LECTURE NOTES 3 FOR 254B

TERENCE TAO

1. Bochner-Riesz summation

We all know that given the Fourier transform

$$\hat{f}(\xi) = \int_{\mathbf{R}^n} e^{-2\pi i x \cdot \xi} f(x)\, dx$$

of a function, one can reconstruct it via the Fourier inversion formula

$$f(x) = \int_{\mathbf{R}^n} e^{2\pi i x \cdot \xi} \hat{f}(\xi)\, d\xi,$$

at least if f is smooth. But if f is not smooth - for instance, if f is merely in L^p - it is not clear whether the Fourier inversion formula is true, or even make sense.

To make things more precise, let us define the partial Fourier integrals $S_R f(x)$ by

$$S_R f(x) = \int_{|\xi| \leq R} e^{2\pi i x \cdot \xi} \hat{f}(\xi)\, d\xi,$$

and ask whether

$$S_R f \to f \text{ in } L^p \text{ norm}$$

for arbitrary $f \in L^p$. There are other types of convergence, such as pointwise convergence, but they are harder to deal with. (Note that weak convergence is not a problem here).

The operators S_R are known as disc multipliers (because they are Fourier multipliers corresponding to the characteristic function of the disc $|\xi| \leq R$), and the problem was known as the disc multiplier problem (at least, until the problem was solved).

The answer to the disc multiplier problem is completely known. In one dimension $n = 1$, one has convergence in L^p norm for all $1 < p < \infty$. However in higher dimensions one only has convergence in L^2; for $p \neq 2$ convergence fails! This phenomenon was a great shock at the time it was discovered (by Fefferman in the late 70s), and the reason has to do with Besicovitch sets, which we will return to later in this course.

To begin with, we first recall that L^p convergence questions are basically equivalent to an L^p norm estimate. (This is one reason why this type of convergence question is easier than other types).

Lemma 1.1 (Uniform boundedness principle). *Let $1 < p < \infty$, and suppose that T_R is a sequence of linear operators such that $T_R f \to f$ in L^p norm as $R \to \infty$ for test functions f. Then in order for $T_R f$ to converge to f in L^p norm for all L^p functions f (i.e. not just for test functions) it is a necessary and sufficient condition that we have the estimate*

$$\|T_R f\|_p \lesssim \|f\|_p \tag{1}$$

for all sufficiently large R, where the constants are independent of R.

Proof This should be in any decent real analysis book. (Me, I was brought up on Rudin's "real and complex analysis".) ∎

Thus to settle the question of whether $S_R f$ converges in L^p to f, it suffices to show the bound

$$\|S_R f\|_p \lesssim \|f\|_p$$

uniformly in R. Actually, it suffices to prove this for $R = 1$, since a simple scaling argument shows that the $R = 1$ case automatically implies the general case. More generally, if a multiplier operator of the form

$$\widehat{Tf}(\xi) = m(\xi)\hat{f}(\xi)$$

is bounded on L^p, and L is any affine transformation whatsoever, then

$$\widehat{T_L f}(\xi) = m(L\xi)\hat{f}(\xi)$$

is also bounded on L^p with exactly the same operator norm. (Exercise).

So we have to see whether

$$\|S_1 f\|_p \lesssim \|f\|_p.$$

Let's first consider the one-dimensional case. In this case we see that

$$S_1 f = \mathcal{F}^{-1}(\chi_{[-1,1]}\hat{f})$$

where $\mathcal{F}^{-1}$ is the inverse Fourier transform. Using the identity

$$\chi_{[-1,1]}(x) = \frac{1}{2}(\operatorname{sgn}(x+1) - \operatorname{sgn}(x-1))$$

we see that it suffices to prove the L^p boundedness of the operator

$$f \mapsto \mathcal{F}^{-1}(\operatorname{sgn}(x \pm 1)\hat{f}).$$

By the invariance of multipliers under affine transformations remarked upon earlier, it thus suffices to prove

$$f \mapsto \mathcal{F}^{-1}(\operatorname{sgn}(x)\hat{f}).$$

But this operator is just $\frac{i}{\pi}H$, where H is the Hilbert transform

$$Hf(x) = p.v. \int f(x-y) \frac{dy}{y}.$$

The result now follows from the classical L^p boundedness result of the Hilbert transform.

Now let's look at the n-dimensional case. When $p = 2$ we have boundedness directly from Plancherel, since the multiplier corresponding to S_1 is bounded. What about $p \neq 2$? It turns out that we do not have boundedness here!

Proposition 1.2 (Fefferman). *For any $N > 0$ and $p > 2$ there exists a non-zero test function f such that*

$$\|S_1 f\|_p \geq N \|f\|_p. \tag{2}$$

We'll delay the proof of Fefferman's result for a while, and ask how to get around the lack of convergence. One way to do this is to make the frequency cutoff more gentle. For any $\delta \geq 0$, define the Bochner-Riesz mean S_R^δ by

$$S_R f(x) = \int_{|\xi| \leq R} (1 - \frac{|\xi|^2}{R^2})^\delta e^{2\pi i x \cdot \xi} \hat{f}(\xi)\, d\xi.$$

When $\delta = 0$ this is just the disc multiplier S_R. When $\delta = 1$ these operators are essentially Fejer means. The point is that the larger δ gets, the smoother the multiplier gets at $|\xi| = R$, which helps convergence.

These operators were first systematically studied in one dimension by Riesz, and in higher dimensions by Bochner.

