Notes on harmonic analysis

Mark Williams

January 11, 2022

Contents

1 Introduction

The purpose of these notes is to present a direct path through some of the central topics of harmonic analysis. We assume familiarity with distribution theory and the Fourier transform, and the basics of measure theory and functional analysis. It will help if the reader has already seen the Hardy-Littlewood maximal function applied to differentiation theory in a measure theory class or text. We want to get quickly to the Calderon-Zygmund theory of singular integrals, Littlewood-Paley theory, H^1-BMO duality, paraproducts, the $T(1)$ theorem, Fourier transform restriction, and to see some applications. The selection and ordering of topics is based on logical connections rather than on the historical development of the subject. We discuss just a few of the simpler results on Fourier series.[1] The emphasis is on real variable methods.

Warnings: We often switch between $\mathbb{R}^n$ and $\mathbb{R}^d$ as notations for euclidean space. We work on $\mathbb{R}^n$ with Lebesgue measure, except when we indicate otherwise, and we use $|S|$ to denote the $n-$dimensional Lebesgue measure of S when $S \subset \mathbb{R}^n$. In formulas involving the Fourier transform and its inverse, we usually ignore accompanying factors of 2π.

2 The Hardy-Littlewood maximal function

The Hardy-Littlewood maximal function is an indispensable tool that will be used in many different contexts in these notes.

First we state a couple of important lemmas.

Lemma 2.1 (Chebyshev). *If $0 < p < \infty$ and $f \in L^p(\mu)$, then*

$$\mu\left(\{x : |f(x)| > \alpha\}\right) \leq \left(\frac{|f|_p}{\alpha}\right)^p.$$

For the proof write $|f|_p^p \geq \int_{|f|>\alpha} |f|^p d\mu$.

Lemma 2.2 (Wiener covering lemma). *Let $\{B_1, \ldots, B_k\}$ be a finite collection of open balls in $\mathbb{R}^n$. Then there exists a finite subcollection $\{B_{j_1}, \ldots, B_{j_l}\}$ of pairwise disjoint balls such that*

$$| \cup_{i=1}^k B_i| \leq 3^n \sum_{r=1}^{l} |B_{j_r}|.$$

Proof. Reindex so that $|B_1| \geq |B_2| \geq \cdots$. Let $B_{j_1} = B_1$. Having chosen $\{B_{j_1}, \ldots, B_{j_k}\}$, let $B_{j_{k+1}} = B_s$, where s is the smallest index $> j_k$ such that B_s is disjoint from $B_{j_1} \cup \cdots \cup B_{j_k}$. This process terminates after finitely many steps. If some B_m is not selected, consider *all* the selected balls whose indices are $< m$; then B_m must intersect one of those balls, say B_{j_r} (or B_m would have been selected). Since B_m is smaller than B_{j_r} we must have $B_m \subset 3B_{j_r}$. Thus the union of the unselected balls is contained in the union of the triples of the selected balls.

□

Definition 2.3. *Let $0 < p \leq \infty$ and $0 < q \leq \infty$. Given $L^p(\mu)$ and $L^q(\nu)$, we say that an operator T (not necessarily linear) is strong (p, q) (or strong-type (p, q)) if*

$$|Tf|_q \lesssim |f|_p.$$

If $0 < q < \infty$, we say T is weak (p, q) if for all $\alpha > 0$[2]

$$\nu\{|Tf| > \alpha\} \lesssim \left(\frac{|f|_p}{\alpha}\right)^q.$$

Definition 2.4 (Hardy-Littlewood maximal function of f). *Let $f : \mathbb{R}^n \to \mathbb{C}$. The uncentered and centered Hardy-Littlewood maximal functions of f are respectively,*

$$M(f)(x) = \sup_{\delta>0, |x-y|<\delta} Av_{B(y,\delta)}|f| \text{ and } \mathcal{M}(f)(x) = \sup_{\delta>0} Av_{B(x,\delta)}|f|.$$

[1] For example, see section 5.1, Lemma 5.14, and Propositions 14.1 and 14.4.

[2] The importance of weak (p, q) is largely tied to the Marcinkiewicz interpolation theorem; see section 12. We sometimes write the condition defining weak (p, q) as $|Tf|_{L^{q,\infty}} \lesssim |f|_p$.

Theorem 2.5. *The uncentered HL maximal operator $f \to M(f)$ is weak type $(1,1)$ (with constant 3^n) and strong type (p,p) for $1 < p < \infty$. (Trivially $M : L^\infty \to L^\infty$ with norm ≤ 1.). The same holds for the centered maximal operator.*

Proof. We show M is weak type $(1,1)$ and apply Marcinkiewicz interpolation. For $\alpha > 0$ let

$$E_\alpha := \{x : M(f)(x) > \alpha\}.$$

Suppose $K \subset E_\alpha$ is compact. For each $x \in K$ there exists an open ball B_x (in fact, $B_x \subset E_\alpha$) such that

$$Av_{B_x}(f) > \alpha \Leftrightarrow |B_x| < \frac{1}{\alpha} \int_{B_x} |f|.$$

Cover K with such balls, use compactness, the covering lemma, and *disjointness* of the balls given by the covering lemma to show

$$|K| \leq \frac{3^n}{\alpha} \int |f| dx.$$

□

Remark 2.6. *For $\alpha > 0$ the* definition *of the maximal function immediately enables one to cover the set $E_\alpha := \{x : M(f) > \alpha\}$ by balls B_x such that*

$$|B_x| < \frac{1}{\alpha} \int_{B_x} |f|,$$

and this leads directly to a weak $(1,1)$ estimate.

2.1 Maximal operators and pointwise convergence

The next theorem illustrates the connection between maximal operators and pointwise convergence. In this theorem we consider operators mapping functions on (X, μ) to functions on (Y, ν).

Theorem 2.7 (Theorem 2.1.14 of G1)**.** *Suppose $0 < p, q < \infty$. Let $T_* f := \sup_{\epsilon > 0} |T_\epsilon f|$ and assume*

$$|T_* f|_{L^{q,\infty}} \leq B |f|_{L^p}. \tag{2.1}$$

Assume also that for all $f \in D$, a dense subspace of L^p, we have $\lim_{\epsilon \to 0} T_\epsilon f$ is finite a.e. and defines Tf. Then for all $f \in L^p$ that pointwise limit exists and defines an operator T that satisfies [3]

$$|Tf|_{L^{q,\infty}} \leq B |f|_{L^p}.$$

Proof. The idea is to use twice the maximal operator to control the "oscillation" of $T_\epsilon(f)(x)$ as $\epsilon \to 0$, when there is a dense subspace of functions g for which $T_\epsilon g \to h := Tg$ a.e., that is, for which the oscillation is zero a.e..

For $f \in L^p$ consider the oscillation of $T_\epsilon(f)$

$$O_f(y) = \limsup_{\epsilon, \theta \to 0} |T_\epsilon(f)(y) - T_\theta(f)(y)|,$$

where $O_g(y) = 0$ $\nu-$a.e for $g \in D$, a dense subset of $L^p(X, \mu)$. Fix $\eta > 0$ and choose $g \in D$ such that $|f - g|_p < \eta$. By the triangle inequality, we have $O_f(y) \leq O_{f-g}(y)$ $\nu-$a.e.. Since T_* is weak (p,q) we may conclude that for any $\delta > 0$:

$$\nu(\{y : 2T_*(f - g)(y) > \delta\}) \leq \left(\frac{2B\eta}{\delta}\right)^q.$$

[3] Initially, an operator T_ϵ is given for each $\epsilon > 0$.

Let $\eta \to 0$ to conclude $\nu(\{y : O_f(y) > \delta\}) = 0$. This gives pointwise convergence of $T_\epsilon(f)(y) \to T(f)(y)$ $\nu-$a.e. (and defines $T(f)$) and implies T satisfies (2.1). $\square$

Remark 2.8. *The weak (p, q) property of T_* and the pointwise convergence for $f \in D$ is all that is needed to conclude pointwise convergence of $T_\epsilon f$ for $f \in L^p$.*

We will give two corollaries of Theorem 2.7. In both the identity operator plays the role of "T". For these we need the following result, which shows that the Hardy-Littlewood maximal function controls the averages of a function with respect to any radially decreasing L^1 function (as, e.g., in (2.4)).

Theorem 2.9. *Let $K(x) = k(|x|)$ be integrable on $\mathbb{R}^n$, where $k(r)$ is decreasing and continuous. Then*

$$\sup_{\epsilon>0}(|f| * K_\epsilon)(x) \leq |K|_1 \mathcal{M}f(x) \tag{2.2}$$

for all $f \in L^1_{loc}$.

Proof. It is enough to prove (2.2) for $x = 0$ (consider translates of f) and for K with compact support in $B(0, R)$ for some $R > 0$ (take limits as $R_j \to \infty$). Fix $f \in L^1_{loc}$ and define

$$F(r) = \int_{S^{n-1}} |f(r\theta)| d\theta \text{ and } G(r) = \int_0^r F(s)s^{n-1}ds.$$

Since K is radial

$$\begin{aligned}\int_{\mathbb{R}^n} |f(y)|K_\epsilon(-y)dy = \int_0^\infty \int_{S^{n-1}} |f(r\theta)|K_\epsilon(re_1)r^{n-1}d\theta dr = \int_0^{\epsilon R} F(r)r^{n-1}K_\epsilon(re_1)dr = \\ G(\epsilon R)K(\epsilon Re_1) - G(0)K_\epsilon(0) - \int_0^{\epsilon R} G(r)dK_\epsilon(re_1) = \int_0^\infty G(r)d(-K_\epsilon(re_1)) := A.\end{aligned} \tag{2.3}$$

With $\nu_n = |B(0, 1)|$ we have

$$G(r) = \int_0^r F(s)s^{n-1}ds = \int_{|y|\leq r} |f(y)|dy \leq \mathcal{M}f(0)\nu_n r^n.$$

Thus,

$$A \leq \mathcal{M}f(0)\nu_n \int_0^\infty r^n d(-K_\epsilon(re_1)) = \mathcal{M}f(0) \int_0^\infty nr^{n-1}\nu_n K(re_1)dr = \mathcal{M}f(0)|K|_1.$$

$\square$

Remark 2.10. *This theorem generalizes to functions with L^1 radially decreasing majorants, a fact used later in the discussion of maximal SIOs (Cotlar's inequality).*

Corollary 2.11. *(a) Let $P(x) = c_n(1 + |x|^2)^{-(n+1)/2}$ be the Poisson kernel on $\mathbb{R}^n$ and set $P_\epsilon(x) = \epsilon^{-n}P(x/\epsilon)$. For $1 \leq p < \infty$ and $f \in L^p$, we have $f * P_\epsilon \to f$ a.e., where P_ϵ is the Poisson kernel.*[4]

(b) ***Lebesgue Differentiation Theorem.*** *For $f \in L^1_{loc}$,*

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

[4] This can be viewed as a statement about pointwise convergence of solutions $u(t, x) = (P_t * f)(x)$ to the Dirichlet problem on a half-space to $f(x)$ a.e. as $t \searrow 0$.

Proof. **a.** We have by Theorem 2.9

$$\sup_{\epsilon>0}(|f| * P_\epsilon)(x) \leq |P|_{L^1}\mathcal{M}(f)(x). \tag{2.4}$$

Since $\mathcal{M}(f)(x)$ is weak $(1,1)$ and strong (p,p) for $1 < p < \infty$, by (2.4) the same applies to

$$\sup_{\epsilon>0}|T_\epsilon f|, \text{ where } T_\epsilon f = f * P_\epsilon.$$

This gives (2.1) above for $p = q$ so can apply Theorem 2.7 taking D to be the set of continuous functions with compact support.

b. We have

$$\lim_{r\to 0}\frac{1}{|B(x,r)|}\int_{B(x,r)} f(y)dy = \lim_{\epsilon\to 0}(k_\epsilon * f)(x) = f(x) \text{ for a.e. } x,$$

where $k = \nu_n^{-1}\chi_{B(0,1)}$. Let $T_\epsilon f = k_\epsilon * f$ and use Theorem 2.7 as in part a.

□

3 Singular integral operators of convolution type

First, some examples:[5]

3.1 Hilbert and Riesz transforms

The Hilbert transform on $\mathbb{R}$ is given by convolution with $\frac{1}{\pi}p.v.\frac{1}{x}$. Equivalently, it is given by the Fourier multiplier $-i\text{sgn}\,\xi$.

Let $W_j = C(n)p.v.\ \frac{x_j}{|x|^{n+1}} \in \mathcal{S}'(\mathbb{R}^n)$. The Riesz transform R_j is defined by $R_j(f)(x) = (f * W_j)(x)$. R_j is also given by the multiplier $-\frac{i\xi_j}{|\xi|}$. We have

$$-I = \sum_j R_j^2 \text{ and } \partial_j\partial_k u = -R_jR_k(\Delta u).$$

3.2 Homogeneous SIOs

This class includes the Hilbert and Riesz transforms. Let Ω be integrable on S^{n-1} with mean zero. Let

$$\begin{aligned} K_\Omega(x) &:= \frac{\Omega(x/|x|)}{|x|^n}, x \neq 0, \\ W_\Omega &= \ p.v.\ K_\Omega \in \mathcal{S}', \\ T_\Omega(f)(x) &= (f * W_\Omega)(x) = \lim_{\epsilon\to 0, N\to\infty} T_\Omega^{(\epsilon,N)}(f)(x), \end{aligned} \tag{3.1}$$

where $T_\Omega^{(\epsilon,N)}(f)(x) = \int_{\epsilon\leq|y|\leq N} K_\Omega(y)f(x-y)dy$.

Below we will show that these operators are weak $(1,1)$ and strong (p,p) for $1 < p < \infty$.

[5]Later we consider nonconvolution-type SIOs (such as paraproducts, pseudodifferential operators,...).

3.3 Calderon-Zygmund kernels and operators

Next we define a class of convolution type operators that includes the homogenous SIOs.

Definition 3.1. *Let $K : \mathbb{R}^d \setminus 0 \to \mathbb{C}$ satisfy for some B:*
i) $|K(x)| \leq B|x|^{-d}$ (the size condition)
ii) $\int_{|x|>2|y|} |K(x) - K(x-y)|dx \leq B$ for all $y \neq 0$ (the Hörmander condition)
iii) $\int_{r<|x|<s} K(x)dx = 0$ for all $0 < r < s$. (the cancellation condition).
We call K a Calderon-Zygmund kernel. *For $f \in \mathcal{S}(\mathbb{R}^d)$ define the CZ operator with kernel K by*

$$Tf(x) = \lim_{\epsilon\to 0} \int_{|x-y|>\epsilon} K(x-y)f(y)dy = \lim_{\epsilon\to 0} \int_{|y|>\epsilon} K(y)f(x-y)dy. \tag{3.2}$$

Observe

$$\int_{|y|>\epsilon} K(y)f(x-y)dy = \int_{1>|y|>\epsilon} K(y)[f(x-y) - f(x)]dy + \int_{|y|\geq 1} K(y)f(x-y)dy. \tag{3.3}$$

Write $f(x-y) - f(x) = -\int_0^1 \nabla f(x-sy)ds \cdot y = g(x,y)y$, where

$$g(x,y) = \int_0^1 \nabla f(x-sy)ds \tag{3.4}$$

The pointwise limit as $\epsilon \to 0$ of (3.3) is

$$\int_{|y|<1} K(y)g(x,y)ydy + \int_{|y|\geq 1} K(y)f(x-y)dy,$$

a C^∞ function of x. [6]

Remark 3.2. *a) When we speak of a CZ operator as being, for example, "bounded on L^2", we refer to the extension defined using density of $\mathcal{S}$ in L^2 which exists by virtue of an estimate*

$$|Tf|_2 \leq C|f|_2 \text{ for all } f \in \mathcal{S}.$$

b) Observe that for $f \in \mathcal{S}$ we have

$$Tf = W * f, \text{ where } W = \text{p.v. } K \in \mathcal{S}'. \tag{3.5}$$

c) The Hörmander condition, (ii) above, played no role in the above argument for $f \in \mathcal{S}$. It is used, for example, in proving L^2 boundedness and the weak $(1,1)$ estimate.

3.4 Brief overview: proving boundedness of SIOs on L^p

The boundedness of SIOs on L^2 can be proved using the Fourier transform.[7] The behavior of SIOs on L^1 is more subtle - weak type (1,1) not strong type (1,1). Knowing L^r boundedness for some $1 < r \leq \infty$ (usually, one takes $r = 2$), one can use the Calderon-Zygmund decomposition to prove a weak type (1,1) estimate for a large class of SIOs including those of the previous section.[8] One can then prove strong type (p,p) estimates for $1 < p < \infty$ by interpolation and duality.

[6] Differentiate under the integral signs.
[7] It can also be proved using "almost orthogonality" arguments - see Cotlar's Lemma in section 8.
[8] The operators considered now need not commute with dilations.

For maximal SIOs we will prove Cotlar's inequality

$$T^*(f) \leq M(T(f)) + CM(f)$$

where M is Hardy-Littlewood maximal operator. This implies L^p boundedness of T^* for $1 < p < \infty$ and thus corresponding pointwise convergence results. We give such results for the Hilbert and Riesz tranforms in Corollary 3.10.

3.5 Calderon-Zygmund decomposition and "stopping time" arguments

A "stopping time argument" is the sort of argument in which a selection procedure stops when it is exhausted at a certain scale and is then repeated at the next scale. The Calderon Zygmund decomposition of an L^1 function f is based on such an argument. For a given $\alpha > 0$, at each scale a cube is selected or not depending on whether $Av_Q(f) := \frac{1}{|Q|}\int_Q f > \alpha$ or not. We proceed from one scale to the next by bisecting the sides of each cube at the first scale. Selected cubes are not subdivided.

The Calderon-Zygmund decomposition at height $\alpha > 0$. Let $f \in L^1(\mathbb{R}^n)$, $\alpha > 0$. We want to decompose $f = g + b$, where g is "good" ($\in L^1 \cap L^\infty$, hence $\in L^p$ for $p \geq 1$) and b is "bad" (contains singular part of f; in general $b \notin L^\infty$).

A *dyadic cube* in $\mathbb{R}^n$ is a cartesian product

$$\prod_{j=1}^{n} [2^k m_j, 2^k(m_j + 1)), \text{ where } k, m_j \in \mathbb{Z}.$$

Observe that two dyadic cubes are either disjoint or related by inclusion.

Theorem 3.3 (Calderon-Zygmund decomposition)**.** *Let $f \in L^1(\mathbb{R}^n)$ and $\alpha > 0$. There exist functions g and b such that $f = g + b$ where:*

(a) $|g|_1 \leq |f|_1$ (so $|b|_1 \leq 2|f|_1$) and $|g|_\infty \leq 2^n\alpha$.

(b) $b = \sum_j b_j$, where each b_j is supported in a dyadic cube Q_j. The cubes Q_j and Q_k are disjoint when $j \neq k$.

(c) $\int_{Q_j} b_j dx = 0$.

(d) $|b_j|_{L^1} \leq 2^{n+1}\alpha|Q_j|$.

(e) $\sum_j |Q_j| \leq \alpha^{-1}|f|_{L^1}$, so $\sum_j |b_j|_{L^1} \leq 2^{n+1}|f|_{L^1}$.

Remark 3.4. *Part (a) implies that g, being integrable and bounded, lies in L^p for all $1 \leq p \leq \infty$. The function b need not be L^∞, but it has* mean zero*.*

Proof. **1.** Decompose $\mathbb{R}^n$ into a union of disjoint dyadic cubes such that $|Q| \geq \frac{1}{\alpha}|f|_1$; use $|f|_1$ finite. Call these cubes "of zero generation". Divide each cube of zero generation into 2^n congruent dyadic subcubes by bisecting each of its sides. Call these smaller cubes "of generation one". Each cube of generation one has a unique "parent" in generation zero. *Select* a cube of generation one if

$$\frac{1}{|Q|}\int_Q |f(x)|dx > \alpha. \tag{3.6}$$

Let $S^{(1)}$ be the set of all selected cubes of generation one. Subdivide each nonselected cube of generation one into 2^n congruent dyadic subcubes and call these new subcubes of generation two. Let $S^{(2)}$ be the set of all cubes of generation two such that (11.30) holds. Continue indefinitely. The set of Q_j in the statement of the theorem is the countable set $\cup_{m=1}^{\infty} S^{(m)}$; this set could be empty. Selected cubes were never subdivided, so the Q_j are disjoint.

2. Define

$$b_j = (f - Av_{Q_j}(f))\chi_{Q_j},\ b = \sum_j b_j,\ g = f - b.$$

Note the b_j have *mean zero.* Also

$$g = \begin{cases} f, \text{ on } \mathbb{R}^n \setminus \cup_j Q_j \\ Av_{Q_j}(f) \text{ on } Q_j \end{cases}. \tag{3.7}$$

3. Each Q_j has a unique nonselected parent Q' of twice the sidelength. Thus,

$$\frac{1}{|Q_j|}\int_{Q_j} |f|dx \le \frac{1}{|Q_j|}\int_{Q'} |f|dx \le \frac{2^n}{|Q'|}\int_{Q'} |f|dx \le 2^n\alpha, \tag{3.8}$$

so by the triangle inequality $\int_{Q_j} |b_j|dx \le 2\int_{Q_j} |f|dx \le 2^{n+1}\alpha|Q_j|$, giving (d). Moreover,

$$\sum_j |Q_j| \le \frac{1}{\alpha}\sum_j \int_{Q_j} |f|dx \le \frac{1}{\alpha}|f|_1$$

giving (e).

4. On Q_j we have $|g| \le 2^n\alpha$ by (3.8). If $x \in \mathbb{R}^n \setminus \cup_j Q_j$, then $\{x\} = \cap_k \overline{Q_x^{(k)}}$, where $Q_x^{(k)}$ is the unique nonselected cube of the kth generation that contains x. So we can apply the Lebesgue differentiation theorem to show $|f| \le \alpha$ a.e. on $\mathbb{R}^n \setminus \cup_j Q_j$. Thus $|g| \le \alpha$ a.e. on this set. The formula (3.7) for g implies $|g|_1 \le |f|_1$, so this gives (a).

□

3.6 A weak type $(1,1)$ estimate assuming L^2 boundedness.

We now apply the Calderon-Zygmund decomposition to estimate singular integral operators.

Theorem 3.5. *Let T be a CZ operator with kernel K as in Definition 3.1. Suppose T is strong $(2,2)$.*[9] *Then T is weak $(1,1)$ and strong (p,p) for $1 < p < \infty$.*

Proof. **1. Weak** $(1,1)$. For a fixed $\alpha > 0$ we must show $|\{x : |T(f)(x)| > \alpha\}| \lesssim \frac{1}{\alpha}|f|_1$. Use Theorem 3.3 at height α to write $f = g + b$. Then $g \in L^1 \cap L^\infty \subset L^2$, so $Tg \in L^2$. Thus, using Theorem 3.3(a) we obtain

$$\{x : |T(g)(x)| > \alpha/2\} \le \frac{2^2}{\alpha^2}|Tg|_2^2 \lesssim \frac{1}{\alpha^2}|g|_2^2 \lesssim_n \frac{1}{\alpha}|f|_1.$$

It remains to control

$$|\{x : |Tb(x) > \alpha/2\}| \le |\cup_j Q_j^*| + |\{x \notin \cup_j Q_j^*, |Tb(x)| > \alpha/2\}| = A + B,$$

where Q_j^* has same center, y_j, as Q_j but larger side length by a factor $2\sqrt{n}$. The term A is controlled by Theorem 3.3(a).

2. We have[10]

$$|B| \le \frac{2}{\alpha}\int_{(\cup_j Q_j^*)^c} |Tb(x)|dx \le \frac{2}{\alpha}\sum_j \int_{(Q_j^*)^c} |Tb_j(x)|dx. \tag{3.9}$$

[9] One could equally well assume here that T is strong (r,r) for some $1 < r \le \infty$.

[10] Here use

$$\sum_j \int_{(Q_j^*)^c} |Tb_j|dx \ge \sum_j \int_{(\cup_j Q_j^*)^c} |Tb_j|dx \ge \int_{(\cup_j Q_j^*)^c} |Tb|dx.$$

Using the mean zero property of b_j, changing the order of integration, and using the Hörmander condition (since $|x-y_j| \geq 2|y-y_j|$) we have

$$\int_{(Q_j^*)^c} |Tb_j(x)|dx = \int_{Q_j} |b_j(y)| \int_{(Q_j^*)^c} |K(x-y)-K(x-y_j)|dxdy \lesssim B|b_j|_1. \tag{3.10}$$

Use this in (3.9) with part (e) of Theorem 3.3 to finish the weak $(1,1)$ estimate.

3. Use strong $(2,2)$ and weak $(1,1)$ with Marcinkiewicz interpolation to get strong (p,p) for $1<p<2$. Then use duality to get strong (p,p) for $2<p<\infty$. Here use $K^*(x) = \overline{K(-x)}$ to see that T^* is strong (p,p) for $1<p<2$.

□

Observe that the function K appears in the above argument only in (3.10), and that x and y are well separated there. Thus, the same argument proves the following more general result, which we will apply later, for example, in proving the Mihlin multiplier theorem.

Theorem 3.6. *Let $T : \mathcal{S} \to \mathcal{S}'$ be an operator bounded on L^2 such that for $f \in L^2_{comp}$ and $x \notin \operatorname{supp} f$*

$$Tf(x) = \int K(x-y)f(y)dy, \tag{3.11}$$

where K is a locally integrable function on $\mathbb{R}^n \setminus 0$ that satisfies the Hörmander condition. Then T is weak $(1,1)$ and strong (p,p) for $1<p<\infty$.

3.7 Sufficient conditions for L^2 boundedness

In this section we show that the size, cancellation, and Hörmander (or smoothness) conditions of Definition 3.1 on the kernel K (with Plancherel) imply L^2 boundedness of the associated operator.

Theorem 3.7. *Let T be a CZ operator with kernel K as in Definition 3.1. Then T is strong $(2,2)$.*

Proof. **1.** Consider restricted kernel $K_{r,s}(x) = K(x)\chi_{r<|x|<s}$ and its Fourier transform

$$m_{r,s}(\xi) = \int e^{-ix\xi}K_{r,s}(x)dx = \int_{|x|<2\pi|\xi|^{-1}} + \int_{|x|>2\pi|\xi|^{-1}} = A + B.$$

By Plancherel it's enough to bound $\sup_{0<r<s} |m_{s,r}|_\infty$. The conclusion then follows from "the standard Fatou argument", Prop. 14.3.

2. Using the cancellation and size conditions we obtain[11]

$$A \leq \left| \int_{|x|<2\pi|\xi|^{-1}} (e^{-ix\xi}-1)K(x)ds \right| \lesssim \int_{|x|<2\pi|\xi|^{-1}} |x||\xi||x|^{-n}dx \lesssim_n 1.$$

3. To treat B first write $B =$

$$\int_{|x|>2\pi|\xi|^{-1}} e^{-ix\xi}K(x)dx = -\int_{|x|>2\pi|\xi|^{-1}} K(x)e^{-i\left(x+\frac{\pi\xi}{|\xi|^2}\right)\cdot\xi}dx = -\int_{\left|x-\frac{\pi\xi}{|\xi|^2}\right|>2\pi|\xi|^{-1}} K\left(x-\frac{\pi\xi}{|\xi|^2}\right)e^{-ix\xi}dx.$$

[11] Here and below we write, for example, $|x|<2\pi|\xi|^{-1}$ for the set $\{|x|<2\pi|\xi|^{-1}\} \cap \{r<|x|<s\}$.

Thus,

$$
\begin{aligned}
2B &= \int_{|x|>2\pi|\xi|^{-1}} K(x)e^{-ix\xi}dx - \int_{\left|x-\frac{\pi\xi}{|\xi|^2}\right|>2\pi|\xi|^{-1}} K\left(x-\frac{\pi\xi}{|\xi|^2}\right)e^{-ix\xi}dx, \text{ so} \\
|2B| &\le \int_{|x|>2\pi|\xi|^{-1}} \left|K(x)-K(x-\frac{\pi\xi}{|\xi|^2})\right| dx + CB \le C'B.
\end{aligned}
\tag{3.12}
$$

where the CB term reflects the obvious change in regions of integration in one of the integrals. The estimate of the CB term uses the size condition on K and $\int_{|x|\lesssim|\xi|^{-1}} |\xi|^d dx = O(1)$. The final inequality uses the Hörmander condition.[12]

□

3.8 Maximal SIOs, Cotlar's inequality.

Recall the pointwise convergence results of section 2.1 for the Poisson kernel P_ϵ and for k_ϵ, where $k = \nu_n^{-1}\chi_{B(0,1)}$.[13] This section shows that similar results hold for the convolution operators given by SIOs like the Hilbert and Riesz transforms.[14]

Theorem 3.8 (Cotlar's inequality)**.** *Suppose K satisfies the usual size and cancellation conditions (as in Defn. 3.1) and the smoothness condition ($x \in \mathbb{R}^n$)*[15]

$$|K(x)-K(x-y)| \le A|y|^\delta |x|^{-n-\delta} \text{ for } |x| > 2|y| > 0. \tag{3.13}$$

Let $K^\epsilon(x) = K(x)\chi_{|x|\ge\epsilon}$. Define $W \in \mathcal{S}'$ by $\langle W, f\rangle := \lim_{\epsilon\to 0}\langle K^\epsilon, f\rangle$, so[16]

$$Tf(x) := (W*f)(x) = \lim_{\epsilon\to 0}\int_{|x-y|>\epsilon} K(x-y)f(y)dy = \lim_{\epsilon\to 0}(K^\epsilon * f)(x).$$

Define the associated maximal SIO

$$T^*f = \sup_{\epsilon>0}|K^\epsilon * f|.$$

Then for $1<p<\infty$ there exists $C = C(p,n,K)$ such that (M is HL maximal operator)

$$|T^*f| \le M(Tf) + CMf. \tag{3.14}$$

Thus T^* is strong (p,p) since T and M are.

Proof. **1.** Let ϕ be a radially decreasing smooth function with integral 1 supported in $B(0,1/2)$. For any function g set $g_\epsilon(x) = \epsilon^{-n}g(\epsilon^{-1}x)$. Write the truncated function $K^\epsilon = K^{(\epsilon)}$ for now. Note K_ϵ satisfies the same three conditions as K uniformly in ϵ. A computation shows

$$f * K^{(\epsilon)} = f * \left((K_{\epsilon^{-1}})^{(1)}\right)_\epsilon = f * W * \phi_\epsilon + f * \left((K_{\epsilon^{-1}})^{(1)} - W_{\epsilon^{-1}} * \phi\right)_\epsilon.$$

Note $\sup_{\epsilon>0}|f * W * \phi_\epsilon| \le |\phi|_{L^1}M(Tf)$ by Theorem 2.9.

[12] We might refer to this as an argument that uses "cancellation in $e^{-ix\xi}$." A similar argument is used in [CS1] for an estimate involving Dirichlet kernels.

[13] The latter result *is* the Lebesgue differentiation theorem.

[14] Each of these transforms is convolution with a certain distribution $W \in \mathcal{S}'$.

[15] Condition 3.13 implies the Hörmander condition.

[16] We say W "coincides with" K on $\mathbb{R}^n \setminus 0$.

2. We claim that

$$|A| := \left|\left((K_{\epsilon^{-1}})^{(1)} - W_{\epsilon^{-1}} * \phi\right)(x)\right| \lesssim (1+|x|)^{-n-\delta}. \tag{3.15}$$

Indeed, for $|x| \geq 1$

$$A = |\int (K_{\epsilon^{-1}}(x) - K_{\epsilon^{-1}}(x-y))\phi(y)dy.$$

Since $|x| \geq 2|y|$ on the support of the integrand, (3.13) easily implies the claim. When $|x| < 1$

$$|A| = |W_{\epsilon^{-1}} * \phi(x)| \text{ and } W_{\epsilon^{-1}} * \phi(x) = \lim_{\delta\to 0}\int_{|x-y|\geq\delta} K_{\epsilon^{-1}}(x-y)\phi(y)dy = I_1 + I_2,$$

where

$$I_1 = \int_{|x-y|\geq 1} K_{\epsilon^{-1}}(x-y)\phi(y)dy \text{ and } I_2 = \lim_{\delta\to 0}\int_{\delta<|x-y|\leq 1} K_{\epsilon^{-1}}(x-y)(\phi(y)-\phi(x))dy.$$

We have $|x-y| \leq 2$ on the support of the integrals, so the size condition on K yields $|I_1| + |I_2| \lesssim 1$. □

Remark 3.9. *(a)With Theorem 2.7 this theorem yields pointwise convergence results for* $\lim_{\epsilon\to 0} T_\epsilon f$*,* $f \in L^p$ *for* $1 < p < \infty$*.*

(b)Theorem 4.3.5 of [G1] shows that if K satisfies size and Hörmander conditions and if *the associated T^* is strong* $(2,2)$*, then T^* is weak* $(1,1)$*. That is an analogue for maximal SIOs of the earlier result, Theorem 3.5, that if an SIO T is L^2 bounded, then T is weak* $(1,1)$*.*

Corollary 3.10 (Pointwise convergence for the Hilbert and Riesz transforms)**.** *Define the maximal Riesz transforms by $R_j^* f(x) = \sup_{\epsilon>0}|R_j^\epsilon f(x)|$, where*

$$R_j^\epsilon f(x) = c_n \int_{|y|\geq\epsilon} \frac{y_j}{|y|^{n+1}} f(x-y)dy.$$

Then for $1 \leq p < \infty$ and $f \in L^p$ we have $\lim_{\epsilon\to 0} R_j^\epsilon f(x) = R_j f(x)$ *a.e. The same applies to the Hilbert transform, where we define*

$$H^\epsilon f(x) = \frac{1}{\pi}\int_{|y|\geq\epsilon} \frac{f(x-y)}{y} dy.$$

Proof. For $1 < p < \infty$ this follows from Cotlar's inequality and Theorem 2.7. For $p = 1$ one needs to use the fact that H^* and R_j^* are weak $(1,1)$, a consequence of Theorem 4.3.5 of [G1]. □

3.9 Application: Mihlin multiplier theorem.

Definition 3.11. *A function $m(\xi)$ is an "L^p multiplier" if the map $f \to Tf := \mathcal{F}^{-1}(m\hat{f})$ is bounded on L^p.*

If m is bounded, it is an L^2 multiplier. When is a bounded function $m(\xi)$ an "L^p multiplier"? If $K = \check{m} \in \mathcal{S}'$, then m is an L^1 multiplier if and only if $K \in L^1$. If m is an L^2 multiplier and an L^1 multiplier, then it is an L^p multiplier for $1 < p < \infty$ (by interpolation and duality). There are L^p multipliers for $1 < p < \infty$ which are not L^1 multipliers, for example, the Hilbert transform.

Remark 3.12. *The characteristic function of the unit disk is not an L^p multiplier on $\mathbb{R}^n$ when $n > 1$ and $p \neq 2$; see section 10.1 [G1].*

The following multiplier theorem will be proved as an application of Theorem 3.6 using a dyadic partition of unity.

Theorem 3.13 (Mihlin multiplier theorem)**.** *Let* $m : \mathbb{R}^d \setminus 0 \to \mathbb{C}$ *satisfy*

$$|\partial^\gamma m| \lesssim |\xi|^{-|\gamma|}$$

away from 0 *for* $|\gamma| \leq d+2$. *Then the operator* $f \to \check{m} * f$ *is weak* $(1,1)$ *and strong* (p,p) *for* $1 < p < \infty$. *That is,* m *is an* L^p *multiplier for these* p.

Proof. **1.** Let χ be a test function such that $\chi = 1$ for $|\xi| \leq 1$, 0 for $|\xi| \geq 2$, and set $\psi(\xi) = \chi(\xi) - \chi(2\xi)$. This gives a dyadic partition of unity:

$$\sum_{j \in \mathbb{Z}} \psi(2^{-j}\xi) = 1 \text{ for } \xi \neq 0, \text{ and we can write } m = \sum_j m_j,\ m_j(\xi) = \psi(2^{-j}\xi) m(\xi).$$

Let $S_N = \sum_{-N}^{N} K_j$, $K_j := \check{m}_j$. Then $S_N \to \check{m}$ in $\mathcal{S}'$ (by the dominated convergence theorem). We claim that S_N, which is C^∞, converges absolutely away from $x = 0$ (and uniformly on $\delta \leq |x| \leq \delta^{-1}$ for any fixed $\delta > 0$) to a function $S(x)$ that satisfies Hörmander's condition. The conclusion then follows from Theorem 3.6.

