{2}}\varphi(\frac{x}{\lambda})$. Then $\|\varphi_\lambda\|_{L^{\frac{2d}{d-2}}} = \|\varphi\|_{L^{\frac{2d}{d-2}}}$, and in fact $\|\varphi_\lambda\|_{\dot{H}^1} = \|\varphi\|_{\dot{H}^1}$. Also,

$$\|\varphi_\lambda\|_{L^2} = \lambda^{-\frac{d-2}{2}}\lambda^{\frac{d}{2}}\|\varphi\|_{L^2} = \lambda\|\varphi\|_{L^2}$$

which is $\lesssim 1$ for $\lambda << 1$. We have $\varphi_\lambda \rightharpoonup 0$ as $\lambda \to 0$, but φ_λ cannot converge to 0 in $L^{\frac{2d}{d-2}}(\mathbb{R}^d)$.

12.4 The Refined Fatou's Lemma

The last result that we include in this chapter is an improvement on Fatou's lemma. It will be important for establishing so-called *asymptotic decoupling* in $L^p(\mathbb{R}^d)$ when we study concentration compactness in Chapters 13 and 14.

Lemma 12.13 (Brezis-Lieb). *Fix $1 \le p < \infty$, and let f_n be an L^p-bounded sequence of functions. Suppose that $f_n \to f$ almost everywhere. Then*

$$\int \big|\,|f_n|^p - |f - f_n|^p - |f|^p\,\big|\,dx = 0.$$

Proof. First, recall that for any real numbers a and b,

$$\Big|\,|a+b|^p - |a|^p\,\Big| \le \varepsilon|a|^p + C_\varepsilon|b|^p \tag{12.1}$$

for any $\varepsilon > 0$, where C_ε is some constant dependent on ε and p. Let

$$g_n^\varepsilon := \left(\Big|\,|f_n|^p - |f_n - f|^p - |f|^p\,\Big| - \varepsilon|f_n - f|^p\right)^+.$$

Because $f_n \to f$ almost everywhere, $g_n^\varepsilon \to 0$ almost everywhere. By the triangle inequality and (12.1) with $a = f_n - f$ and $b = f$,

$$\begin{aligned} g_n^\varepsilon &\le \left(\Big|\,|f_n|^p - |f_n - f|^p\Big| + |f|^p - \varepsilon|f_n - f|^p\right)^+ \\ &\le (C_\varepsilon|f|^p + |f|^p)^+ \\ &= (C_\varepsilon + 1)\,|f|^p. \end{aligned}$$

As $f \in L^p$ (the sequence f_n is L^p-bounded!), $|f|^p$ is integrable and so by the Dominated convergence theorem $\int g_n^\varepsilon\,dx \to 0$. Thus,

$$\begin{aligned} &\int \big|\,|f_n|^p - |f - f_n|^p - |f|^p\,\big|\,dx \\ &\qquad \le \int \left(\Big|\,|f_n|^p - |f_n - f|^p - |f|^p\,\Big| - \varepsilon|f_n - f|^p\right)^+ dx + \int \varepsilon|f_n - f|^p\,dx \\ &\qquad\qquad = \int g_n^\varepsilon\,dx + \varepsilon\int |f_n - f|^p\,dx. \end{aligned}$$

For any $\varepsilon > 0$, the first quantity goes to 0. As f_n is L^p-bounded by assumption, the second term goes to 0 uniformly in ε. This completes the proof. ☐

Chapter 13

The Sharp Gagliardo-Nirenberg Inequality

As alluded to in the introduction to Chapter 12, in some sense, the next three chapters are in some sense devoted to the notion of *concentration compactness*. Broadly, concentration compactness studies the defects of compactness, and describes how to deal with these defects. This will be described formally in great deal in the future.

Although concentration compactness is the underlying theme of the next three chapters, it is contextualized by various *sharp* inequalities. Indeed, our primary use for the principle of concentration compactness will be to determine optimal constants for the Gagliardo-Nirenberg inequality, the Sobolev embedding theorem, and the Strichartz inequality.