It's easy to see that $S_R^\delta f$ converges to f uniformly if f is a test function. Thus, we may pose the same question as before, namely whether $S_R^\delta f$ converges to f in L^p norm for all f in L^p. The question reduces as before to verifying whether

$$\|S_1^\delta f\|_p \lesssim \|f\|_p. \tag{3}$$

A necessary condition was found by Herz:

Theorem 1.3 (Herz). *In order for* (3) *to hold, one must have*

$$|\frac{1}{p} - \frac{1}{2}| < \frac{2\delta + 1}{2n}. \tag{4}$$

In particular, we see that the larger δ gets, the larger the interval on which we have a shot at convergence. This doesn't quite give Fefferman's theorem, by the way, but it does give the weaker observation that (2) holds whenever $p \geq 2n/(n-1)$ or $p \leq 2n/(n+1)$.

Let's see how the proof goes. We first observe that S_1^δ is a Fourier multiplier operator with multiplier $(1-|\xi|^2)_+^\delta$. This multiplier is real, so the operator is self-adjoint. Thus boundedness on L^p is equivalent to boundedness on $L^{p'}$. (Actually, even if the multiplier was not real, this last statement would still be true - exercise!). So it suffices to consider the $p < 2$ case.

Let f be a Schwarz function whose Fourier transform is identically 1 on the unit ball. (To find this function, first choose $\hat{f}$). Testing (3) with this value of f, we see that we must have

$$\|\hat{m}_\delta\|_p \lesssim 1.$$

where

$$m_\delta(\xi) = (1-|\xi|^2)_+^\delta.$$

We claim that

$$\hat{m}_\delta(x) = \frac{Ce^{2\pi i|x|} + Ce^{-2\pi i|x|} + o(1)}{|x|^{\frac{n+1}{2}+\delta}} \tag{5}$$

as $|x| \to \infty$; inserting this estimate into the previous we obtain a contradiction unless

$$\frac{n}{p} < \frac{n+1}{2} + \delta;$$

and this is just a re-arrangement of (4).

We'll defer the proof of (5) to an appendix. Note however that it is very similar to the formula for $\widehat{d\sigma}$. Morally speaking, $\widehat{d\sigma}$ is comparable to m_{-1}; this is akin to the heuristic that the delta function is "of the same strength as" the distribution $1/x$. Note that every time δ is lowered by 1, (5) predicts that $\hat{m}_\delta$ is multiplied by roughly $|x|$; this is consistent with the heuristic observation that the derivative of m_δ is roughly comparable to $m_{\delta-1}$. Also, this formula matches up with the ones for the classical Dirichlet and Fejer kernels, which correspond to the cases $(n,\delta) = (1,0)$ and $(1,1)$ respectively. As one last heuristic, note that m_δ is in L^2 when $\delta > -1/2$, and the same goes for the right-hand side of (5). None of this adds up to a proof, but it can help in remembering the numerology of (5).

Given the necessary conditions of Herz and Fefferman, it is natural to conjecture

Conjecture 1.4 (Bochner-Riesz conjecture). *Let $\delta > 0$ and $1 \leq p \leq \infty$ be such that (4) holds. Then $S_R^\delta f$ converges to f in L^p norm for all $f \in L^p$.*

This conjecture has only been completely settled for $n \leq 2$. There are some easy cases we can dispose of first. When $p = 2$ the conjecture is true, since S_R^δ is given by a bounded multiplier. When $\delta > (n-1)/2$ the formula (5) implies that the convolution kernel of S_R^δ is integrable (note that there is no singularity near zero, in fact $\hat{m}_\delta$ must be analytic since m_δ is compactly supported), so the claim follows from Young's inequality.

It is possible to interpolate between these results using (for instance) complex interpolation. However, using restriction theory one can do better.

Theorem 1.5 (Fefferman). *Let S be the unit sphere. Suppose that p is such that $R_S(2 \to p)$ holds. Then the Bochner-Riesz conjecture is true for this exponent p.*

Proof Fix δ such that (4) holds. We have to show that

$$\|f * K_\delta\|_p \lesssim \|f\|_p$$

where K_δ is the Fourier transform of m_δ.

This looks like the estimate we used to prove the Tomas-Stein theorem, except that it's (L^p, L^p) instead of $(L^p, L^{p'})$. As before, we break up K as

$$K_\delta = \phi K_\delta + \sum_{k>0} \psi_k K_\delta;$$

we won't need any fancy moment conditions on ϕ or ψ this time. This is because we have strict inequality in (4) and don't have to do anything subtle. (People have studied the endpoint, because one still gets weak-type (p, p) estimates there even though strong-type estimates fail. For the endpoint one does need to be more delicate. See [3], [6]).

The function ϕK_δ is a bump function, and so convolving with this guy is clearly bounded on L^p (e.g. by Young's inequality). So as before we are left with showing that

$$\|\sum_{k>0} f * \psi_k K_\delta\|_p \lesssim \|f\|_p.$$

Since we have a bit of room in (4), we will just use the triangle inequality. Actually, we're going to show

$$\|f * \psi_k K_\delta\|_p \lesssim 2^{(n(\frac{1}{p}-\frac{1}{2})-\frac{1}{2}-\delta)k}\|f\|_p. \tag{6}$$

When (4) holds, the constant happens to decay exponentially in k, so we get what we want from the triangle inequality.