2. We show first that

$$|\check{m}_j(x)| \lesssim_B 2^{j(d-k)} |x|^{-k} \text{ and } |D\check{m}_j(x)| \lesssim_B 2^{j(d+1-k)} |x|^{-k} \text{ for } 0 \leq k \leq d+2. \tag{3.16}$$

Indeed, our assumption on m implies $|\partial^\gamma m_j|_\infty \lesssim 2^{-j|\gamma|}$ and thus

$$|\partial^\gamma m_j|_1 \lesssim 2^{-j|\gamma|} 2^{jd} \text{ for } |\gamma| \leq d+2. \tag{3.17}$$

Similarly,

$$|\partial^\gamma (\xi_i m_j)|_1 \lesssim 2^{-j(|\gamma|-1)} 2^{jd} \text{ for } |\gamma| \leq d+2.$$

These estimates imply

$$|x^\gamma \check{m}_j(x)|_\infty \lesssim_B 2^{j(d-|\gamma|)} \text{ and } |x^\gamma D\check{m}_j(x)|_\infty \lesssim_B 2^{j(d+1-|\gamma|)}. \tag{3.18}$$

Since $|x|^k \lesssim \sum_{|\gamma|=k} |x^\gamma|$, the estimates (3.16) follow. Now sum the $\check{m}_j$ over $2^j < |x|^{-1}$ (take $k = 0$) and $2^j > |x|^{-1}$ (take $k = d+2$) to get $|S(x)| \lesssim_B |x|^{-d}$ for $x \neq 0$. For a fixed $\delta > 0$ this argument gives uniform convergence of S_N to $S(x)$ on $\delta \leq |x| \leq \delta^{-1}$. Essentially the same argument gives $|\nabla S(x)| \lesssim_B |x|^{-d-1}$, so S satisfies the Hörmander condition.

□

This result, like the Marcinkiewicz multiplier theorem, is not L^p sensitive; it yields a multiplier for *all* L^p, $1 < p < \infty$, or none.

4 Singular integral operators acting on Banach space-valued functions

Sometimes a nonlinear inequality can be viewed as a linear norm estimate for an operator acting on and taking values in suitable Banach spaces. This applies, for example, to Littlewood-Paley inequalities involving square functions. We introduce square functions now to clarify this and provide some motivation for this section.

Let $\hat{\chi}$ be a test function such that $\hat{\chi} = 1$ for $|\xi| \le 1$, 0 for $|\xi| \ge 2$, and set $\hat{\psi}(\xi) = \hat{\chi}(\xi) - \hat{\chi}(2\xi)$. So $\hat{\psi}$ is supported in the annulus $\frac{1}{2} \le |\xi| \le 2$ and gives a dyadic partition of unity:

$$\sum_{j \in \mathbb{Z}} \hat{\psi}(2^{-j}\xi) = 1 \text{ for } \xi \neq 0.$$

For $f : \mathbb{R}^n \to \mathbb{C}$ set

$$\Delta_j f = f * \psi_{2^{-j}}, \text{ where } \psi_{2^{-j}}(x) = 2^{jn}\psi(2^j x) \text{ and so } \widehat{\psi_{2^{-j}}}(\xi) = \hat{\psi}(2^{-j}\xi). \tag{4.1}$$

The Fourier transform of $\Delta_j f$ is supported in an annulus $2^{j-1} \le |\xi| \le 2^{j+1}$. The *square function of* f is defined as

$$f \to \left(\sum_{j \in \mathbb{Z}} |\Delta_j f|^2 \right)^{1/2} = |(\Delta_j f)|_{\ell^2} = S(f). \tag{4.2}$$

We can think of the square function (for $1 < p < \infty$) as arising from a *linear* map $S_L(f)(x) = (\Delta_j f(x))$ as follows:[17]

$$Sf(x) = |S_L f(x)|_{\ell^2}. \tag{4.3}$$

The Littlewood-Paley inequality (proved below) says that $S_L : L^p(\mathbb{R}^n, \mathbb{C}) \to L^p(\mathbb{R}^n, \ell^2)$ and

$$|Sf|_p \sim |f|_p, \text{ or in other words } |S_L f|_{L^p(\mathbb{R}^n, \ell^2)} \sim |f|_p. \tag{4.4}$$

In one of our proofs of (4.4) we will write $S_L f$ as a Banach space-valued singular integral and apply an analogue of Theorem 3.5.

4.1 Singular integrals with $L(\mathcal{B}_1, \mathcal{B}_2)-$valued kernels

First we give a little background on Bochner integrals. Let (X, μ) be $\sigma-$finite and suppose $F : (X, \mu) \to \mathcal{B}$ is $\mathcal{B}-$measurable.[18] This implies that the function $x \to |F(x)|_{\mathcal{B}}$ is measurable, so the space $L^1(X, \mathcal{B})$ is well-defined. The case $\mathcal{B} = \ell^2$ will be important for us.

If $F = \sum_{j=1}^N f_j u_j \in L^1(X) \otimes \mathcal{B}$ set

$$\int F(x) d\mu := \sum_j \left(\int f_j d\mu \right) u_j \in \mathcal{B}.$$

This extends uniquely to $I_F : L^1(X, \mathcal{B}) \to \mathcal{B}$, the Bochner integral. The spaces $L^p(X, \mathcal{B})$ are defined in the obvious way.

Remark 4.1. *Let* $1 \le p \le \infty$*.*

(a) We have

$$|F|_{L^p(\mathbb{R}^n, \mathcal{B})} = \sup_{|G|_{L^{p'}(\mathbb{R}^n, \mathcal{B}^*)} \le 1} \left| \int \langle G(x), F(x) \rangle dx \right|,$$

so the space $L^p(\mathbb{R}^n, \mathcal{B})$ *isometrically embeds in* $(L^{p'}(\mathbb{R}^n, \mathcal{B}^*))^*$*.*

[17] We often write sequences $(a_j)_{j=1}^{\infty}$ as just (a_j).

[18] We say $F : (X, \mu) \to \mathcal{B}$ is $\mathcal{B}-$measurable if for all $u \in B^*$ the map $x \to u(F(x)) \in \mathbb{C}$ is measurable and there exists a set X_0 of full measure such that $F(X_0)$ is contained in some separable subspace of $\mathcal{B}$.

(b) Similarly,

$$|G|_{L^{p'}(\mathbb{R}^n,\mathcal{B}^*)} = \sup_{|F|_{L^p(\mathbb{R}^n,\mathcal{B})}\le 1} \left| \int \langle G(x), F(x)\rangle dx \right|$$

so the space $L^{p'}(\mathbb{R}^n,\mathcal{B}^*)$ isometrically embeds in $(L^p(\mathbb{R}^n,\mathcal{B}))^*$.

Let $\vec{K}(x)$ for $x \in \mathbb{R}^n \setminus 0$ be an $L(\mathcal{B}_1,\mathcal{B}_2)-$valued kernel. Assume K is $L(\mathcal{B}_1,\mathcal{B}_2)$ measurable and locally integrable outside $x=0$. Then the integral

$$\vec{T}(F)(x) = \int_{\mathbb{R}^n} \vec{K}(x-y)F(y)dy \in \mathcal{B}_2 \tag{4.5}$$

is well-defined for $F \in L^2_{comp}(\mathbb{R}^n,\mathcal{B}_1)$ when x lies outside the support of F. An important case for Littlewood-Paley theory is when $\mathcal{B}_1 = \mathbb{C}$ and $\mathcal{B}_2 = \ell^2$.

The main result of this section is the following vector analogue of Theorem 3.6:

Theorem 4.2. *Suppose $\vec{T} : L^2(\mathbb{R}^n,\mathcal{B}_1) \to L^2(\mathbb{R}^n,\mathcal{B}_2)$*[19] *is such* (4.5) *holds that for $F \in L^2_{comp}(\mathbb{R}^n,\mathcal{B}_1)$ and $x \notin \operatorname{supp} F$, where $\vec{K}$ satisfies Hörmander's condition*

$$\int_{|x|>2|y|} |\vec{K}(x-y) - \vec{K}(x)|_{\mathcal{B}_1\to\mathcal{B}_2} dx \le A < \infty.$$

Then $\vec{T}$ is strong type (p,p) for $1<p<\infty$ and weak type $(1,1)$.

Proof. **1.** The weak $(1,1)$ estimate is proved by applying the CZ decomposition directly to the function on $\mathbb{R}^n$ given by $x \to |F(x)|_{\mathcal{B}_1}$ and repeating the arguments in the proof of Theorem 3.5. The proof is finished by using Marcinkiewicz interpolation to treat $1<p<2$, and interpolation and duality to treat $p>2$.

2. $p>2$. Use that $\vec{K}^*(x)$ is an operator from $\mathcal{B}_2^*$ to $\mathcal{B}_1^*$ which also satisfies the Hörmander condition, and that $\vec{T}^* : L^2(\mathbb{R}^n,\mathcal{B}_2^*) \to L^2(\mathbb{R}^n,\mathcal{B}_1^*)$. The latter fact is a consequence of Remark 4.1(b). Marcinkiewicz interpolation then implies that T^* is strong (p',p') for $1<p'<2$. To finish we apply Remark 4.1(a) to conclude T is strong (p,p) for $2<p<\infty$. □

5 Littlewood-Paley theory

In this section we give two proofs of the Littlewood-Paley theorem. First we discuss some "precursors", where we use that word in a logical rather than a temporal sense.

5.1 Precursors: Lacunary Fourier series, Khinchine's inequality

The next proposition shows that for functions that can be expressed as linear combinations of Rademacher-type functions, L^p norms are comparable as p varies.

Proposition 5.1 (Khinchine-type inequality)**.** *Let (Ω,μ) be a probability space and let $\{r_k, k = 1,2,\ldots\}$ be a family of independent real-valued random variables such that $|r_k|_2 = 1$, $|r_k|_\infty = A$, and $\int r_k^n \mu = 0$ for n odd (and all k).*[20] *For $p \in (0,\infty)$ we have*

$$|\sum c_k r_k|_p \sim_p (\sum |c_k|^2)^{1/2} = |\sum c_k r_k|_2. \tag{5.1}$$

[19] Here one could equally well assume that T is strong (r,r) for some $1<r\le\infty$

[20] The Rademaker functions satisfy these conditions.

Proof. **1.** First consider the direction $\lesssim$. The r_k are orthonormal by independence since $\int r_k = 0$ and $|r_k|_2 = 1$, so (5.1) holds for $p = 2$. For $p < 2$ one can use Hölder:

$$|\sum_k c_k r_k|_p = |1 \cdot \sum_k c_k r_k|_p \leq |1|_r \cdot |\sum c_k r_k|_2 \lesssim |(c_k)|_{\ell^2}, \text{ where } \frac{1}{p} = \frac{1}{2} + \frac{1}{r}.$$

2. If we treat $p = 2q$, where $q \geq 2$ is an integer, we can apply R-T interpolation to the map $T((c_k)) == \sum_k c_k r_k$. The main thing to observe is that by independence

$$\int \prod_{j=1}^{p} r_{k_j} d\mu = 0 \tag{5.2}$$

unless each k_j appears an even number of times. We illustrate the argument when the c_k are real:

$$\begin{aligned} \int \Big(\sum c_k r_k\Big)^p &= \sum_{k_1,\dots,k_p} \prod_{j=1}^{p} c_{k_j} \int \prod_{j=1}^{p} r_{k_j} d\mu \leq C(p) \sum_{n_1,\dots,n_q} \prod_{i=1}^{q} |c_{n_i}|^2 \int \prod_{i=1}^{q} r_{n_i}^2 d\mu \leq \\ & A^{2q} C(p) \sum_{n_1,\dots,n_q} \prod_{i=1}^{q} |c_{n_i}|^2 \sim \Big(\sum_k |c_k|^2\Big)^q = \Big(\sum_k |c_k|^2\Big)^{p/2}. \end{aligned} \tag{5.3}$$

3. For the $\gtrsim$ direction assume $1 \leq p < \infty$. It suffices to assume $1 \leq p \leq 2$, in fact, that $p = 1$. The inequality then follows easily from Hölder and the previous case; see [CS1], p. 114. □

Remark 5.2. *So the independence of the r_k implies that for functions that can be expressed as $\sum_k a_k r_k$, L^p norms are equivalent to L^2 norms. Independence implies that in* (5.2)*, the integral of the product can be separated into a certain product of the integrals. In the L^p setting where $p \neq 2$, independence can sometimes serve as a substitute for orthogonality.*

Note the similarities of the above proof to the proof of the following result on lacunary Fourier series, where again L^p norms turn out to be equivalent to L^2 norms. Consider $f(x) = \sum_{k=1}^{\infty} \hat{f}(\lambda_k) e^{2\pi i \lambda_k x}$, where the λ_k are positive integers such that for some $A > 1$ we have $\lambda_{k+1} \geq A\lambda_k$ for all k.

Theorem 5.3 (Lacunary Fourier series)**.** *For $1 \leq p < \infty$ we have*

$$|f|_p \sim_{p,A} |f|_1.$$

Consequently, we can replace $|f|_1$ by $|f|_2$ on the right if we like.

Sketch of proof. First treat $p \geq 2$ of the form $p = 2m$ for some $m \in \mathbb{N}$. Pick a positive integer r such that $A^r > m$. It is easy to check that we can break up f into a sum of r functions ϕ_s, $s = 1, ..., r$, where for each s, ϕ_s has Fourier coefficients vanishing except possibly on the lacunary set[21]

$$\Lambda_s := \{\lambda kr + s : k = 0, 1, 2, \dots\} := \{\mu_1, \mu_2, \dots\}. \tag{5.4}$$

For example,

$$\phi_s(x) = \sum_{\mu_k \in \Lambda_s} \hat{\phi}_s(\mu_k) e^{2\pi i \mu_k x}. \tag{5.5}$$

[21]It is technically helpful to truncate f to $f_N := \sum_{k=1}^{N} ...$, and to prove estimates for f_N that are uniform in N.

The sequences Λ_s are lacunary with constant A^r, and the functions ϕ_s are orthogonal. For each s we have

$$\int_0^1 |\phi_s(x)|^{2m} dx = \sum_{\mu_{j_1}+\cdots+\mu_{j_m}=\mu_{k_1}+\cdots+\mu_{k_m}} \hat{\phi}_s(\mu_{j_1})\ldots\hat{\phi}_s(\mu_{j_m})\overline{\hat{\phi}_s(\mu_{k_1})}\ldots\overline{\hat{\phi}_s(\mu_{k_m})}. \tag{5.6}$$

The key observation is that since $A^r > m$ we have[22]

$$\mu_{j_1}+\cdots+\mu_{j_m} = \mu_{k_1}+\cdots+\mu_{k_m} \Rightarrow \{\mu_{j_1},\ldots,\mu_{j_m}\} = \{\mu_{k_1},\ldots,\mu_{k_m}\}. \tag{5.7}$$

Thus,

$$\int_0^1 |\phi_s(x)|^{2m} dx = \sum_{j_1}\cdots\sum_{j_m} |\hat{\phi}_s(\mu_{j_1})|^2\cdots|\hat{\phi}_s(\mu_{j_m})|^2 = (|\phi_s|_{L^2}^2)^m,$$

which implies

$$|\phi_s|_{L^{2m}} = |\phi_s|_{L^2} \text{ for } s = 1,\ldots,r.$$

Thus,

$$|f_N|_{L^p} \le |f_N|_{L^{2m}} \le \sum_{s=1}^r |\phi_s|_{L^{2m}} \le \sqrt{r}\left(\sum_{s=1}^r |\phi_s|_{L^{2m}}^2\right)^{1/2} = \sqrt{r}\left(\sum_{s=1}^r |\phi_s|_{L^2}^2\right)^{1/2} = \sqrt{r}|f_N|_{L^2}. \tag{5.8}$$

Finish using Fatou to let $N \to \infty$ and interpolation for $1 < p < 2$.

□

Remark 5.4. *Each function ϕ_s is lacunary, like the original f, but ϕ_s is lacunary with a possibly much bigger constant A^r. If one tried to prove Theorem 5.3 by a direct application of Proposition 5.1 with the functions $e^{2\pi i\lambda_k x}$ playing the role of the r_k (or rather, say, functions $\cos(2\pi\lambda_k x)$), one should expect trouble since A can be quite close to 1, and so the family $\{\cos(2\pi\lambda_k x)\}_k$ would not really "behave independently." Passing to the ϕ_s is a way of increasing A and getting more nearly independent functions.*

Here we use the observation that when λ_{k+1} is much *larger than λ_k, one does have approximate equalities like*

$$\int_0^1 \cos^2(2\pi\lambda_{k+1}x)\cos^2(2\pi\lambda_k x)dx \approx \int_0^1 \cos^2(2\pi\lambda_{k+1}x)dx \cdot \int_0^1 \cos^2(2\pi\lambda_k x)dx.$$

To see this break up the integral on the left into a sum of integrals over periods of $\cos^2(2\pi\lambda_{k+1}x)$. On each period the other function is nearly constant. Approximate $\int_0^1 \cos^2(2\pi\lambda_k x)dx$ by the obvious Riemann sum to finish.

Next we consider the relation of these results to the Littlewood-Paley estimate. Let $f = \sum \hat{f}(n)e^{inx}$ and set

$$\Delta_k f(x) = \begin{cases} \sum_{2^k\le|n|<2^{k+1}} \hat{f}(n)e^{inx}, & k \ge 1 \\ \sum_{|n|\le 1} \hat{f}(n)e^{inx}, & k = 0 \end{cases}. \tag{5.9}$$

and define the *square function*

$$Sf = \left(\sum_{k=0}^\infty |\Delta_k f|^2\right)^{1/2}. \tag{5.10}$$

[22] See [G1] p. 241 for details.

Theorem 5.5 (Littlewood-Paley for Fourier series)**.** *For $1 < p < \infty$ and $f \in L^p([0, 2\pi])$ we have $|Sf|_p \sim |f|_p$.*

Remark 5.6. *In the case when $2^k \le \lambda_k < 2^{k+1}$ for all k, Theorem 5.3 on lacunary Fourier series (with $|f|_2$ on the right in the statement) follows immediately from Theorem 5.5, since $|\Delta_k f(x)| = |\hat f(\lambda_k)|$. In a similar way Prop. 5.1 (Khinchine) in the case of Rademacher functions follows from the LP theorem for the Haar square function; see section 13.1.*

The proof of Theorem 5.5 given in section 5.3, like the proof of the corresponding result on $\mathbb{R}^d$, uses the observation that one can use Khinchine to prove LP if one has a sufficiently good L^p multiplier theorem. Khinchine gives the pointwise *inequality*

$$|Sf(x)|^p = \left(\sum_j |P_j f(x)|^2\right)^{p/2} \le \mathbb{E}\left|\sum_j r_j(p)(P_j f)(x)\right|^p,$$

which is then integrated $\int dx$. To get rid of the P_j and obtain $|f|_p^p$, a multiplier theorem is used.

The multiplier result used for Theorem 5.5 is derived from the Mihlin theorem using a transference *lemma, Lemma 5.14.*

A related question is whether Littlewood-Paley theorems remain valid when the square function is defined in terms of *discontinuous* cutoffs; there are positive results in dimension one. See [S1], p. 104.

5.2 Proof of Theorem 5.7 using vector-valued singular integrals

Recall from the beginning of section 4 that we constructed a test function $\hat\psi$, supported in the annulus $\frac{1}{2} \le |\xi| \le 2$, which gives a dyadic partition of unity:

$$\sum_{j\in\mathbb{Z}} \hat\psi(2^{-j}\xi) = 1 \text{ for } \xi \ne 0.$$

For $f : \mathbb{R}^n \to \mathbb{C}$ we set

$$\Delta_j f = f * \psi_{2^{-j}}, \text{ where } \psi_{2^{-j}}(x) = 2^{jn}\psi(2^j x) \text{ and so } \widehat{\psi_{2^{-j}}}(\xi) = \hat\psi(2^{-j}\xi). \tag{5.11}$$

With $S_L f(x) = (\Delta_j f(x))$, the *square function of f* was defined as

$$Sf(x) = \left(\sum_{j\in\mathbb{Z}} |\Delta_j f|^2\right)^{1/2} = |S_L f(x)|_{\ell^2}. \tag{5.12}$$

Theorem 5.7 (LP)**.** *The operator S is strong type (p, p) for $1 < p < \infty$ and weak type $(1, 1)$. In fact for $1 < p < \infty$ and $f \in L^p$ one has $|S(f)|_p \sim |f|_p$.*

Remark 5.8. *(a) If the functions f_j have Fourier transforms with disjoint supports then clearly*

$$|\sum_j f_j|_{L^2}^2 = \sum_j |f_j|_{L^2}^2. \tag{5.13}$$

If L^2 is replaced by L^p with $p \ne 2$, the two sides may not be comparable; see Ex. 5.1.8, 5.1.9 [G1][23]

[23] To see this pick $\phi \in \mathcal{S}'$ with $\hat\phi \ge 0$ and supported in $|\xi| \le 1/4$. Consider $f_j = e^{2\pi i j x}\phi$, so the $\hat f_j(\xi) = \hat\phi(\xi - j)$ have disjoint supports. Then $\sum_{j=1}^N |f_j|_{L^p}^p$ grows like $N + 1$ while $|\sum_{j=0}^N f_j|_{L^p}^p \gtrsim (N+1)^{p-1}$, so desired inequality fails for $p > 2$.

(b) The estimate $|f|_2^2 \lesssim |Sf|_2^2$ in Theorem 5.7 holds for nearly the same reason. The terms $\Delta_j f$ have Fourier transforms that are nearly mutually disjoint. A given $\xi \neq 0$ lies in at most three annuli of the form $2^{j-1} \leq |\xi| \leq 2^{j+1}$, so by Plancherel

$$|f_2^2 = \left| \sum_j \psi(2^{-j}\xi)\hat{f}(\xi) \right|_2^2 \leq 4 \sum_j |\hat{\psi}(2^{-j}\xi)\hat{f}(\xi)|_2^2 = 4|Sf|_2^2. \tag{5.14}$$

The reverse inequality follows directly from $\sum_j |\hat{\psi}(2^{-j}\xi)|^2 \lesssim 1$.

Proof of Theorem 5.7. **1.** The case $p = 2$ was done in Remark 5.8(b). We will apply Theorem 4.2 to treat other p. Note that

$$\vec{T}(f)(x) = S_L(f)(x) = \left(\int \psi_{2^{-j}}(x-y)f(y)dy \right) = \int \vec{K}(x-y)f(y)dy \tag{5.15}$$

where for each $x \in \mathbb{R}^n$ the operator $\vec{K}(x) \in L(\mathbb{C}, \ell^2)$ is given by

$$\vec{K}(x)a = (\psi_{2^{-j}}(x)a)\,. \tag{5.16}$$

Clearly $|K(x)|_{\mathbb{C}\to\ell^2} = \sum_j |\psi_{2^{-j}}(x)|^2$, so the main work is to check that $\vec{K}$ satisfies Hörmander's condition.

2. For $|x| \geq 2|y|$ we have for some $\theta \in [0,1]$

$$|\psi_{2^{-j}}(x-y) - \psi_{2^{-j}}(x)| \lesssim 2^{n+1)j}|\nabla\psi(2^j(x-\theta y))||y| \lesssim 2^{(n+1)j}|y|(1+2^j|x|)^{-n-1} \lesssim 2^{(n+1)j}|y|. \tag{5.17}$$

We also have (use $2^{j-1} < 2^j$)

$$|\psi_{2^{-j}}(x-y) - \psi_{2^{-j}}(x)| \lesssim 2^{nj}|\psi(2^j(x-y))| + 2^{nj}|\psi(2^j(x))| \lesssim 2^{nj}(1+2^{j-1}|x|)^{-n-1}. \tag{5.18}$$

Taking the geometric mean of two of the previous estimates gives

$$|\psi_{2^{-j}}(x-y) - \psi_{2^{-j}}(x)| \lesssim 2^{(n+\frac{1}{2})j}|y|^{1/2}(1+2^{j-1}|x|)^{-n-1}. \tag{5.19}$$

Using the last term in (5.17) when $2^j < 2/|x|$ and (5.19) when $2^j \geq 2/|x|$ gives for $|x| \geq 2|y|$:

$$\begin{aligned} |\vec{K}(x-y) - \vec{K}(x)|_{\mathbb{C}\to\ell^2} &= \left(\sum_j |\psi_{2^{-j}}(x-y) - \psi_{2^{-j}}(x)|^2 \right)^{1/2} \leq \sum_j |\psi_{2^{-j}}(x-y) - \psi_{2^{-j}}(x)| \lesssim \\ &|y| \sum_{2^j < 2/|x|} 2^{n+1)j} + |y|^{1/2} \sum_{2^j \geq 2/|x|} 2^{(n+\frac{1}{2})j}(1+2^{j-1}|x|)^{-n-1} \lesssim |y||x|^{-n-1} + |y|^{1/2}|x|^{-n-\frac{1}{2}}. \end{aligned} \tag{5.20}$$

Thus, $\int_{|x|>2|y|} |\vec{K}(x-y) - \vec{K}(x)|_{\mathbb{C}\to\ell^2} dx \lesssim 1$, so we may apply Theorem 4.2 to show S is strong (p,p) for $1 < p < \infty$ and weak $(1,1)$.

3. For $1 < p < \infty$ we obtain the reverse inequality by duality. Choose $\hat{\phi}$ also smooth and compactly supported in $\mathbb{R}^n \setminus 0$ such that $\hat{\phi}\hat{\psi} = \hat{\psi}$, and define $\tilde{\Delta}_j$ in the obvious way using ϕ. Then since $\tilde{\Delta}_j \Delta_j = \Delta_j$:

$$|\langle f, g \rangle| = \left| \sum_j \langle \Delta_j f, \tilde{\Delta}_j g \rangle \right| \leq |Sf|_p |\tilde{S}g|_{p'} \lesssim |Sf|_p |g|_{p'}.$$

Here we used the fact that $\tilde{S}$ satisfies the forward inequality. Thus, $|f|_p \lesssim |Sf|_p$.

$\square$

Remark 5.9. *(a) Theorem 5.7 uses that $\hat\psi$ is smooth. When $\hat\psi$ is replaced by the* characteristic function *of an annulus this theorem doesn't work when $n > 1$. See Remark 3.12.*

(b) Let $1 < p < \infty$. If S_t is the "square function" defined using the characteristic functions of a rectangular "tiling" of $\mathbb{R}^n$, then $|S_t f|_{L^p(\mathbb{R}^n} \sim |f|_{L^p}$.[24] *This doesn't work for annular tiling.*

(c) The decomposition of $\mathbb{R}^n \setminus \{0\}$ into dyadic rectangles R_j with disjoint interiors is referred to as a dyadic decomposition of $\mathbb{R}^n$.

(d) ***Dual LP estimate.*** *When $\hat\psi$ is real-valued the operators Δ_j are self-adjoint. If $\vec{T}f = (\Delta_j f)$, Theorem 5.7 shows $\vec{T} : L^p(\mathbb{R}^n, \mathbb{C}) \to L^p(\mathbb{R}^n, \ell^2)$. By duality for $1 < p < \infty$ we have $\vec{T}^* : L^{p'}(\mathbb{R}^n, \ell^2) \to L^{p'}(\mathbb{R}^n, \mathbb{C})$, so the dual inequality holds (use $\vec{T}^*(g_j) = \sum_j \Delta_j g_j$):*

$$|\sum_j \Delta_j g_j|_{p'} \lesssim ||(g_j)|_{\ell^2}|_{p'}\,. \tag{5.21}$$

Consider also the continuous version of the LP theorem. With ψ as in Theorem 5.7 define the continuous version of the square function

$$Sf(x) = \left(\int_0^\infty |(\psi_t * f)(x)|^2 \frac{dt}{t}\right)^{1/2}. \tag{5.22}$$

Theorem 5.10. *Let ψ be as in Theorem 5.7. For $1 < p < \infty$ and $f \in L^p(\mathbb{R}^n)$ we have*

$$|Sf|_p \sim |f|_p.$$

There is also a weak (1,1) inequality as in Theorem 5.7.

The proof is much like that of Theorem 5.7, but now define $S_L(f)(x) = ((\psi_t * f)(x)) \in H$ and replace $L(\mathbb{C}, \ell^2)$ by $L(\mathbb{C}, H)$, where $H := L^2(\mathbb{R}^+, \frac{dt}{t})$. For each x the operator $\vec{K}(x) \in L(\mathbb{C}, H)$ is given by

$$\vec{K}(x)a = \psi_t(x)a.$$

Remark 5.11. *Formally, we can write*

$$\begin{aligned}\int_0^\infty |\psi_t * f(x)|^2 \frac{dt}{t} &= \sum_j \int_{2^{-j}}^{2^{-j+1}} |\psi_t * f(x)|^2 \frac{dt}{t} \sim \sum_j \int_{2^{-j}}^{2^{-j+1}} |\psi_{2^{-j}} * f(x)|^2 \frac{dt}{t} \sim \\ &\sum_j |\psi_{2^{-j}} * f(x)|^2 \cdot 2^{-j} \cdot 2^j = \sum_j |\psi_{2^{-j}} * f(x)|^2.\end{aligned} \tag{5.23}$$

One can remove the second $\sim$ and do the t integration to get

$$\sum_j \int_{2^{-j}}^{2^{-j+1}} |\psi_{2^{-j}} * f(x)|^2 \frac{dt}{t} = \sum_j |\psi_{2^{-j}} * f(x)|^2 \cdot \log 2 = \sum_j |\psi_{2^{-j}} * f(x)|^2 \cdot \log 2.$$

In a sense this "explains" the definition of the continuous square function.

[24] The rectangular tiling of $\mathbb{R}^n \setminus 0$ is a disjoint union of sets $R_j = I_{j_1} \times \cdots \times I_{j_n}$, $j \in \mathbb{Z}^n$, where $I_k := [2^k, 2^{k+1}) \cup (-2^{k+1}, -2^k]$.

5.3 Proof of LP using the Mihlin multiplier theorem and Khinchine's inequality

We use the same notation as in the previous section.

Theorem 5.12 (LP)**.** *For $1 < p < \infty$ and $f \in L^p(\mathbb{R}^d)$ we have $|Sf|_p \sim |f|_p$, with constants $C(p,d)$.*

Proof. For $f \in \mathcal{S}$ let $S_N f = \left(\sum_{-N}^{N} |\Delta_j f|^2\right)^{1/2}$. Let let r_j be the sequence of Rademacher functions, or a fair coin-tossing sequence of iid random variables, or some "equivalent" sequence defined on some measure space Ω. For $p \in \Omega$ let $m_N(\xi) = \sum_{-N}^{N} r_j(p)\hat{\psi}(2^{-j}\xi)$ and note that

$$\begin{aligned} Sf(x) &= |(\Delta_j f)(x)|_{\ell^2} = |r_j(p)(\Delta_j f)(x)|_{\ell^2} = |a_j(x)|_{\ell^2}, \text{ where } a_j = \Delta_j f \\ |S^N f(x)|^p &= \left(\sum_{-N}^{N} |a_j(x)|^2\right)^{p/2} \leq \mathbb{E}\left|\sum_{-N}^{N} r_j(p)(\Delta_j f)(x)\right|^p, \end{aligned} \tag{5.24}$$

where we have applied Khinchine in the second line and $\mathbb{E}$ refers to p-space.[25] By Fubini we obtain

$$\int_{\mathbb{R}^d} |Sf(x)|^p dx \lesssim \limsup_{N\to\infty} \mathbb{E}\int_{\mathbb{R}^d} \left|\sum_{-N}^{N} r_j(p)(\Delta_j f)(x)\right|^p dx = \limsup_{N\to\infty} \mathbb{E}|(T_{m_N} f)|_{L^p}^p \lesssim |f|_{L^p}^p. \tag{5.25}$$

For the last inequality we used the Mihlin theorem, after observing that m_N satisfies the hypotheses of that theorem uniformly in N and $p \in \Omega$.[26]

This gives $|Sf|_p \lesssim |f|_p$. For the reverse direction use duality as before. Extend to $f \in L^p$ by density and Fatou. That is, choose $f_k \in \mathcal{S}$, $f_k \to f \in L^p$ and show $\lim_{k\to\infty} |Sf_k - Sf|_p = 0$ by Fatou. □

Remark 5.13. *The first line in* (5.24) *takes advantage of the fact that the square function is obtained from a sum of* squares. *The second line treats $P_j f(x)$ as a constant c_j for each x. One does not have*

$$\left(\sum_{-N}^{N} |\Delta_j f(x)|^2\right)^{p/2} \leq \left|\sum_{-N}^{N} (\Delta_j f)(x)\right|^p$$

but does have

$$\left(\sum_{-N}^{N} |\Delta_j f(x)|^2\right)^{p/2} \leq \mathbb{E}\left|\sum_{-N}^{N} r_j(p)(\Delta_j f)(x)\right|^p$$

by Khinchine because of the independence of the r_j. Finally we use that the family of multipliers parametrized by p given by $m_N(\xi)$ have corresponding kernels K_N that satisfy kernel estimates uniformly with respect to p.

Now we prove the corresponding theorem for Fourier series, Theorem 5.5.

Proof of Theorem 5.5. **1.** For $n \in \mathbb{Z}$ let

$$\psi_j(n) = \begin{cases} \chi_{2^j \leq |n| < 2^{j+1}}(n), & j \geq 1 \\ \chi_{0 \leq |n| < 2}(n), & j = 0 \end{cases},$$

[25] Probability comes in exactly here. We use the iid property and the fact that $r_j(p) = \pm 1$ with probability $1/2$ each.

[26] The multiplier is strong (p,p) just for $1 < p < \infty$.

and let Δ_j be the operator associated to this multiplier. Let Ω and r_j be as in the above proof, but now set

$$m_N(n) = \sum_{j=0}^{N} r_j(p)\psi_j(n) \text{ for each } p \in \Omega.$$

As before Khinchine implies for $x \in [0,1]$:

$$\begin{aligned} Sf(x) &= |(\Delta_j f)(x)|_{\ell^2} = |r_j(p)(\Delta_j f)(x)|_{\ell^2} \\ |S^N f(x)|^p &= \left(\sum_{-N}^{N} |\Delta_j f(x)|^2\right)^{p/2} \leq \mathbb{E}\left|\sum_{-N}^{N} r_j(p)(\Delta_j f)(x)\right|^p. \end{aligned} \tag{5.26}$$

By Fubini we obtain

$$\int_0^1 |Sf(x)|^p dx \lesssim \limsup_{N\to\infty} \mathbb{E} \int_0^1 \left|\sum_{-N}^{N} r_j(p)(\Delta_j f)(x)\right|^p dx = \limsup_{N\to\infty} \mathbb{E}|(T_{m_N} f)|^p_{L^p} \lesssim |f|^p_{L^p}, \tag{5.27}$$

provided $m_N(n)$ is an L^p multiplier uniformly with respect to N and $p \in \Omega$. We discuss this key point next.

2. The property we need follows from the *transference* lemma, Lemma 5.14, which is of great interest in itself. Assuming the lemma for now, for any N and $p \in \Omega$, take $b(n) = m_N(n)$, and note that one can easily find a smooth multiplier $b_N(\xi)$ such that $m_N(n) = b_N(n)$, where the multipliers b_N satisfy the Mihlin hypotheses uniformly with respect to N and $p \in \Omega$. Here we just use the fact that the dyadic sets

$$|n| < 2,\ 2^j \leq |n| < 2^{j+1},\ \ j \geq 1$$

are well separated. Apply Lemma 5.14 to finish. □

Lemma 5.14 (Transference of multipliers)**.** *Let $b(\xi)$ be a continuous function which defines an L^p multiplier on $\mathbb{R}^d$ for some $1 < p < \infty$.*[27] *Then the sequence $(b(n))$, $n \in \mathbb{Z}^d$ is an L^p multiplier on $\mathbb{T}^d$ and the corresponding operators satisfy*

$$|T_{(b(n))}|_{p\to p} \leq |T_{b(\xi)}|_{p\to p}.$$

Proof. **1.** Let T and S be the operators corresponding to $b(\xi)$ and $b(m)$ respectively. Suppose P, Q are trigonometric polynomials on $\mathbb{T}^d$ and let $L_\epsilon(x) = e^{-\pi\epsilon|x|^2}$. We claim that for $\alpha, \beta > 0$ such that $\alpha + \beta = 1$, we have

$$\lim_{\epsilon\to 0} \epsilon^{n/2} \int_{\mathbb{R}^d} T(PL_{\epsilon\alpha})(x)\overline{Q(x)L_{\epsilon\beta}(x)}dx = \int_{\mathbb{T}^d} S(P)(x)\overline{Q(x)}dx. \tag{5.28}$$

By linearity it suffices to prove this for $P(x) = e^{2\pi imx}$, $Q(x) = e^{2\pi ikx}$. By Parseval

$$\int_{\mathbb{T}^d} S(P)(x)\overline{Q(x)}dx = \sum_r b(r)\hat{P}(r)\overline{\hat{Q}(r)} = \begin{cases} b(m),\ m = k \\ 0,\ m \neq k \end{cases}.$$

On the other hand the integral on the left of (5.28) is

$$\epsilon^{n/2} \int b(\xi)\widehat{PL_{\epsilon\alpha}}(\xi)\overline{\widehat{QL_{\epsilon\beta}}(\xi)}d\xi = (\epsilon\alpha\beta)^{-n/2} \int b(\xi)e^{-\pi\frac{|\xi-m|^2}{\epsilon\alpha}} e^{-\pi\frac{|\xi-k|^2}{\epsilon\beta}}d\xi.$$

[27]The result is also true for $p = 1$ by a slightly different proof.