In this chapter, our focus is the Gagliardo-Nirenberg inequality. For good measure, we recall the inequality here.

Theorem 13.1 (Gagliardo-Nirenberg). *Fix $d \geq 1$. Let p satisfy $2 \leq p < \infty$ if $d = 1$ or $d = 2$, and $2 \leq p < \frac{2d}{d-2}$ if $d \geq 3$. Then for $f \in H^1(\mathbb{R}^d)$,*

$$\|f\|_{L^p} \lesssim \|f\|_{L^2}^{\frac{2d-p(d-2)}{2p}} \|\nabla f\|_{L^2}^{\frac{d(p-2)}{2p}}.$$

We will prove the sharp version of inequality first without using concentration compactness, and then spend the rest of the chapter developing the principle of concentration compactness to provide and alternate proof. Chapters 14 and 15 develop analogous concentration compactness principles for the Sobolev embedding theorem and Strichartz inequality, respectively.

13.1 The Focusing Cubic NLS

Before we determine the optimal constant in the Gagliardo-Nirenberg inequality, we feel it necessary to provide a suitable source of motivation. After all, we have spent the entirety of this text generously employing estimates involving $\lesssim$, with no regard for explicit constants. Unsurprisingly, the optimal constant in Theorem 13.1 is important in the study of certain nonlinear Schrödinger equations. In particular, we will consider the **focusing cubic nonlinear Schrödinger equation** in $d = 2$, given by

$$\begin{cases} i\partial_t u + \Delta u = -|u|^2 u \\ u(0) = u_0 \end{cases}. \tag{13.1}$$

Here, $t \in \mathbb{R}$ and $x \in \mathbb{R}^2$, and for now we take the initial data u_0 to be a complex-valued Schwartz function. This is one of the most commonly studied nonlinear Schrödinger equations.

Chapter 14

The Sharp Sobolev Embedding Theorem

Chapter 15

Concentration Compactness in Partial Differential Equations

References

[1] F.M. Christ and M.I. Weinstein. "Dispersion of small amplitude solutions of the generalized Korteweg-de Vries equation". In: *Journal of Functional Analysis* 100.1 (1991), pp. 87–109. ISSN: 0022-1236.

[2] J. Duoandikoetxea. *Fourier Analysis*. Crm Proceedings & Lecture Notes. American Mathematical Soc., 2001. ISBN: 9780821883846.

[3] L. Grafakos. *Classical Fourier Analysis*. Second Edition. 489 pp: Springer, 2008.

[4] L.I. Hedberg. "On Certain Convolution Inequalities". In: *Proceedings of the American Mathematical Society* 36.2 (1972), pp. 505–510.

[5] R.A. Hunt. "An extension of the Marcinkiewicz interpolation theorem to Lorentz spaces". In: *Bull. Amer. Math. Soc.* 70.6 (Nov. 1964), pp. 803–807.

[6] E.H. Lieb and M. Loss. *Analysis*. Second Edition. 348 pp: American Mathematical Society, 2001.

[7] E. Stein and R. Shakarchi. *Fourier Analysis: An Introduction*. 326 pp: Princeton University Press, 2003.

[8] E. Stein and R. Shakarchi. *Functional Analysis: Introduction to Further Topics in Analysis*. 423 pp: Princeton University Press, 2011.

[9] E. Stein and R. Shakarchi. *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*. 402 pp: Princeton University Press, 2005.

[10] E.M. Stein and T.S. Murphy. *Harmonic Analysis: Real-variable Methods, Orthogonality, and Oscillatory Integrals*. Monographs in harmonic analysis. Princeton University Press, 1993.

[11] T. Tao. "Recent progress on the restriction conjecture". In: *ArXiv Mathematics e-prints* (Nov. 2003). eprint: math/0311181.