The kernel $\psi_k K_\delta$ is supported on an annulus $\{x : |x| \sim 2^k\}$. So this operator is somewhat localized; the values of f at a point x only influence points which are close by (within $O(2^k)$). Now we invoke a useful lemma that reduces the study of such "local" operators to a compact set.

Lemma 1.6. *Suppose T is a linear operator taking functions on $\boldsymbol{R}^n$ to functions of $\boldsymbol{R}^n$. Suppose T is local in the sense that the support of Tf always remains within R of the support of f for some $R > 0$. Then for any $1 \leq p \leq q \leq \infty$, the bound*

$$\|Tf\|_q \lesssim \|f\|_p \text{ for all } f \in L^p(\boldsymbol{R}^n) \tag{7}$$

is equivalent to the bound

$$\|Tf\|_{L^q(B(x,2R)} \lesssim \|f\|_p \text{ for all } f \in L^p(B(x,R)) \tag{8}$$

holding uniformly in x. In other words, to show (7) *it suffices to test it for functions supported on an R-ball.*

Intuitively, the idea is that functions on distinct R-balls basically do not interfere with each other.

Proof Clearly (7) implies (8). Now suppose that (8) holds uniformly in x. Let f be a function in $L^p(\mathbf{R}^n)$. We cover $\mathbf{R}^n$ by a finitely overlapping collection $\{B\}$ of R-balls, and make a partition of unity $1 = \sum_B \psi_B$ subordinate to this partition. We write

$$\|Tf\|_q^q = \int |T(\sum_B \psi_B f)|^q = \int |\sum_B T(\psi_B f)|^q.$$

Since T is local, the functions $T(\psi_B f)$ are just supported on the double $2B$ of B. These balls are finitely overlapping, so we have the pointwise estimate

$$|\sum_B T(\psi_B f)|^q \lesssim \sum_B |T(\psi_B f)|^q.$$

Putting this back into the previous estimate and simplifying, we obtain

$$\|Tf\|_q \lesssim (\sum_B \|T(\psi_B f)\|_q^q)^{1/q}.$$

Applying (8) we get

$$\|Tf\|_q \lesssim (\sum_B \|\psi_B f\|_p^q)^{1/q}.$$

We now observe the elementary inequality

$$(\sum_B a_B^q)^{1/q} \leq (\sum_B a_B^p)^{1/p}$$

which follows from the assumption $q \geq p$. This is easily established by checking the easy cases $q = p$ and $q = \infty$ and then using Hölder's inequality to interpolate. Thus we have

$$\|Tf\|_q \lesssim (\sum_B \|\psi_B f)\|_p^p)^{1/p}.$$

Since the balls B are finitely overlapping, it is easy to see that the right-hand side is essentially $\|f\|_p$. ∎

The condition that $q \geq p$ is absolutely necessary. This is an example of Littlewood's principle: "the higher exponents are always on the left". More precisely:

Lemma 1.7. *Let T be a non-zero translation invariant operator. Then the estimate $\|Tf\|_q \lesssim \|f\|_p$ is only possible if $q \geq p$.*

Proof Let ϕ be any bump function such that $T\phi$ is non-zero. Let $N > 0$ be a large number, and let $x_1, \dots x_N$ be N very widely separated points. Define f by

$$f(x) = \sum_{i=1}^{N} \phi(x - x_i).$$

If the above estimate held for f, we would have

$$\| \sum_{i=1}^{N} T\phi(x - x_i)\|_q \lesssim \| \sum_{i=1}^{N} \phi(x - x_i)\|_p$$

since T is translation invariant. Now if the x_i are sufficiently far apart, the summands are essentially disjointly supported. The right-hand side grows in N like $N^{1/p}$, and the left-hand side grows like $N^{1/q}$. Thus we must have $N^{1/q} \lesssim N^{1/p}$ for all large N, which implies $q \geq p$ as desired. ■

Let's return to the proof of (6). From the above discussion we just have to prove

$$\|f * \psi_k K_\delta\|_{L^p(B(x,C2^k))} \lesssim 2^{(n(\frac{1}{p}-\frac{1}{2})-\frac{1}{2}-\delta)k}\|f\|_p$$

for all f supported on a ball $B(x, \frac{1}{2}C2^k)$. By translation invariance we may take $x = 0$.

We're supposed to be using the Tomas-Stein restriction theorem, which is an (L^p, L^2) estimate. But we don't have an L^2 norm on the right, just an L^p norm. But we can fix that easily enough, using Hölder's inequality and the fact that our L^p norm is on a compact set. More precisely, we have

$$\|F\|_{L^p(B(0,C2^k))} = \|F\chi_{B(0,C2^k)}\|_p \leq |B(0, C2^k)|^{\frac{1}{p}-\frac{1}{2}}\|F\|_2 \lesssim 2^{nk(\frac{1}{p}-\frac{1}{2})}\|F\|_2;$$

note that the hypothesis $R_S(2 \to p)$ implies that $p < 2$. So we only need to show that

$$\|f * \psi_k K_\delta\|_2 \lesssim 2^{-k/2}2^{-\delta k}\|f\|_p$$

for all f supported on $B(0, C2^k)$.