When $m = k$ this equals $(\epsilon\alpha\beta)^{-n/2}\int b(\xi)e^{-\pi\frac{|\xi-m|^2}{\epsilon\alpha\beta}}d\xi$, which approaches $b(m)$ as $\epsilon \to 0$. If $m \neq k$ then for every ξ we have either $|\xi - m| \geq \frac{1}{2}$ or $|\xi - k| \geq \frac{1}{2}$. Use this to break up the integral into two terms, pull out $|b|_\infty$, and integrate out the remaining exponentials. Each term is seen to approach 0 as $\epsilon \to 0$.

2. Show $S : L^p(\mathbb{T}^d) \to L^p(\mathbb{T}^d)$. Using (5.28) we have for P, Q trigonometric polynomials:

$$\begin{aligned} \left|\int_{\mathbb{T}^d} S(P)(x)\overline{Q(x)}dx\right| &\lesssim |T|_{p\to p}\limsup_{\epsilon\to 0}\epsilon^{n/2}|PL_{\epsilon/p}|_{L^p(\mathbb{R}^d)}|QL_{\epsilon/p'}|_{L^{p'}(\mathbb{R}^d)} = \\ &|T|_{p\to p}\left(\int_{\mathbb{T}^d}|P(x)|^p dx\right)^{1/p}\left(\int_{\mathbb{T}^d}|Q(x)|^{p'}dx\right)^{1/p'}, \end{aligned} \tag{5.29}$$

provided that for any continuous periodic g on $\mathbb{R}^d$ we have

$$\lim_{\epsilon\to 0}\epsilon^{d/2}\int_{\mathbb{R}^d} g(x)e^{-\pi\epsilon|x|^2}dx = \int_{\mathbb{T}^d} g(x)dx. \tag{5.30}$$

3. Proof of (5.30). Use the the periodicity of g and the Poisson summation formula, Prop. 14.4, on the exponential to write the left side of (5.28) as

$$\begin{aligned} \epsilon^{n/2}\sum_{k\in\mathbb{Z}^d}\int_{\mathbb{T}^d} g(x-k)e^{-\pi\epsilon|x-k|^2}dx &= \int_{\mathbb{T}^d} g(x)\epsilon^{n/2}\sum_{k\in\mathbb{Z}^d} e^{-\pi\epsilon|x-k|^2}dx = \\ \int_{\mathbb{T}^d} g(x)\sum_{k\in\mathbb{Z}^d} e^{-\pi|k|^2/\epsilon}e^{2\pi ikx}dx &= \int_{\mathbb{T}^d} g(x)dx + A_\epsilon, \end{aligned} \tag{5.31}$$

where $|A_\epsilon| \leq |g|_\infty \sum_{|k|\geq 1} e^{-\pi|k|^2/\epsilon} \to 0$ as $\epsilon \to 0$.

□

6 Some applications of LP theory

6.1 A simple multiplier result

Proposition 6.1. *Suppose $\{m_n\}$ is bounded and constant on each region $\{n : 2^k \leq |n| < 2^k\}$. Then (m_n) is an $L^p(\mathbb{T})$ Fourier multiplier for all $p \in (1, \infty)$.*

Proof. With $f(x) = \sum \hat{f}(n)e^{2\pi inx}$, let $Tf(x) = \sum m_n\hat{f}_n e^{i2\pi nx}$. Let m_k^* denote the constant value of m_n on $2^k \leq |n| < 2^k$. Then for $\Delta_k f$ as in (5.9), we have the pointwise estimate

$$\Delta_k(Tf) = m_k^*\Delta_k f \Rightarrow \sum_k |\Delta_k(Tf)|^2 \leq \sup_k |m_k^*| \sum_k |\Delta_k f|^2 \Rightarrow S(Tf) \leq \sup_n |m_n| \cdot Sf.$$

Now Theorem 5.5 implies the result.

□

Remark 6.2. *This result can also be proved using the Mihlin theorem and transference. See the proof of Theorem 5.5 in section 5.3.*

6.2 Characterization of function spaces: Hölder, Sobolev, Besov, Triebel, Hardy

We let $\psi_{2^{-j}}$ and Δ_j be as in the proof of Theorem 5.12.

Definition 6.3. *For $0 < \alpha < 1$ the Hölder space $C^\alpha(\mathbb{R}^d)$ consists of f such that*

$$|f|_\alpha = |f|_\infty + [f]_\alpha, \ \text{where } [f]_\alpha := \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x-y|^\alpha}.$$

Theorem 6.4 (Hölder spaces)**.** *Let $|f| \le 1$. Then $f \in C^\alpha(\mathbb{R}^d)$ if and only if*

$$A := \sup_{j \in \mathbb{Z}} 2^{j\alpha} |\Delta_j f|_\infty < \infty.$$

In that case $A \sim [f]_\alpha := \sup_{x\neq y} \frac{|f(x)-f(y)|}{|x-y|^\alpha}$.

Proof. **1.** ($\Rightarrow$) With $\psi_{2^{-j}} := \phi_j$, we have

$$|\Delta_j f(x)| \le \int |f(x-y) - f(x)||\phi_j(y)|dy \le \int [f]_\alpha |y|^\alpha |\phi_j(y)| dy = 2^{-j\alpha}[f]_\alpha \int |y|^\alpha |\psi(y)| dy,$$

where the last equality follows by the change of variable $z = 2^j y$. This gives $A \lesssim [f]_\alpha$.

2. ($\Leftarrow$) We let $g_l(x) = \sum_{-l}^{l} \Delta_j f(x)$ and show for all y that

$$\sup_l |g_l(x-y) - g_l(x)| \lesssim A|y|^\alpha. \tag{6.1}$$

We have (use Bernstein (14.3) $|\nabla \Delta_j f|_\infty \lesssim 2^j |\Delta_j f|_\infty$):

$$\begin{aligned} |g_l(x-y) - g_l(x)| \le \sum_{2^j \le |y|^{-1}} |\Delta_j f(x-y) - \Delta_j f(x)| + \sum_{|y|^{-1} < 2^j} 2|\Delta_j f|_\infty \lesssim \\ \sum_{2^j \le |y|^{-1}} A 2^{j(1-\alpha)}|y| + \sum_{|y|^{-1}<2^j} A 2^{-j\alpha} \lesssim A|y|^\alpha. \end{aligned} \tag{6.2}$$

3. Now (6.1) implies $|g_l(x) - g_l(0)| \lesssim |x|^\alpha$ uniformly wrt l. Also Arzela-Ascoli implies a subsequence $g_l - g_l(0)$ converges to some g uniformly on compacts, where $[g] \lesssim CA$ by (6.1). With the uniform growth of order $|x|^\alpha$, this implies $g_l - g_l(0) \to g$ in $\mathcal{S}'$, hence in $\mathcal{S}'_P$, the space of tempered distributions mod polynomials (see #12 of section 14). Now $g_l \to f$ in $\mathcal{S}'_P$ and $g_l - g_l(0) = g_l$ in $\mathcal{S}'_P$. So $f = g$ in $\mathcal{S}'_P$. Since $g(0) = 0$ we have

$$|(f-g)(x)| \le |f|_\infty + |g(x) - g(0)| \lesssim 1 + A|x|^\alpha.$$

Since $\alpha < 1$, we must have $f = g$ equals a constant. So $[f]_\alpha \lesssim A$. $\square$

Let ψ as in (5.11) give a dyadic partition of unity $\sum_{j\in\mathbb{Z}} \hat\psi(2^{-j}\xi) = 1$ for $\xi \neq 0$. Set

$$\hat\Phi(\xi) = \begin{cases} \sum_{j \le 0} \hat\psi(2^{-j}\xi), \ \xi \neq 0 \\ 1, \ \xi = 0 \end{cases}, \tag{6.3}$$

and let $\mathcal{P}f = f * \Phi$. Observe that $\hat\Phi$ is a test function.

Definition 6.5 (Inhomogeneous Sobolev spaces). *(a) Let $s \in \mathbb{R}$. A tempered distribution $u \in \mathcal{S}'$ is an element of the Sobolev space $W^{s,p}(\mathbb{R}^n)$ if*

$$|f|_{s,p} := |\mathcal{F}^{-1}(\langle\xi\rangle^s \hat{f})|_{L^p} < \infty.$$

(b)Set $f_s = \mathcal{F}^{-1}(\langle\xi\rangle^s \hat{f})$, so $|f|_{s,p} = |f_s|_p$.

Theorem 6.6. *(a) If a tempered distribution $f \in \mathcal{S}'$ is an element of the Sobolev space $W^{s,p}(\mathbb{R}^n)$ then*

$$|\mathcal{P}f|_p + \left||(2^{js}\Delta_j f)|_{\ell^2}\right|_{L^p(\mathbb{R}^n)} \lesssim |f|_{s,p}. \tag{6.4}$$

(b) If a tempered distribution f satisfies $|\mathcal{P}f|_p + \left||(2^{js}\Delta_j f)|_{\ell^2}\right|_{L^p(\mathbb{R}^n)} < \infty$, then $f \in W^{s,p}$ and

$$|f|_{s,p} \lesssim |\mathcal{P}f|_p + \left||(2^{js}\Delta_j f)|_{\ell^2}\right|_{L^p(\mathbb{R}^n)}. \tag{6.5}$$

Proof. **1. (b)** Write $f_s = (\hat{\Phi}\hat{f}_s)^\vee + ((1-\hat{\Phi})\hat{f}_s)^\vee$; we'll show each term has an L^p norm dominated by the right side of (6.5). Choose a test function $\eta_0(\xi)$ equal to one on supp $\hat{\Phi}$. Then the test function $\langle\xi\rangle^s\eta_0(\xi)$ is clearly an L^p multiplier, so

$$|(\hat{\Phi}\hat{f}_s)^\vee|_p = |(\langle\xi\rangle^s\eta_0(\xi)\hat{\Phi}\hat{f})^\vee|_p \lesssim |\mathcal{P}f|_p. \tag{6.6}$$

Next choose a C^∞ function η_∞ equal to zero near 0 and equal to one on a neighborhood of supp $1-\hat{\Phi}$. Mihlin implies $m(\xi) := \langle\xi\rangle^s|\xi|^{-s}\eta_\infty(\xi)$ is an L^p multiplier, so

$$|((1-\hat{\Phi})\hat{f}_s)^\vee|_p = |(m(\xi)|\xi|^s(1-\hat{\Phi})\hat{f})^\vee|_p \lesssim |g_s|_p, \text{ where } \widehat{g_s} := |\xi|^s(1-\hat{\Phi})\hat{f}. \tag{6.7}$$

2. Choose a real-valued test function $\hat{\zeta}$ supported away from 0 and equal to one on supp $\hat{\psi}$, so $\hat{\zeta}\hat{\psi} = \hat{\psi}$. Define $\hat{\theta}(\xi) = |\xi|^s\hat{\zeta}(\xi)$ and set $\Delta_j^\theta f = f * \theta_{2^{-j}}$. We have

$$\widehat{g_s} = \sum_{j\geq 1}|\xi|^s\hat{\psi}(2^{-j}\xi)\hat{\zeta}(2^{-j}\xi)\hat{f}(\xi) = \sum_{j\geq 1}2^{js}\hat{\psi}(2^{-j}\xi)\hat{\theta}(2^{-j}\xi)\hat{f}(\xi) \tag{6.8}$$

This gives $g_s = \sum_{j\geq 1}\Delta_j^\theta(2^{js}\Delta_j f)$, so the dual LP estimate (5.21) implies

$$|g_s|_p \lesssim ||(h_j)|_{\ell^2}|_{L^p(\mathbb{R}^n)}, \text{ where } h_j = \begin{cases} 2^{js}\Delta_j f, \ j \geq 1 \\ 0, \ j < 1 \end{cases}. \tag{6.9}$$

3. (a) Now assume $|f_s|_p < \infty$, where $\hat{f}_s = \langle\xi\rangle^s\hat{f}$. We have

$$\widehat{\mathcal{P}f} = \hat{\Phi}\hat{f} = \langle\xi\rangle^{-s}\hat{\Phi}\hat{f}_s \Rightarrow |\mathcal{P}f|_p \lesssim |f_s|_p = |f|_{s,p}. \tag{6.10}$$

Set $\hat{\sigma} = |\xi|^{-s}\hat{\psi}(\xi)$ and $\Delta_j^\sigma g := g * \sigma_{2^{-j}}$. Since $\hat{\Phi}(\xi)\hat{\sigma}(2^{-j}\xi) = 0$ for $j \geq 2$, we have

$$2^{js}\hat{\psi}(2^{-j}\xi)\hat{f} = \hat{\sigma}(2^{-j}\xi)|\xi|^s\hat{f} = \hat{\sigma}(2^{-j}\xi)|\xi|^s(1-\hat{\Phi}(\xi))\hat{f} \Rightarrow 2^{js}\Delta_j f = \Delta_j^\sigma g_s. \tag{6.11}$$

Thus, Theorem 5.7 implies[28]

$$\left|\left(\sum_{j\geq 2}2^{js}\Delta_j f|^2\right)^{1/2}\right|_p \lesssim |g_s|_p \lesssim |f_s|_p = |f|_{s,p}.$$

Since test functions are L^p multipliers, we have for $j = 1$: $|2^s\Delta_1 f|_p \lesssim |g_s|_p \lesssim |f|_{s,p}$.

∎

[28]For the second inequality we use the Mihlin theorem.

Remark 6.7. *(a)* Homogeneous *Sobolev and Hölder/Lipschitz spaces can be characterized using dyadic partitions of unity* $\sum_{j\in\mathbb{Z}} \widehat{\psi_{2^{-j}}}(\xi) = 1$ *for* $\xi \neq 0$. *The nonhomogeneous spaces can be characterized using* $\hat{\Phi}(\xi) + \sum_{j\geq 1} \widehat{\psi_{2^{-j}}}(\xi) = 1$ *for all* ξ. *Note that the condition* $\sup_{j\in\mathbb{Z}} 2^{j\alpha}|\Delta_j f|_\infty < \infty$ *in Theorem 6.4, which is satisfied by any polynomial, characterizes* $[f]_\alpha$ *finite, not* $|f|_{C^\alpha}$ *finite. For the latter use the analogous condition on the* Φ, ψ_j, $j \geq 1$. *This is the difference between Theorem 6.4 and the characterization of Hölder spaces in, e.g, [AG]. (Alinhac-Gerard). When* f *is a polynomial of degree* ≥ 1, $|\mathcal{P}f|_\infty = \infty$, *so* $|\mathcal{P}f|_\infty < \infty$ *implies* f *is not a polynomial of degree* ≥ 1. *On the other hand*

$$\sup_{j\geq 1} 2^{j\alpha}|\Delta_j f|_\infty < \infty \text{ and } |\mathcal{P}f|_\infty < \infty$$

implies $|f|_\infty$ *finite.*

(b) The proof of Theorem 6.6 used the estimate $|Sf|_p \leq |f|_p$ *and its dual estimate* (5.21), *the Mihlin theorem, the simple fact that test functions are* L^p *multipliers for* $1 < p < \infty$, *and support properties of the* $\hat{\psi}(2^{-j}\xi)$.

Definition 6.8. ***Inhomogeneous Besov-Lipschitz and Triebel-Lizorkin norms.*** *Let* $\alpha \in \mathbb{R}$ *and* $0 < p, q \leq \infty$. *For* $f \in \mathcal{S}'(\mathbb{R}^n)$ *set*

$$\begin{aligned} |f|_{B_p^{\alpha,q}} &= |\mathcal{P}f|_p + \left|\left(\left|2^{j\alpha}|\Delta_j f|\right|_p\right)_{j=1}^\infty\right|_{\ell^q} \\ |f|_{F_p^{\alpha,q}} &= |\mathcal{P}f|_p + \left|\left|\left(2^{j\alpha}|\Delta_j f|\right)_{j=1}^\infty\right|_{\ell^q}\right|_p. \end{aligned} \tag{6.12}$$

Homogeneous Besov-Lipschitz and Triebel-Lizorkin norms. *Let* $\alpha \in \mathbb{R}$ *and* $0 < p, q \leq \infty$. *For* $f \in \mathcal{S}'(\mathbb{R}^n)/\mathcal{P}$ *set*

$$\begin{aligned} |f|_{\dot{B}_p^{\alpha,q}} &= \left|\left(\left|2^{j\alpha}|\Delta_j f|\right|_p\right)_{j\in\mathbb{Z}}\right|_{\ell^q} \\ |f|_{\dot{F}_p^{\alpha,q}} &= \left|\left|\left(2^{j\alpha}|\Delta_j f|\right)_{j\in\mathbb{Z}}\right|_{\ell^q}\right|_p. \end{aligned} \tag{6.13}$$

Remark 6.9. *(a) The above* B *and* F *norms differ only in the order of the* L^p *and* ℓ^q *norms occurring within them.*

(b) We have shown that $W^{s,p} = F_p^{s,2}$ *for* $1 < p < \infty$ *and* $C^\alpha = B_\infty^{\alpha,\infty}$ *for* $0 < \alpha < 1$.

(c) On p. 93 of [G2] it is said that $\dot{F}_1^{0,2} = H^1$. *This is not correct; for example,* $1 \notin H^1$. *Instead we have* $H^1 \subset \dot{F}_1^{0,2}$ *and if* $f \in \dot{F}_1^{0,2}$, *then there exists a unique polynomial* Q *such that* $f - Q \in H^1$.

There are several equivalent characterizations of Hardy spaces H^p, $0 < p \leq 1$. We are mainly interested in H^1; see Chapter 2 of [G2] for equivalence proofs. The reader can regard part (a) of the next proposition as a definition.

Proposition 6.10 (The Hardy space H^1)**.**

(a) The space $H^1(\mathbb{R}^d)$ *is the proper subspace of* L^1 *consisting of functions such that*

$$|f|_{H^1} := |f|_{L^1} + |Sf|_{L^1} \sim |f|_{L^1} + \sum_{j=1}^d |R_j f|_{L^1} \sim \left|\sup_{t>0} |P_t * f|\right|_{L^1} < \infty.$$

Here the R_j *are Riesz transforms if* $d \geq 2$, *the Hilbert transform if* $d = 1$, *and* P *is the Poisson kernel.*[29]

(b) Elements of H^1 *satisfy* $\int f dx = 0$.[30]

[29] As usual $P_t(x) = t^{-n}P(x/t)$.

[30] This follows from $R_j f \in L^1 \Rightarrow \widehat{R_j f}$ is continuous at 0.

6.3 Singular integrals on C^α.

Theorem 6.11. *Let K be a CZ kernel as in Definition 3.1 and let $0 < \alpha < 1$. For any $f \in L^2 \cap C^\alpha(\mathbb{R}^d)$ we have $Tf \in C^\alpha$ and*

$$[Tf]_\alpha \leq C(\alpha, d)B[f]_\alpha, \text{ for } B \text{ as in Defn. 3.1.}$$

Proof. First note

$$|Tf|_{L^\infty} \lesssim B([f]_\alpha + |f|_{L^2}).$$

(Proof: Consider $\lim_{\epsilon\to 0}\int_{R>|y|>\epsilon} K(y)[f(x-y) - f(x)]dy + \int_{|y|>R} K(y)f(x-y)dy$.)

So enough to show

$$\sup_j 2^{j\alpha}|\Delta_j Tf|_{L^\infty} \leq CB[f]_\alpha. \tag{6.14}$$

Let $\tilde{\Delta}_j f$ be multiplier associated to $\hat{\phi}(2^{-j}\xi)$, where $\hat{\phi} \in C_0^\infty(\mathbb{R}^d \setminus 0)$ and $\hat{\phi}\hat{\psi} = \hat{\psi}$. So

$$|\Delta_j Tf|_\infty = |\tilde{\Delta}_j T\Delta_j f|_\infty| \leq \|\tilde{\Delta}_j T\|_{\infty\to\infty}|\Delta_j f|_\infty \leq C[f]_\alpha 2^{-j\alpha}.$$

The kernel of $\tilde{\Delta}_j T$ is $2^{jd}\phi(2^j\cdot) * K$, whose L^1 norm is same as

$$|T_j\phi|_{L^1} := |\phi * 2^{-jd}K(2^{-j}\cdot)|_{L^1} \leq CB.$$

For the latter inequality we used the fact that $2^{-jd}K(2^{-j}\cdot)$ satisfies the conditions of Defn. 3.1 with the same B uniformly in j, and that if T is an SIO associated to such a kernel, then

$$\eta \in \mathcal{S}, \int \eta = 0 \Rightarrow |T\eta|_{L^1} \leq C(\eta)B \tag{6.15}$$

We sketch the proof of (6.15) below.

□

Sketch of proof of (6.15). **1.** Write $\eta = \sum_{l=1}^\infty c_l a_l$ where the a_l are *"atoms"* which satisfy

$$\int a_l = 0, \operatorname{supp} a_l \subset B(0, l),\ |a_l|_{L^\infty} \leq l^{-d}$$

and $\sum_l |c_l| \leq C(\eta)$.

2. For each a_l we have $|Ta_l|_1 \lesssim B$. To see this we estimate

$$\begin{aligned} &\int_{|x|\leq 2l} |Ta_l(x)|dx \lesssim |Ta_l|_2 l^{d/2} \lesssim B|a_l|_2 l^{d/2} \lesssim B \\ &\int_{|x|>2l} |Ta_l(x)|dx \leq \int\int_{|x|>2|y|} |K(x-y) - K(x)|dx|a_l(y)|dy \leq B|a_l|_1 \lesssim B. \end{aligned} \tag{6.16}$$

3. Now use step **2** to estimate

$$\sum_l |Tc_l a_l|_1 \leq \sum_l |c_l|CB \leq CB.$$

Note that $S_N := \sum_{l=1}^N c_l a_l \to \eta$ in L^1 (in fact all L^p, $1 \le p \le \infty$). Use the fact that T is weak $(1,1)$ to show[31]

$$T\eta = \sum_l T(c_l a_l) \text{ a.e. } .$$

□

Corollary 6.12 (Schauder estimate)**.** *Let $f \in C_0^{2,\alpha}$, $0 < \alpha < 1$. Then*

$$[f_{x_i x_j}]_\alpha \le C(\alpha, d)[\Delta f]_\alpha.$$

Proof. Apply Theorem 6.11 to R_{ij}, the double Riesz transform, using $f_{x_i x_j} = R_{ij}(\Delta f)$. □

There is an L^p analogue as well.

6.4 Sobolev embedding of $\dot{H}^s$ for $0 \le s < d/2$.

First some preparation:

Definition 6.13. *Let $s > -\frac{d}{2}$. The homogeneous Sobolev space $\dot{H}^s(\mathbb{R}^d)$ is the completion of $\mathcal{S}$ in the norm*

$$|u|_{\dot{H}^s} = ||\xi|^s \hat{u}|_2.$$

We have

$$|u|_{\dot{H}^s}^2 \sim \sum_{j\in\mathbb{Z}} |\Delta_j u|_{\dot{H}^s}^2. \tag{6.17}$$

Generalized triangle inequality. For any $1 \le q \le p \le \infty$ we have (where $(f_j) := (f_j)_{j\in\mathbb{Z}}$)

$$||(f_j)|_{\ell^q}|_{L^p(\mu)} \le \left|(|f_j|_{L^p(\mu)})\right|_{\ell^q}. \tag{6.18}$$

An important case is where $q = 2 \le p$. Then

$$\left|\left(\sum_j |f_j|^2\right)^{1/2}\right|_{L^p(\mu)} \le \left(\sum_j |f_j|_{L^p(\mu)}^2\right)^{1/2}.$$

Applied to the LP square function this gives

$$|Sf|_p^2 \le \sum_j |\Delta_j f|_{L^p}^2. \tag{6.19}$$

This is used in proving:

[31] For $\delta > 0$ write

$$\left|\left\{T\eta - \sum_1^\infty T(c_l a_l) > \delta\right\}\right| \le \left|\left\{T\eta - \sum_1^N T(c_l a_l) > \delta/2\right\}\right| + \left|\left\{\sum_{N+1}^\infty T(c_l a_l) > \delta/2\right\}\right|.$$

Now use $S_N \to \eta$ in L^1 and $\sum |c_l| < \infty$.

Proposition 6.14. *For $0 \le s < d/2$ with*

$$\frac{1}{2} - \frac{1}{p} = \frac{s}{d}, \tag{6.20}$$

one has

$$|f|_{L^p} \le C(s,d)|f|_{\dot{H}^s} \tag{6.21}$$

for $f \in \mathcal{S}$.

Remark 6.15. *The norms in the inequality* (6.21) *each "scale" in an obvious way, and that leads to the restriction* (6.20)*. Explicitly, with $f_\lambda(x) = f(\lambda x)$,*

$$|f_\lambda|_{\dot{H}^s} = \lambda^{s-\frac{d}{2}}|f|_{\dot{H}^s} \text{ and } |f_\lambda|_{L^p} = \lambda^{-\frac{d}{p}}|f|_{L^p}.$$

Proof. If $\hat{f}$ has support on $|\xi| \sim R$, then Bernstein (14.3) gives

$$|f|_p \lesssim R^{d(\frac{1}{2}-\frac{1}{p})}|f|_2 \sim |f|_{\dot{H}^s}.$$

Thus, if $f = \sum_j \Delta_j f$ we have

$$|\Delta_j f|_p \lesssim |\Delta_j f|_{\dot{H}^s} \Rightarrow \sum_j |\Delta_j f|_p^2 \lesssim \sum_j |\Delta_j f|_{\dot{H}^s}^2 \sim |f|_{\dot{H}^s}^2.$$

The LP theorem gives $|f|_p^2 \le |Sf|_p^2$ and $|Sf|_p^2 \le \sum_j |\Delta_j f|_p^2$ by (6.19), so (6.21) follows. □

Compare the following easier result involving the nonhomogeneous Sobolev space $H^s = W^{s,2}$.

Proposition 6.16. *Suppose $s > d/2$ and $2 \le p \le \infty$. Then*

$$|f|_{L^p} \le C(s)|f|_{H^s}. \tag{6.22}$$

Proof. The case $p = 2$ is trivial, the case $p = \infty$ is easily done with the Fourier transform: write

$$u(x) = \int e^{ix\xi}\hat{u}(\xi)d\xi = \int e^{ix\xi}\hat{u}(\xi)\langle\xi\rangle^s\langle\xi\rangle^{-s}d\xi,$$

and apply Cauchy-Schwarz. Now interpolate. □

7 BMO and the Hardy space H^1

Definition 7.1. *(a). For $f \in L^1_{loc}(\mathbb{R}^d)$ let $f_Q = Av_Q f = \frac{1}{|Q|}\int_Q f(x)dx$ and define the* sharp maximal function

$$f^\sharp(x) = \sup_{\textit{all cubes } Q \ni x} Av_Q|f - f_Q|.$$

(b) $f \in BMO$ if $|f|_{BMO} := |f^\sharp|_{L^\infty} < \infty = \sup_Q \frac{1}{|Q|}\int_Q |f(x) - f_Q|dx$.

Clearly $L^\infty \subset BMO$ ($|f|_{BMO} \le 2|f|_{L^\infty}$), but BMO is strictly larger.

Proposition 7.2. *Let $f^{\sharp\sharp}(x) := \sup_{Q\ni x} \inf_c Av_Q|f-c|$. Then*

$$f^{\sharp\sharp}(x) \le f^{\sharp}(x) \le 2f^{\sharp\sharp}(x).$$

Proof. The first inequality is clear. For the second write $|f - f_Q| \le |f-c| + |c - f_Q|$. ◻

Example 7.3. *For example, $\log|x| \in BMO$. Use $\log|rx| = \log|x| + \log r$ to reduce to considering balls B of measure 1. Look at two cases. If $B \subset B(0,2)$, take $c_B = 0$. If $B \subset \{1 \le r < |x| < r+1\}$ take $c_B = \log r$. Here we use that $f \in BMO$ if there exists A such that for each ball B there exists c_B such that*[32]

$$\frac{1}{|B|}\int_B |f - c_B| \le A.$$

Cutoffs of BMO functions are not necessarily BMO. Consider $h(x) = \chi_{x>0}\log(1/x)$, and observe that $\frac{1}{2\epsilon}\int_{-\epsilon}^{\epsilon} |h - h_{B(0,\epsilon)}|dx$ is unbounded as $\epsilon \to 0$. The cutoff introduces violent "oscillation" at 0.

The following lemma is used in the proof of the John-Nirenberg inequality.

Lemma 7.4. *Suppose the cubes Q_j are nested and the side length of Q_j is 2^{-1} sidel length Q_{j-1}. Then*

$$|f_{Q_n} - f_{Q_0}| \le 2^d n |f|_{BMO}. \tag{7.1}$$

Proof. $|f_{Q_1} - f_{Q_0}| \le \frac{1}{|Q_1|}\int_{Q_1} |f - f_{Q_0}| \le \frac{1}{|Q_1|}\int_{Q_0} |f - f_{Q_0}| \le 2^d |f|_{BMO}$. Now use the triangle inequality.

◻

7.1 John-Nirenberg inequality

Theorem 7.5 (John-Nirenberg inequality)**.** *Let $f \in BMO(\mathbb{R}^n)$. Then there are absolute constants C, c such that for any cube Q one has*

$$|\{x \in Q : |f - f_Q| > \lambda\}| \le C|Q| \exp(-c\lambda/|f|_{BMO}).$$

Proof. **1.** The inequality is not altered if f and λ are multiplied by the same constant, so assume $|f|_{BMO} = 1$ and fix a cube Q and a $b > 1$ to be chosen. Recall that if $B \subset \mathbb{R}^n$ is a ball and $m \in \mathbb{N}$

$$|Av_B f - Av_{2^m B}| \le 2^n m |f|_{BMO}. \tag{7.2}$$

Set $Q^0 = Q$ and apply the CZ decomposition to the function $f - Av_Q f$ on Q at height b (so Q not selected since $1 = |f|_{BMO} < b$) to obtain a countable collection of cubes $\{Q_j^{(1)}\}$ of first generation. Fix a $Q_j^{(1)}$ of first generation and perform CZ to $f - Av_{Q_j^{(1)}} f$ at height b on that cube. Since $|f|_{BMO} = 1 < b$, $Q_j^{(1)}$ does not satisfy the new selection criterion. Do this for all cubes of first generation to obtain the cubes $Q_j^{(2)}$ of second generation. Repeat the process indefinitely to get a doubly indexed collection $\{Q_j^{(k)}\}$.[33] These cubes satisfy:

(A-k) The interior of every $Q_j^{(k)}$ is contained in a unique $Q_{j'}^{(k-1)}$.

[32]See Prop. 7.2.

[33]In this proof we apply the full CZ decomposition countably many times. Each time the ambient (or original) space is a fixed cube, rather than all of $\mathbb{R}^n$. The meaning of "$n-$th generation" here is not the same as in the original proof of the CZ decomposition.

(B-k) $b < |Q_j^{(k)}|^{-1} \int_{Q_j^{(k)}} |f - f_{Q_{j'}^{(k-1)}}|dx \le 2^n b$.

(C-k) $|f_{Q_j^{(k)}} - f_{Q_{j'}^{(k-1)}}| \le 2^n b$.

(D-k) $\sum_j |Q_j^{(k)}| \le \frac{1}{b} \sum_{j'} |Q_{j'}^{(k-1)}|$.

(E-k) $|f - f_{Q_{j'}^{(k-1)}}| \le b$ a.e. on $Q_{j'}^{(k-1)} \setminus \cup_j Q_j^{(k)}$.

The second inequality in (B-k) holds since $Q_{j'}^{(k-1)}$ does not satisfy the selection criterion for cubes of generation k. (C-k) follows from (7.2), and the other properties, except (D-k), follow from the CZ decomposition (or its proof).

2. Proof of (D-k). We have

$$\begin{aligned}\sum_j |Q_j^{(k)}| &\le \frac{1}{b} \sum_j \int_{Q_j^{(k)}} |f - f_{Q_{j'}^{(k-1)}}|dx = \frac{1}{b} \sum_{j'} \sum_{j \text{ corresponding to } j'} \int_{Q_j^{(k)}} |f - f_{Q_{j'}^{(k-1)}}|dx \le \\ &\frac{1}{b} \sum_{j'} \int_{Q_{j'}^{(k-1)}} |f - f_{Q_{j'}^{(k-1)}}|dx \le \frac{1}{b} \sum_{j'} |Q_{j'}^{(k-1)}| \|f\|_{BMO} = \frac{1}{b} \sum_{j'} |Q_{j'}^{(k-1)}|.\end{aligned} \tag{7.3}$$

3. We claim

$$\begin{aligned}&(a) \sum_j |Q_j^{(k)}| \le b^{-k} |Q^{(0)}| \\ &(b) \{x \in Q : |f - Av_{Q^{(0)}} f| > 2^n k b\} \subset \cup_j Q_j^{(k)} \ (a.e.)\end{aligned} \tag{7.4}$$

Part (a) is proved by iterating (D-k). To prove (b) we use induction to show for all $l \ge 1$:

$$|f - f_{Q^{(0)}}| \le 2^n l b \text{ a.e. on } Q^{(0)} \setminus \cup_s Q_s^{(l)}. \tag{7.5}$$

Clearly this holds for $l = 1$. Assume (7.5) for a fixed l. Since for each r

$$\begin{aligned}&|f - f_{Q_r^{(l)}}| \le b \text{ a.e. on } Q_r^{(l)} \setminus \cup_s Q_s^{(l+1)} \\ &|f_{Q_r^{(l)}} - f_{Q^{(0)}}| \le 2^n l b,\end{aligned} \tag{7.6}$$

we obtain

$$|f - f_{Q^{(0)}}| \le 2^n (l+1) b \text{ a.e. on } Q_r^{(l)} \setminus \cup_s Q_s^{(l+1)}. \tag{7.7}$$

With (7.5) we see that the estimate (7.7) holds a.e. on $Q^{(0)} \setminus \cup_s Q_s^{(l+1)}$.

4. Fix $\lambda > 0$. If $2^n k b < \lambda \le 2^n (k+1) b$ for some $k \ge 0$ then

$$\begin{aligned}|\{x \in Q : |f - Av_Q f| > \lambda\}| &\le |\{x \in Q : |f - Av_Q f| > 2^n k b\}| \le \sum_j |Q_j^{(k)}| \le \frac{1}{b^k} |Q| = \\ &|Q| e^{-k \log b} \le |Q| b e^{-\lambda \log b / (2^n b)}.\end{aligned} \tag{7.8}$$

Take $b = e$ to finish.

$\square$

Here is a frequently used consequence of (7.5).

Proposition 7.6. *For $0 < p < \infty$ and $f \in BMO$ we have*

$$\sup_Q \left(\frac{1}{|Q|} \int_Q |f - f_Q|^p \right)^{1/p} \lesssim_{p,n} |f|_{BMO}. \tag{7.9}$$

For $1 < p < \infty$ and $f \in L^1_{loc}$, the left and right of (7.9) are comparable.

Proof. Write

$$\begin{aligned} \frac{1}{|Q|} \int_Q |f - f_Q|^p &= \frac{p}{|Q|} \int_0^\infty \alpha^{p-1} |\{x \in Q : |f - f_Q| > \alpha\}| d\alpha \\ &\lesssim p \int_0^\infty \alpha^{p-1} e^{-c\alpha/|f|_{BMO}} d\alpha \sim \Gamma(p) |f|^p_{BMO}. \end{aligned} \tag{7.10}$$

When $1 < p < \infty$ the sup on the left is $\gtrsim$ the corresponding sup with $p = 1$, which is $|f|_{BMO}$.

□

7.2 BMO as a substitute for L^∞.

Proposition 7.7 (SIO mapping property)**.** *Let T be associated to a CZ kernel K as in Definition 3.1. Then*

$$|Tf|_{BMO} \leq CB|f|_{L^\infty}$$

for $f \in L^\infty \cap L^2$. (use L^2 so Tf well-defined).[34]

Proof. Let $B_0 = B(x_0, R)$, $B_0^* = B(x_0, 2R)$. Let

$$c_{B_0} := \int_{|y - x_0| > 2R} K(x_0 - y) f(y) dy.$$

Then $\int_{B_0} |Tf(x) - c_{B_0}| dx \leq$

$$\int_{B_0} \int_{|y-x_0|>2R} |K(x - y) - K(x_0 - y)||f(y)| dy dx + \int_{B_0} |T(\chi_{B_0^*} f)(x)| dx \leq C|B_0||f|_{L^\infty}. \tag{7.11}$$

Here we used the Hörmander condition on the first integral (set $x' = x_0 - y$)) and L^2 boundedness of T on the second.[35]

□

To discuss BMO as a substitute for L^∞ in interpolation, we need some preparation. Next we introduce a maximal function $M_d f$ whose L^p norm is dominated by that of $f^\sharp$ in some circumstances; see Prop. 7.13. We need that proposition to interpolate with BMO.

Definition 7.8 (Dyadic maximal function)**.** *Let $\mathcal{Q}_{dyad} = \{2^k[0,1)^d + 2^k\mathbb{Z}^d, k \in \mathbb{Z}\}$. For $f \in L^1_{loc}$ set*

$$M_d f(x) = \sup_{Q \ni x, Q \in \mathcal{Q}_{dyad}} Av_Q|f| = \sup_{Q \ni x, Q \in \mathcal{Q}_{dyad}} \frac{1}{|Q|} \int_Q |f| dy.$$

[34] More generally, if $f \in L^\infty$, then Tf has a well-defined action on compactly supported L^2 functions with integral zero; see Definition 9.7. With this definition one can show $Tf \in BMO$.

[35] The constant B is from Defn. 3.1.

Remark 7.9. *Clearly we have the pointwise estimate $M_d f \le Mf$ (M is the Hardy-Littlewood maximal function). Thus, M_d is weak $(1,1)$ and strong (p,p) for $1 < p \le \infty$. We can't expect the reverse pointwise estimate to hold, but Prop. 7.13 gives a reverse estimate with L^p norms.*

Proposition 7.10 (Good-λ inequality)**.** *For $\lambda > 0$, $\gamma > 0$*

$$|\{x \in \mathbb{R}^d : M_d f > 2\lambda, f^\sharp \le \gamma\lambda\}| \le \gamma 2^d |\{M_d f > \lambda\}|.$$

Proof. Suppose $|M_d f > \lambda\}| = |\Omega_\lambda| < \infty$. Write $\Omega_\lambda = \cup_j Q_j$ as in above proof (the Q_j are maximal such that $Av_{Q_j}|f| > \lambda$). It suffices to show for each j that

$$|\{x \in Q_j : M_d f > 2\lambda, f^\sharp \le \gamma\lambda\}| \le \gamma 2^d |Q_j|.$$

For $x \in Q_j$ such that $M_d f > 2\lambda$, by maximality of Q_j, if Q_j' is unique dyadic containing Q_j with twice the side length,

$$M_d((f - Av_{Q_j'} f)\chi_{Q_j})(x) \ge M_d(f\chi_{Q_j})(x) - |Av_{Q_j'} f| > 2\lambda - \lambda = \lambda. \tag{7.12}$$

Since M_d is weak $(1,1)$,

$$|M_d f > 2\lambda| \le |M_d((f - Av_{Q_j'} f)\chi_{Q_j}) > \lambda| \le \frac{1}{\lambda}\int_{Q_j} |(f - Av_{Q_j'} f| \le \frac{2^d|Q_j|}{\lambda} f^\sharp(\xi) \tag{7.13}$$

for all $\xi \in Q_j$. To finish note that we may assume $f^\sharp(\xi) \le \gamma\lambda$ for some $\xi \in Q_j$ (otherwise there is nothing to prove).

□

Remark 7.11. *There is no need to use a CZ decomposition in the proof of Prop. 7.10 (as was done in [CS1]).*

Remark 7.12. *Since $f^\sharp \lesssim 2Mf$ we have for $1 < p < \infty$*

$$|f^\sharp|_p \lesssim |Mf|_p \lesssim |f|_p \lesssim |M_d f|_p.$$

The next result shows these norms are sometimes all comparable.

Proposition 7.13 (Compare $M_d f$, $f^\sharp$)**.** *For $1 \le p_0 \le p < \infty$ suppose $f \in L^1_{loc}$ and $M_d f \in L^{p_0}$. Then $|M_d f|_p \lesssim |f^\sharp|_p$.*

Proof. Using Proposition 7.10 and making obvious changes of variable (e.g. in second inequality) we get

$$\begin{aligned}&\int_0^{\lambda_0} |M_d f > 2\lambda| p\lambda^{p-1} d\lambda \le \int_0^{\lambda_0} |M_d f > 2\lambda, f^\sharp \le \lambda\gamma| p\lambda^{p-1} d\lambda + \int_0^{\lambda_0} |f^\sharp > \gamma\lambda| p\lambda^{p-1} d\lambda \le \\ &\quad 2^{d+p}\gamma \int_0^{\lambda_0} |M_d f > 2\lambda| p\lambda^{p-1} d\lambda + \gamma^{-p}|f^\sharp|_p^p.\end{aligned} \tag{7.14}$$

Let $\gamma = 2^{-d-p-1}$ and let $\lambda_0 \to \infty$ to finish.[36]

□

[36] To absorb the first term on the left of the last inequality we need it to be finite. That is a consequence of $M_d f \in L^{p_0}$, $p \ge p_0$: write $\lambda^{p-1} = \lambda^{p-p_0}\lambda^{p_0-1} \le \lambda_0^{p-p_0}\lambda^{p_0-1}$.

Proposition 7.14 (Interpolation property of BMO)**.** *Suppose T linear is bounded on L^{p_0} for some $1 \le p_0 < \infty$ and bounded from $L^\infty \to BMO$. Then T is bounded on L^p for any $p_0 < p < \infty$.*

Proof. Let $Sf := (Tf)^\sharp$; then S is sublinear. Claim: S bounded on L^{p_0} and on L^∞. With M the HL maximal function and using $L^{1,\infty}$ in the first three terms if $p_0 = 1$, we have

$$|Sf|_{p_0} = |(Tf)^\sharp|_{p_0} \le |M(Tf)|_{p_0} \lesssim |Tf|_{p_0} \lesssim |f|_{p_0}$$

and

$$|Sf|_\infty = |(Tf)^\sharp|_\infty = |Tf|_{BMO} \lesssim |f|_\infty.$$

Now apply Marcinkiewicz interpolation to get Sf bounded on L^p for $p_0 < p < \infty$. If $f \in L^p \cap L^{p_0}$ then $Tf \in L^{p_0}$, so $M_d(Tf) \in L^{p_0}$ and

$$|Tf|_p \le |M_d(Tf)|_p \lesssim |(Tf)^\sharp|_p \le |f|_p.$$

Here we used $|Tf| \le M_d(Tf)$ a.e. (by the Lebesgue differentiation theorem) and Prop. 7.13.

□

7.3 Haar functions, dyadic harmonic analysis

Define the set of dyadic intervals on $[0,1]$,

$$D = \{[(k-1)2^{-n}, k2^{-n}) : n \ge 0, 0 \le k \le 2^n\},$$

and let A_n be the subset of intervals of length 2^{-n}. For $I \in D$ the Haar function associated to I is

$$h_I = (\chi_{I_L} - \chi_{I_R})|I|^{-1/2},$$

where L, R denote the left and right halves of I. The Haar system is $\{h_I : I \in D\}$. Identify the intervals with the nodes of the obvious binary tree, each node is associated to the corresponding h_I.[37] So the root $[1,0)$ is associated to

$$h_{[0,1)} = \chi_{[0,1/2)} - \chi_{[1/2,1)}.$$

This is h_1 in the $\{h_l : l = 0,1,2,\dots\}$ labelling, where $h_0 = \chi_{[0,1)}$.

Let F_n be the $\sigma-$algebra generated by A_n (note F_n is finite and its elements are unions of elements of A_n). An orthonormal basis for the subspace of $L^2([0,1])$ consisting of F_n- measurable functions is $\{\chi_I|I|^{-1/2}, I \in A_n\}$. Let E_n denote orthogonal projection of L^2 onto this subspace. It is easy to see

$$E_n f = \sum_{I \in A_n} \chi_I Av_I f.$$

The operator E_n extends to L^1 (consider $\min(f,n)$, $n \to \infty$) and is uniquely characterized by the property that for $f \in L^1$, $E_n(f)$ is the unique F_n-measurable function such that[38]

$$\int_B E_n f dx = \int_B f dx \text{ for all } B \in F_n.$$

[37] Each A_n gives a row of the tree.
[38] Thus, $E_n(f)$ is the "conditional expectation" $E(f|F_n)$.

If E is the expectation operator $Ef = \int_0^1 f$ we have[39]

$$
\begin{aligned}
&a) E(E_n f) = Ef \\
(7.15)\qquad &b) f \geq 0 \Rightarrow E_n f \geq 0 \\
&c) |E_n f|_p \leq |f|_p.
\end{aligned}
$$

We have (easily, see p 213 of [CI])

$$
\begin{aligned}
&|E_n f - f|_\infty \to 0 \text{ for } f \in C([0,1]) \\
(7.16)\qquad &|E_n f - f|_p \to 0 \text{ for } 1 \leq p < \infty.
\end{aligned}
$$

Also, an inductive argument shows (set $D_n = \cup_{k \leq n} A_k$) for $f \in L^1([0,1])$

$$
\begin{aligned}
&E_n f - E_{n-1} f = \sum_{I \in A_{n-1}} \langle f, h_I \rangle h_I \\
(7.17)\qquad &E_n f = \int_0^1 f + \sum_{I \in D_{n-1}} \langle f, h_I \rangle h_I.
\end{aligned}
$$

With (7.16) this implies $\{1, h_I, I \in D\}$ is an ON basis of L^2. Set

$$P_l f = \sum_{k=0}^{l} \langle f, h_l \rangle h_l.$$

This is a conditional expectation with respect to an obvious $\sigma-$algebra ($\mathcal{P}_l$ in [CS1]) so obtain $\{1, h_I, I \in D\}$ is a Schauder basis of $L^p([0,1])$ for $1 \leq p < \infty$.[40]

To see that the basis is unconditional for $1 < p < \infty$, consider the operator T defined (initially) on functions $f \in L^1$ with finite Haar expansion by

$$Tf = \sum_{I \in D} x_I \langle f, h_I \rangle h_I, \tag{7.18}$$

where $|x_I| \leq 1$ for all I. Enough to show T is strong (p,p) for $1 < p < \infty$. Strong $(2,2)$ is clear, so just need weak $(1,1)$ to finish (via interpolation and duality). Use CZ decomposition and "standard argument", p. 216 of [CS1]. One point is to show, after writing $f = g + b$, $b = \sum_{J \in \mathcal{B}} (f - f_J)\chi_J$, that

$$T((f - f_J)\chi_J) = \sum_{I \in D} x_I \langle (f - f_J)\chi_J, h_I \rangle h_I$$

is supported in J.[41]

Remark 7.15. *A corollary of this argument is that if T is a bounded linear operator on $L^2([0,1])$ and if Th_I is supported in I for all $I \in D$, then T is weak $(1,1)$ and thus strong (p,p) for $1 < p \leq 2$. This corollary is useful in the discussion of paraproducts.*

Burkholder's inequality gives another way to check that the Haar basis is unconditional for $1 < p < \infty$; see p. 217 [CS1]. The proof is remarkably clever, but (for me) not illuminating.

Theorem 7.16. *Suppose $|b_k| \leq |a_k|$ for all $k \geq 0$. Then*

$$\left| \sum_k b_k h_k \right|_p \leq C(p) \left| \sum_k a_k b_k \right|_p.$$

[39]See section 13.

[40]Use $\|P_l\|_{p \to p} \leq 1$, density of $C([0,1])$ in L^p for $1 \leq p < \infty$,...

[41]Fix J and consider cases $J \subsetneq I$, $I \subsetneq J$. In second case support of $I-$th term is contained in J. In first case use h_I constant on J and $(f - f_J)\chi_J$ has mean zero.

7.4 Littlewood-Paley inequality for the Haar square function.

Definition 7.17 (Square function)**.** *For $f \in L^1([0,1])$ with mean zero, write $f = \sum_I a_I h_I$, and set*

$$Sf = (\sum a_I^2 h_I^2)^{1/2}.$$

Remark 7.18. *If $f_n = \sum_{I \in A_n} a_I h_I$ then $Sf = (\sum_n f_n^2)^{1/2}$. To see this use that if $I \neq J \in A_n$, then $I \cap J = \emptyset$.*

Theorem 7.19. *Let $f \in L^p([0,1])$, $1 < p < \infty$, mean zero. Then for $1 < p < \infty$*

$$|Sf|_p \sim |f|_p.$$

There is obviously equality when $p = 2$.

Proof. Write $f = \sum_I a_I h_I$. Burkholder gives (r_I is Rademacher function)

$$|f|_p \sim |\sum r_I(s) a_I h_I|_p \Rightarrow |f|_p \sim \int_0^1 |\sum r_I(s) a_I h_I|_p^p ds \sim |Sf|_p^p,$$

the last $\sim$ by Khinchine's inequality. To see this use Khinchine to get

$$\int_0^1 |\sum r_I(s) a_I h_I|^p ds = E(|\sum r_I(s) a_I h_I|^p) \sim (Sf)^p.$$

Then integrate dt. $\square$

Remark 7.20. *Theorem 7.19 immediately implies Khinchine's inequality when the r_k there are Rademacher functions. See the appendix on probability.*

7.5 Dyadic H^1-dyadic BMO duality.

In this section H^1 is dyadic $H^1([0,1])$ and BMO is dyadic $BMO([0,1])$. The atomic decomposition of H^1 and $H^1 - BMO$ duality are easier in the dyadic setting, so one can view this section as a warm-up for the next two sections.

Definition 7.21. *Let Sf be the Haar square function.*
a) $H^1([0,1]) = \{f \in L^1 : Avf = 0, Sf \in L^1\}$
b) $BMO = \{f \in L^2 : Avf = 0, |f|_{BMO} := \sup_{I \in D} \left(Av_I|f - f_I|^2\right)^{1/2} < \infty$.

Proposition 7.22 (Carleson condition)**.** *Let $f \in L^2$ with mean zero. Then $f \in BMO$ if and only if*

$$\sup_{I \in D} |I|^{-1} \sum_{J \subset I, J \in D} a_J^2 = \sup_{I \in D} Av_I|f - f_I|^2 = |f|_{BMO}^2, \tag{7.19}$$

where $f = \sum_{I \in D} a_I h_I$.

Proof. Let $m_I f = Av_I f$ for now and observe for $J, I \in D$ that

$$m_I h_J = \begin{cases} 0, J \subset I \\ h_J(t) \text{ for } t \in I, I \subsetneq J \end{cases} . \tag{7.20}$$

In the first case use $Avh_J = 0$; in the second case use h_J constant on I. For each $I \in D$ this implies

(7.21) $$\sum_{J \subset I} a_I h_I = (f - m_I f)\chi_I,$$

since (with $f = \sum_J a_J h_J$)

(7.22) $$(m_I f)\chi_I = \sum_{I \subsetneq J} a_J h_J \chi_I, \text{ while } f\chi_I = \sum_{I \subsetneq J} a_J h_J \chi_I + \sum_{J \subset I} a_J h_J \chi_I.$$

Using the orthogonality of the Haar system, this gives (7.19).

□

Remark 7.23. *(a) This shows that BMO is "invariant under changes of signs" in $\sum_I a_I h_I$. The same obviously holds for H^1. Also*

(7.23) $$|g|_{H^1} \le |f|_{H^1} \text{ and } |g|_{BMO} \le |f|_{BMO}$$

whenever coefficients of f (in Haar expansion) dominate those of g.

(b) Could also define BMO by the equivalent norm $|f|_{BMO} = \sup_{I \in D} Av_I |f - f_I|$. Proof uses John-N inequality and writes the average using the distribution function of $f - f_I$.

(c) Similarly can replace 2 by p and $1/2$ by $1/p$ in Definition 7.21 when $1 < p < \infty$ and get an equivalent BMO norm; see [G2] p165.

We will use an atomic decomposition of H^1 to prove $H^1 - BMO$ duality, in particular, Fefferman's inequality.

Definition 7.24 (atom). *We say a is an H^1 atom if for some $I \in D$:*

$$\operatorname{supp} a \subset I, \quad Av_I a = 0, \quad |a|_2 |I|^{1/2} \le 1.$$

Proposition 7.25. *If a is an atom, then $|a|_{H^1} \le 1$ and $|a|_1 \le 1$.*

Proof. For I as above write $a = \sum_{J \subset I} c_J h_J$. Then

$$\int (Sa)^2 = \sum_{J \subset I} |c_J|^2 = |a|_2^2 \le |I|.$$

Use C-Schwarz to finish. □

Theorem 7.26 (Atomic decomposition). *Every $f \in H^1$ can be written $f = \sum_{j=1}^{\infty} c_j a_j$ with convergence in H^1, where the a_j are atoms and*

(7.24) $$\sum_j |c_j| \le C|f|_{H^1}.$$

Remark 7.27. *The convergence in H^1 is absolute since*

$$|\sum_j c_j a_j|_{H^1} \le \sum_j |c_j||a_j|_{H^1} \le \sum |c_j| \le C|f|_{H^1} = C|Sf|_1.$$

Proof. The proof is another example of a stopping-time argument.

Write $f = \sum_I b_I h_I$ and normalize so $\int_0^1 Sf = 1$. To each node of the Haar tree associate b_I. Let

$$B([0,1]) = \{x \in [0,1] : Sf(x) > 2^{n_1}\}, \tag{7.25}$$

where 2^{n_1} is the smallest power of 2 such that $|\mathcal{B}([0,1])| \leq \frac{1}{4}$. Let $F([0,1])$ be the set of maximal dyadic intervals in $B([0,1])$.[42] If $F([0,1]) = \emptyset$ let $T([0,1]) = D$ be the full tree, set $c_1 = 2^{n_1}$, $f_{[0,1]} = f = \sum_{I \in T([0,1])} b_I h_I$, and $a_1 = 2^{-n_1} f_{[0,1]}$. Then $f = c_1 a_1$ is the desired decomposition. Note that

$$|a_1|_2 = 2^{-n_1}|f_{[0,1]}|_2 = 2^{-n_1}|Sf_{[0,1]}|_2 \leq 1, \tag{7.26}$$

since $F([0,1]) = \emptyset$ implies $Sf_{[0,1]} \leq 2^{n_1}$ on $[0,1]$, and thus $|Sf_{[0,1]}|_2^2 \leq 2^{2n_1}|[0,1]|$.[43]

Suppose then that $F([0,1]) = \{I_{1,j}\} \neq \emptyset$. Each $I_{1,j}$ is the root of a tree $T_{1,j}([0,1])$. Remove these from D and define $T([0,1])$ to be the tree that is left. Set $f_{[0,1]} = \sum_{I \in T([0,1])} b_I h_I$. Define c_1, a_1 as indicated above.

Now repeat the process on each $I_{1,j}$ and corresponding $T_{1,j}([0,1])$. Define $B(I_{1,j}) = \{x \in I_{1,j} : Sf(x) > 2^{n_{1,j}}\}$, where $2^{n_{1,j}}$ is the smallest power such that $|B(I_{1,j})| < |I_{1,j}|/4$. Let $F(I_{1,j}) = \{I_{2,k}\}$ be the set of maximal dyadic intervals in $B(I_{1,j})$. Each $I_{2,k}$ is the root of a tree $T_{2,k}(I_{1,j})$ that we remove from $T_{1,j}([0,1])$ to yield a tree $T(I_{1,j})$. Now define

$$f_{I_{1,j}} = \sum_{I \in T(I_{1,j})} b_I h_I,$$

which by construction satisfies $|f_{I_{1,j}}|_2^2 \leq 2^{2n_{1,j}}|I_{1,j}|$.

Continue... Get a family of pairwise disjoint subtrees T_l (such as $T([0,1])$, the $T(I_{1,j})$, the $T(I_{2,k})$,...) with roots I_l (such as $[0,1]$, the $I_{1,j}$, the $I_{2,k}$,...) and associated functions $f_l = \sum_{I \in T_l} b_I h_I$ (such as $f_{[0,1]}$, the $f_{I_{1,j}}$,...), which satisfy (with $n_l = n_1, n_{1,1}, n_{1,2}...$)

$$|f_l|_2 \leq 2^{n_l}|I_l|^{1/2} \text{ and } |I_l| \leq 4|\{x \in I_l : Sf(x) > 2^{n_l - 1}\}|,$$

the latter because of the minimality property of 2^{n_l}.

Note that if $I_l \subsetneq I_{l'}$ then $n_l > n_{l'}$.[44] Now define

$$c_l = 2^{n_l}|I_l|, \;\; a_l = c_l^{-1} f_l. \tag{7.27}$$

Then $f = \sum c_l f_l$ and

$$\begin{aligned}\sum_l c_l = \sum_l 2^{n_l}|I_l| \leq 4\sum_l 2^{n_l}|\{x \in I_l : Sf(x) > 2^{n_l-1}\} \leq \\ 4\sum_{k \in \mathbb{N}} 2^k|\{x \in [0,1] : Sf(x) > 2^{k-1}\}| \leq C|Sf|_1 \leq C.\end{aligned} \tag{7.28}$$

For the third inequality we used the fact that if $I_l \cap I_{l'} \neq \emptyset$, then $n_l \neq n_{l'}$.

□

[42]Note: the elements of $F([0,1])$ are disjoint.

[43]Proof: If $Sf_{[0,1]}(x) > 2^{n_1}$ for some $x \in [0,1]$, then $\sum_{I \in D} b_I^2 \chi_I(x) > 2^{2n_1}$. Some finite partial sum $\sum_{I \in \mathcal{F}} b_I^2 \chi_I(x)$ must then be $> 2^{2n_1}$, which implies that some dyadic interval $J \subset B([0,1])$, contradiction.

[44]Proof: Since $I_l \subsetneq I_{l'}$, we have $l < l'$ and $I_l \subset B(I_{l'})$. Say $n_{l'} \geq n_l$. Since $Sf(x) > 2^{n_{l'}}$ for $x \in I_l \subsetneq I_{l'}$, we have $Sf(x) > 2^{n_l}$ for $x \in I_l$ and so $I_l \subset B(I_l)$, contradiction.

We use this decomposition to prove Fefferman's inequality.

Proposition 7.28 (Fefferman's inequality)**.** *For $f \in H^1$ and $h \in BMO$ we have*

$$|\langle f, h\rangle| \leq C|f|_{H^1}|h|_{BMO}.$$

Proof. By previous theorem it suffices to prove this when $f \in H^1$ is an atom. Suppose a is an H^1 atom, so

$$\text{supp } a \subset I, \quad Av_I a = 0, \quad |a|_2|I|^{1/2} \leq 1.$$

Then (with $h_I = Av_I h$ now)

$$|\int_0^1 ah| = |\int_I a(h - h_I)| \leq |a|_2|(h - h_I)\chi_I|_2 \leq |h|_{BMO}.$$

Recall $|a|_{H^1} \leq 1$.

□

Proposition 7.29. *$BMO = (H^1)^*$.*

Proof. It remains to show that any $L \in (H^1)^*$ is given by a BMO function. Write $f = \sum_0^\infty f_n$ with $f_n = \sum_{I\in A_n}\langle f, h_I\rangle h_I$. Then

$$|f|_{H^1} = |Sf|_1 = (\sum_n f_n^2)^{1/2} = (f_n)_{\ell^2}. \tag{7.29}$$

Thus, the map $f \to (f_n)$ isometrically identifies H^1 with a subspace of $L^1([0,1], \ell^2)$. By Hahn-Banach extend L to $\tilde{L} \in L^1([0,1], \ell^2)^* = L^\infty([0,1], \ell^2)$ with the same norm as L. So there exists $(\phi_n) \in L^\infty([0,1], \ell^2)$ such that for all $(g_n) \in L^1([0,1], \ell^2)$ we have

$$\tilde{L}((g_n)) = \int_0^1 \sum_{n=0}^\infty g_n\phi_n. \tag{7.30}$$

So

$$Lf = \int_0^1 \sum_{n=0}^\infty \phi_n f_n = \int_0^1 \sum_{n=0}^\infty \phi_n \sum_{A_n}\langle f, h_I\rangle h_I = \langle f, h\rangle, \tag{7.31}$$

where

$$h = \sum_{n=0}^\infty \sum_{I\in A_n}\langle \phi_n, h_I\rangle h_I.$$

To see $h \in BMO$ compute[45]

$$\sup_{J\in D}|J|^{-1}\sum_{n=0}^\infty \sum_{I\subset J, I\in A_n}|\langle h_I, \phi_n\rangle|^2 \leq \sup_{J\in D}|J|^{-1}\int_J \sum_{n=0}^\infty |\phi_n(x)|^2 dx \leq |(\phi_n)|^2_{L^\infty([0,1],\ell^2)} = |\tilde{L}|. \tag{7.32}$$

So $|h|_{BMO} \leq |L|$. The reverse inequality follows from Fefferman's inequality.

□

[45] For the second inequality use Cauchy-Schwarz, $|h_I|_2 = 1$, and disjointness of the $I \in A_n$.

7.6 Atomic decomposition of (classical) H^1 functions

We return now to the classical Hardy space, $H^1(\mathbb{R}^n)$, which was characterized in several ways in Proposition 6.10. First, we state without proof some properties and alternative characterizations of H^1. Detailed proofs may be found in Chapter 2 of [G2].

Proposition 7.30 (The Hardy space H^1)**.**
(a) The space $H^1(\mathbb{R}^n)$ is the proper subspace of L^1 consisting of functions such that

$$|f|_{H^1} := |f|_{L^1} + |Sf|_{L^1} \sim |f|_{L^1} + \sum_{j=1}^{d} |R_j f|_{L^1} \sim \left| \sup_{t>0} |P_t * f| \right|_{L^1} < \infty.$$

Here S is the LP square function, the R_j are Riesz transforms if $n \geq 2$, the Hilbert transform if $n = 1$, and P is the Poisson kernel.[46]
(b) Elements of H^1 satisfy $\int f dx = 0$. [47]

For atomic decompositions we need another characterization of H^1 in terms of nontangential maximal functions u_A^*.

Definition 7.31. *Let $\mathbb{R}_+^{n+1} = \{(y,t) : y \in \mathbb{R}^n, t > 0\}$ and for $x \in \mathbb{R}^n$ define the "cone based at x of aperture $A > 0$" to be $\Gamma_{x,A} = \{(y,t) \in \mathbb{R}_+^{n+1} : |x - y| < At\}$. Define the nontangential maximal function*

$$u_A^*(x) = \sup_{(y,t)\in\Gamma_{x,A}} |u(y,t)|.$$

Proposition 7.32. *Let $f \in H^1$ and let $u(x,t) = (P_t * f)(x)$ be the Poisson integral of f for $(x,t) \in \mathbb{R}^{n+1}$. Then $f \in H^1$ if and only if for each $A > 0$ we have $u_A^* \in L^1(\mathbb{R}_+^{n+1})$. Moreover,*

$$|f|_{H^1(\mathbb{R}^n)} \sim_A |u_A^*|_{L^1(\mathbb{R}_+^{n+1})}.$$

Remark 7.33. *Here are a few other useful facts:*
(a) Let $1 < q \leq \infty$. Then $\{g \in L_c^q : \int g = 0\} \subset H^1$. Here the subscript c indicates compact support. See Exercise 2.1.7 of [G2]; this can be done using Riesz transforms.[48]
(b) $\{f \in \mathcal{S} : \operatorname{supp} \hat{f}$ is away from a neighborhood of zero$\}$ is dense in H^1. See Stein [S1], p. 231 for a proof using Riesz transforms.
(c) The space H^1 is complete, and the injection $i : H^1 \to \mathcal{S}'$ is continuous.
(d) For $1 \leq r \leq \infty$ $L^r \cap H^1$ is dense in H^1.

Definition 7.34 (L^2 atoms for H^1)**.** *A function A is an L^2 atom for H^1 if there exists a cube Q such that:*
(a) $\operatorname{supp} A \subset Q$*,*
(b) $|A|_2 \leq |Q|^{-1/2}$*,*
(c) $\int A(x)dx = 0$*.*

We now show that L^2 atoms for H^1 are elements of $H^1(\mathbb{R}^n)$.

Proposition 7.35. *Every L^2 atom A for H^1 satisfies $|A|_{H^1} \sim |A|_{L^1} + |SA|_{L^1} \lesssim_n 1$. Note: the constant on the right is independent of A.*

[46] As usual $P_t(x) = t^{-n}P(x/t)$.
[47] So $H^1 \subsetneq L^1$ and, in fact, "most" test functions $\phi \in C_c^\infty$ are not in H^1!
[48] See p. 101 of handwritten notes, orange notebook.

Proof. **1.** Clearly $|A|_{L^1} \lesssim_n 1$. Let A be an atom supported in a cube Q, whose center we can suppose is the origin. Let $Q^* = 2\sqrt{n}Q$. We esimate $|SA|_{L^1}$ by estimating $|SA|_{L^1(Q^*)}$ and $|SA|_{L^1((Q^*)^c)}$. Using Cauchy-Schwarz and the fact that $|Sf|_2 \lesssim |f|_2$ we obtain

$$\int_{Q^*} SA(x)dx \leq |SA|_2|Q^*|^{1/2} \lesssim |A|_2|Q^*|^{1/2} \lesssim 1. \tag{7.33}$$

2. Outside Q^*. We have for $0 \leq \theta \leq 1$:

$$\begin{aligned} \Delta_j A(x) = \int_Q A(y)\psi_{2^{-j}}(x-y)dy = 2^{jn}\int_Q A(y)[\psi(2^j x - 2^j y) - \psi(2^j x)]dy = \\ 2^{jn}\int_Q A(y)\nabla\psi(2^j x - 2^j\theta y)\cdot(-2^j y)dy. \end{aligned} \tag{7.34}$$

Since $|x-\theta y| \geq |x|/2$ when $x \notin Q^*$, $y \in Q$ we obtain

$$|\Delta_j A(x)| \lesssim 2^{jn}\int_Q |A(y)|\frac{|2^j y|}{(1+2^j|x|)^N}dy \leq \frac{2^{j(n+1)}}{(1+2^j|x|)^N}|A|_2|y|_{L^2(Q)} \lesssim \frac{2^{j(n+1)}}{(1+2^j|x|)^N}|Q|^{1/n}. \tag{7.35}$$

For $x \notin (Q^*)^c$ this implies

$$\left(\sum_j |\Delta_j A(x)|^2\right)^{1/2} \lesssim |Q|^{1/n}\left(\sum_j \frac{2^{2j(n+1)}}{(1+2^j|x|)^{2N}}\right)^{1/2} \lesssim |Q|^{1/n}|x|^{-(1+n)}. \tag{7.36}$$

To get the last inequality we took N large enough and summed separately over $2^j \leq |x|^{-1}$ and $2^j > |x|^{-1}$. Since $|x| \gtrsim |Q|^{1/n}$ this easily implies $|SA|_{L^1((Q^*)^c)} \lesssim_n 1$.

□

Proposition 7.36 (Atomic decomposition of H^1 functions)**.** *Let $f \in H^1(\mathbb{R}^n)$. There exist atoms A_j and scalars λ_j such that*

$$S_N := \sum_{j=1}^N \lambda_j A_j \to f \text{ in } H^1. \tag{7.37}$$

Moreover, $|f|_{H^1} \sim \inf_{\{\text{all such sequences } (\lambda_j)\}} |(\lambda_j)|_{\ell^1}$.[49]

Remark 7.37. *(a) Suppose $f \in H^1(\mathbb{R}^n)$ and (7.37) holds for some sequence (λ_j) such that $\sum_{j=1}^\infty |\lambda_j| < \infty$. Prop. 7.35 implies $\sum_{j=1}^N |\lambda_j A_j|_{H^1} \lesssim \sum_{j=1}^N |\lambda_j|$. Thus,*

$$|f|_{H^1} \leq |f - \sum_{j=1}^N \lambda_j A_j|_{H^1} + \sum_{j=1}^N |\lambda_j|, \tag{7.38}$$

so (7.37) implies $|f|_{H^1} \leq \sum_{j=1}^\infty |\lambda_j|$.

(b) To finish the proof of Prop. 7.36 it will be enough to show that if $f \in H^1$, then atoms A_j and scalars λ_j exist such that $S_N \to f$ in $\mathcal{S}'$ and

$$\sum_{j=1}^\infty |\lambda_j| \leq |f|_{H^1}. \tag{7.39}$$

Indeed, since $\sum_j |\lambda_j| < \infty$, the sequence (S_N) is Cauchy in H^1, so $S_N \to f$ in H^1; see Remark 7.33(c).

[49] See Theorem 2.3.12 of [G2].

We begin by proving a special case of the assertion in Remark 7.37(b) when $n = 1$. The proof of Prop. 7.38 contains the main idea of the proof of Prop. 7.36.

Proposition 7.38. *Let $f \in H^1(\mathbb{R}) \cap L^2$. There exist atoms A_j and scalars λ_j such that*

$$S_N := \sum_{j=1}^{N} \lambda_j A_j \to f \text{ in } \mathcal{S}'. \tag{7.40}$$

Moreover, $\sum_{j=1}^{\infty} |\lambda_j| \le |f|_{H^1}$.

Proof. We follow the argument of [W] and also the exposition of that argument in [GR].[50] Take f to be real-valued for convenience.[51] In step 1 we will give a concise overview of the argument containing claims for which technical details are omitted. In subsequent steps we fill in those details.

1. Choose $\psi \in C_c^\infty(\mathbb{R}^n)$ to be real, radial, supported in $\{|x| \le 1\}$, and satisfying $\int \psi(x)dx = 0$ and

$$\int_0^\infty e^{-\theta}\hat{\psi}(\theta)d\theta = -1. \tag{7.41}$$

Set $\psi_t(y) = t^{-1}\psi(y/t)$. Let $u = P_t * f$ be the Poisson integral of f. We claim that the following limit holds:

$$f(x) = \lim_{\epsilon\to 0} \int_\epsilon^\infty \int \partial_t u(y,t)\psi_t(x-y)dydt \text{ in } \mathcal{S}'. \tag{7.42}$$

For $k \in \mathbb{Z}$ define the open set

$$E^k = \{x \in \mathbb{R} : u_2^*(x) > 2^k\} = \cup_{j=1}^{\infty} I_j^k,$$

where the I_j^k are the maximal disjoint component open intervals. For any interval $I \subset \mathbb{R}$ define the associated "tent" over I by[52]

$$\hat{I} = \{(y,t) \in \mathbb{R}^2_+ : (y-t, y+t) \subset I\}.$$

Define

$$\hat{E}^k = \cup_j \hat{I}_j^k \text{ and } T_j^k = \hat{I}_j^k \setminus \hat{E}_j^k.$$

Observe that the T_j^k are disjoint, and the union of their (compact) closures is $\mathbb{R}^2_+$.[53]

We claim that the following equations (which are "formally obvious", given (7.42)) hold in $\mathcal{S}'$:

$$f = \sum_{k,j} \int_{T_j^k} \int \partial_t u(y,t)\psi_t(x-y)dydt := \sum_{j,k} g_j^k = \sum_{j,k} \lambda_j^k a_j^k, \tag{7.43}$$

where for some C to be chosen $\lambda_j^k = C2^k|I_j^k|$, and where the functions a_j^k thereby defined are (we claim) atoms for that choice of C. The a_j^k clearly inherit the property $\int a_j^k dx = 0$ from ψ and satisfy supp $a_k^j \subset \overline{I_j^k}$. Moreover, we claim

$$\sum_{j,k} \lambda_j^k \lesssim \int u_2^* dx \lesssim |f|_{H^1}. \tag{7.44}$$

[50] Thanks to J. M. Wilson for quickly and clearly answering my questions about his paper [W].

[51] The same argument can be applied to a complex-valued $f \in H^1 \cap L^2$.

[52] If $I = (a,b)$, then the tent $\hat{I}$ is an open triangle with base I and apex given by the point $(\frac{a+b}{2}, \frac{b-a}{2})$.

[53] A typical T_j^k is a tent over I_k^j with some smaller triangles (which are contained in $\hat{E}^{k+1}$) removed.