Okay, this is looking more like Tomas-Stein, but we don't have a Fourier transform on the right-hand side like we do in a restriction estimate. But that's okay, we just use Plancherel:

$$\|\hat{f}\widehat{\psi_k K_\delta}\|_2 \lesssim 2^{-k/2}2^{-\delta k}\|f\|_p.$$

Since K_δ is the Fourier transform of m_δ, this is

$$\|\hat{f}(\hat{p}si_k * m_\delta)\|_2 \lesssim 2^{-k/2}2^{-\delta k}\|f\|_p.$$

We will shortly show the key estimate

$$|\hat{\psi}_k * m_\delta(\xi)| \lesssim 2^{-\delta k}(1 + 2^k d(\xi, S))^{-N} \tag{9}$$

for any N. Assuming this for the moment, we see that we just need to show that

$$\|\hat{f}(1 + 2^k d(\xi, S))^{-N}\|_2 \lesssim 2^{-k/2}\|f\|_p$$

to finish the proof. We square this as

$$\int \frac{|\hat{f}|^2}{(1 + 2^k d(\xi, S))^{2N}}\ d\xi \lesssim 2^{-k}\|f\|_p^2.$$

We first deal with the portion when $d(\xi, S) > 1/2$. This is an error term, and we shall be very crude. We may estimate $\hat{f}$ pointwise by $2^{Ck}\|f\|_p$ for some C, just from the definition of $\hat{f}$, Hölder's inequality, and the fact that f is compactly supported. On the other hand, the denominator in the integral is 2^{-Nk} for any N, and rapidly decreasing as $\xi \to \infty$. This decay beats all other factors, and the bound is easy to prove. So we just need to show

$$\int_{1/2 \leq |\xi| \leq 3/2} \frac{|\hat{f}|^2}{(1 + 2^k d(\xi, S))^{2N}}\ d\xi \lesssim 2^{-k}\|f\|_p^2.$$

We use polar co-ordinates, discarding the bounded Jacobian, and rewrite this as

$$\int_{1/2}^{3/2} (1 + 2^k|r-1|)^{-2N} (\int_{rS^{n-1}} |\hat{f}|^2\ d\omega)\ dr \lesssim 2^{-k}\|f\|_p^2$$

Now we recall the hypothesis $R_{S^{n-1}}(p \to 2)$. By earlier remarks we therefore have $R_{rS^{n-1}}(p \to 2)$ uniformly for $1/2 \leq r \leq 3/2$. So we can estimate the left-hand side of this as

$$\int_{1/2}^{3/2} (1 + 2^k|r-1|)^{-2N}\|f\|_p^2\ dr$$

and the claim follows. (To evaluate the r integral, do the interval where $|r-1| \leq 2^{-k}$ and $|r-1| > 2^{-k}$ separately.)

We now have to show (9). The idea is as follows. Since ψ_k vanishes at the origin to infinite order, $\hat{\psi}_k$ has vanishing moments to infinite order. On the other hand, m_δ is quite smooth (except on the unit sphere). The combination of the two should lead to much cancellation.

We first smooth out m_δ a bit to remove this annoying singularity at the sphere. Write

$$m_\delta(\xi) = m_\delta(\xi)\phi(2^k(|\xi|-1)) + m_\delta(\xi)(1 - \phi(2^k(|\xi|-1))) = m_\delta^0 + m_\delta^1$$

where ϕ is a bump function which equals 1 near the origin (it doesn't matter whether this is the same ϕ we used to define ψ_k, but it may as well be). The second term here is nice and smooth, but let's look at the first term for the moment. It's supported in the annulus of thickness $C2^{-k}$ around the unit sphere, and is quite small there, only about $2^{-\delta k}$. In particular, we see that

$$m_\delta^0 \lesssim \int_{|r-1|\lesssim 2^{-k}} 2^{-\delta k} d\sigma_r dr$$

where $d\sigma_r$ is surface measure on the sphere of radius r. Thus we can estimate the contribution of m_δ^0 to (9) as

$$\int_{|r-1|\lesssim 2^{-k}} 2^{-\delta k} d\sigma_r * |\widehat{\psi}_k| dr$$

However, from our previous work we know that

$$d\sigma_r * |\widehat{\psi}_k|(x) \lesssim 2^k(1 + 2^k d(x, S))^{-N};$$

(r is so close to 1 that it may as well be 1). Thus we see that the contribution of this case is acceptable.

Now we turn to the main term, which is

$$\widehat{\psi}_k * m_\delta^1(\xi).$$

To handle this we will (implicitly) take advantage of the moment conditions of $\widehat{\psi}_k$ by writing it as a derivative of a similar function.

Since ψ_0 vanishes near the origin, we may write $\psi_0(x) = |x|^{2N}\eta(x)$ for some bump function η and some large N. This implies that

$$\psi_k(x) = 2^{-2Nk}|x|^{2N}\eta(2^{-k}x)$$

so that

$$\hat{\psi}_k(\xi) = 2^{-2Nk}\Delta^N\varphi_k(\xi)$$

where

$$\varphi_k(\xi) = 2^{nk}\hat{\eta}(2^k\xi).$$

Since derivatives can be passed between the two factors of a convolution via integration by parts, we thus see that

$$\widehat{\psi}_k * m_\delta^1 = 2^{-2Nk}\varphi_k * \Delta^N m_\delta^1.$$

We now observe that φ_k is an approximation to the identity of thickness 2^{-k}, so that

$$|\varphi_k(x)| \lesssim 2^{nk}(1 + 2^k|x|)^{-N}.$$

Also, repeated applications of the product rule show that

$$|2^{-2Nk}\Delta^N m_\delta^1(\xi)| \lesssim 2^{-\delta k}(1 + 2^k d(x, S))^{-N+\delta}.$$

So we are basically back to the same situation we had for the m_δ^0 term, except that the analogue of m_δ^0 is not perfectly supported on the annulus of thickness $C2^{-k}$ but has a rapidly decreasing tail. But this tail causes no extra difficulty, as can be easily verified (it is convenient to divide into the cases $d(\xi, S) \leq 1/2$ and $d(\xi, S) > 1/2$ respectively). This completes the proof of (9) and thus of the theorem. ■

Thus we have proven the Bochner-Riesz conjecture for all $1 \leq p \leq 2(n+1)/(n+3)$. The Bochner-Riesz problem is self-adjoint, so we also have proven it for $p \geq 2(n+1)/(n-1)$.