Then it remains just to show that $|a_j^k| \le |I_j^k|^{-1/2}$ or equivalently

$$|g_j^k|_2 \le C2^k |I_j^k|^{1/2} \text{ for } C \text{ as above.} \tag{7.45}$$

We do this by duality. For $\phi \in L^2(\mathbb{R})$ with $|\phi|_2 = 1$ we have (use Cauchy-Schwarz and Plancherel)

$$\begin{aligned} &|\int \phi(x) g_j^k(x)dx| = \left| \int_{T_j^k} \partial_t u(y,t)(\phi * \psi_t)(y)dydt \right| \le \\ &\left(\int_{T_j^k} t|\nabla u|^2 dtdy\right)^{1/2} \left(\int_{\mathbb{R}^2_+} |(\phi * \psi_t)(y)|^2 \frac{dtdy}{t}\right)^{1/2} \lesssim C_1|\phi|_2 \left(\int_{T_j^k} t|\nabla u|^2 dtdy\right)^{1/2}. \end{aligned} \tag{7.46}$$

By Green's theorem the last integral is bounded by[54]

$$\int_{\partial T_j^k} \left(|u|t \left|\frac{\partial u}{\partial \nu}\right| + \frac{1}{2}u^2 \left|\frac{\partial t}{\partial \nu}\right| \right) ds \tag{7.47}$$

where ν denotes the outward unit normal. Clearly, $\left|\frac{\partial t}{\partial \nu}\right| \le 1$ and $|\partial T_j^k| \le C_2 |I_j^k|$. Moreover, we claim that for some absolute constant C_3 both

$$|u| \le C_3 2^k \text{ and } |t\nabla u| \le C_3 2^k \text{ on } \partial T_j^k. \tag{7.48}$$

This gives (7.45) with "C" determined by C_i, $i = 1, 2, 3$. Next we justify the several claims made above.

2. (7.42) **and toward** (7.43)**.** Since $\hat{P}_t(\xi) = e^{-t|\xi|}$, if $\phi \in \mathcal{S}$ we have by Parseval

$$\begin{aligned} &\int_\epsilon^\infty \int \partial_t u(y,t)(\phi * \psi_t)(y)dydt = \int_\epsilon^\infty \int (-|\xi|)e^{-t|\xi|}\hat{f}(\xi)\hat{\psi}(t|\xi|)\hat{\phi}(-\xi)d\xi dt = \\ &\int \int_{\epsilon|\xi|}^\infty e^{-s}\hat{\psi}(s)\hat{f}(\xi)\hat{\phi}(-\xi)dsd\xi \to \int \hat{f}(\xi)\hat{\phi}(-\xi)d\xi \text{ as } \epsilon \to 0. \end{aligned} \tag{7.49}$$

Note that the first integral in (7.49) is absolutely convergent. Thus, if we let

$$T_j^{k(\epsilon)} = \{(y,t) \in T_j^k : t \ge \epsilon\} \text{ and } g_j^{k(\epsilon)}(x) = \int_{T_j^{k(\epsilon)}} \partial_t u(y,t)\psi_t(x-y)dydt,$$

the dominated convergence theorem implies for *fixed* ϵ that

$$\left\langle \sum_{j,k} g_j^{k(\epsilon)}, \phi \right\rangle = \int_\epsilon^\infty \int \partial_t u(y,t)(\phi * \psi_t)(y)dydt. \tag{7.50}$$

From this and (7.49) we conclude[55]

$$g^\epsilon := \sum_{j,k} g_j^{k(\epsilon)} \text{ satisfies } g^\epsilon \to f \text{ in } \mathcal{S}' \text{ as } \epsilon \to 0. \tag{7.51}$$

3. Toward (7.45)**.** First we perform the estimates (7.46), (7.47) with g_j^k, T_j^k replaced by $g_j^{k(\epsilon)}$ and $T_j^{k(\epsilon)}$. The application of Green's theorem to $T_j^{k(\epsilon)}$ is clearly valid now. We claim (7.48) holds on

[54] At this point there is some ambiguity about the meaning of $t\partial_\nu u$ on the part of ∂T_j^k where $t = 0$. Thus, there is a question about the correctness of this application of Green's theorem. We clarify this below.

[55] The assertion (7.51) is not the same as (7.43).

$\partial T_j^{k(\epsilon)}$ uniformly for small ϵ. The essential observation is that "because of the 2 in u_2^*", and because the slopes of the nonhorizontal segments of $\partial T_j^{k(\epsilon)}$ are ± 1, every point of $\partial T_j^{k(\epsilon)}$ is contained in a cone $\Gamma_{z,2}$ for some $z \notin E^{k+1}$.[56] It helps to draw a picture to see this (or see the figure on p. 262 of [GR]). This observation immediately implies $|u| \le C_3 2^k$.

To estimate $t\partial_t u(x,t)$ at $(x,t) \in \partial T_j^{k(\epsilon)}$, or in fact at any point in the closure of $T_j^{k(\epsilon)}$, we use the mean value property of harmonic functions to write

$$\partial_t u(x,t) = |B_{x,t}|^{-1} \int_{B_{x,t}} \partial_t u dy \sim t^{-2} \int_{\partial B_{x,t}} u\nu_t ds, \tag{7.52}$$

where $B_{x,t}$ is a ball centered at (x,t) of radius t/c, where c is a geometric constant big enough to guarantee that $B_{x,t}$ is contained in $\Gamma_{z,2}$ for some $z \notin E^{k+1}$.[57] Using (7.52) we obtain the estimate

$$|t\partial_t u(x,t)| \lesssim t^{-1} \int_{\partial B_{x,t}} |u| ds \lesssim 2^k \text{ for } (x,t) \in \overline{T_j^{k(\epsilon)}} \text{ uniformly for } \epsilon \text{ small.} \tag{7.53}$$

The estimate of $t\partial_y u$ is similar, so (7.48) holds on $\partial T_j^{k(\epsilon)}$ uniformly for $\epsilon > 0$ small, and we have

$$|g_j^{k(\epsilon)}|_2 \le C 2^k |I_j^k|^{1/2} \text{ uniformly for } \epsilon \text{ small.} \tag{7.54}$$

4. Definition of g_j^k and (7.45). Define $g_j^k \in \mathcal{S}'$ by

$$\langle g_j^k, \phi \rangle = \lim_{\epsilon \to 0} \int_{T_j^{k(\epsilon)}} \int \partial_t u(y,t) \psi_t(x-y) \phi(x) dx dy dt = \int_{T_j^k} \int \partial_t u(y,t) (\psi_t * \phi)(y) dy dt \tag{7.55}$$

for $\phi \in \mathcal{S}$. To see that this limit defines a distribution write

$$\begin{aligned} (\psi_t * \phi)(y) &= \int t^{-1}\psi(x/t)\phi(y-x)dx = \int t^{-1}\psi(x/t)[\phi(y-x) - \phi(y)]dx = \\ &\int t^{-1}\psi(x/t)\nabla\phi(y-\theta x)\cdot(-x)dx, \quad \theta \in [0,1]. \end{aligned} \tag{7.56}$$

Substituting into (7.55) and setting $u = \frac{x}{t}$, we have $-xdx = -t^2 du$, so one can use (7.53) to finish. This shows $g_j^{k(\epsilon)} \to g_j^k$ in $\mathcal{S}'$. The estimate (7.54) implies that some subsequence of $g_j^{k(\epsilon)}$ has a weak limit in L^2, so this limit must be g_j^k and we have[58]

$$|g_j^k|_2 \le C 2^k |I_j^k|^{1/2}. \tag{7.57}$$

Note that the sequence $(\int_{T_j^{k(\epsilon)}} t|\nabla u|^2 dt dy)^{1/2}$ is bounded above and increasing as $\epsilon \to 0$. Thus, it is Cauchy. This implies that the sequence $g_j^{k(\epsilon)}$ is Cauchy in L^2.[59] Thus,

$$g_j^{k(\epsilon)} \to g_j^k \text{ in } L^2 \text{ as } \epsilon \to 0. \tag{7.58}$$

---

[56] Recall Defn. 7.31.

[57] Here again we take advantage of "aperture 2 versus slope ± 1." Observe also that if $(y,t) \in T_j^{k(\epsilon)}$, then some point of $(y-t, y+t)$ fails to lie in E^{k+1}.

[58] We could omit this observation about the weak limit, since we have a strong limit.

[59] Estimate $|\langle g_j^{k(\epsilon)} - g_j^{k(\eta)}, \phi \rangle|$, where $|\phi|_2 = 1$, as in (7.46) to see this.

5. $g_j^{k(\epsilon)} \to g_j^k$ in H^1. The functions $g_j^{k(\epsilon)}$ and g_j^k have vanishing first moment since ψ does, and clearly have support in $\overline{I_j^k}$. Thus, $G_j^k := g_j^k |g_j^k|_2^{-1} |I_j^k|^{-1/2}$ is an atom. Proposition 7.35 implies for some absolute constant C:

$$|G_j^k|_{H^1} \leq C \text{ and thus } |g_j^k|_{H^1} \leq |g_j^k|_2 |I_j^k|^{1/2} \leq C2^k |I_j^k|. \tag{7.59}$$

Similar inequalities hold for $g_j^{k(\epsilon)}$ and for the differences $g_j^{k(\epsilon)} - g_j^k$, so with (7.58) we obtain

$$g_j^{k(\epsilon)} \to g_j^k \text{ in } H^1 \text{ as } \epsilon \to 0. \tag{7.60}$$

6. The claim (7.44)**.** We have $\sum_{j,k} \lambda_j^k = C \sum_{k=-\infty}^{\infty} 2^k |E^k|$, so (7.44) is a consequence of the following general fact: If $f : \mathbb{R} \to [0, \infty]$ is measurable, then

$$\int f(x)dx \sim \sum_{-\infty}^{\infty} 2^k |\{x : f(x) > 2^k\}| \tag{7.61}$$

with absolute comparability constants.

Indeed, setting $E^k = \{x : f(x) > 2^k\}$ we have

$$\sum_{-\infty}^{\infty} 2^k |\{x : f(x) > 2^k\}| = \sum_{-\infty}^{\infty} 2^k |E^k| = \int \sum_{-\infty}^{\infty} 2^k \chi_{E^k} dx. \tag{7.62}$$

To finish we show $\sum_{-\infty}^{\infty} 2^k \chi_{E^k}(x) \sim f(x)$. If $0 < f(x) < \infty$ there exists a unique j such that $2^j < f(x) \leq 2^{j+1}$, so

$$\sum_{-\infty}^{\infty} 2^k \chi_{E^k}(x) = \sum_{-\infty}^{j} 2^k \sim 2^j \sim f(x).$$

The cases $f(x) = 0$ and $f(x) = \infty$ are clear.

7. Establish (7.43)**: $\sum_{j,k} g_j^k = f$ in $\mathcal{S}'$.** First we show that for some $g \in H^1$:

$$\sum_{j,k} g_j^k = g \text{ in } H^1. \tag{7.63}$$

Indeed, by (7.59) $\sum_{j,k} |g_j^k|_{H^1} \lesssim \sum_{j,k} 2^k |I_j^k| \lesssim |f|_{H^1}$.

The same estimates show that for each ϵ the sum $\sum_{j,k} g_j^{k(\epsilon)}$, called g^ϵ in (7.51), converges to an element of H^1. That is,

$$\sum_{j,k} g_j^{k(\epsilon)} = g^\epsilon \text{ in } H^1.$$

In (7.51) we showed $g^\epsilon \to f$ in $\mathcal{S}'$ as $\epsilon \to 0$. To finish then, it is enough to show $g^\epsilon \to g$ in H^1 (and thus in $\mathcal{S}'$) as $\epsilon \to 0$. But on the one hand by (7.59) we have:

$$|g - g^\epsilon|_{H^1} \leq \sum_{j,k} |g_j^k - g_j^{k(\epsilon)}|_{H^1} \lesssim 2|f|_{H^1} \text{ uniformly for small } \epsilon,$$

and on the other by (7.60), $g_j^{k(\epsilon)} \to g_j^k$ in H^1 as $\epsilon \to 0$. Thus, $g^\epsilon \to g$ in H^1 as $\epsilon \to 0$.

□

We now finish the proof of Proposition 7.36.

Proof of Proposition 7.36. **1. The case $\mathbb{R}$.** Most of the work for this case is done in Prop. 7.38, where we took $f \in H^1(\mathbb{R}) \cap L^2$. Now suppose just that $f \in H^1(\mathbb{R})$. By Remark 7.33(d) we can find $f_n \in H^1 \cap L^2$ such that $f_n \to f$ in H^1. Taking subsequences we can arrange both that $f_n(x) \to f(x)$ a.e. and

$$|f_{n+1} - f_n|_{H^1} \leq 2^{-n}|f|_{H_1} \text{ for all } n. \tag{7.64}$$

By Prop. 7.38 for each n we can write (taking $f_0 = 0$)

$$f_{n+1} - f_n = \sum_i \lambda_{ni} a_{ni} \text{ in } H^1, \text{ where } \sum_i |\lambda_{ni}| \leq C2^{-n}|f|_{H^1} \tag{7.65}$$

and the a_{ni} are atoms. Since

$$\sum_i |\lambda_{1i}| + \sum_{n=1}^{\infty} \sum_i |\lambda_{n,i}| \lesssim |f|_{H^1},$$

we have

$$f = \sum_{n=0}^{\infty}(f_{n+1} - f_n) = \sum_i \lambda_{1i} a_{1i} + \sum_{n=1}^{\infty} \sum_i \lambda_{n,i} a_{ni} \text{ in } H^1, \tag{7.66}$$

and the sums can be rearranged as a single series converging to f in H^1.

2. The case $\mathbb{R}^n$, $n > 1$. The proof is essentially the same as in the case $n = 1$. The main difference is that we have to give a way of decomposing the open set

$$E^k = \{x \in \mathbb{R}^n : u^*_{30n} > 2^k = \cup_{j=1}^{\infty} \Omega_j^k$$

into nonoverlapping cubes Ω_j^k. We do this using the Whitney decomposition of E^k, Theorem 14.5. With $l(\Omega_j^k)$ the side length of Ω_j^k, define tents

$$\hat{\Omega}_j^k = \{(y, t) : y \in \Omega_j^k, 0 < t < l(\Omega_j^k)\},$$

and set

$$\hat{E}^k = \cup_{j,k} \hat{\Omega}_j^k \text{ and } T_j^k = \hat{\Omega}_j^k \setminus \hat{E}_j^{k+1}.$$

Now proceed as before.[60] □

7.7 $H^1 - BMO$ duality

In this section we show $(H^1)^* = BMO$. Recall that $BMO(\mathbb{R}^n)$ is the set of all locally integrable functions f (mod constants) such that

$$\sup_{\text{cubes } Q} |f - f_Q|_Q < \infty,$$

where f_Q denotes the average of f on Q. A consequence of Proposition 7.6 is that

$$f \in BMO \Rightarrow f \in L^p_{loc} \text{ for } 1 \leq p < \infty. \tag{7.67}$$

[60] More detail is provided in [GR], p. 279.

Definition 7.39 (Action of $b \in BMO$ on H^1)**.** ***1.*** *Let $b \in BMO$ and let $H^1_0 \subset H^1$ be the dense subset of finite linear combinations of atoms. Then by (7.67)*

$$L_b(g) := \int b(x)g(x)dx \tag{7.68}$$

is well-defined for $g \in H^1_0$, and the integral depends on b only mod constants, as desired.

2. *If $b \in L^\infty \subset BMO$, then (7.68) is well-defined on H^1; that is, b defines an element of $L_b \in (H^1)^*$. We claim that in this case*

$$|L_b|_{H^1 \to \mathbb{C}} \le C_n |b|_{BMO} = C(n,b). \tag{7.69}$$

This follows easily after writing $L_b f$ using Prop. 7.36. Use Cauchy-Schwarz and properties of atoms.

3. *We define L_b for general $b \in BMO$ as the weak limit $\tilde{L}_b$ of a subsequence of $b_M := b\chi_{|b| \le M}$, $M = 1, 2, \ldots$. Here use $|b_M|_{BMO} \le \frac{9}{4}|b|_{BMO}$, (7.69), and Banach-Alaoglu.*[61] *To see this is well defined check that $\tilde{L}_b(g) = L_b(g)$ on the dense subspace H^1_0.*

This procedure defines L_b for any $b \in BMO$.

Theorem 7.40 ($(H^1)^* = BMO$)**.** *(a) If $b \in BMO$, then $L_b \in (H^1)^*$ and $|L_b| \le |b|_{BMO}$. Moreover, the map $b \to L_b$ is injective.*

(b) Given $L \in (H^1)^*$ there exists $b \in BMO$ such that $L_b(g) = L(g)$ for $g \in H^1_0$; thus $L_b = L$. Moreover,

$$|b|_{BMO} \lesssim_n |L_b|. \tag{7.70}$$

Proof. **1. Part (a)** Let $b \in BMO$; so $b \in L^2_{loc}$. To see $b \to L_b$ is injective, suppose $L_b = 0$ and let Q be any cube. Now $b \in L^2(Q)$ satisfies

$$\int_Q b(x)g(x)dx = 0 \tag{7.71}$$

for all $g \in L^2_Q$ with $\int_Q g = 0$, since such $g \in H^1$ by Remark 7.33(b). Let $g = b - b_Q$ and observe

$$\int_Q g^2 = 0,$$

so $b = b_Q$ a.e. on Q. That is, $b = 0$ a.e. on Q as an element of BMO.

2. Part (b) Let $L \in (H^1)^*$ and let Q be any cube. Let $L^2(Q)$ denote the space of square integrable functions on $\mathbb{R}^n$ supported in Q, and let $L^2_0(Q)$ be the closed subspace of $L^2(Q)$ of functions with mean zero. We know $L^2_0(Q) \subset H^1$; in fact

$$|g|_{H^1} \lesssim_n |Q|^{1/2}|g|_{L^2}, \text{ for } g \in L^2_0(Q). \tag{7.72}$$

Thus, L is also a bounded functional on $L^2_0(Q) \subset H^1$ with

$$|L|_{L^2_0(Q) \to \mathbb{C}} \lesssim_n |Q|^{1/2}|L|_{H^1 \to \mathbb{C}}. \tag{7.73}$$

By Riesz there exists $F^Q \in L^2_0(Q)$ such that for all $g \in L^2_0(Q)$

$$L(g) = \int_Q F^Q(x)g(x)dx, \text{ where } |F^Q|_{L^2(Q)} = |L|_{L^2_0(Q) \to \mathbb{C}}. \tag{7.74}$$

[61] See Exercise 3.1.4 of [G2]. Write $f_{KL} = \min\left(K + \max(0, f - K), L + \min(0, f - L)\right)$.

This associates to *any* cube Q a function $F^Q \in L^2_0(Q)$ such that (7.74) holds. If $Q \subset Q'$ it is easily checked that $F^Q - F^{Q'}$ is a constant on Q.

3. Now let Q_m denote the cube centered at the origin of side length m, $m = 1, 2, \ldots$. Set

$$b(x) := F^{Q_m}(x) - \frac{1}{|Q_1|}\int_{Q_1} F^{Q_m}(t)dt \text{ for } x \in Q^m.$$

This gives a well-defined function on $\mathbb{R}^n$ such that for any cube Q there exists c_Q such that

$$F^Q = b - c_Q \text{ on } Q. \tag{7.75}$$

Now (7.74) and (7.75) imply $b \in L^2_{loc}$ satisfies

$$L(g) = \int_Q b(x)g(x)dx \text{ for all } g \in L^2_0(Q). \tag{7.76}$$

This implies $L_b(g) = L(g)$ for all $g \in H^1_0$. To see that $b \in BMO$ and that (7.70) holds, show

$$\sup_Q \frac{1}{|Q|}\int_Q |b - b_Q| \lesssim_n |L|_{H^1 \to \mathbb{C}}$$

using (7.75), Cauchy-Schwarz, (7.73), and (7.74).

□

8 Almost orthogonality

For motivation let H be a Hilbert space and suppose $T : H \to H$ has the representation $T = \sum_j T_j$, where $\mathrm{Ran}(T_j) \perp \mathrm{Ran}(T_k)$ and $\mathrm{Ran}(T_j^*) \perp \mathrm{Ran}(T_k^*)$ for $j \neq k$. Thus, $T_k^* T_j = 0$ and $T_k T_j^* = 0$ for $j \neq k$. Let P_j denote orthogonal projection onto $\ker(T_j)^\perp = \overline{\mathrm{Ran} T_j^*}$ and set $P_j f = f_j$. Then for $f \in H$ we have

$$Tf = \sum_j T_j f_j \Rightarrow |Tf|^2 = \sum_j |T_j f_j|^2 \leq \sup_j |T_j|^2 \sum_j |f_j|^2 = M^2 |f|^2. \tag{8.1}$$

Thus $|T|_{H \to H} \leq M = \sup_j |T_j|$. Cotlar's lemma replaces this too strong vanishing requirement on $T_k^* T_j$ and $T_k T_j^*$ by a condition that gives sufficient "off diagonal" decay, that is, decay in $|j - k|$.

8.1 Cotlar's lemma

Lemma 8.1 (Cotlar's Lemma)**.** *Let $T_j : H \to H$, $j = 1, \ldots, N$ be operators on a Hilbert space H such that for some function $\gamma : \mathbb{Z} \to \mathbb{R}^+$ we have*

$$|T_j^* T_k| \leq \gamma^2(j-k) \text{ and } |T_j T_k^*| \leq \gamma^2(j-k)$$

for any $1 \leq j, k \leq N$. Suppose

$$\sum_{l=-\infty}^{\infty} \gamma(l) := A < \infty.$$

Then $|\sum_{j=1}^N T_j| \leq A$.

Proof. **1.** Let $T = \sum_{j=1}^{N} T_j$. Then

$$(T^*T)^n = \sum_{j_1,\dots,j_n,k_1,\dots,k_n=1}^{N} T_{j_1}^* T_{k_1} \cdots T_{j_n}^* T_{k_n}, \text{ and} \tag{8.2}$$

$$\begin{aligned}
|T_{j_1}^* T_{k_1} \cdots T_{j_n}^* T_{k_n}| &\leq |T_{j_1}||T_{k_n}| \prod_{i=1}^{n-1} |T_{k_i} T_{j_{i+1}}^*| \\
|T_{j_1}^* T_{k_1} \cdots T_{j_n}^* T_{k_n}| &\leq \prod_{i=1}^{n} |T_{j_i}^* T_{k_i}|, \text{ so} \\
|T_{j_1}^* T_{k_1} \cdots T_{j_n}^* T_{k_n}| &\leq (|T_{j_1}||T_{k_n}|)^{1/2} | \prod_{i=1}^{n-1} |T_{k_i} T_{j_{i+1}}^*|^{1/2} \prod_{i=1}^{n} |T_{j_i}^* T_{k_i}|.^{1/2}
\end{aligned} \tag{8.3}$$

We took the geometric mean of the first two estimates to get the third.

2. For any bounded operator $S : H \to H$ we have $|S|^2 = |SS^*| = |S^*S|$, so our assumptions imply $|T_j| \leq \sqrt{A}$ for all j. If $S = S^*$ we get $|S|^2 = |S^2|$ and this implies $|S|^n = |S^n|$ when n is a power of two. So $|T^*T|^n = |T|^{2n}$ when n is a power of two.

3. Let n be a power of two. Steps 1 and 2 give: $|T|^{2n} = |T^*T|^n \leq$

$$\begin{aligned}
&\sum_{j_1,\dots,j_n,k_1,\dots,k_n=1}^{N} |T_{j_1}|^{1/2} |T_{j_1}^* T_{k_1}|^{1/2} |T_{k_1} T_{j_2}^*|^{1/2} \cdots |T_{k_{n-1}} T_{j_n}^*|^{1/2} |T_{j_n}^* T_{k_n}|^{1/2} |T_{k_n}|^{1/2} \leq \\
&\sum_{j_1,\dots,j_n,k_1,\dots,k_n=1}^{N} \sqrt{A} \gamma(j_1 - k_1) \gamma(k_1 - j_2) \gamma(j_2 - k_2) \cdots \gamma(k_{n-1} - j_n) \gamma(j_n - k_n) \sqrt{A} \leq N A^{2n}.
\end{aligned} \tag{8.4}$$

For the last inequality we summed first over j_1, then over k_1, then over j_2, etc.. Thus, $|T| \leq N^{1/2n} A$. Let $n \to \infty$ to finish.

□

Remark 8.2. *The proof shows that $j - k$ can be replaced by $k - j$ (or not) wherever it occurs in the statement of the lemma.*

Lemma 8.3 (Schur's lemma)**.** *Suppose $Tf(x) = \int_Y K(x,y) f(y) d\nu(y)$. Then*

$$\begin{aligned}
&(a)|T|_{1\to 1} \leq \sup_y \int_X |K(x,y)| d\mu(x) := A \\
&(b)|T|_{\infty\to\infty} \leq \sup_x \int_Y |K(x,y)| d\nu(y) := B \\
&(c)|T|_{p\to p} \leq C(A,B) \\
&(d)|T|_{1\to\infty} \leq |K|_\infty.
\end{aligned} \tag{8.5}$$

The proof follows easily from the definitions; for (c) use interpolation.

8.2 L^2 boundedness of SIOs of convolution type

Here we give an application of Cotlar's lemma to reprove (a slightly weaker version of) an earlier result. This illustrates how the lemma can be used to avoid dependence on the Fourier transform.

Corollary 8.4 (L^2 boundedness of CZ SIOs)**.** *Let K be a CZ kernel as in Defn. 3.1 such that in addition $|\nabla K(x)| \leq B|x|^{-d-1}$. Then the associated SIO T satisfies*

$$|T|_{2\to 2} \leq CB.$$

Proof. **1.** Let $\sum_{j\in\mathbb{Z}} \psi(2^{-j}x) = 1$ be a radial, dyadic partition of unity (as in proof of Thm. 3.13). Let $K_j(x) = K(x)\psi(2^{-j}x)$ and T_j the associated operator:

$$T_j f(x) = \int K_j(x-y)f(y)dy.$$

Then with $\tilde{K}_j(x) = \overline{K_j}(-x)$ by Schur we have:

$$|T_j^* T_k|_{2\to 2} \leq |\tilde{K}_j * K_k|_1 \text{ and } |T_j T_k^*|_{2\to 2} \leq |K_j * \tilde{K}_k|_1.$$

2. We record a few properties of K_j:

$$\begin{aligned} &(a) \int K_j(x)dx = 0 \text{ and } |\nabla K_j|_\infty \lesssim 2^{-j}2^{-jd} \\ &(b) \int |K_j| \lesssim 1, \int |x||K_j(x)|dx \lesssim 2^j. \end{aligned} \tag{8.6}$$

The first part of (a) follows from the cancellation condition in Defn. 3.1 and the fact that ψ is *radial.* The other properties follow easily from size conditions and $|x| \sim 2^j$ on supp $\psi(2^{-j}x)$.

3. Estimate $|T_j^* T_k|_{2\to 2}$. Since $|T_j^* T_k| = |T_k^* T_j|$ it will suffice to estimate $|\tilde{K}_j * K_k|_{1\to 1}$ for $j \geq k$. Using both parts of (8.6) we get

$$\begin{aligned} &|(\tilde{K}_j * K_k)(x)| = |\int \overline{K_j}(y-x)K_k(y)dy| = \left|\int \left(\overline{K_j}(y-x) - \overline{K_j}(-x)\right) K_k(y)dy\right| \leq \\ &\int |\nabla K|_\infty |y||K_k(y)|dy \lesssim 2^{-j}2^{-jd}2^k. \end{aligned} \tag{8.7}$$

Since supp $\tilde{K}_j * K_k \subset B(0, C2^j)$ for some $C > 0$, we obtain $|\tilde{K}_j * K_k|_1 \lesssim 2^{k-j} = 2^{-|j-k|}$.

4. We apply Cotlar with $\gamma^2(l) = C2^{-l}$ to get

$$|\sum_{-N}^{N} T_j|_{2\to 2} \lesssim C.$$

Let $S_N = \sum_{-N}^N T_j$. Then for $f \in \mathcal{S}$, $S_N f \to Tf$ pointwise in x.[62] Now apply the standard Fatou argument of Proposition 14.3 to finish.

□

8.3 Calderon-Vaillancourt theorem

Here is another nice application of Cotlar's Lemma.

[62]Proof: Look for example at $\int_{|y|<1} K(y)g(x,y)ydy$ below (3.3), replace $K(y)$ there by $k_N(y) := \sum_{-N}^N K(y)\psi(2^{-j}y)$, and note $|k_N(y)| \leq C|K(y)|$. For each x apply the dominated convergence theorem ($N \to \infty$) to the integral $\int_{|y|<1}$ in y.

Theorem 8.5 (Calderon-Vaillancourt). *Let $a(x,\xi) \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d)$ and for $f \in \mathcal{S}$ set*

$$Tf(x) = \int e^{ix\xi} a(x,\xi)\hat{f}(\xi)d\xi,$$

where

$$\sup_{x,\xi} \left(|\partial_x^\alpha a| + |\partial_\xi^\alpha a|\right) \leq B \text{ for } |\alpha| \leq 2d+1.$$

Then $|T|_{2\to 2} \lesssim 1$.

Proof. We can take $B = 1$. Construct a "lattice partition of unity" in $\mathbb{R}^d$, $\sum_{k\in\mathbb{Z}^d} \chi(\xi - k) = 1$, for an appropriate test function χ. Set $\chi_{kl}(x,\xi) = \chi(x-k)\chi(\xi - l)$, $a_{kl} = a\chi_{kl}$, and for $f \in \mathcal{S}$ set (with $f(\xi)$ not $\hat{f}(\xi)$; that's ok by Plancherel)

$$T_{kl}f(x) = \int e^{ix\xi} a_{kl}(x,\xi)f(\xi)d\xi.$$

Using f here allows us to take the kernel of T_{kl} to be $K_{kl}(x,\xi) = e^{ix\xi}a_{kl}(x,\xi)$. Schur now gives

$$\sup_{k,l} |T_{kl}|_{2\to 2} \lesssim C. \tag{8.8}$$

Next show

$$\begin{aligned} &(a)|T^*_{k'l'}T_{k,l}|_{2\to 2} \lesssim \langle k'-k\rangle^{-2d-1}\langle l'-l\rangle^{-2d-1} \\ &(b)|T_{k'l'}T^*_{k,l}|_{2\to 2} \lesssim \langle k'-k\rangle^{-2d-1}\langle l'-l\rangle^{-2d-1}. \end{aligned} \tag{8.9}$$

Then Cotlar gives for any N that $|\sum_{|k|<N,|l|<N} T_{kl}f|_2 \lesssim |f|_2$. As $N \to \infty$ these sums converge pointwise to Tf for $f \in \mathcal{S}$, so the conclusion follows by Fatou (Prop. 14.3).

To check (8.9)(a) consider the kernel of $T^*_{kl}T_{k'l'}$:

$$K_{klk'l'}(\xi,\eta) = \int e^{-ix(\xi-\eta)}\overline{a_{kl}(x,\xi)}a_{k'l'}(x,\eta)dx.$$

This involves the product $\chi(x-k)\chi(\xi-l)\chi(x-k')\chi(\eta-l')$ so the integrand vanishes, and hence $T^*_{kl}T_{k'l'} = 0$ if $|k-k'| \geq$ some C. To prove the decay in $\langle l-l'\rangle$ we can suppose this is big, and note that $\xi \sim l$, $\eta \sim l'$ on the support of the integrand, so $|\xi - \eta| \sim |l-l'|$. Integrate by parts $2d+1$ times to generate $|\xi-\eta|^{-2d-1}$ and thereby obtain

$$|K_{klk'l'}(\xi,\eta)| \leq C\langle l-l'\rangle^{-2d-1}.$$

The support in ξ or η of $K_{klk'l'}(\xi,\eta)$ is of size $\leq C$, so Schur implies

$$|T^*_{kl}T_{k'l'}|_{2\to 2} \leq C\langle l-l'\rangle^{-2d-1}.$$

This proves (8.9)(a); (b) is similar.

$\square$

8.4 Hardy's inequality

Theorem 8.6 (Hardy's inequality)**.** *For any* $0 \le s < d/2$

$$||x|^{-s} f|_2 \lesssim |f|_{\dot{H}^s}.$$

Remark 8.7. *(a) This inequality is scale-invariant. That is, if it holds for f it holds for f_λ, where $f_\lambda(x) = f(\lambda x)$, $\lambda > 0$. We have*

$$|f_\lambda|_{\dot{H}^s} = \lambda^{s-\frac{d}{2}} |f|_{\dot{H}^s}$$

and similarly for the left side.

(b). A step in the proof's motivation is to show (for the usual dyadic operators P_k) that

$$||x|^{-s} P_k f|_2 \lesssim 2^{ks} |P_k f|_2. \tag{8.10}$$

"By scaling" it suffices to prove the case $k = 0$:

$$||x|^{-s} P_0 f|_2 \lesssim |P_0 f|_2.$$

This is because

$$\begin{aligned} ||x|^{-s} P_k f|_2 &= 2^{k(s-\frac{d}{2})} ||x|^{-s} P_0 f_{2^{-k}}|_2 \text{ and} \\ |P_k f|_2 &= 2^{-k\frac{d}{2}} |P_0 f_{2^{-k}}|_2. \end{aligned} \tag{8.11}$$

Proof of Theorem 8.6. . Assume $s > 0$. Localize dyadically in x and ξ. Let

$$\chi_j(x) = \psi(2^{-j} x), \ \widehat{P_k f} = \psi(2^{-j}\xi) \hat{f}(\xi).$$

We have, using only the triangle inequality:

$$||x|^{-s} f|_2^2 \lesssim \sum_l 2^{-2ls} |\chi_l f|_2^2 \lesssim \sum_l 2^{-2ls} \left(\sum_{k+l \le 0} |\chi_l P_k f|_2 \right)^2 + \sum_l 2^{-2ls} |\chi_l P_{>-l} f|_2^2 = A + B. \tag{8.12}$$

Next estimate

$$|\chi_l P_k f|_2 \lesssim 2^{ld/2} |P_k f|_\infty \lesssim 2^{d(k+l)/2} |P_k f|_2, \tag{8.13}$$

where the first inequality is trivial and the second uses Bernstein (14.4). Substitute (8.13) to get

$$A \lesssim \sum_l \left(\sum_{k+l \le 0} 2^{(\frac{d}{2}-s)(l+k)} 2^{sk} |P_k f|_2 \right)^2 \lesssim \sum_k 2^{2ks} |P_k f|_2^2 \lesssim |f|_{\dot{H}^s}^2, \tag{8.14}$$

where we used Schur for sums. For B use almost orthogonality and change order of summation to get

$$B \lesssim \sum_l 2^{-2ls} |P_{>-l} f|_2^2 \lesssim \sum_l \sum_{k > -l} 2^{-2(l+k)s} 2^{2sk} |P_k f|_2^2 \lesssim \sum_k 2^{2sk} |P_k f|_2^2 \sum_{l > -k} 2^{-2(l+k)s} \lesssim |f|_{\dot{H}^s}^2. \tag{8.15}$$

□

Remark 8.8. *See [CS1] for discussion of motivation for the proof and connections to the uncertainty principle. In particular the decision to use Bernstein for $k + l \le 0$ and almost orthogonality for $k + l \ge 0$ is motivated by considering the support of $\hat{\chi}_l * \widehat{P_k f}$. The support of $\hat{\chi}_l$ is near 2^{-l} while that of $\widehat{P_k f}$ is near 2^k. For fixed l and $k >> -l$ the supports of the $\hat{\chi}_l * \widehat{P_k f}$ will be separated.*

9 Singular integral operators not of convolution type

Definition 9.1. *A* kernel *on $\mathbb{R}^n$ is a function $K : \mathbb{R}^n \times \mathbb{R}^n \setminus \{x = y\} \to \mathbb{C}$, which is locally integrable away from the diagonal. A kernel on $\mathbb{R}^n$ is said to satisfy* standard estimates *if there exist positive constants C, δ such that for all distinct x, y and all z such that $|x - y| > 2|x - z|$ we have*

(i) $|K(x, y)| \le A|x - y|^{-n}$

(ii) $|K(x, y) - K(z, y)| \le A\frac{|x-z|^\delta}{|x-y|^{n+\delta}}$

(iii) $|K(y, x) - K(y, z)| \le A\frac{|x-z|^\delta}{|x-y|^{n+\delta}}$.

Given a standard kernel we can't associate an operator to it simply by setting $Tf(x) = \int K(x, y)f(y)dy$. This makes no sense when $K = |x - y|^{-n}$ for example.

Definition 9.2. *A continuous linear operator $T : \mathcal{D} \to \mathcal{D}'$ is said to be* associated to a kernel K *if for all test functions f, g with disjoint supports,*

$$\langle Tf, g\rangle = \int\int K(x, y)f(y)g(x)dydx. \tag{9.1}$$

Remark 9.3. *(a) By the Schwartz kernel theorem a continuous linear operator $T : \mathcal{D} \to \mathcal{D}'$ is* associated to a kernel K *if and only if its Schwartz kernel $W \in \mathcal{D}'(\mathbb{R}^n \times \mathbb{R}^n)$ satisfies*

$$\langle Tf, g\rangle = \langle W, f \otimes g\rangle$$

for all test functions f, g. We say that the distribution W coincides with K *off the diagonal.*

(b) The formula (9.1) *still holds when f and g are any L^∞ functions with disjoint compact supports. The routine proof uses approximation in L^2 of f and g by test functions; see [G2].*

(c) The last two properties of Defn. 9.1 are implied by the stronger but simpler condition $|\nabla K(x, y)| \le C|x - y|^{-n-1}$.