The region $2(n+1)/(n+3) < p < 2(n+1)/(n-1)$ is much more difficult, except when $p = 2$. By interpolation one can get some results, but they are not optimal (we do not get the sharp value of δ. Further progress has been made on this problem, but surprisingly little. When $n \leq 2$ the whole conjecture is completely proven. When $n = 3$ we know a little bit more than the above results, which give the conjecture for $1 \leq p \leq 4/3$ or $p \geq 4$; we actually know the conjecture to be true for $p \leq 26/19$ and $p \geq 26/7$. As one can see we are still a ways from settling the whole thing.

2. Fefferman's disc multiplier counterexample

We now prove Proposition 1.2. By duality we may restrict ourselves to the case $p > 2$.

We first observe a nice observation about multiplier operators, which lets us start with a boundedness result for a high-dimensional operator and deduce the corresponding result for a low dimensional "slice" of the operator. (It would be very nice if we could somehow reverse this direction of implication and use the low-dimensional theory to understand the high-dimensional theory!)

Lemma 2.1 (De Leeuw). *Suppose that m is a smooth function on $\mathbf{R}^n$, and that the operator T defined by*

$$\widehat{Tf}(\xi) = m(\xi)\hat{f}(\xi)$$

is bounded on $L^p(\mathbf{R}^n)$. Then the operator T_0 defined by

$$\widehat{T_0 g}(\underline{\xi}) = m(\underline{\xi}, 0)\hat{g}(\underline{\xi})$$

for $\underline{\xi} \in \mathbf{R}^{n-1}$, is bounded on $L^p(\mathbf{R}^{n-1})$.

Proof From the invariance of L^p multiplier bounds under affine transformations, we see that we may replace $m(\xi)$ by

$$m_R(\underline{\xi}, \xi_n) = m(\underline{\xi}, \xi_n/R)$$

in the definition of T without affecting the L^p boundedness property. Letting $R \to \infty$ and taking limits, we may replace m by

$$m_\infty(\underline{\xi}, \xi_n) = m(\underline{\xi}, 0).$$

In other words, the operator

$$\widehat{T_\infty f}(\underline{\xi}, \xi_n) = m(\underline{\xi}, 0)\hat{f}(\underline{\xi}, \xi_n)$$

is bounded on $L^p(\mathbf{R}^n)$. If we now apply this fact to a function of the form

$$f(\underline{\xi}, \xi_n) = g(\underline{\xi})\psi(\xi_n)$$

and observe that

$$\hat{f}(\underline{\xi}, \xi_n) = \hat{g}(\underline{\xi})\hat{\psi}(\xi_n)$$

we obtain the desired result. ∎

Because of De Leeuw's principle, it suffices to prove Fefferman's result for $n = 2$.

Let $1 \ll a \ll R$ be large numbers to be chosen later. (a is going to end up at about 100, and R is going to be sent off to infinity). Let ψ be a non-negative bump function on the square $[0,1] \times [0,1]$ which is strictly positive on the sub-square $[1/3, 2/3] \times [1/3, 2/3]$, and consider the function

$$f(x_1, x_2) = e^{2\pi i x_1}\psi(\frac{x_1}{R^2}, \frac{x_2}{aR}).$$

Thus this function is supported on the rectangle $[0, R^2] \times [0, aR]$, and has a certain oscillation. Its Fourier transform is given by

$$\hat{f}(\xi_1, \xi_2) = aR^3\hat{\psi}(R^2(\xi_1 - 1), aR\xi_2).$$

Since $\hat{\psi}$ is rapidly decreasing, we see that $\hat{f}$ is concentrated near the point $(1, 0)$. Since

$$\widehat{S_1 f}(\xi) = \chi_{|\xi| \leq 1}(\xi)\hat{f}(\xi),$$

we expect that it will not be too wasteful to approximate $\chi_{|\xi|\leq 1}$ by $\chi_{\xi_1 \leq 1}$:

$$\widehat{S_1 f}(\xi) = \chi_{\xi_1 \leq 1}(\xi)\hat{f}(\xi) - \chi_{\xi_1 \leq 1, |\xi| \geq 1}\hat{f}(\xi).$$

We think of the first term as the main term, and the second term as an error term.

Let's first deal with the error term. The contribution to $S_1 f(x)$ is

$$-\int e^{2\pi i x\xi}\chi_{\xi_1 \leq 1, |\xi| \geq 1} aR^3\hat{\psi}(R^2(\xi_1 - 1), aR\xi_2)\ d\xi.$$

We'll estimate this crudely[1] by absolute values:

$$\int_{\xi_1 \leq 1, |\xi| \geq 1} aR^3|\hat{\psi}(R^2(\xi_1 - 1), aR\xi_2)|\ d\xi_1 d\xi_2.$$

[1] Getting estimated by absolute values is a common fate for error terms. The subtle analysis methods tend to get reserved for the main terms.

We first deal with the case where $\xi_1 \leq 0$. In this case the $\hat{\psi}$ term is incredibly small: it is $O(R^{-N})$ for any N, and decays rapidly at infinity. So it's easy to see that the total contribution of this part is $O(R^{-N})$. As for the remainder, we observe that

$$\xi_1 \geq 0, |\xi| \geq 1 \implies \xi_1 > 1 - \frac{\xi_2^2}{2}.$$

Thus it suffices to estimate

$$\int_{1-\frac{\xi_2^2}{2} \leq \xi_1 \leq 1} aR^3 |\hat{\psi}(R^2(\xi_1 - 1), aR\xi_2)| \, d\xi_1 d\xi_2.$$

We make the change of variables $x = R^2(\xi_1 - 1)$, $y = aR\xi_2$:

$$\int_{-\frac{y^2}{2a} \leq x \leq 0} |\hat{\psi}(x, y)| \, dx dy.$$

Since $\hat{\psi}$ is rapidly decreasing, we certainly have

$$|\hat{\psi}(x, y)| \lesssim \frac{1}{(1 + y^2)^1 0}$$

and so we can estimate this as

$$\int_{-\infty}^{\infty} \frac{y^2}{2a} \frac{1}{(1 + y^2)^{10}} \, dy = O(1/a).$$

Now let's look at the main term. We rewrite it on the Fourier side as

$$\chi_{\xi_1 \leq 1}(\xi) \hat{f}(\xi) = aR^3 (\hat{\psi} \chi_H)(R^2(\xi_1 - 1), aR\xi_2)$$

where $H = \{(\xi_1, \xi_2) : \xi_1 \leq 0\}$. The inverse Fourier transform of this is

$$e^{2\pi i x_1} (\psi * \mathcal{F}^{-1} \chi_H)(\frac{x_1}{R^2}, \frac{x_2}{aR}).$$

Since

$$\chi_H(\xi_1, \xi_2) = \frac{1}{2} - \frac{1}{2} \mathrm{sgn}(\xi_1)$$

we have (in the sense of distributions)

$$\mathcal{F}^{-1} \chi_H(x_1, x_2) = \frac{1}{2} \delta(x_1) \delta(x_2) - \frac{i}{2\pi} \frac{1}{x_1} \delta(x_2)$$

where the distribution $\frac{1}{x_1}$ is interpreted in the principal value sense. Thus

$$\psi * \mathcal{F}^{-1} \chi_H(x_1, x_2) = \frac{1}{2} \psi(x_1, x_2) + \frac{1}{2\pi i} p.v. \int \psi(x_1 - y, x_2) \frac{dy}{y}.$$

In particular, since ψ is non-negative on $[1/3, 2/3] \times [1/3, 2/3]$ we see that

$$|\psi * \mathcal{F}^{-1} \chi_H(x_1, x_2)| \sim 1$$

on the rectangle $[10, 11] \times [1/3, 2/3]$. Thus the main term is comparable to 1 in magnitude on the rectangle $[10R^2, 11R^2] \times [R/3, 2R/3]$. Since this is much bigger than the error term if a is big enough, we see that $S_1 f$ is comparable to 1 on this rectangle.

By rotating and translating this example, we obtain

Lemma 2.2. *Let T be a $R^2 \times aR$ rectangle in the plane with arbitrary position and orientation, and let T' be the rectangle of the same length and one third of the width which is shifted over by $10R^2$ in the direction of the long axis of T. Then there exists a function f_T supported on T such that $|f_T| \leq 1$ on T, and $|S_1 f_T| \sim 1$ on T'.*

In other words, we can make a function supported on T, such that the disc multiplier essentially shifts this function by ten units to a slightly different rectangle T'.

So far we have not obtained (2), except for small N. The idea is to get several rectangles T and pile them together. The key geometrical fact is:

Lemma 2.3. *For any $M > 0$, there exists an $R > 0$ and a collection of $R^2 \times aR$ rectangles T, such that the rectangles T are all disjoint, and*

$$|\bigcup T'| \leq M^{-1} |\bigcup T|. \tag{10}$$

Let's see how this implies Fefferman's result. Let M be a large number to be chosen later, and let $\{T\}$ be the rectangles given by the above lemma. Define the corresponding functions f_T from Lemma 2.2. We define f by

$$f = \sum_T \varepsilon_T f_T$$

where ε_T are iid random signs. Since the f_T have disjoint supports, we have

$$\|f\|_p = (\sum_T \|f_T\|_p^p)^{1/p}$$

regardless of the choice of signs ε_t. Each f_T has size about 1 on T, so we have

$$\|f\|_p \lesssim (\sum_T |T|)^{1/p} = |\bigcup_T T|^{1/p}.$$

What about $S_1 f$? We have

$$S_1 f = \sum_T \varepsilon_T S_1 f_T$$

so by Khinchin's inequality we have

$$\mathbf{E}(\|S_1 f\|_p^p) \sim \|(\sum_T |S_1 f_T|^2)^{1/2}\|_p^p.$$

On the other hand, $|S_1 f_T| \sim 1$ on T'. Thus we have

$$\mathbf{E}(\|S_1 f\|_p^p) \gtrsim \|(\sum_T |\chi_{T'}|^2)^{1/2}\|_p^p.$$

The right-hand side simplifies to

$$\| \sum_T \chi_{T'} \|_{p/2}^{p/2}.$$

From Hölder's inequality we have

$$\| \sum_T \chi_{T'} \|_1 \leq \| \sum_T \chi_{T'} \|_{p/2} | \bigcup_T T |^{1-2/p}.$$

Since the left-hand side of this is $\sum_T |T'| \sim |\bigcup_T T|$, this can be re-arranged using (10) as

$$\| \sum_T \chi_{T'} \|_{p/2}^{p/2} \gtrsim M^{p/2-1} | \bigcup_T T |.$$

Combining all our estimates together, we see that

$$\|S_1 f\|_p \gtrsim M^{p/2-1} \|f\|_p;$$

since $p > 2$, Fefferman's result is thus proved by taking M sufficiently large.