Example 9.4. *An important special case is when K is antisymmetric, that is,*

$$K(y, x) = -K(x, y) \text{ when } x \ne y.$$

Then the integral

$$\langle Tf, g\rangle = \frac{1}{2}\int\int K(x, y)[f(y)g(x) - f(x)g(y)]dydx$$

converges absolutely for any *test functions f and g, and agrees with the previous formula when f, g have disjoint supports. In this case we call T the operator* canonically associated to K.

Definition 9.5. *A* singular integral operator *is a continuous linear map from $\mathcal{D}$ to $\mathcal{D}'$ which is associated to a standard kernel.*

Observe that the operators on $\mathbb{R}$ given by $f \to f'$ and $f \to af$, where $a \in L^\infty$ are both associated to $K = 0$.

Definition 9.6. *A singular integral operator T is said to be* bounded *if it extends to a bounded operator on $L^2(\mathbb{R}^n)$. Such operators are called (by [G2] for example) Calderon-Zygmund operators.*

Definition 9.7 (Action of singular integral operators on $f \in C^\infty \cap L^\infty$ and $f \in L^\infty$)**.**

Let T be a singular integral operator (Defn. 9.5).

(a) Let $\mathcal{D}_0$ be the set of test functions with integral zero. For $f \in C^\infty \cap L^\infty$ we define $Tf \in \mathcal{D}_0'$ as follows. Let $\phi \in \mathcal{D}_0$, let B be a ball containing the support of ϕ, and let $5B$ denote the obvious

concentric ball. Write $f = f\chi_{5B} + f(1-\chi_{5B})$. Then $Tf\chi_{5B}$ is already defined. Let $x_0 \in \operatorname{supp}\phi$ and set

$$\langle Tf(1-\chi_{5B}), \phi\rangle := \langle f(1-\chi_{5B}), T^t\phi\rangle = \int_{|y-x|>2|x-x_0|}\int_{x\in B}(K(x,y)-K(x_0,y))f(y)\phi(x)dxdy.$$

For (x,y) in the domain of integration we have

$$|K(x,y)-K(x_0,y)| \lesssim \frac{|x-x_0|^\delta}{|x-y|^{n+\delta}} \lesssim |y-x_0|^{-n-\delta} \lesssim (1+|y|)^{-n-\delta},$$

so the integral is absolutely convergent.

(b) One checks this definition is well-defined and consistent with the earlier definition when $f \in \mathcal{D}$. Moreover, if T is a CZ operator, i.e., if it has a bounded extension to L^2, then this definition works for any $f \in L^\infty$, since $f\chi_{5B} \in L^2$. Moreover, the same estimates show that if $f \in L^\infty$ and $\phi \in L^2_c$ with integral zero, then

$$|\langle Tf,\phi\rangle| \lesssim |T(f\chi_{5B})|_2|\phi|_2 + |f|_\infty|\phi|_1. \tag{9.2}$$

Remark 9.8. *(a) Functions $f \in BMO$ (which are defined up to constants) define elements of $\mathcal{D}_0'$.*

(b). CZ operators have bounded extensions that are weak (1,1) and strong (p,p) for $1 < p < \infty$. The proof, which uses the CZ decomposition, is similar to the proof in the convolution type case; see section 3.6 or [G2], Thm. 4.2.2.

Theorem 9.9. *Let T be a CZ operator (so bounded on L^2); in [G2] notation $T \in CZO(\delta, A, B)$, where $B = \|T\|_{2\to 2}$. Then T has an extension satisfying*

$$|Tf|_{L^1} \le C_{n,\delta}(A+B)|f|_{H^1}.$$

The proof ([G2], p232) first considers L^2 atoms in H^1. Then express $f \in H^1$ in terms of atoms.

Theorem 9.10. *Let $T \in CZO(\delta, A, B)$. Then for any $f \in L^\infty$, Tf can be identified with a BMO function such that*

$$|Tf|_{BMO} \le C_{n,\delta}(A+B)|f|_{L^\infty}.$$

Proof. Let $L^2_{0,c}$ denote the set of compactly supported L^2 functions with integral zero. It is dense in H^1, since it contains L^2 atoms.

The main step is to show that for all $f \in L^\infty$ and $\phi \in L^2_{0,c}$ we have

$$\langle Tf,\phi\rangle = \int T^t\phi(x)f(x)dx. \tag{9.3}$$

In that case one has using Theorem (9.9)

$$|\langle Tf,\phi\rangle| \le |T^t\phi|_{L^1}|f|_{L^\infty} \lesssim |\phi|_{H^1}|f|_{L^\infty},$$

and the result follows easily since $(H^1)^* = BMO$; see Theorem 3.2.2 of [G2]. The proof of (9.3) uses Definition 9.7 and Remark 9.3(b).

□

9.1 Carleson measures and BMO

In preparation for the $T(1)$ theorem we define Carleson measures and discuss how they can arise from BMO functions.

Definition 9.11 (Carleson measures)**.** *Let $B = B(x_0, R) \subset \mathbb{R}^n$. Define the (cylindrical) tent over B to be*

$$T(B) = \{(x,t) : x \in B, 0 < t \leq R.\}$$

A positive measure μ on $\mathbb{R}^{n+1}_+$ is called a Carleson measure *if*

$$\|\mu\| := \sup_B \frac{\mu(T(B))}{|B|} < \infty,$$

where the sup is over all balls in $\mathbb{R}^n$.

Remark 9.12. *Lebesgue measure on $\mathbb{R}^{n+1}_+$ is not a Carleson measure. Let P be any plane in $\mathbb{R}^{n+1}$. For $A \subset \mathbb{R}^{n+1}_+$ define $\mu(A)$ to be the $n-$dimensional Lebesgue measure of $A \cap P$. Then μ is a Carleson measure.*

The next theorem is used, for example, in estimating paraproducts and in the proof of the $T1$ theorem; see [G2], Thm. 3.3.7.

Theorem 9.13. *Let ϕ on $\mathbb{R}^n$ satisfy*

$$|\phi(x)| \leq A(1+|x|)^{-n-\delta}, \text{ for some } 0 < A, \delta < \infty$$

and let μ be a Carleson measure. Then

$$\int_{\mathbb{R}^{n+1}_+} |(\phi_t * f)(x)|^p d\mu \lesssim_{n,p} \|\mu\| \int |f|^p dx.$$

For the proof we need some preparation.

Definition 9.14. *For a measurable function u on $\mathbb{R}^{n+1}_+$ we define the* nontangential maximal function

$$x \to Nu(x) = \sup_{|y-x|<t} |u(y,t)|.$$

Remark 9.15. *For ϕ and f as in Theorem 9.13, if $u(x,t) = \phi_t * f(x)$ then*

$$N(x) = \sup_{t>0} \sup_{|y-x|<t} |\phi_t * f(y)| \lesssim Mf(x).$$

To see this use Cor. 2.1.12 of [G1] and Ex. 3.3.4 of [G2]. In Ex. 3.3.4 note that

$$\psi_t(x-u) \geq \phi_t(y-u) \text{ for } |y-x| < t,$$

since $y - u = x - u - (x-y)$.[63]

[63] Here our ϕ is the Φ in Ex. 3.3.4.

Proof of Theorem 9.13. Let $u(x,t) = \phi_t * f(x)$. The result follows directly from

$$\mu\{(x,t) : |u(x,t)| > \lambda\} \lesssim \|\mu\| |\{x : Nu > \lambda\}| \tag{9.4}$$

and Remark 9.15.[64]

Indeed, if $|u(y,t)| > \lambda$ then $B(y,t) \subset \{x : Nu > \lambda\}$, and $\{x : Nu > \lambda\}$ is the union of all such balls. If $|\{x : Nu > \lambda\}|$ we can apply the Vitali covering lemma to obtain pairwise disjoint balls $B_j = B(y_j, t_j)$ such that $\{Nu > \lambda\} \subset \cup_j B(y_j, 3t_j)$. If $|u(y,t)| > \lambda$, then $B(y,t)$ must intersect some B_j with $t_j \geq t/2$.[65] Thus, $(y,t) \in B(y_j, 3t_j) \times (0, 2t_j) := T_j$, so

$$\{(x,t) : |u(x,t)| > \lambda\} \subset \cup_j T_j.$$

Since μ is a Carleson measure this implies

$$\mu(\{|u(x,t)| > \lambda\}) \leq \sum_j \mu(T_j) \lesssim \|\mu\| \sum_j |B_j| \lesssim \|\mu\| |Nu > \lambda|.$$

$\square$

The next theorem provides a class of Carleson measures that will be used in proving the $T1$ theorem.

Theorem 9.16 (G2). *Let $b \in BMO(\mathbb{R}^n)$ and let ψ be a measurable function on $\mathbb{R}^n$ that satisfies*

$$|\psi(x)| + |\nabla\psi(x)| \leq A(1 + |x|)^{-n-\delta}, \text{ for some } 0 < A, \delta < \infty, \tag{9.5}$$

and which has mean zero. For $t > 0$ set $\psi_t(x) := t^{-n}\psi(x/t)$. Then the measure μ on $\mathbb{R}^{n+1}_+$ defined by

$$d\mu(x,t) = |(\psi_t * b)(x)|^2 dx \frac{dt}{t} \tag{9.6}$$

is a Carleson measure and $\|\mu\| \lesssim_{\delta,n,A,\psi} |b|^2_{BMO}$.

The proof below uses Plancherel and

$$\int_0^\infty |\hat{\psi}(t\xi)|^2 \frac{dt}{t} \leq B < \infty, \tag{9.7}$$

where B is uniform in $\xi \neq 0$.[66] This gives

$$\left| \int_0^\infty |(\psi_t * g)(x)|^2 \frac{dt}{t})^{1/2} \right|_2 \lesssim |g|_2. \tag{9.8}$$

We also need the following lemma.

Lemma 9.17. *Let $f \in BMO(\mathbb{R}^n)$. For any ball B and $m \in \mathbb{N}$ we have*

(a) $|f_{2^m B} - f_B| \leq 2^n m |f|_{BMO}$.

(b) For any $\delta > 0$ there exists $C_{n,\delta}$ such that for all x_0 and $R > 0$,

$$R^\delta \int \frac{|f(y) - f_{B(x_0,R)}|}{(R + |y - x_0|)^{n+\delta}} dy \leq C_{n,\delta} |f|_{BMO}.$$

[64] Here we follow [C1].

[65] This comes from the Vitali selection criterion of [S1], e.g.

[66] If ψ is radial, B is independent of $\xi \neq 0$ by a scaling argument; consider $\hat{\psi}(tr\xi)$ for some $r > 0$, set $s = tr$, and so on.

Proof. (b) With $B = B(0,1)$ reduce to

$$(9.9)\quad \int \frac{|f(y) - f_{B(0,1)}|}{(1+|y|)^{n+\delta}}dy \leq \int_B \frac{|f(y) - f_{B(0,1)}|}{(1+|y|)^{n+\delta}}dy + \sum_0^\infty \int_{2^{k+1}B \setminus 2^k B} \frac{|f - f_{2^{k+1}B}| + |f_{2^{k+1}B} - f_B|}{(1+|y|)^{n+\delta}}dy.$$

Use part (a); the $2^{-k\delta}$ permits summing over k in the second term on the right.

□

Proof of Theorem 9.16. Let $Q_t f := \psi_t * f$ and for any cube Q let Q^* be the concentric cube with side length $3\sqrt{n}\,\ell(Q)$. Write

$$b = (b - b_Q)\chi_{Q^*} + (b - b_Q)\chi_{Q^{*c}} + b_Q,$$

and note $Q_t b_Q = 0$. Thus

$$(9.10)\quad \begin{aligned} &\mu(T(Q)) = \int_0^{\ell(Q)} \int_Q |Q_t b(x)|^2 dx \frac{dt}{t} \leq \\ &\int_0^\infty \int_{\mathbb{R}^n} |Q_t(b - b_Q)\chi_{Q^*}(x)|^2 dx \frac{dt}{t} + \int_0^{\ell(Q)} \int_Q |Q_t(b - b_Q)\chi_{Q^{*c}}(x)|^2 dx \frac{dt}{t} = A + B. \end{aligned}$$

By Plancherel and (9.7) we have

$$A \lesssim \int_{Q^*} |b - b_Q|^2 dx \lesssim |b|_{BMO}^2 |Q|.$$

Use Lemma 9.17(a) and Prop. 7.6 to get the second inequality. To treat B note first that

$$|Q_t(b - b_Q)\chi_{Q^{*c}}(x)| \lesssim \int_{Q^{*c}} \frac{t^\delta |b - b_Q|}{(t + |x-y|)^{n+\delta}} dy$$

and $t + |x - y| \geq |x - y| \geq \frac{1}{2}\left(|c_Q - y| + \frac{\sqrt{n}}{2}\ell(Q)\right)$ for $x \in Q$, $y \in Q^{*c}$. Substituting into B gives

$$(9.11)\quad B \lesssim \int_0^{\ell(Q)} t^{2\delta - 1} \int_Q \left(\int_{\mathbb{R}^n} \frac{|b(y) - b_Q|}{(\ell(Q) + |c_Q - y|)^{n+\delta}} dy \right)^2 dx dt \lesssim |Q||b|_{BMO}^2.$$

To get the last inequality we integrated in t and used Lemma 9.17(b).

□

9.2 Continuous paraproducts

Let $\phi, \psi \in \mathcal{S}(\mathbb{R}^n)$ with $\int \phi = 1, \int \psi = 0$. With $\phi_t(x) = t^{-n}\phi(x/t)$ set

$$P_t f = f * \phi_t, \quad Q_t f = f * \psi_t.$$

Definition 9.18. *Formally define the (continuous) paraproduct*

$$\pi(f, b) = \pi_b(f) := \int_0^\infty Q_t(Q_t b \cdot P_t f) \frac{dt}{t}.$$

Remark 9.19. *(a). The Fourier transform of $Q_t(Q_tb \cdot P_tf)$ at ξ is $H_t(\xi) := \hat{\psi}(t\xi)[\widehat{(\psi_t * b)} * \widehat{(\phi_t * f)}](\xi)$. Since the outside factor $\hat{\psi}(t\xi)$ is concentrated near $|\xi| \sim t^{-1}$, in the $d\eta$ integral defining $[\widehat{(\psi_t * b)} * \widehat{(\phi_t * f)}](\xi)$ the $\eta-$frequencies of f that contribute to $H_t(\xi)$ are lower than the $(\xi - \eta)-$frequencies of b with which they are "paired". We will compare this later to analogous "discrete" paraproducts.*

(b) The outside factor of Q_t enforces a sort of almost orthogonality among the contributions to the paraproduct. This is used in the proof of Lemma 9.23 below, and will be more clear when we look at the analogue of this lemma for discrete paraproducts in section 9.5.

The next lemma shows that when $b \in BMO$, $|\pi_b|_{2\to 2} \lesssim |b|_{BMO}$. Note that $f \to bf$ generally fails to be bounded on L^2.

Lemma 9.20. *If $f, g \in L^2$ and $b \in BMO$, then*

$$\lim_{\epsilon\to 0, N\to\infty} \left\langle \int_\epsilon^N Q_t(Q_tb \cdot P_tf)\frac{dt}{t}, g \right\rangle$$

exists and the inner product is $\lesssim |b|_{BMO}|f|_2|g|_2$ uniformly wrt ϵ, N.

Proof. We have

(9.12)
$$\left|\left\langle \int_\epsilon^N Q_t(Q_tb \cdot P_tf)\frac{dt}{t}, g \right\rangle\right| \leq \left\langle \left(\int_\epsilon^N |Q_t^t g|^2 \frac{dt}{t}\right)^{1/2}, \left(\int_\epsilon^N |Q_tb \cdot P_tf|^2 \frac{dt}{t}\right)^{1/2} \right\rangle \lesssim |b|_{BMO}|f|_2|g|_2.$$

Here we used that $|Q_tb|^2\, dx\frac{dt}{t}$ is a Carleson measure, Theorem 9.16, as well as Theorem 9.13.

The existence of the limit follows since as $\epsilon \to 0$, $N \to \infty$,

$$\int\int_{t\notin[\epsilon,N]} |Q_t^t g|^2 \frac{dt}{t} dx \to 0.$$

□

Remark 9.21. *Similarly for $p \in (1, \infty)$ π_b extends to a map satisfying $|\pi_b|_{p\to p} \lesssim |b|_{BMO}$. The proof uses L^2 boundedness and the fact that π_b has a standard kernel.*

Lemma 9.22. *If $b \in BMO$ then the kernel of π_b is associated to a standard kernel.*[67]

Proof. We use $|Q_tb|_\infty \lesssim |b|_{BMO}$ (see the proof of Prop. 9.33). The kernel is

$$K(x, y) = \int_0^\infty \int \psi_t(x - z)Q_tb(z)\phi_t(z - y)dz\frac{dt}{t},$$

so

$$|K(x, y)| \lesssim |b|_{BMO} \int_0^\infty \int t^{-n}(1+t^{-1}|x-z|)^{-n-1} \times t^{-n}(1+t^{-1}|z-y|)^{-n-1}dz\frac{dt}{t} = C|b|_{BMO}|x-y|^{-n},$$

where the last estimate is proved using the change of variable $s = \frac{z-y}{t}$. □

The next proposition describes mapping properties of π_g for $g \in H^s$. Again, the analogue fails for ordinary products. A similar result holds for the scale of Hölder spaces or L^p Sobolev spaces.

[67]See section 9.5 for the discrete analogue.

Proposition 9.23. *For all $s \geq 0$, $f \in L^\infty$, and $g \in H^s$ one has*

$$|\pi_g(f)|_{H^s} \lesssim |f|_\infty |g|_{H^s}.$$

Proof. We have

$$|\pi_g(f)|^2_{H^s} \lesssim \int_0^\infty t^{-2s}|Q_t g \cdot P_t f|_2^2 \frac{dt}{t} \lesssim \int_0^\infty |P_t f|^2_\infty t^{-2s}|Q_t g|_2^2 \frac{dt}{t} \lesssim |f|^2_\infty |g|^2_{H^s}. \tag{9.13}$$

For the first inequality we used Plancherel and the boundedness of Q_t on L^2. □

9.3 The $T(1)$ theorem

Definition 9.24. *A normalized bump is a smooth function supported in $B(0, 10)$ such that*

$$|\partial^\alpha \phi(x)| \leq 1 \text{ for all } |\alpha| \leq 2[n/2] + 2. \tag{9.14}$$

For $x_0 \in \mathbb{R}^n$, $R > 0$, set

$$\tau^{x_0} f_R(x) = R^{-n} f\left(\frac{x - x_0}{R}\right). \tag{9.15}$$

Definition 9.25. *A singular integral operator T is* weakly bounded *if there exists $C > 0$ such that for all normalized bumps f, g, all $x_0 \in \mathbb{R}^n$, and all $R > 0$ we have*

$$|\langle T(\tau^{x_0} f_R), \tau^{x_0} g_R\rangle| \leq C R^{-n}. \tag{9.16}$$

Denote the smallest constant C in (9.16) by $|T|_{wb}$.

Example 9.26. *The operator T canonically associated to an antisymmetric kernel is weakly bounded. Write*

$$\phi(x)\psi(y) - \phi(y)\psi(x) = \phi(x)[\psi(y) - \psi(x)] + \psi(x)[\phi(x) - \phi(y)] = O(|x - y|).$$

Observe that if T is bounded on L^2, then it is weakly bounded; for this use

$$|\tau^{x_0} f_R|_2 = R^{-n/2}|f|_2. \tag{9.17}$$

Theorem 9.27. *Let T be a singular integral operator associated to a standard kernel (Defn. 9.5). If T is weakly bounded and satisfies $T(1) \in BMO$, $T^t(1) \in BMO$, then T extends to a bounded operator on L^2.*

The necessity of these conditions for L^2 boundedness has been established; recall Theorem 9.10.

Proof of Theorem 9.27. **1.** The strategy is to decompose T as

$$T = \pi_b + (\pi_{b'})^t + S,\ b = T(1),\ b' = T^t(1),$$

where the π operators are essentially paraproducts and S is associated to a "quasiclassical kernel", which is relatively easy to handle.[68] Our presentation follows both [C1] and [G2], but differs from [G2] by showing that the proof there amounts to using this decomposition, and differs from [C1] by giving a much simpler treatment (as in [G2]) of S - no need for Cotlar's lemma; see also Remark 9.28.

[68] The operator π_b, for example, is actually a sum of n operators each of which is essentially a paraproduct; this is clarified below.

The "paraproducts" are estimated as in Lemma 9.20 using the above results on Carleson measures, namely, Theorems 9.13 and 9.16.

2. Let ϕ be a smooth radial function supported in $B(0,1/2)$ such that $\int \phi = 1$. Let

$$P_t f = \phi_t * f, \text{ where } \phi_t(x) = t^{-n}\phi(x/t), \text{ so } \widehat{\phi_t}(\xi) = \hat{\phi}(t\xi).$$

Let $f \in \mathcal{S}$ have $\hat{f}$ supported away from a neighborhood of 0. Observe that

$$\begin{aligned} Tf &= \lim_{s\to 0} P_s^2 T P_s^2 f \text{ in } \mathcal{S}' \\ 0 &= \lim_{s\to\infty} P_s^2 T P_s^2 f \text{ in } \mathcal{S}'/\mathcal{P}. \end{aligned} \tag{9.18}$$

So we have in $\mathcal{S}'/\mathcal{P}$:

$$\begin{aligned} Tf = &-\lim_{\epsilon\to 0}\int_\epsilon^{1/\epsilon} s\frac{d}{ds}(P_s^2 T P_s^2 f)\frac{ds}{s} = \\ &-\lim_{\epsilon\to 0}\int_\epsilon^{1/\epsilon}\left[s(\frac{d}{ds}P_s^2)TP_s^2 f\frac{ds}{s} + P_s^2 T(s\frac{d}{ds}P_s^2)f\right]\frac{ds}{s} \end{aligned} \tag{9.19}$$

Take the Fourier transform in x, compute the $s-$derivatives, then go back to x variables to rewrite Tf as

$$-\lim_{\epsilon\to 0}\sum_{k=1}^n\left[\int_\epsilon^{1/\epsilon}\tilde{Q}_{k,s}(Q_{k,s}TP_s)P_s f\frac{ds}{s} + \int_\epsilon^{1/\epsilon}P_s(P_s T Q_{k,s})\tilde{Q}_{k,s} f\frac{ds}{s}\right], \tag{9.20}$$

where (ignoring harmless constants)

$$\begin{aligned} Q_{k,s}g &= g * (\psi_k)_s, \text{ where } \psi_k = 2D_{x_k}\phi, \\ \tilde{Q}_{k,s}g &= g * (\theta_k)_s, \text{ where } \theta_k = x_k\phi. \end{aligned} \tag{9.21}$$

The functions ψ_k, θ_k are smooth, odd, bumps supported in $B(0,1/2)$ with integral zero. Their oddness implies

$$(Q_{k,s})^t = -Q_{k,s} \text{ and } (\tilde{Q}_{k,s})^t = -\tilde{Q}_{k,s}, \text{ while } (P_s)^t = P_s. \tag{9.22}$$

Set

$$T_{k,s} := Q_{k,s}TP_s \text{ and note } P_s T Q_{k,s} = -((T^t)_{k,s})^t. \tag{9.23}$$

Dropping k, we will estimate one of the terms in the first sum of (9.20),

$$\int_0^\infty \tilde{Q}_s T_s P_s f\frac{ds}{s}. \tag{9.24}$$

Terms in the second sum can be handled similarly using (9.23).

3. Kernel of T_s. For $h \in \mathcal{S}$ we can write (dropping k):

$$(T_s h)(x) = \int \langle T(\tau^y\phi_s), \tau^x\psi_s\rangle h(y)dy, \tag{9.25}$$

Thus, T_s has kernel (use $\int\psi = 0$ for the second equality)

$$K_s(x,y) = \langle T(\tau^y\phi_s), \tau^x\psi_s\rangle = \int\int \phi_s(v-y)[K(u,v) - K(x,v)]\psi_s(u-x)dudv. \tag{9.26}$$

We claim (for A and δ as in the standard kernel estimates):

$$|K_s(x,y)| \lesssim_{n,\delta} (|T|_{wb} + A)p_s(x-y), \text{ where } p(u) = (1+|u|)^{-n-\delta}. \tag{9.27}$$

When $|x-y| \leq 5s$ one uses weak boundedness. Otherwise $\tau^y \phi_s$, $\tau^x \psi_s$ have disjoint supports and one uses the integral in (9.26) along with the second condition in Definition 9.1.[69]

Observe also that since $\int \phi = 1$ we have

$$T_s(1)(x) = \int K_s(x,y)dy = \langle T^t(\tau^x \psi_s), 1\rangle = (\psi_s * T(1))(x) = Q_s b(x). \tag{9.28}$$

4. The intuition for this step ([G2]) is that T_s is an averaging operator at scale s, and $P_s f$, being an average on that scale, is essentially constant on that scale. So $T_s P_s f$ should be close to $P_s f T_s(1)$; the latter term will give rise to a Carleson measure. We write

$$T_s P_s f = T_s(1) P_s f + [T_s P_s f - T_s(1) P_s f]. \tag{9.29}$$

The contribution from the second term will be estimated using (9.27). Using (9.29) and reinstating k, we write

$$\int_0^\infty \tilde{Q}_{k,s} T_{k,s} P_s f \frac{ds}{s} = \int_0^\infty \tilde{Q}_{k,s} T_{k,s}(1) P_s f \frac{ds}{s} + \int_0^\infty \tilde{Q}_{k,s}[T_{k,s} P_s f - T_{k,s}(1) P_s f] \frac{ds}{s}. \tag{9.30}$$

Set

$$\begin{aligned} \pi_{k,b}(f) &= \int_0^\infty \tilde{Q}_{k,s} T_{k,s}(1) P_s f \frac{ds}{s} \text{ and } \pi_b(f) = \sum_{k=1}^n \pi_{k,b}(f) \\ S_{k,1}(f) &= \int_0^\infty \tilde{Q}_{k,s}[T_{k,s} P_s f - T_{k,s}(1) P_s f] \frac{ds}{s} \text{ and } S_1(f) = \sum_{k=1}^n S_{k,1}(f). \end{aligned} \tag{9.31}$$

Noting that

$$\langle \int_0^\infty P_s(P_s T Q_{k,s}) \tilde{Q}_{k,s} f \frac{ds}{s}, g\rangle = \langle f, \int_0^\infty \tilde{Q}_{k,s}(T^t)_{k,s} P_s g \frac{ds}{s}\rangle, \tag{9.32}$$

we similarly define

$$\begin{aligned} \pi_{k,b'}(g) &= \int_0^\infty \tilde{Q}_{k,s}(T)^t_{k,s}(1) P_s g \frac{ds}{s} \text{ and } \pi_{b'}(g) = \sum_{k=1}^n \pi_{k,b'}(g) \\ S_{k,2}(g) &= \int_0^\infty \tilde{Q}_{k,s}[(T)^t_{k,s} P_s g - (T)^t_{k,s}(1) P_s g] \frac{ds}{s} \text{ and } S_2(g) = \sum_{k=1}^n S_{k,2}(g). \end{aligned} \tag{9.33}$$

From (9.28) and Definition 9.18, we see that $\pi_{k,b}$ and $\pi_{k,b'}$ are essentially paraproducts. Setting $S = S_1 + S_2^t$ we have

$$T = \pi_b + (\pi_{b'})^t + S. \tag{9.34}$$

Applying Lemma 9.20 we have

$$|\langle \pi_b(f), g\rangle| \lesssim |b|_{BMO} |f|_2 |g|_2 \tag{9.35}$$

[69] For more detail see [G2], p.244.

for $g \in \mathcal{S}_0$ (i.e., $g \in \mathcal{S}$ and $\hat{g}$ vanishes to infinite order at 0), and $f \in \mathcal{S}$ such that supp $\hat{f}$ is disjoint from a neighborhood of 0. The operator $(\pi_{b'})^t$ satisfies the same estimate, so it remains to estimate $\langle Sf, g\rangle$.

5. We estimate using Cauchy-Schwartz, (9.8), and (9.27):

$$|\langle S_{k,1}f, g\rangle| \lesssim |g|_2 \left(\int_0^\infty \int\int p_s(x-y)|P_sf(y) - P_sf(x)|^2 dydx\frac{ds}{s}\right)^{1/2}. \tag{9.36}$$

Setting $u = x - y$ and using Plancherel, the square root becomes:

$$\left(\int_0^\infty \int\int p_s(u)|\hat{\phi}(s\xi)(1-e^{2\pi iu\xi})|^2|\hat{f}(\xi)|^2 dud\xi\frac{ds}{s}\right)^{1/2}. \tag{9.37}$$

Since

$$|(1-e^{2\pi iu\xi})|^2 \lesssim |u|^{\delta/2}|\xi|^{\delta/2} = |u/s|^{\delta/2}|s\xi|^{\delta/2},$$

the integrals in s and u produce harmless constants.[70] We conclude

$$|\langle S_{k,1}f, g\rangle| \lesssim |f|_2|g|_2.$$

Functions satisfying the conditions on f and g are dense in L^2, so this concludes the proof.

□

Remark 9.28. *(a) Since our π_b is not* exactly *a paraproduct in the sense of Definition 9.18 with $\int_0^\infty Q_t^2\frac{dt}{t} = I$ as in [C1], we don't (obviously) have $\pi_b(1) = b$, $\pi_{b'}(1) = b'$ as in [C1]. On the other hand $S(1) = 0$ and $S^t(1) = 0$ (use $\int\psi = 0$), so*

$$\begin{aligned} T(1) &= b = \pi_b(1) + \pi_{b'}^t(1) \\ T^t(1) &= b' = \pi_b^t(1) + \pi_{b'}(1). \end{aligned} \tag{9.38}$$

But the transpose terms in (9.38) *are both zero (again use $\int\psi = 0$), so we do have $\pi_b(1) = b$, $\pi_{b'}(1) = b'$, as in [C1].*

(b) The proof of the $T(1)$ given here using continuous paraproducts was obtained by Coifman and Meyer. The original proof of David and Journé used discrete paraproducts and an explicit decomposition like $T = \pi_b + (\pi_{b'})^t + S$; see Remark 9.34.

9.4 Application: the Cauchy integral on a Lipschitz curve

Theorem 9.29 (Calderon, 1977)**.** *Let A be a Lipschitz function on $\mathbb{R}$.[71] The operator*

$$C_Af(x) = \lim_{\epsilon\to 0}\int_{|x-y|>\epsilon}\frac{f(y)}{x-y+i[A(x)-A(y)]}dy$$

is bounded on L^2 when $|A'|_{L^\infty}$ is small enough.

Proof. Write

$$\frac{1}{x-y+i[A(x)-A(y)]} = \frac{1}{x-y}\sum_{m=0}^\infty(-i)^m\left(\frac{A(x)-A(y)}{x-y}\right)^m.$$

[70] Here we take advantage of cancellation. Is almost orthogonality hidden here?

[71] Thus, A is differentiable a.e.

The operator with kernel

$$K_m(x,y = \left(\frac{A(x)-A(y)}{x-y}\right)^m \frac{1}{x-y}$$

is called the $m-th$ *Calderon commutator* and denoted $\mathcal{C}_m$. If L is the Lipschitz constant of A, we show by induction that there exists $R>0$ such that for all m

$$|\mathcal{C}_m|_{2\to 2} \leq R^m L^m. \tag{9.39}$$

The result then follows if $L < R^{-1}$.

Boundedness of the Hilbert transform gives the case $m=0$. Assume (9.39) for m. One shows ([G2], Example 4.1.7) that K_m is a standard kernel with $\delta=1$, $A=8(2m+1)L^m$. Thus, by Theorem 9.10 and assumption (9.39) we have

$$|\mathcal{C}_m|_{L^\infty\to BMO} \lesssim 8(2m+1)L^m + |\mathcal{C}_m|_{2\to 2}. \tag{9.40}$$

By the $T(1)$ theorem

$$||\mathcal{C}_{m+1}|_{2\to 2}|_{2\to 2} \lesssim |\mathcal{C}_{m+1}(1)|_{BMO} + |\mathcal{C}^t_{m+1}(1)|_{BMO} + |\mathcal{C}_{m+1}(1)|_{wb}, \tag{9.41}$$

where $|\mathcal{C}_{m+1}(1)|_{wb} \lesssim 8(2m+1)L^m$ since K_m is an *antisymmetric* standard kernel; see Example 9.26.

To complete the induction step we use $\mathcal{C}_{m+1}(1) = \mathcal{C}_m(A')$, which is almost formally obvious. See Exercise 4.3.4 of [G2]. By (9.40) we get

$$|\mathcal{C}_m(A')|_{BMO} \lesssim |A'|_\infty[8(2m+1)L^m + R^m L^m] \leq L\cdot[8(2m+1)L^m + R^m L^m],$$

and this allows us to modify R if necessary to complete the induction step.

$\square$

9.5 Discrete paraproducts

Let $\psi \in \mathcal{S}(\mathbb{R}^n)$ be radial with $\hat{\psi}$ supported in $1/2 \leq |\xi| \leq 2$ such that

$$\sum_{j\in\mathbb{Z}} \hat{\psi}(2^{-j}\xi) = 1 \text{ for } \xi \neq 0. \tag{9.42}$$

As usual set $\Delta_j f = f * \psi_{2^{-j}}$; Δ_j is self-transpose since ψ is even. Set

$$\hat{\phi}(\xi) = \begin{cases} \sum_{j\leq 0} \hat{\psi}(2^{-j}\xi),\ \xi \neq 0 \\ 1,\ \xi = 0 \end{cases}. \tag{9.43}$$

Then $\hat{\phi}$ is a test function supported in $|\xi| \leq 2$ and equal to one on $|\xi| \leq 1/2$. Set

$$S_j = \sum_{k\leq j} \Delta_j, \text{ so } S_j f(x) = (f * \phi_{2^{-j}})(x). \tag{9.44}$$

Definition 9.30. *For an appropriate functions g, f define*

$$P_g(f) = \sum_{j\in\mathbb{Z}} \Delta_j g \cdot S_{j-3} f = \sum_{j\in\mathbb{Z}} \sum_{k\leq j-3} \Delta_j g \cdot \Delta_k f.$$

Remark 9.31. *(a) For an appropriate choice of $\tilde{\psi}$ such that $\widehat{\tilde{\psi}}\hat{\psi} = \hat{\psi}$ and with $\tilde{\Delta}_j f := f * \tilde{\psi}_{2^{-j}}$, we have*

$$P_g f = \sum_{j\in\mathbb{Z}} \tilde{\Delta}_j(\Delta_j g \cdot S_{j-3} f), \tag{9.45}$$

a formula that more closely resembles the definition of the continuous paraproduct π_g and which makes clear that $P_g f$ is a sum of almost orthogonal terms. Here we have used the fact that the Fourier transform of $\Delta_j g \cdot S_{j-3} f$ is supported in $2^{j-2} \le |\xi| \le 2^{j+2}$. Thus, each ξ belongs to at most 4 annuli of this form. The annuli are "almost disjoint".

(b) The paraproduct contains essentially half (or better, two-fifths of) the product[72]

$$fg = \sum_j \sum_k \Delta_j g \Delta_k f.$$

(c) When $f, g \in \mathcal{S}$ and $\int g = 0$, the series defining $P_g f$ converges absolutely a.e., since $|S_{j-3} f| \lesssim Mf$, and $\sum_j |\Delta_j g|$ converges for all x; see Exercise 4.4.1, [G2].

The following result is parallel to Lemma 9.20. Its proof clearly uses the almost orthogonality of the terms in the definition of $P_g f$.

Proposition 9.32. *For $b \in BMO$ and $f \in L^2$ the series $\sum_{|j|\le M} \Delta_j g S_{j-3} g$ converges in L^2 as $M \to \infty$ to a function $P_b f$. Moreover,*

$$|P_b|_{2\to 2} \lesssim_n |b|_{BMO}.$$

Proof. Use (9.45), let $\tilde{\Delta}_j$ play the role of the external Q_t in the continuous paraproduct, and repeat the proof of Lemma 9.20. Now the relevant Carleson measure is μ, where[73]

$$d\mu(x,t) = \sum_j |\Delta_j b(x)|^2 dx \delta_{2^{-(j-3)}}(t)$$

for which

$$\int F(x,t)d\mu = \sum_j \int F(x, 2^{-(j-3)}) |\Delta_j b(x)|^2 dx.$$

□

With Prop. 9.32 the next result shows that when $b \in BMO$, the operator P_b is a Calderon-Zygmund operator.

Proposition 9.33. *Let $b \in BMO$. The operator P_b, being L^2 bounded has a Schwarz kernel W_b, and W coincides with a standard kernel off the diagonal.*

Proof. **1.** For (a) we consider the kernel L_j of each operator $f \to \Delta_j b \cdot S_{j-3} f$. We have

$$L_j(x,y) = (b * \psi_{2^{-j}})(x) 2^{(j-3)n} \phi(2^j(x-y)). \tag{9.46}$$

[72] See Prop. 10.3.