3. The Besicovitch set construction

We now prove Lemma 2.3.

Consider a the right-angled triangle T with vertices $(0,0)$, $(1,0)$, $(0,1)$. Divide the base into N intervals from $(i/N, 0)$ to $((i+1)/N)$, where $i = 0, 1, \ldots, N-1$. This divides the big triangle T into N triangles T_i of base $1/N$ and height 1. Currently the union of these triangles has measure 1/2.

The key lemma is

Lemma 3.1. *Let N be a power of 2. There exists a decreasing sequence of numbers $a_0 \geq \ldots \geq a_{N-1}$ such that the union of the translated triangles $T_i + a_i e_1$ has measure $O(\log\log N / \log N)$.*

Assuming this lemma for the moment, let R_i be a $1/10x1/(10N)$ rectangle inscribed inside $T_i + a_i e_1$, and let R_i' be defined as in the previous section, with the orientation chosen so that R_i' is below the x-axis. Then it is easy to see that the R_i' are all disjoint (because of the monotonicity of the a_i), so that $|\bigcup R_i'| \sim 1$. Lemma 2.3 then follows by setting N roughly equal to R/a and rescaling by R^2.

It remains to prove Lemma 3.1. Let us first try to understand the case $N = 2$. Suppose we try the sequence $a_0 = 0$, $a_1 = -\delta$, where $0 < \delta \ll 1$ is some number we shall choose later. This pushes the triangles in a little bit. It's easy to see that the new union of triangles consists of a triangle similar to T but $1 - \delta$ times as small, together with a bowtie shaped object whose area is only $O(delta^2)$ the size of the original triangle T.

Now let's go back to the general case. Suppose we take the sequence $a_i = -\lfloor \frac{i+1}{2} \rfloor \delta/N$. Look at what this does to each pair (T_{2n-1}, T_{2n}) of adjacent triangles. By similar considerations to the above, these triangles become squished together to form a triangle T'_n similar to $T_{2n-1} \cup T_{2n}$ but $1-\delta$ times as small, together with a bow-tie of area about $O(\delta^2)$ times the size of $T_{2n-1} \cup T_{2n}$. Also, the triangles T'_n as $n = 0, \ldots, N/2-1$ are arranged to form a single triangle which is similar to T but is $1-\delta$ times as small.

In short, we have adjusted the original configuration $T_0 \bigcup \ldots \bigcup T_{N-1}$ to become a new configuration $T'_0 \bigcup \ldots \bigcup T'_{N/2-1}$ of $1-\delta$ the size, and half as many slices, together with a bunch of bow-ties whose total area is at most $O(\delta^2)$.

Now take the new triangles $T'_0, \ldots, T'_{N/2-1}$ and repeat the same procedure, translating each triangle T'_i by $-\lfloor \frac{i+1}{2} \rfloor (1-\delta)\delta/(N/2)$. This induces a further translation on the component triangles T_{2i-1}, T_{2i} which make up T'_i, and also move the old bow-ties around. By doing this we create some new triangles $T''_0, \ldots, T''_{N/4}$, which are arranged in a triangle which is similar to but $(1-\delta)^2$ times the size of the original triangle T, plus some new bow-ties whose total area is at most $O(\delta^2(1-\delta)^2)$.

We keep repeating this procedure until the number of slices goes down to 1. We then end up with a small triangle which is similar to but $(1-\delta)^{\log N}$ times the size of T, plus a large number of bow-ties whose total area is at most

$$\delta^2 + \delta^2(1-\delta)^2 + \delta^2(1-\delta)^4 + \ldots = O(\delta).$$

Thus we have achieved a total area of

$$O((1-\delta)^{\log N} + \delta),$$

and the result follows by choosing δ so that $\log(1-\delta) = \log\log N/\log(N)$.

■

There are other nice constructions which work too. One of the simplest is the following example of Kahane. Let E be the middle-halves Cantor set, that is the set of all reals in $[0,1]$ whose base four expansion only consists of 0s and 1s. Let E_N be the N^{th} stage of this set, i.e. the set whose first N digits in the base four expansion consists of only 0s and 1s. Now consider the set formed by taking all line segments between $\{0\} \times E_N$ and $\{1\} \times E_N$. This set contains about N $1/10xN/10$ rectangles, whose translated counterparts are all disjoint, but has total measure about $\log\log N/\log N$.

The above construction is closely related to the Kakeya problem. Define a Besicovitch set in $\mathbf{R}^n$ to be a set which contains a line segment in every direction. The above construction can be adapted to show that Besicovitch sets can have arbitrarily small measure, or even zero measure for $n = 2$, and a slight modification gives the same result for all $n \geq 2$. The Kakeya conjecture asserts that even though Besicovitch sets can have zero measure, they must still have Hausdorff dimension n. In a later set of notes we shall study this conjecture and see why it is so closely related to restriction and Bochner-Riesz type problems. For now, we shall digress to non-linear Schrodinger equations.