[73] Compare $d\mu(x,t) = |(\psi_t * b)(x)|^2 dx \frac{dt}{t}$ in the earlier proof.

We have

$$(9.47)\qquad \begin{aligned} &(a)|L_j(x,y) \lesssim_n |b|_{BMO}\frac{2^{nj}}{(1+2^j|x-y|)^{n+1}} \\ &(b)|\partial_x^\alpha \partial_y^\beta L_j(x,y)| \lesssim_{n,N,\alpha,\beta} \quad |b|_{BMO}\frac{2^{j(n+|\alpha|+|\beta|)}}{(1+2^j|x-y|)^{n+1+N}} \end{aligned}$$

for $N \geq |\alpha| + |\beta|$. For (a) we use the obvious estmate involving ϕ and

$$(9.48)\qquad |b * \psi_{2^{-j}}|_{L^\infty} \lesssim_n |b|_{BMO} \text{ for all } j.$$

Part (b) is then obvious.

2. To prove (9.48) (without appeal to $H^1 - BMO$ duality) we write

$$\Delta_j b(x) = \int \psi_{2^{-j}}(x-y)[b(y) - b_{B(x,2^{-j})}]dy$$

and estimate:

$$(9.49)\qquad |\psi_{2^{-j}}(x-y)| \lesssim \frac{2^{jn}}{(1+2^j|x-y|)^{n+1}} = 2^{-j}\frac{1}{(2^{-j}+|x-y|)^{n+1}}.$$

Now use Lemma 9.17(b).

3. To finish we first show that $|L_b(x,y)| \leq \sum_j |L_j(x,y)| \lesssim \frac{|b|_{BMO}}{|x-y|^n}$ for $x \neq y$. Letting $u = |x-y| > 0$ we show

$$(9.50)\qquad \sum_j \frac{2^{jn}u^n}{(1+2^j u)^{n+1}} \lesssim 1$$

by summing $\sum_j \frac{1}{1+2^j u}$ over j such that $2^j u \geq 1$, and summing $\sum_j 2^j u$ over j such that $2^j u < 1$.

Similarly, we have $|\partial_x^\alpha \partial_y^\beta L_b| \lesssim_{n,\alpha,\beta} \frac{|b|_{BMO}}{|x-y|^{n+|\alpha|+|\beta|}}$ for $x \neq y$,

□

Remark 9.34. *(a) We have $P_b(1) = b$, $P_b^t(1) = 0$. This is formally clear. For example, $P_b(1) = b$ since $S_j(1) = 1$ for all j and $\sum_{j\in\mathbb{Z}} \Delta_j b = b$. See [G2], p. 263 for details.*

(b)The original proof by David and Journé of the $T1$ theorem is discussed in [G2], section 4.5.3. One defines $R = T - P_b - (P_{b'})^t$ for $b = T(1)$, $b' = T^t 1$, uses Proposition 9.32 (which involves Carleson measures) to estimate the paraproducts, and uses Cotlar's lemma to estimate R. Proposition 9.33 implies that R is associated to a standard kernel. Roughly, one writes $R = \sum_j R_j$, where the kernels of the R_j satisfy cancellation conditions like $\int K_j(x,y)dy = 0 = \int K_j(x,y)dx$ as a consequence of $R(1) = 0 = R^t(1)$.

10 Applications of dyadic decompositions and paraproducts

In this section we apply paraproducts to analyze H^s regularity of products and compositions. Similar results hold for Hölder and $W^{s,p}$ spaces.

10.1 Products

We now change notation slightly (to avoid mild headaches).[74] Let $\psi \in C^\infty(\mathbb{R}^n)$ with $\psi = 1$ for $|\xi| \le 1/2$ and $\psi = 0$ for $|\xi| \ge 1$. We set $\phi(\xi) = \psi(\xi/2) - \psi(\xi)$, so ϕ is supported in the annulus $1/2 \le |\xi| \le 2$. Then in the telescoping sum

$$1 = \psi(\xi) + \sum_{p \ge 0} \phi(2^{-p}\xi),$$

there are never more than two nonzero terms for a given ξ. For $u \in \mathcal{S}'(\mathbb{R}^n)$ let

$$u_{-1} = S_0 u = \psi(D)u, \; u_p = \phi(2^{-p}D)u, \text{ so } u = S_0 u + \sum_{p \ge 0} u_p. \tag{10.1}$$

We will write $|u|_{H^s} = |u|_s$, so now $|u|_{L^2} = |u|_0$, and with some abuse we continue to write $|u|_{L^\infty} = |u|_\infty$. Recall the dyadic characterization of H^s.

Proposition 10.1. *Let $s \in \mathbb{R}$. Suppose $u \in \mathcal{S}'(\mathbb{R}^n)$ and set $c_p(u) := 2^{ps}|u_p|_0$.*
(a) If $u \in H^s$ then $(c_p) \in \ell^2$ and $|(c_p)|_{\ell^2} \lesssim |u|_s$.
(b) If $(c_p) \in \ell^2$ then $u \in H^s$ and $|u|_s \lesssim |(c_p)|_{\ell^2}$.

Let us denote the partial sums

$$S_p u = \sum_{q=-1}^{p-1} u_q.$$

Definition 10.2 (paraproduct)**.** *For $u, v \in \mathcal{S}'$ define the paraproduct $T_u v = \sum_{p \ge 2} S_{p-2}u \cdot v_p$.*

Observe that we can now decompose the product uv as

$$uv = T_u v + T_v u + R, \text{ where } R = \sum_{|p-q| \le 2} u_p v_q. \tag{10.2}$$

The next proposition is the main result of this subsection.

Proposition 10.3. *For $s, t \in \mathbb{R}$ suppose $u \in H^s(\mathbb{R}^n)$, $v \in H^t(\mathbb{R}^n)$.*
(a) Then for any $\epsilon > 0$

$$\begin{cases} T_u v \in H^t, \text{ if } s > n/2 \text{ (or if just } u \in L^\infty) \\ T_u v \in H^{t+s-\frac{n}{2}} \text{ if } s < n/2 \\ T_u v \in H^{t+s-\frac{n}{2}-\epsilon} = H^{t-\epsilon} \text{ if } s = \frac{n}{2} \end{cases}. \tag{10.3}$$

(b) Suppose $s, t \in \mathbb{R}$ satisfies $s + t - \frac{n}{2} > 0$. Then $uv = T_u v + T_v u + R$, where $R \in H^{s+t-\frac{n}{2}}$.
(c) In particular, if $s + t - \frac{n}{2} > \min(s, t) := \sigma$, we have $uv \in H^\sigma$.

For the proof we need some preparation.

Lemma 10.4. *Let $s \in \mathbb{R}$ and $u \in H^s$. Then*[75]

$$|u_p|_\infty \lesssim 2^{p(\frac{n}{2}-s)} 2^{ps} |u_p|_0 = 2^{p(\frac{n}{2}-s)} c_p(u). \tag{10.4}$$

[74] This notation is similar to that used in [AG].
[75] If $s > \frac{n}{2}$ and $s - \frac{n}{2} \notin \mathbb{N}$, this estimate implies the Sobolev injection $H^s \subset C^{s-\frac{n}{2}}$.

Proof. Write $u_p = \int e^{ix\xi}\hat{u}_p(\xi)d\xi$. Thus,

$$|u_p|_\infty \lesssim \int_{|\xi|\lesssim 2^p} |\hat{u}_p(\xi)d\xi \lesssim |u_p|_0 2^{pn/2} = 2^{p(\frac{n}{2}-s)}c_p(u).$$

□

The next lemma is used to analyze the remainder term R in Proposition 10.3. It provides an analogue of Proposition 10.1(b) when the summands of $u = \sum_q a_q$ have supports in dyadic balls rather than annuli *and* $s > 0$.

Lemma 10.5. *Let $(a_q)_{q\geq -1}$ be a sequence of tempered distributions such that for some constant C:*

$$\text{supp } \hat{a}_q \subset \{\xi : |\xi| \leq C2^q\}.$$

Suppose for some positive *s that $c_q(u) := 2^{qs}|a_q|_0$ satisfies $(c_q(u)) \in \ell^2$. Then*

$$u = \sum_q a_q \in H^s \textit{ and } |u|_s \lesssim |(c_q(u))|_{\ell^2}.$$

Proof. Looking at supports of Fourier transforms, we see that for some N the dyadic blocks $u_p = \phi(2^{-p}D)u$ satisfy $u_p = \sum_{q\geq p-N}(a_q)_p$.[76] Thus,

$$|u_p|_0 \lesssim \sum_{q\geq p-N} |(a_q)_p|_0 \lesssim \sum_{q\geq p-N} |a_q|_0 = \sum_{q\geq p-N} 2^{-qs}c_q(u).$$

Now multiply through by 2^{ps} and apply Young's inequality for convolutions of sequences to estimate $|u|_s$.

□

Proof of Proposition 10.3. **(a)**. We have $T_uv = \sum_{p\geq 2} S_{p-2}uv_p := \sum w_p$, where spec $w_p \subset \{|\xi| \sim 2^p\}$.[77] If $s > \frac{n}{2}$ we have $u \in L^\infty$, so

$$2^{pt}|w_p|_0 \lesssim 2^{pt}|S_{p-2}u|_\infty|v_p|_0 \lesssim 2^{pt}|u|_\infty|v_p|_0 = |u|_\infty c_p(v).$$

Thus, $|T_uv|_t \lesssim |u|_\infty|v|_t$.

If $s < \frac{n}{2}$ by Lemma 10.4 we have $|u_q|_\infty \lesssim 2^{q(\frac{n}{2}-s)}c_q(u) \lesssim 2^{q(\frac{n}{2}-s)}|(c_q(u))|_{\ell^\infty}$, so

$$|S_{p-2}u|_\infty \lesssim 2^{p(\frac{n}{2}-s)}|(c_q(u))|_{\ell^\infty}. \tag{10.5}$$

Thus,

$$2^{p(t+s-\frac{n}{2})}|w_p|_0 \lesssim 2^{p(t+s-\frac{n}{2})} \cdot 2^{p(\frac{n}{2}-s)}|(c_q(u))|_{\ell^\infty}2^{-pt}c_p(v) = |(c_q(u))|_{\ell^\infty}c_p(v),$$

which implies $|T_uv|_{t+s-\frac{n}{2}} \lesssim |u|_s|v|_t$.[78]

When $s = n/2$, instead of (10.5) we obtain $|S_{p-2}u|_\infty \lesssim p|(c_q(u))|_{\ell^\infty}$, so the above argument yields

$$|T_uv|_{t+s-\frac{n}{2}-\epsilon} \leq C_\epsilon|u|_s|v|_t.$$

(b). Write $R = \sum_{|p-q|\leq 2} u_pv_q := \sum_p u_pV_p$, where spec $(u_pV_p) \subset \{|\xi| \lesssim 2^p\}$. We have

$$|u_pV_p|_0 \leq |u_p|_\infty|V_p|_0 \lesssim 2^{p(\frac{n}{2}-s)}c_p(u)2^{-pt}c_p(V) \Rightarrow 2^{p(s+t-\frac{n}{2})}|u_pV_p| \lesssim c_p(u)c_p(V).$$

[76]Here $(a_q)_p$ denotes the p component of the single term a_q!

[77]The spectrum of w_p, spec w_p, is the support of $\hat{w}_p$.

[78]Here we used $|(c_q(u))|_{\ell^\infty} \lesssim |(c_q(u))|_{\ell^2} = |u|_s$, which can be proved, for example, by the closed graph theorem.

Applying Lemma 10.5 and using $|(c_p(u)c_p(V))|_{\ell^2} \le |(c_p(u)c_p(V))|_{\ell^1} \lesssim |u|_s|v|_t$, we obtain

$$|R|_{s+t-\frac{n}{2}} \lesssim |u|_s|v|_t.$$

Part (c) follows immediately from parts (a), (b). □

The next result is a corollary of the previous proof.

Corollary 10.6. *Let $s > 0$ and suppose $u, v \in H^s \cap L^\infty$. Then $uv \in H^s$ and*[79]

$$|uv|_s \lesssim |u|_\infty|v|_s + |v|_\infty|u|_s.$$

Proof. From above we have $|T_u v|_s \lesssim |u|_\infty|v|_s$ and $|T_v u|_s \lesssim |v|_\infty|u|_s$. Writing $R = \sum u_p V_p$ as above we have

$$|u_p V_p|_0 \le |u_p|_\infty|V_p|_0 \lesssim |u|_\infty|V_p|_0.$$

Multiply through by 2^{ps} and apply Lemma 10.5 to obtain $|R|_s \lesssim |u|_\infty|v|_s$. □

10.2 Compositions

The next proposition is the main result of this subsection.

Proposition 10.7. *Let $F : \mathbb{R} \to \mathbb{C}$ be C^∞ with $F(0) = 0$. Suppose $s > 0$ and $u \in H^s \cap L^\infty$. Then $F(u) \in H^s \cap L^\infty$, and*

$$|F(u)|_s \le C_s(|u|_\infty)\ |u|_s.$$

We will need the following lemma (from [Met]) in the proof of Proposition 10.7. It provides an analogue of Proposition 10.1(b) in which spectral localization is replaced by estimates that mimic this localization.

Lemma 10.8. *Let $M \in \mathbb{N}$ and suppose $0 < s < M$. Suppose (f_k) is a sequence of functions in $H^M(\mathbb{R}^n)$ such that for all $|\alpha| \le M$:*

$$|\partial^\alpha f_k|_0 \le 2^{k(|\alpha|-s)}c_k,\ \text{where } (c_k) \in \ell^2.$$

Then $f = \sum f_k \in H^s$ and $|f|_s \lesssim |(c_k)|_{\ell^2}$.

Proof. First note that the series $\sum_k f_k = f$ in L^2.

Let $\theta_{-1}(\xi) = \psi(\xi)$ and $\theta_j(\xi) = \phi(2^{-j}\xi)$, $j \ge 0$. We have

$$\begin{aligned} &|\theta_j(D)f_k|_0 \le C|f_k|_0 \le C2^{-ks}c_k \\ &|\theta_j(D)f_k|_0 \le C2^{-Mj}|f_k|_M \le C2^{-ks}2^{M(k-j)}c_k. \end{aligned} \tag{10.6}$$

We use the first estimate when $k \ge j$, the second when $k < j$. Thus,

$$2^{js}|\theta_j(D)f|_0 \le C(c'_j + c''_j),\ \text{where } c'_j = \sum_{k\ge j} 2^{(j-k)s}c_k,\ \ c''_k = \sum_{k<j} 2^{(M-s)(k-j)}c_k. \tag{10.7}$$

By Young's inequality we have $|(c'_k)|_{\ell^2} \le |(c_j)|_{\ell^2}$ and $|(c''_j)|_{\ell^2} \le |(c_k)|_{\ell^2}$. Proposition 10.1(b) now gives the result.

□

[79] The result is trivially true for $s = 0$.

Proof of Proposition 10.7. **1.** Suppose $s > 0$.[80] We will use the following two estimates:

$$(a)|\partial^\alpha u_p|_0 \lesssim 2^{p(|\alpha|-s)} c_p(u) \qquad (10.8)$$
$$(b)|\partial^\alpha S_p u|_\infty \lesssim 2^{p|\alpha|} |u|_\infty.$$

The first follows from spec $u_p \subset \{|\xi| \sim 2^p\}$ and the second is a Bernstein inequality.

2. Since $S_p u \to u$ in L^2 and the S_p are uniformly bounded in L^∞, we have $F(S_p u) \to F(u)$ in L^2. Thus, in L^2

$$F(u) = F(S_0 u) + \sum_{p \geq 0} [F(S_{p+1}u) - F(S_p u)] = F(S_0 u) + \sum_{p \geq 0} m_p u_p, \qquad (10.9)$$
$$\text{where } m_p = \int_0^1 F'(S_p u + t u_p) dt.$$

The estimate (10.8)(b) implies $|\partial^\alpha m_p|_\infty \leq C(\alpha, |u|_\infty) 2^{p|\alpha|}$. With (10.8)(a) this implies

$$|\partial^\alpha (m_p u_p)|_0 \leq C_\alpha 2^{p(|\alpha|-s)} c_p(u),$$

so the result now follows from Lemma 10.8.

$\square$

10.3 Paralinearization

In this subsection we prove the following refinement of Proposition 10.7, an example of "paralinearization" due to Bony [B]. Again, we follow [Met].

Proposition 10.9. *Let $F : \mathbb{R} \to \mathbb{C}$ be C^∞ with $F(0) = 0$. If $u \in H^s(\mathbb{R}^n)$, where $\rho := s - \frac{n}{2} > 0$, then*

$$F(u) - T_{F'(u)} u \in H^{s+\rho}(\mathbb{R}^n).$$

Proof. **1.** We can suppose $F'(0) = 0$, since if $F'(0) = a \neq 0$ we can replace $F(u)$ by $F(u) - au$ (and use $T_a u = au$). With this reduction Propositions 10.3 and 10.7 imply $T_{F'(u)} u \in H^s$.[81]

2. By definition we have, setting $g = F'(u)$,

$$T_{F'(u)} u = \sum_{p \geq 2} S_{p-2} g \cdot u_p = S_0 g \cdot u_2 + S_1 g \cdot u_3 + \dots. \qquad (10.10)$$

We will compare this with (10.9): $F(u) = F(S_0 u) + \sum_{p \geq 0} m_p u_p$. The terms $F(S_0 u)$ and $m_0 u_0 + m_1 u_1$ lie in H^∞, so it remains to show

$$\sum_{p \geq 2} (m_p - S_{p-2} g) \cdot u_p \in H^{s+\rho}. \qquad (10.11)$$

The main step is to prove

$$|\partial^\alpha (m_p - S_{p-2} g)|_\infty \leq C_\alpha 2^{(|\alpha|-\rho)p}. \qquad (10.12)$$

With (10.8)(a) this implies

$$|\partial^\alpha [(m_p - S_{p-2} g) u_p]|_0 \leq C_\alpha 2^{p(|\alpha|-(s+\rho))} c_p(u), \qquad (10.13)$$

[80] The case $s = 0$ is easy.
[81] Since $\rho > 0$ we have $u \in L^\infty$.

so Lemma 10.8 implies (10.11).

3. To prove (10.12) we write

$$m_p - S_{p-2}g = [m_p - F'(S_{p-2}u)] + [F'(S_{p-2}u) - S_{p-2}g]$$

and show

$$(10.14)\qquad \begin{aligned} &(a)|\partial^\alpha[m_p - F'(S_{p-2}u)]|_\infty \le C_\alpha 2^{(|\alpha|-\rho)p}\\ &(b)|\partial^\alpha[F'(S_pu) - S_pg]|_\infty \le C_\alpha 2^{(|\alpha|-\rho)p}. \end{aligned}$$

and examine these pieces separately.

4. For (10.14)(a) we write

$$(10.15)\qquad \begin{aligned} &F'(S_pu + tu_p) - F'(S_{p-2}u) = \mu_p w_p, \text{ where}\\ &\mu_p = \int_0^1 F''(S_{p-2}u + \tau w_p)d\tau, \quad w_p = u_{p-2} + u_{p-1} + tu_p. \end{aligned}$$

For μ_p we have an estimate like that for m_p in the previous proof:

$$(10.16)\qquad |\partial^\alpha \mu_p|_\infty \le C(\alpha, |u|_\infty)2^{p|\alpha|}.$$

For w_p we use the Bernstein inequality (14.4) to get

$$(10.17)\qquad |\partial^\alpha w_p|_\infty \le C_\alpha 2^{p(|\alpha|+\frac{n}{2})}|w_p|_0 \le C_\alpha 2^{p(|\alpha|+\frac{n}{2})}2^{-ps}c_p(u) \le C'_\alpha 2^{p(|\alpha|+\frac{n}{2})}2^{-ps} = C'_\alpha 2^{p(|\alpha|-\rho)}.$$

Together these imply $|\partial^\alpha(\mu_p w_p)|_\infty \le C_\alpha 2^{p(|\alpha|-\rho)}$ and thus (10.14)(a).

5. To analyze (10.14)(b) we set $G = F'$ and write $F'(S_pu) - S_pg =$

$$G(S_pu) - S_pG(u) = G(S_pu) - S_p(G(S_pu)) - S_p(G(u) - G(S_pu)).$$

Bernstein (14.4) implies

$$(10.18)\qquad |\partial_x^\alpha S_p(G(u) - G(S_pu))|_\infty \lesssim 2^{(|\alpha|+\frac{n}{2})p}|S_p(G(u) - G(S_pu))|_0.$$

With

$$|S_p(G(u) - G(S_pu))|_0 \lesssim |G(u) - G(S_pu)|_0 \lesssim |u - S_pu|_0 \lesssim 2^{-ps},$$

we see that the left side of (10.18) satisfies (10.14)(b).

6. To finish we need the following estimate for $a \in H^\sigma$ when $|\alpha| < \sigma - \frac{n}{2}$:

$$(10.19)\qquad |\partial_x^\alpha(a - S_pa)|_\infty \le C2^{p(|\alpha|-\sigma+\frac{n}{2})}|a|_\sigma.$$

To see this write $a - S_pa = \sum_{j\ge p} a_j$, use (14.4) to estimate

$$|\partial_x^\alpha a_j|_\infty \lesssim 2^{j(|\alpha|+\frac{n}{2})}|a_j|_0 \sim 2^{j(|\alpha|+\frac{n}{2})}2^{-j\sigma}|a_j|_\sigma \lesssim 2^{j(|\alpha|-\sigma+\frac{n}{2})}|a|_\sigma,$$

and observe that $\sum_{j\ge p}|\partial^\alpha a_j|_\infty$ converges if $|\alpha| < \sigma - \frac{n}{2}$.

7. Now we claim that for N large enough:

$$(10.20)\qquad |\partial_x^\alpha[G(S_pu) - S_pG(S_pu)]|_\infty \lesssim 2^{p(|\alpha|-s-N+\frac{n}{2})}|G(S_pu)|_{s+N}.$$

Indeed, this follows by taking $a = G(S_pu))$ and $\sigma := s + N$ in (10.19). Since

$$|S_pu|_{s+N} \le C2^{pN} \text{ and } |S_pu|_\infty \le C$$

with C independent of p, Proposition 10.7 implies $|G(S_pu)|_{s+N} \le C_{s+N}(|u|_\infty)2^{pN}$. With (10.20) this completes the verification of (10.14)(b). □

11 Fourier transform restriction

Let S be a smooth compact hypersurface in $\mathbb{R}^n$, $n \geq 2$. In this section we discuss estimates of the form

$$|\hat{f}|_S|_{L^q(S)} \lesssim_{n,p,S} |f|_{L^p(\mathbb{R}^n)}. \tag{11.1}$$

Theorem 11.1 (Tomas-Stein)**.** *The estimate* (11.1) *holds for the set*

$$\left\{(p,q) : \frac{1}{q} = \frac{n+1}{n-1}\frac{1}{p'}, \ \ 1 \leq p \leq \frac{2(n+1)}{n+3}\right\}.$$

Observe that if $f \in L^p \cap L^1$, then $\hat{f}$ is continuous so its restriction to S is defined pointwise. The estimate in the case $p = 1$, $q = \infty$ is obvious. We shall prove the "endpoint" case $q = 2$, $p = \frac{2n+2}{n+3}$. One can then treat the other cases using interpolation and the fact that S has finite measure.[82] An example due to Knapp shows that the $q = 2$ estimate fails for $p > \frac{2(n+1)}{n+3}$.

In the next two sections we do some preparation.

11.1 Stationary phase and the role of nonvanishing Gauss curvature

Consider the *oscillatory integral*

$$I(\lambda) = \int_{\mathbb{R}^n} e^{i\lambda\phi(x)} a(x)dx, \ \ \lambda >> 0, \tag{11.2}$$

where $a \in \mathcal{D}$ and ϕ is real and C^∞.

Lemma 11.2 (Stationary phase)**.** *(a) Nonstationary case: Suppose* $\nabla\phi \neq 0$ *on* supp a. *Then for all* $N \geq 1$

$$|I(\lambda)| \lesssim_{a,\phi,N} \lambda^{-N} \ \ \textit{as } \lambda \to \infty.$$

(b) Stationary case: Suppose $\nabla\phi(x_0) = 0$ *for some* $x_0 \in$ supp a, $\nabla\phi \neq 0$ *away from* x_0, *and the hessian* $\det D^2\phi(x_0) \neq 0$. [83]. *Then for* $\lambda \geq 1$

$$|I(\lambda)| \lesssim_{a,\phi,n} \lambda^{-n/2}.$$

More precisely one has

$$I(\lambda) = e^{i\lambda\phi(\xi_0)} w(\lambda) \ \ \textit{where } |\partial_\lambda^k w| \lesssim \lambda^{-\frac{n}{2}-k}. \tag{11.3}$$

Proof. In part (a) integrate by parts repeatedly using $L(e^{i\lambda\phi}) = e^{i\lambda\phi}$, where

$$L = \frac{1}{i\lambda}\frac{\phi_\xi}{|\phi_\xi|^2} \cdot \partial_\xi, \ \ L^t = \frac{i}{\lambda}\partial_\xi \cdot \frac{\phi_\xi}{|\phi_\xi|^2} \cdot .$$

For (b) take $\xi_0 = 0$ and write

$$\phi(\xi) = \phi(0) + \frac{1}{2}D^2\phi(0)\xi \cdot \xi + O(|\xi|^3).$$

[82] Interpolation gives the points on the segment GD in Fig. 5.10 of [G2].

[83] We then say that the critical point at x_0 is *nondegenerate*.

Nondegeneracy implies $|\phi_\xi| \gtrsim |\xi|$ on supp a.[84] Choose $\chi \in \mathcal{D}$ and break the integral $I(\lambda)$ into $A+B$, where A includes $\chi(\lambda^{1/2}\xi)$ and B includes $1-\chi(\lambda^{1/2}\xi)$. We have

$$|A| \leq \int |\chi(\lambda^{1/2}\xi)|d\xi \lesssim \lambda^{-n/2}.$$

To estimate B use

$$\left|\partial_\xi^\alpha \left(\frac{\phi_\xi}{|\phi_\xi|^2}\right)\right| \lesssim |\xi|^{-1-|\alpha|} \text{ on supp } a$$

and $(L^t)^N$ for $N > n$ to get

$$|B| \lesssim \lambda^{-N} \int_{\lambda^{-1/2}}^{\infty} (\lambda^{N/2} r^{-N} + r^{-2N}) r^{n-1} dr \lesssim \lambda^{-n/2}.$$

To prove (11.3) use $|\phi(\xi) - \phi(0)| \lesssim |\xi|^2$ and argue similarly, taking $N > n + 2k$ now.

□

We now apply this lemma to estimate the Fourier transform of a localized surface measure on a smooth hypersurface surface $S \subset \mathbb{R}^n$. The observed decay when S has nonvanishing Gauss curvature leads to nontrivial Fourier restriction results.[85]

Proposition 11.3. *Let $S \subset \mathbb{R}^n$ be a smooth hypersurface with induced surface measure σ, and let $\mu := \phi\sigma$ for some $\phi \in \mathcal{D}(\mathbb{R}^n)$.*[86] *Assume that S has nonvanishing Gauss curvature on the support of ϕ. Then*

$$|\hat{\mu}(\xi)| \lesssim_{S,\mu} |\xi|^{-(n-1)/2} \text{ for } |\xi| \geq 1,\ \xi \in \mathbb{R}^n. \tag{11.4}$$

Proof. **1.** For $\xi \neq 0$ write $\xi = \lambda\nu$ for some $\nu \in S^{n-1}$. Then[87]

$$\hat{\mu}(\xi) = \int_{\mathbb{R}^d} e^{-ix\cdot\xi}\phi(x)d\sigma = \int_{\mathbb{R}^d} e^{-i\lambda\nu\cdot x}\phi(x)d\sigma. \tag{11.5}$$

Let $x = (x', x_d)$ and suppose also that $\phi|_S$ is supported in a coordinate patch where S is given as the graph of $x_d = f(x')$. Then in that patch

$$\sigma = \delta(x_d - f(x'))\sqrt{1 + |\nabla f(x')|^2}, \tag{11.6}$$

so

$$\int_{\mathbb{R}^n} e^{-i\lambda\nu\cdot x}\phi(x)d\sigma = \int_{\mathbb{R}^{n-1})} e^{-i\lambda[\nu'\cdot x' + \nu_d f(x')]}\phi(x', f(x'))\sqrt{1+|\nabla f(x')|^2}dx'. \tag{11.7}$$

2. The phase $\phi(x') := \nu' \cdot x' + \nu_d f(x')$ has a critical point at x_0' if and only if ν is perpendicular to $T_{x_0}S$, where $x_0 = (x_0', f(x_0'))$. Any such critical point is nondegenerate, and thus isolated (by the inverse function theorem), if and only if the Gauss curvature of S is nonzero at x_0. The result follows by applying the stationary phase lemma (both parts) to the integral on the right in (11.7). One considers the case when ϕ is supported in a patch with no critical points, and the case when ϕ is supported in a patch with one critical point. □

[84] Here use that $D^2\phi(0)$ is nonsingular and the homogeneity of $\frac{1}{2}D^2\phi(0)\xi \cdot \xi$.

[85] The case $p = 1$, $q = \infty$ is regarded as trivial.

[86] The measure $\phi\sigma$ can be viewed either as a measure on S or as a measure on $\mathbb{R}^n$ supported on S. We use both interpretations below. Authors sometimes write $\phi d\sigma$ for $\phi\sigma$.

[87] The integrals on the left in (11.5) and (11.7) are sometimes written as $\int_S \ldots$, because the measure $\sigma \in \mathcal{S}'$ is supported on S.

11.2 Fractional integration, Riesz potentials

In the proof of the Tomas-Stein theorem we will use the following result on "fractional integration":[88]

Proposition 11.4 (Fractional integration)**.** *Let α be a real number such that $0 < \alpha < n$ and suppose $1 < p < q < \infty$ satisfy*

$$\frac{1}{p} - \frac{1}{q} = \frac{\alpha}{n}. \tag{11.8}$$

Then for $I_\alpha(f)(x) := \int_{\mathbb{R}^n} |x-y|^{\alpha-n} f(y)dy$ we have

$$|I_\alpha f|_q \lesssim_{n,\alpha,p} |f|_p. \tag{11.9}$$

Proof. With $f \in \mathcal{S}$ and $R(x)$ to be chosen, write

$$I_\alpha f(x) = A + B = \int_{|y|<R(x)} f(x-y)|y|^{\alpha-n}dy + \int_{|y|\geq R(x)} f(x-y)|y|^{\alpha-n}dy$$

Now[89]

$$A = R(x)^s (f * K_{R(x)})(x) \text{ where } K(y) = |y|^{\alpha-n}\chi_{|y|<1},$$

so Theorem 2.9 implies

$$|A| \leq R(x)^s Mf(x) \int_{|y|<1} |y|^{\alpha-n}dy \lesssim R(x)^s Mf(x).$$

Since $(n-\alpha)p' = n + \frac{p'n}{q} > n$, we can apply Hölder to get

$$|B| \leq \left||y|^{\alpha-n}\chi_{|y|\geq R(x)}\right|_{p'} |f|_p \lesssim R(x)^{-n/q}|f|_p.$$

This gives

$$|I_\alpha f(x)| \lesssim R(x)^s Mf(x) + R(x)^{-n/q}|f|_p \leq (Mf(x))^{p/q}|f|_p^{1-\frac{p}{q}},$$

where we took $R(x) = |f|_p^{p/n}(Mf(x))^{-p/n}$ to minimize the sum. Take the L^q norm of both sides to finish.

□

11.3 Proof of the Tomas-Stein theorem

Proof. **1.** Let σ be the surface measure on S. Using $\langle g, \overline{h}\rangle = \langle \hat{g}, \overline{\hat{h}}\rangle$, where $\langle\cdot,\cdot\rangle$ is the distributional pairing on $\mathbb{R}^n$, we have

$$|\hat{f}|^2_{L^2(S)} = \int_S \overline{\hat{f}}(\xi)\hat{f}(\xi)d\sigma = \langle \hat{f}\sigma, \overline{\hat{f}}\rangle = \langle \widehat{f*\check{\sigma}}, \overline{\hat{f}}\rangle = \langle f*\check{\sigma}, \overline{f}\rangle \leq |f|_p|f*\check{\sigma}|_{p'}. \tag{11.10}$$

To finish it is enough to show for $f \in \mathcal{S}(\mathbb{R}^n)$ that

$$|f*\check{\sigma}|_{p'} \lesssim_n |f|_p, \text{ when } p = \frac{2(n+1)}{n+3}. \tag{11.11}$$

[88]The *Riesz potential* of order α is $I_\alpha = (-\Delta)^{-\alpha/2}$, where the equality holds up to a constant factor.
[89]Recall $K_\epsilon(x) := \epsilon^{-n}K(x/\epsilon)$.

2. Let $\phi \in \mathcal{D}(\mathbb{R}^n)$, and set $\mu = \phi\sigma$. We take $\phi|_S$ to be supported in a coordinate patch where S is the graph of a smooth function $\tau = h(\xi')$. We have reduced to showing

$$|f * \check{\mu}|_{p'} \lesssim_n |f|_p, \text{ when } p = \frac{2(n+1)}{n+3}. \tag{11.12}$$

Now for $K = \check{\mu}$ we may write

$$(f * \check{\mu})(x) = \int_{\mathbb{R}^n} K(x' - y', t - s) f(y', s) dy' ds. \tag{11.13}$$

By Proposition 11.3 we have

$$|K(x' - y', t - s)| \lesssim \langle x' - y', t - s\rangle^{-(n-1)/2} \leq \langle t - s\rangle^{-(n-1)/2} \text{ for all } x', y'. \tag{11.14}$$

Writing

$$(U(t)g)(x') = \int_{\mathbb{R}^{n-1}} K(x' - y', t) g(y') dy' \text{ for } g \in \mathcal{S}(\mathbb{R}^{n-1}),$$

we have

$$|U(t)g|_\infty \lesssim \langle t\rangle^{-(n-1)/2} |g|_1 \text{ and } |U(t)g|_2 \lesssim |g|_2,$$

where we used $|\hat{K}(\xi', t)| \lesssim 1$ uniformly in t. The latter follows from

$$K(x', t) = \int e^{ix'\xi' + ith(\xi')} \phi(\xi', h(\xi')) \sqrt{1 + |\nabla h(\xi')|^2} d\xi'. \tag{11.15}$$

Interpolating between the two estimates of $U(t)g$ gives for $1 \leq p \leq 2$:

$$|U(t)g|_{p'} \lesssim_{p,n} \langle t\rangle^{-\alpha} |g|_p \text{ where } \alpha = (n-1)\left(\frac{1}{p} - \frac{1}{2}\right). \tag{11.16}$$

To conclude we apply fractional integration in t to obtain

$$\begin{aligned} |f * \check{\mu}|_{L^{p'}(\mathbb{R}^n)} \leq \left| \int_{-\infty}^{\infty} |U(t-s) f(s)|_{L^{p'}(\mathbb{R}^{n-1})} ds \right|_{L^{p'}(\mathbb{R})} \lesssim \\ \left| \int_{-\infty}^{\infty} |t-s|^{-\alpha} |f(s)|_{L^p(\mathbb{R}^{n-1})} ds \right|_{L^{p'}(\mathbb{R})} \lesssim |f|_{L^p(\mathbb{R}^n)}. \end{aligned} \tag{11.17}$$

By Proposition 11.4 the final inequality requires

$$\frac{1}{p} - \frac{1}{p'} = 1 - \alpha.$$

With (11.16) this implies $p = \frac{2(n+1)}{n+3}$ and $\alpha = \frac{n-1}{n+1}$.

□

11.4 Application: Strichartz estimates and a nonlinear Schrödinger equation.

Consider the Schrödinger initial value problem on $\mathbb{R}^{1+d}$:

$$\begin{aligned} \partial_t u + i\Delta u &= h \\ u|_{t=0} &= f. \end{aligned} \tag{11.18}$$

Definition 11.5. *A pair* (p,q) *is* Strichartz admissible *if*

$$\frac{2}{p}+\frac{d}{q}=\frac{d}{2} \text{ and } 2\le p\le\infty,\quad (p,q)\ne(2,\infty).$$

Theorem 11.6. *Let* (a,b) *and* (p,q) *be Strichartz admissible with* $a>2$, $p>2$ *and suppose* u *satisfies* (11.18) *with* h, f *Schwartz functions. Then*[90]

$$|u|_{L_t^pL_x^q}\lesssim |f|_{L_x^2}+|h|_{L_t^{a'}L_x^{b'}}. \tag{11.19}$$

The norms of u *and* h *can be restricted to a time interval* $I\ni 0$.

Remark 11.7. *(a) If* u, f, $h=0$ *satisfy* (11.18), *then the rescaled functions*

$$u_\lambda(x,t)=u(x/\lambda,t/\lambda^2),\ f_\lambda(x)=f(x/\lambda),\ \lambda>0$$

also satisfy (11.18). *If the Strichartz estimate* (11.19) *holds for solutions, then the relation* $\frac{2}{p}+\frac{d}{q}=\frac{d}{2}$ *must be satisfied by* (p,q) *and* (a,b), *because of the way the norms scale. We say the estimate* (11.19) *is* scaling invariant *under the above rescaling.*

(b) The restriction $p>2$ *is technical. The "endpoint estimate" for* $p=2$, $d\ge 3$ *was proved by Keel and Tao.*

Proof of Theorem 11.6. **1.** We can write the solution (with obvious abuse of notation)

$$u(t)=e^{-i\Delta t}f+\int_0^t e^{-i\Delta(t-s)}h(s)ds. \tag{11.20}$$

When $h=0$ this is equivalent for $t>0$ to[91]

$$u(x,t)=t^{-d/2}\int_{\mathbb{R}^d}e^{i|x-y|^2/4t}f(y)dy=\int e^{ix\xi}e^{-it|\xi|^2}\hat f(\xi)d\xi. \tag{11.21}$$

The formulas (11.21) imply the easy estimates (in obvious notation)

$$\begin{aligned}&|U(t)f|_{L_x^\infty}\lesssim |t|^{-d/2}|f|_{L_x^1}\\ &|U(t)f|_{L_x^2}=|f|_{L_x^2}.\end{aligned} \tag{11.22}$$

For $1\le q'\le 2$ Riesz-Thorin interpolation gives

$$|U(t)f|_{L_x^q}\lesssim |t|^{(-d/2)(\frac{1}{q'}-\frac{1}{q})}|f|_{L_x^{q'}}. \tag{11.23}$$

2. Let $(Uf)(x,t)=(U(t)f)(x)$ and observe that

$$\begin{aligned}&(a)(U^*F)(x)=\int_{-\infty}^\infty U(-s)F(s)ds \text{ whence}\\ &(b)(U\circ U^*F)(t)=\int_{-\infty}^\infty U(t-s)F(s)ds.\end{aligned} \tag{11.24}$$

[90]Note the primes on a,b.

[91]Here we probably omitted some constant factors out front and in the exponents.

By the TT^* principle (see #21 of section 14), the following estimates are equivalent:[92]

$$\begin{aligned}(a)&|Uf|_{L^p_tL^q_x} \lesssim |f|_{L^2_x}\\ (b)&|U^*F|_{L^2_x} \lesssim |F|_{L^{p'}_tL^{q'}_x}\\ (c)&|U\circ U^*F|_{L^p_tL^q_x} \lesssim |F|_{L^{p'}_tL^{q'}_x}.\end{aligned} \tag{11.25}$$

We will show these estimates hold when (p,q) is Strichartz admissible with $p>2$ by proving (c). Then applying (a) with $f=U^*F$ and (b) with (a,b) in place of (p,q) gives

$$|U\circ U^*F|_{L^p_tL^q_x} \lesssim |F|_{L^{a'}_tL^{b'}_x}. \tag{11.26}$$

3. Using (11.23) we obtain

$$|U\circ U^*F(t)|_{L^q_x} \lesssim \int_{-\infty}^{\infty} |t-s|^{(-d/2)(\frac{1}{q'}-\frac{1}{q})}|F(s)|_{L^{q'}_x}ds \tag{11.27}$$

We can apply fractional integration in t (Prop. 11.4) to conclude (11.25)(c) provided

$$1+\frac{1}{p}=\frac{1}{p'}+\frac{d}{2}\left(\frac{1}{q'}-\frac{1}{q}\right) \text{ and } 0<\frac{d}{2}\left(\frac{1}{q'}-\frac{1}{q}\right)<1. \tag{11.28}$$

The first condition is equivalent to $\frac{2}{p}+\frac{d}{q}=\frac{d}{2}$ and the second is equivalent to $p>2$. Thus, the estimates (11.25)(a)-(c) hold. This treats the case $h=0$.

4. To treat the case $f=0$, one could use (11.25)(c) and the Christ-Kiselev lemma for the retarded Strichartz estimate; see [Tao2]. Instead, as in [CS1] write

$$u(t)=\int_0^t U(t-s)h(s)ds=\int_{-\infty}^{\infty}\chi_{0<s<t}U(t-s)h(s)ds. \tag{11.29}$$

We claim

$$\begin{aligned}&|u|_{L^p_tL^q_x}\lesssim |h|_{L^{p'}_tL^{q'}_x}\\ &|u|_{L^p_tL^q_x}\lesssim |h|_{L^1_tL^2_x}.\end{aligned} \tag{11.30}$$

The first estimate is proved just like (11.25)(c). To prove the second write

(11.31)
$$|u(t)|_{L^q_x}\lesssim\int_{-\infty}^{\infty}|U(t-s)h(s)|_{L^q_x}ds \Rightarrow |u|_{L^p_tL^q_x}\lesssim\int_{-\infty}^{\infty}|U(t-s)h(s)|_{L^p_tL^q_x}ds\le\int_{-\infty}^{\infty}|h(s)|_{L^2_x}ds.$$

To obtain (11.19) in the case $f=0$, interpolate between the estimates in (11.30).[93] The time-localized estimates can be proved using the formula (11.20), which shows that $u|_I$ depends just on $h|_I$. [94]

□

[92] We also need estimate (b) with (a,b) in place of (p,q).
[93] Check this works.
[94] Check. Note the retarded estimate is different from the time-localized estimate.

Remark 11.8. *One could use the Fourier transform representation of u in* (11.21) *and the Tomas-Stein theorem to obtain the first estimate in* (11.22)*. This is the connection between decay of $\hat{\mu}$ and Strichartz estimates, where μ is (almost) the surface measure of the Schrödinger paraboloid. Note that the solution to the Schrödinger problem* (11.18) *with $h = 0$ is*

$$u(x,t) = \mathcal{F}^{-1}_{\tau,\xi}(\hat{f}\mu),$$

where $\mu = \delta(\tau - |\xi|^2)$.

Nonlinear Schrödinger equation. Next we apply Strichartz estimates to the problem

$$i\partial_t\psi + \Delta\psi = \lambda|\psi|^{4/d}\psi, \quad \psi|_{t=0} = \psi_0, \quad \lambda \in \mathbb{R}, x \in \mathbb{R}^d \tag{11.32}$$

with small data in L^2. The exponent here is the unique one for which the rescaled functions

$$\lambda^{d/2}\psi(\lambda x, \lambda^2 t)$$

are solutions. This rescaling leaves the L^2 norm of the data unchanged.

Inspired by Duhamel we consider the integral integration

$$\psi(t) = e^{it\Delta}\psi_0 + i\lambda\int_0^t e^{i(t-s)\Delta}|\psi(s)|^{4/d}\psi(s)ds. \tag{11.33}$$

Definition 11.9. *A global weak solution of* (11.32) *is a solution of the integral equation* (11.33) *that lies in the space*[95]

$$X := (C_t \cap L^\infty_t)([0,\infty) : L^2_x) \cap L^{p_0}_t([0,\infty) : L^{2p_0}_x), \ \ \textit{where } p_0 = 1 + \frac{4}{d}. \tag{11.34}$$

We'll look at the case $d = 2$, so $p_0 = 3$. Observe that the pairs $(\infty, 2)$ and $(3, 6)$ are Strichartz admissible.

Proposition 11.10 (Global existence for small data)**.** *There exists $\epsilon_0(\lambda, d)$ such that the problem* (11.32) *admits a unique weak solution provided $|\psi_0|_2 \le \epsilon_0(\lambda, d)$.*

Proof. **1. Uniqueness.** Let ψ, ϕ be two weak solutions. Then

$$(\psi - \phi)(t) = i\lambda\int_0^t e^{i(t-s)\Delta}[|\psi(s)|^2\psi(s) - |\phi(s)|^2\phi(s)]ds. \tag{11.35}$$

Apply Strichartz estimates to estimate the $X(I)$ norm, $I = [0,t)$:

$$|\psi - \phi|_{X(I)} \le C|\lambda| \ \left||\psi|^2\psi - |\phi|^2\phi\right|_{L^1(I:L^2_x)} \le C|\lambda|(|\psi|_{X(I)} + |\phi|_{X(I)})^2|\psi - \phi|_{X(I)}.$$

Here we used Strichartz to get the first inequality and Hölder to get the second, after writing[96]

$$|\psi|^2\psi - |\phi|^2\phi = |\psi|^2(\psi - \phi) + \phi(|\psi|^2 - |\phi|^2). \tag{11.36}$$

[95] The occurrence of $L^{p_0}_t([0,\infty) : L^{2p_0}_x)$ is natural, since Hölder implies that if $\psi \in X$ we have

$$|\psi(t)|_{L^2_x} \le C\left(|\psi_0|_{L^2} + |\lambda|\int_0^t |\psi(s)|^{p_0}_{2p_0}ds\right) < \infty \text{ when } \psi \in X.$$

[96] For example,

$$\left||\psi|^2(\psi - \phi)\right|_{L^1L^2} \le |\psi|^2_{L^3L^6}|\psi - \phi|_{L^3L^6}.$$

If I is small enough, $C|\lambda|(|\psi|_{X(I)} + |\phi|_{X(I)})^2 < 1$ so $|\psi - \phi|_{X(I)} = 0$ and thus $\psi = \phi$ on I. The nonempty set of t where $\psi = \phi$ is thus open and closed, so equals the whole line.

2. Existence. Define

$$(A\psi)(t) = e^{it\Delta}\psi_0 + i\lambda \int_0^t e^{i(t-s)\Delta}|\psi(s)|^2\psi(s)ds.$$

Using the same estimate as above on the nonlinear term we get

$$|A\psi|_X \leq C(|\psi_0|_{L^2_x} + |\lambda||\psi|^3_X). \tag{11.37}$$

So if $|\psi|_X \leq R$ and $C(|\psi_0|_{L^2_x} + |\lambda||\psi|^3_X) < R$, then A maps an R-ball in X to itself. The second condition will hold if we choose $R = C\epsilon$ and ϵ small enough so that (11.37) holds when $|\psi_0|_{L^2} < \epsilon$. Estimating $|A\psi - A\phi|_X$ as in step **2** and reducing ϵ if necessary, we see that A is a contraction. So there is a unique fixed point in the closed $R = C\epsilon-$ball of X.

□

12 Appendix 1: Interpolation theorems

First we state the special cases of the Riesz-Thorin and Marcinkiewicz theorems that we frequently use in these notes.

Theorem 12.1 (Riesz-Thorin)**.** *Let $1 \leq p_1 < p_2 \leq \infty$. Suppose T is a linear operator that is strong (p_i, p_i) for $i = 1, 2$. Then T is strong (p, p) for all $p_1 \leq p \leq p_2$. It's enough to have T defined initially just on a dense linear subspace of $L^{p_1} \cap L^{p_2}$. This works on quite general measure spaces.*

Theorem 12.2 (Marcinkiewicz)**.** *Same set up as Riesz-Thorin, but now assume T is a linear operator that is weak (p_i, p_i) for $i = 1, 2$. Then T is strong (p, p) for all $p_1 < p < p_2$. This works on quite general measure spaces. T can be just sublinear.*

Remark 12.3. *Both theorems have extensions; see Folland [F], p. 193, for example. The hypotheses of Marcinkiewicz are weaker. In the extended versions the restrictions on p_j, q_j are stronger in Marcinkiewicz.*[97] *Riesz-Thorin gives sharper estimate on the operator norm of T.*

The next theorem is a simplified version of Stein's complex interpolation theorem [S2], in which both operators and spaces are allowed to vary.[98] It contains Riesz-Thorin as the special case $T_z = T$.

Theorem 12.4. *Let X and Y be $\sigma-$finite measure spaces, and let T_z be a family of linear operators depending on $z \in \mathbb{C}$ which map simple functions on X to measurable functions on Y.*[99] *We suppose that for any simple functions f, g on X, Y respectively the function $\psi(z) = \int_Y T_z f \cdot g$ is analytic on the strip $S = \{0 < \mathrm{Re}\ z < 1\}$, and bounded and continuous on $\overline{S}$.*[100]

Let p_0, p_1, r_0, r_1 lie in $[1, \infty]$. Suppose T_z is bounded from $L^{p_0}(X)$ to $L^{r_0}(Y)$ when Re $z = 0$ *and from $L^{p_1}(X)$ to $L^{r_1}(Y)$ when* Re $z = 1$ *with uniform operator-norm bounds B_0, B_1 respectively. Then for $0 \leq t \leq 1$, the operator T_t is bounded from $L^p(X)$ to $L^r(Y)$, where*

$$(\frac{1}{p}, \frac{1}{r}) = t(\frac{1}{p_1}, \frac{1}{r_1}) + (1-t)(\frac{1}{p_0}, \frac{1}{r_0}).$$

Moreover, $|T_t|_{p \to r} \leq B_0^{1-t} B_1^t$.

[97] In the extended versions the assumption in Marcinkiewicz, for example, is that T is weak (p_0, q_0) and weak (p_1, q_1). Similarly for RT, but with weak replaced by strong.

[98] The paper [S2] may be Stein's first paper.

[99] Here we define a "simple function" to be a function that takes only a finite number of nonzero values on sets of *finite* measure.

[100] This bound may depend on f, g.

Proof. **1.** Fix p, r, t satisfying the given convexity relation, and let f, g be simple functions on X, Y satisfying $|f|_p = |g|_{r'} = 1$.[101] Let $\Phi(t) = \int_Y T_t f \cdot g$. It suffices to show

$$|\Phi(t)| \leq B_0^{1-t} B_1^t. \tag{12.1}$$

2. Let $\alpha(z) = \frac{1-z}{p_0} + \frac{z}{p_1}$ and $\beta(z) = \frac{1-z}{r_0} + \frac{z}{r_1}$. Next define analytic families of functions f_z, g_z by setting[102]

$$f_z = \begin{cases} |f|^{p\alpha(z)} \mathrm{sgn}(f), \ p \neq \infty \\ f, \ p = \infty \end{cases} \quad \text{and } g_z = \begin{cases} |g|^{r'(1-\beta(z))} \mathrm{sgn}(g), \ r' \neq \infty \\ g, \ r' = \infty \end{cases}.$$

One checks easily that

$$\begin{aligned} &|f_z|^{p_0} = |f|^p \text{ and } |g_z|^{r_0'} = |g|^{r'} \text{ when Re } z = 0 \\ &|f_z|^{p_1} = |f|^p \text{ and } |g_z|^{r_1'} = |g|^{r'} \text{ when Re } z = 1 \\ &f_z = f \text{ and } g_z = g \text{ when } z = t. \end{aligned} \tag{12.2}$$

3. Define $\Phi(z) = \int_Y T_z f_z \cdot g_z$. We claim that Φ is analytic on S and bounded and continuous on $\overline{S}$. To check this use the linearity of T_z, which implies for example, that if $f = \sum_j c_j \chi_{A_j}$, then

$$T_z f_z = \sum_j |c_j|^{p\alpha(z)} \mathrm{sgn}(c_j) T_z \chi_{A_j}, \text{ when } p \neq \infty.$$

Moreover, when Re $z = 0$ we can use (12.2) to obtain

$$|\Phi(z)| \leq |T_z f_z|_{r_0} |g_z|_{r_0'} \leq B_0 |f_z|_{p_0} |g_z|_{r_0'} = B_0, \tag{12.3}$$

and similarly $|\Phi(z)| \leq B_1$ when Re $z = 1$. The Hadamard three lines lemma now gives (12.1). $\square$

Remark 12.5. *The finiteness of $|T_t|_{p \to r}$ was proved by Stein [S2] under the much weaker assumption on $\psi(z)$ that there exists $A > 0$, $a < \pi$ such that*

$$|\psi(x + iy)| \leq e^{Ae^{ar}} \textit{ for } 0 \leq x \leq 1, |y| \leq r.$$

The Hadamard lemma must now be replaced by a lemma due to Hirschman, and one obtains a more complicated formula than $B_0^{1-t} B_1^t$ for the operator norm bound.

13 Appendix 2: Probability

13.1 Khinchine's inequality

Consider the Rademacher functions $r_j(x) = \mathrm{sgn} \sin(2\pi 2^j x)$ on $[0, 1]$. They are i.i.d. random variables on $[0, 1]$ with Lebesgue measure. We have $E(r_j) = 0$, $E(r_j^2) = 1$. Let $a_j \in \mathbb{C}$, $S_N = \sum_{j=1}^N a_j r_j$, $1 \leq p < \infty$. Then Khinchine says

$$|S_N|_{L^p} \sim |S_N|_{L^2} = \left(\sum_{j=1}^N |a_j|^2 \right)^{1/2} \text{ with constants } C = C(p) \text{ independent of } N. \tag{13.1}$$

[101] Here r' is the Hölder conjugate of r.

[102] For each z, f_z is a simple function on X and g_z is a simple function on Y. Also, if $z = |r|e^{i\theta}$, we set $\mathrm{sgn}(z) = e^{i\theta}$.

The proof uses the sub-Gaussian bound: for all N, $a_j \in \mathbb{C}$, $\lambda > 0$

$$m\left\{x \in [0,1] : |S_N(x)| \geq \lambda(\sum_{j=1}^{N} |a_j|^2)^{1/2}\right\} \leq 4e^{-\lambda^2/2}.$$

If the a_j are all ± 1 this says $m\left\{x \in [0,1] : |S_N(x)| \geq \lambda\sqrt{N}\right\} \leq 4e^{-\lambda^2/2}$. The central limit theorem shows this is sharp; see [CS1], p. 114.

Proof of Khinchine using Littlewood-Paley inequalities for the Haar square function. Divide $[0,1]$ into 2^k disjoint intervals $[a_j, b_j)$ of length 2^{-k}, and let $r_k = 1$ on the first, -1 on the second, etc., with signs alternating. Redefine $h_I = \chi_{I_L} - \chi_{I_R}$. Then

$$r_k = \sum_{|I|=2^{-k+1}} h_I.$$

Let $f = \sum_k c_k r_k$. Then $f = \sum_I a_I h_I$, where $a_I = c_k$ for all $|I| = 2^{-k+1}$. Now

$$\Delta_{k-1} f = \sum_{|I|=2^{-(k-1)}} a_I h_I = c_k \sum_{|I|=2^{-(k-1)}} h_I \Rightarrow |\Delta_{k-1} f(x)|^2 = |c_k|^2.$$

Now Theorem 7.19 gives[103]

$$|f|_p \sim |Sf|_p \sim |(c_k)|_{\ell^2}.$$

$\square$

Remark 13.1. *Khinchine's inequality holds for any sequence of i.i.d. random variables X_j such that $P(X_j = \pm 1) = \frac{1}{2}$. This applies, for example, to Rademacher functions and to the (fair) coin-tossing sequence.*

13.2 Conditional expectation

Let (Ω, Σ, P) be a probability space, X an L^1 random variable, and $\mathcal{F} \subset \Sigma$ a sub-σ-algebra. Note that X is not necessarily $\mathcal{F}-$measurable.

Definition 13.2 (Conditional expectation). *The conditional expectation $E(X|\mathcal{F})$ is an $\mathcal{F}$-measurable function uniquely characterized (it turns out) by the property*

$$\int_B E(X|\mathcal{F})dP = \int_B X dP \text{ for all } B \in \mathcal{F}. \tag{13.2}$$

Remark 13.3. *If $X \in L^2$ then $E(X|\mathcal{F})$ is orthogonal projection onto the closed subspace of $L^2(\Omega)$ of $\mathcal{F}$-measurable functions. For $X \in L^1$ or for $X \geq 0$ not in L^1 can define $E(X|\mathcal{F})$ satisfying* (13.2) *by taking limits; see [CS1], p123.*

We will use the following properties (see [CS1]):

$$\begin{aligned} &(a) E(E[X|\mathcal{F}]) = EX \\ &(b) X \geq 0 \Rightarrow E[X|\mathcal{F}] \geq 0 \\ &(c) |E[X|\mathcal{F}]|_p \leq |X|_p. \end{aligned} \tag{13.3}$$

Part (c) follows from Jensen's inequality and (a).

[103] This argument was suggested in [C1].

14 Appendix 3: Assorted results.

About half of the results below are used at some point in the main text.

1. A linear operator of strong type (p,q) that commutes with translations is given on $\mathcal{S}$ by convolution with an element of $\mathcal{S}'$; Thm. 2.5.2 [G1]. For $1 \le p, q \le \infty$ denote the set of such operators by $\mathcal{M}^{p,q}(\mathbb{R}^n)$. When $p = q$ these spaces include convolution-type SIOs. For $1 \le p < \infty$ $\mathcal{M}_p(\mathbb{R}^n)$ is the set of bounded functions m such that $T_m \in \mathcal{M}^{p,p}$, where $T_m(f) = \mathcal{F}^{-1}(m\hat{f})$. We say T_m is given by the Fourier multiplier m.

2. **Dirichlet and Fejér kernels.** The Dirichlet and Fejér kernels on $\mathbb{T}^1$ are

$$\begin{aligned} D_N(x) &= \sum_{|m|\le N} e^{2\pi i m x} = \frac{\sin((2N+1)\pi x)}{\sin \pi x},\ x \in [0,1] \\ F_N(x) &= \frac{1}{N+1}\sum_{j=0}^{N} D_j(x) = \sum_{|m|\le N}\left(1 - \frac{|m|}{N+1}\right) e^{2\pi i m x} = \frac{1}{N+1}\left(\frac{\sin((N+1)\pi x)}{\sin \pi x}\right)^2 \end{aligned} \tag{14.1}$$

The (square sum) Dirichlet and Fejér kernels on $\mathbb{T}^n$ are

$$\begin{aligned} D(n,N)(x) &= \sum_{|m_j|\le N} e^{2\pi i m x} = \prod_{j=1}^{n} D_N(x_j) \\ F(n,N)(x) &= \prod_{j=1}^{n} F_N(x_j) = \sum_{|m_j|\le N}\left(1 - \frac{|m_1|}{N+1}\right)\cdots\left(1 - \frac{|m_n|}{N+1}\right) e^{2\pi i m x}. \end{aligned} \tag{14.2}$$

Dirichlet kernels fail to be approximate identities. Fejér kernels, which are averages of Dirichlet kernels, are approximate identities; section 3.1.3 [G1]. One can use Fejér kernels to show trigonometric polynomials are dense in $L^p(\mathbb{T}^n)$ for $1 \le p < \infty$; Prop. 3.1.10 [G1]. Here is an example of the use of Fejér kernels.

Proposition 14.1. *If $f, g \in L^1(\mathbb{T}^n)$ and $\hat{f}(m) = \hat{g}(m)$ for all $m \in \mathbb{Z}^n$, then $f = g$ a.e.*

Proof. Suppose $\hat{f}(m) = 0$ for all m. Then $f * F(n,N) = 0$ for all N. But $f * F(n,N) \to f$ in L^1 as $N \to \infty$, so $|f|_1 = 0$ and thus $f = 0$ a.e.

□

Corollary 14.2. *Suppose $f \in L^1(\mathbb{T}^n)$ and $\sum_{m\in\mathbb{Z}^n} |\hat{f}(m)| < \infty$. Then*

$$f(x) = \sum_{m\in\mathbb{Z}^n} \hat{f}(m) e^{2\pi i m x} \ \textit{a.e.}.$$

3. Dirichlet square sums of a continuous function may diverge at a point. Fejér means converge a.e. for $f \in L^1(\mathbb{T}^n)$. There exists an $f \in L^1(\mathbb{T}^1)$ whose Fourier series diverges a.e.

4. Convergence of Fourier series in $L^p(\mathbb{T}^1)$ is equivalent to L^p boundedness of the conjugate function or of either of the two Riesz projections. This works for $1 < p < \infty$; Defn. 3.5.4 [G1]. This extends to square Dirichlet sums for $f \in L^p(\mathbb{T}^n)$, $1 < p < \infty$. There is a similar result for Fourier integrals on the line, where the L^p boundedness of the Hilbert transform is used.

5. Problems of a.e. convergence of Fourier series on the torus can be related to boundedness of certain maximal operators on $\mathbb{R}^n$ (transference). The Carleson-Hunt theorem says that for $1 < p < \infty$ and $f \in L^p(\mathbb{T}^1)$, the Fourier series of f converges pointwise a.e. to f. This also works for $\mathbb{T}^n$; see Theorem 3.6.14 of [G1]. The relevant maximal operator is the Carleson operator

$$\mathcal{C}(f) = \sup_{R>0} \left| \int_{-R}^{R} e^{i2\pi x\xi} \hat{f}(\xi) d\xi \right|.$$

This is bounded on L^p for $1 < p < \infty$; Chapter 11, [G1].

Transference can also be used to derive $L^p(\mathbb{T}^n)$ multiplier results from $L^p(\mathbb{R}^d)$ multiplier results. This is done in one of the proofs of section 5.3.

6. **Hilbert transform.** The Hilbert transform is given by the Fourier multiplier $-i\text{sgn}\ \xi$. It is also given by convolution with $\frac{1}{\pi} p.v. \frac{1}{x}$. It is bounded on L^p for $1 < p < \infty$. The maximal HT $H^*(f)(x) = \sup_{\epsilon>0} |H^\epsilon(f)(x)|$ is bounded on L^p, $1 < p < \infty$. So Thm. 2.1.14 applies to give pointwise result.

7. **Hausdorff-Young inequality.** The Fourier transform is strong (p, p') for $1 \le p \le 2$. Proof: It is strong $(2, 2)$ and strong $(1, \infty)$. Interpolate using Riesz-Thorin.

8. If $0 < p < q < r \le \infty$, then $L^p \cap L^r \subset L^q$ and $|f|_q \le |f|_p^\theta |f|_r^{1-\theta}$, where $\frac{1}{q} = \frac{\theta}{p} + \frac{1-\theta}{r}$. (Proof: Use Hölder with exponents $p/\theta q$, $r/(1-\theta)q$.)

9. Let ϕ be increasing, C^1 on $[0, \infty)$ with $\phi(0) = 0$ and $f \in L^p(X, \mu)$. Then

$$\int_X \phi(|f|) d\mu = \int_0^\infty \phi'(\alpha) d_f(\alpha) d\alpha.$$

Important case is $\phi(\alpha) = \alpha^p$. Here $d_f(\alpha) = |\{x : |f| > \alpha\}|$.

10. (a) **Bernstein's inequality I: general case.** Let $f \in L^p(\mathbb{R}^d)$ have its spectrum contained in $\{|\xi| \le R\}$. Then for $1 \le p \le q \le \infty$ we have

$$|\partial^\alpha f|_q \le C_{\alpha,d} R^{|\alpha| + \frac{d}{p} - \frac{d}{q}} |f|_p. \tag{14.3}$$

Proof: Let $\hat{h}$ be a test function equal to 1 on $B(0, 1)$. Set $h_R(x) = R^d h(Rx)$, so $\widehat{h_R}(\xi) = \hat{h}(\frac{\xi}{R})$. Then $f = f * h_R$. Use Young's inequality to estimate the convolution and its derivatives.

(b) **Bernstein's inequality II: often used cases.**

$$\begin{aligned} &|\partial^\alpha f|_\infty \lesssim R^{|\alpha|} |f|_\infty \\ &|\partial^\alpha f|_\infty \lesssim R^{|\alpha| + \frac{d}{2}} |f|_2. \end{aligned} \tag{14.4}$$

11. **Tempered distributions modulo polynomials.** Let $\mathcal{S}_0$ be the subspace of $\mathcal{S}$ consisting of ϕ such that $\partial^\alpha \hat{\phi}(0) = 0$ for all α. Then $\mathcal{S}'/\mathcal{P} := \mathcal{S}'_P$ is the dual of $\mathcal{S}_0$. If $f \in \mathcal{S}'$ and $\sum_{j\in\mathbb{Z}} \hat{\psi}(2^{-j}\xi) = 1$ is a dyadic partition of unity, then $\sum_{-N}^N P_j f \to f$ in $\mathcal{S}'_P$ as $N \to \infty$. By duality this is equivalent to $\sum_{-N}^N \hat{\psi}(2^{-j}\xi)\hat{\phi} \to \hat{\phi}$ when $\phi \in \mathcal{S}_0$. In the proof consider a sum over $j \ge N$, where infinite order decay of $\hat{\phi}$ is used, and a sum over $j \le -N$ where infinite order vanishing at 0

of $\hat{\phi}$ and the support condition $|\xi| \le 2^{-N}$ are used.

12. **Definition of H^1.** $H^1(\mathbb{R}^d)$ is the completion of $\mathcal{S}$ under $|f|_{H^1} = |f|_{L^1} + \sum_j |R_j f|_{L^1}$, where the R_j are Riesz transforms. Using $(H^1)^* = BMO$, one can then write $BMO = L^\infty + \sum_j R_j L^\infty$, where the action of R_j on L^∞ has to be defined suitably. BMO is the smallest space that contains $T(L^\infty)$ for all SIOs as in Defn. 7.1 of [CS1].

13. The Haar functions span $L^p([0,1])$ for $1 \le p < \infty$ (in the sense of a Schauder basis). The basis is unconditional for $1 < p < \infty$.

14. The series $\sum_{n=1}^{\infty} x_n$ in a Banach space converges unconditionally if $\sum \epsilon_n x_n$ converges for all ϵ_n with $\epsilon_n = \pm 1$. Equivalent conditions: (a) all rearrangements converge, (b) $\sum a_n x_n$ converges whenever $(a_n) \in \ell^\infty$. Unconditional convergence does not imply absolute convergence: $\sum c_n e_n$ converges unconditionally in ℓ^p when $(c_n) \in \ell^p$ for $1 \le p < \infty$, but converges absolutely only when $(c_n) \in \ell^1$.

15. **Uniform boundedness of T_N implies boundedness of T: standard Fatou argument.**

Proposition 14.3. *Let $T : \mathcal{S} \to \mathcal{D}'$ be a linear operator and let T_N be a family of linear operators such that $T_N f \to Tf$ pointwise for $f \in \mathcal{S}$. Suppose for some $1 \le p < \infty$ that there exists a C such that $|T_N|_{p \to p} \le C$. Then T extends to an operator on L^p such that $|T|_{p \to p} \le C$.*

Proof. Let $f \in \mathcal{S}$. By Fatou

$$\int |Tf|^p(x)dx = \int \liminf_{N \to \infty} |T_N f(x)|^p dx \le \liminf_{N \to \infty} \int |T_N f|^p dx \le C|f|_p^p.$$

□

16. **Generalized triangle inequality.** For any $1 \le q \le p \le \infty$ we have (where $(f_j) := (f_j)_{j \in \mathbb{Z}}$)

$$||(f_j)|_{\ell^q}|_{L^p(\mu)} \le \left|(|f_j|_{L^p(\mu)})\right|_{\ell^q}. \tag{14.5}$$

An important case is where $q = 2 \le p$. Then

$$\left|\left(\sum_j |f_j|^2\right)^{1/2}\right|_{L^p(\mu)} \le \left(\sum_j |f_j|^2_{L^p(\mu)}\right)^{1/2}.$$

Applied to the LP square function this gives[104]

$$|Sf|_p^2 \le \sum_j |P_j f|^2_{L^p(\mu)}.$$

17. **The TT^t (or TT^*) principle.** Suppose B_1, B_2 are Banach spaces, $T : B_1 \to B_2$ and $T^t : B_2' \to B_1'$. (a) Then T is bounded if and only if T^t is bounded and $|T| = |T^t|$. (b) If B_1 is a Hilbert space, so $B_1 = B_1'$, then boundedness of T is equivalent to boundedness of $TT^t : B_2' \to B_2$.

18. **Poisson summation formula**. If $f \in L^1(\mathbb{R}^n)$ and if its Fourier series converges, it is natural to ask what function on $\mathbb{R}^n$ the series represents. Under appropriate conditions it represents the periodization of f, $\sum_{m \in \mathbb{Z}^n} f(x+m)$.

[104] See section 6.4 for an application to Sobolev embeddings.

Proposition 14.4. *Suppose $f, \hat{f} \in L^1(\mathbb{R}^n)$ and*

$$|f(x)| + |\hat{f}(x)| \lesssim (1+|x|)^{-n-\delta} \text{ for some } \delta > 0.$$

Then $f, \hat{f}$ are continuous and for all x we have

$$\sum_{m\in\mathbb{Z}^n} \hat{f}(m)e^{2\pi imx} = \sum_{m\in\mathbb{Z}^n} f(x+m).$$

Proof. Define $F(x) = \sum_m f(x+m) \in L^1(\mathbb{T}^n)$. We have

$$\hat{F}(m) = \int_{\mathbb{T}^n} F(x)e^{-2\pi imx}dx = \sum_{k\in\mathbb{Z}^n} \int_{\mathbb{T}^n+k} f(x)e^{-2\pi imx}dx = \hat{f}(m).$$

Thus, $\sum_m |\hat{F}(m)| < \infty$ and Cor. 14.2 gives the result. □

19. **Poisson kernel.** For $x \in \mathbb{R}^n$ the Poisson kernel is $P(x) = \frac{c_n}{(1+|x|^2)^{(n+1)/2}}$. Harmonic functions on $\mathbb{R}^{n+1}_+$ with boundary value $f \in L^p$, $1 \le p \le \infty$ are given by

$$P_t * f(x) = c_n \int_{\mathbb{R}^n} \frac{t}{|(t, x-y)|^{n+1}} f(y)dy,$$

where $P_t(x) = t^{-n}P(x/t) = c_n \frac{t}{|(t,x)|^{n+1}}$. Then for each $t > 0$, $P_t * f \in L^p$. If $1 \le p < \infty$ we have $P_t * f \to f$ a.e. as $t \to 0$.[105]

20. **Young's inequality** (a) If $f \in L^1$, $g \in L^p$ $(1 \le p \le \infty)$, then $|f * g|_p \le |f|_1 |g|_p$.

(b) General form: Suppose $1 \le p, q, r \le \infty$ and $p^{-1} + q^{-1} = r^{-1} + 1$. If $f \in L^p$ and $g \in L^q$ then $|f * g|_r \le |f|_p |g|_q$.

21. **Whitney decomposition of an open set; [S1], p.167.** By a "cube" we mean a closed cube with edges parallel to the axes.

Theorem 14.5. *Let $\Omega \subset \mathbb{R}^n$ be open. There exists a countable set of cubes Q_k such that*

1) $\cup_k Q_k = \Omega$,

2) The Q_k have disjoint interiors.

3) diam $Q_k \le$ dist $(Q_k, \Omega^c) \le 4$ diam Q_k.

15 References

[AG] Alinhac, S. and Gerard, P., *Pseudodifferential operators and the Nash-Moser theorem*, Grad. Studies in Math., vol. 82, AMS, 2007.

[B] Bony, J.-M., Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires, Ann. Sci. Ec. Norm. Sup. 14 (1981), 209-246.

[C1] Christ, M., *Lectures on Singular Integral Operators*, CBMS Regional Conference Series, number 77, AMS, 1990.

[CS1] Muscalu, C. and Schlag, W., *Classical and Multilinear Harmonic Analysis*, Volume 1, Cambridge Univ. Press, 2013.

[105] This last statement requires a correct choice of c_n of course.

[CS2] Muscalu, C. and Schlag, W., *Classical and Multilinear Harmonic Analysis*, Volume 2, Cambridge Univ. Press, 2013.

[F] Folland, G., *Real Analysis*, Pure and Applied Math., Wiley, 1984.

[G1] Grafakos, L., *Classical Fourier Analysis*, Grad. Texts in Math., 249, second edition, Springer 2008.

[G2] Grafakos, L., *Modern Fourier Analysis*, Grad. Texts in Math., 250, third edition, Springer 2014.

[GR] Garcia-Cuerva, J. and Rubio de Francia, J.L., *Weighted Norm Inequalities and Related Topics*, North Holland Math, Vol 116.

[Met], Metivier, G., *Paradifferential calculus and applications to the Cauchy problem for nonlinear systems*, Notes from a course at the University of Pisa, 2008.

[S1] Stein, E., *Singular Integrals and Differentiability Properties of Functions*, Princeton Math. Series, vol. 30, PUP, 1970.

[S2], Stein, E., *Interpolation of linear operators*, T.A.M.S, 83 (1956), 482-492.

[Tao] Tao, T., Course notes for Math 247A,B (Harmonic Analysis) at UCLA

[Tao2] Tao, T., *Nonlinear Dispersive Equations*, CBMS Regional Conference Series, number 106, AMS, 2006.

[W] Wilson, J. Michael, *On the atomic decomposition for Hardy spaces*, Pacific J. Math, 116, (1), 1985, 201-207.