4. Appendix: the Bochner-Riesz kernel

We now prove the bound (5). The classical way to do this would be to do some clever changes of variables and compute (5) exactly as an expression involving a Bessel function. In fact, the exact formula is

$$\widehat{m_\delta(x)} = \pi^{-\delta}\Gamma(1+\delta)|x|^{-(n/2)-\delta}J_{n/2+\delta}(2\pi|x|).$$

We're going to do this in a more modern way - fuzzier, with not as many exact integrals or special functions - but more robust. I'll be sketchy with the details; those of you who remember the previous quarter's class will find much of this familiar (indeed, this stuff is very close to the computation of the Helmholtz kernel we did last quarter).

Since m_δ is radial, we may take $\xi = \lambda e_n$ for some large λ. We need to evaluate the integral

$$\int_{|\xi|\leq 1} (1-|\xi|^2)^\delta e^{2\pi i \lambda \xi_n}\ d\xi.$$

We smoothly cut this into three pieces: the piece where $|\xi - e_n| \ll 1$ (the north pole), the piece where $|\xi + e_n| \ll 1$ (the south pole), and everything else (the equator), where $|\xi_n| \leq 1-\varepsilon$ for some $\varepsilon > 0$.

Let's deal with the equator first. The core part where $|\xi| \ll 1$ is extremely small by the principle of non-stationary phase, so let's just look at the surface part $|\xi| \sim 1$. In this case we can use polar co-ordinates and reduce to

$$\int_{r\sim 1} (1-r^2)_+^\delta r^{n-1} \int_{S^{n-1}:|\omega_n|\leq 1-\varepsilon} e^{2\pi i \lambda r \omega_n}\ d\omega.$$

The inner integral, and hence the entire integral, is also rapidly decreasing in λ by the principle of non-stationary phase. So this part is also negligible.

Now let's look at the north pole part. We can approximate this part by

$$(1-|\xi|^2)_+^\delta = f d\omega * d\mu + error$$

where the error is smoother than the main term, f is a smooth function on the sphere supported on the northern polar cap, and

$$d\mu(\underline{\xi}, \xi_n) = \delta(\underline{\xi})\eta(\xi_n)(-\xi_n)_+^\delta$$

is a measure supported on the vertical axis, with η a bump function which equals 1 near the origin. Indeed, we can easily work out that

$$f d\omega * d\mu(\underline{\xi}, \xi_n) = f(\underline{\xi}, \Phi(\underline{\xi}))\eta(\xi_n - \Phi(\underline{\xi})J(\underline{\xi})(\Phi(\underline{\xi}) - \xi_n)_+^\delta$$

where J is some Jacobian factor; by choosing f properly one can make this a good approximation to the kernel m^δ near the north pole. The error vanishes to order $\delta + 1$ or more at the sphere; one can do a similar decomposition with this error, with a new error term now vanishing to order $\delta + 2$. We can continue until the error is so smooth that we can easily absorb it into the error term of (5).

We now look at the treatment of the main term. The contribution to (5) of the main term is

$$\widehat{f d\omega}(\lambda e_n)\widehat{d\mu}(\lambda e_n).$$

The first factor is $Ce^{2\pi i\lambda}\lambda^{-(n-1)/2} + o(\lambda^{-(n-1)/2})$ by the discussion used to derive the Fourier transform of surface measure. We now claim that the other factor is $(C + o(1))\lambda^{-1-\delta}$. Since the $\underline{\xi}$ variable is pretty much irrelevant here, this claim is equivalent to

$$\widehat{\eta(-\xi_n)_+^\delta}(\lambda) = (C + o(1))\lambda^{-1-\delta}.$$

If it were not for the η, this would be true, since $(-\xi_n)_+^\delta$ is homogeneous of degree δ, and in k dimensions, the homogeneity of a function and the homogeneity of its Fourier transform always add up to $-k$. The η has the effect of convolving the Fourier transform with a Schwartz function, which is quite harmless and only perturbs the homogeneity slightly.

Anyway, putting the terms together we get the desired estimate (5).

References

1. M. Christ, *Weak type endpoint bounds for Bochner-Riesz multipliers*, Revista Mat. Iberoamericana **3** (1987), 25–31.
2. M. Christ, *Weak type (1,1) bounds for rough operators*, Annals of Mathematics, **128** (1988): 19–42.
3. C. Fefferman, *Inequalities for strongly singular convolution operators*, Acta Math. **124** (1970), 9–36.
4. C. Fefferman, *The multiplier problem for the ball*, Ann. of Math. **94** (1971): 330–336.
5. E. M. Stein, M. H. Taibleson, and G. Weiss, *Weak-type estimates for maximal operators on certain H^p spaces*, Rend. Circ. Mat. Palermo suppl. **1** (1981): 81–97.
6. T. Tao, *Weak-type endpoint bounds for Riesz means*, Proc. Amer. Math. Soc. **124** (1996), 2797–2805.
7. T. Tao, *The Bochner-Riesz conjecture implies the Restriction conjecture*, to appear, Duke Math J.
8. T. Tao, *The weak-type endpoint Bochner-Riesz conjecture and related topics* , to appear, Indiana Math. J.

9. P. Tomas, *A restriction theorem for the Fourier transform*, Bull. Amer. Math. Soc. **81** (1975), 477–478.

Department of Mathematics, UCLA, Los Angeles CA 90095-1555

E-mail address: tao@math.ucla.edu