LECTURE NOTES 1 FOR 254B

TERENCE TAO

1. Introduction

This class will focus on two open problems in harmonic analysis: the restriction conjecture, and the Kakeya conjecture. Solutions to these problems have applications to PDE, spectral theory, and number theory; we will focus on the PDE applications. The techniques used to attack these problems have also been found to be useful elsewhere.

Let's begin with the restriction conjecture. We first ask a very basic question about the Fourier transform

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

of a function f on $\mathbf{R}^n$. Namely, how is the size of $\hat{f}$ related to the size of f? More quantitatively, when do we have estimates of the form

$$\|\hat{f}\|_{L^q(\mathbf{R}^n)} \leq C\|f\|_{L^p(\mathbf{R}^n)} \tag{1}$$

for all functions[1] f, where C is a constant?

From Plancherel we have

$$\|\hat{f}\|_{L^2(\mathbf{R}^n)} = \|f\|_{L^2(\mathbf{R}^n)}$$

while by the triangle inequality $|\int f(x)\ dx| \leq \int |f(x)|\ dx$ we have the easy estimate

$$\|\hat{f}\|_{L^\infty(\mathbf{R}^n)} \leq \|f\|_{L^1(\mathbf{R}^n)}. \tag{2}$$

From the method of real or complex interpolation, we therefore have the Hausdorff Young inequality

$$\|\hat{f}\|_{L^{p'}(\mathbf{R}^n)} \leq \|f\|_{L^p(\mathbf{R}^n)} \text{ for all } 1 \leq p \leq 2,$$

where of course p' is the dual exponent of p:

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

It turns out that the Hausdoff-Young inequalities are the only estimates we have of the form (1). So that is the end of that story.

[1] We will generally assume a priori that all functions f are Schwarz; once we obtain a quantitative estimate, we can remove this assumption by a limiting argument.

Now suppose we are not interested in controlling the Fourier transform $\hat{f}$ on all of $\mathbf{R}^n$, but only on some subset of $\mathbf{R}^n$. For instance, suppose we were looking for an estimate of the form

$$\|\hat{f}\|_{L^q(B)} \leq C\|f\|_{L^p(\mathbf{R}^n)} \tag{3}$$

where $B = B(0,1)$ is the unit ball. Of course, from the Hausdorff-Young inequality we have (3) whenever $q = p'$ and $1 \leq p \leq 2$. But now we have slightly more estimates, because the L^q spaces are nested when the underlying space is compact:

$$\|g\|_{L^q(B)} \lesssim \|g\|_{L^r(B)} \text{ whenever } q \leq r. \tag{4}$$

This is proven by Hölder's inequality (or by interpolating between the easy cases $r = q$ and $r = \infty$. Here and in the sequel, we use $A \lesssim B$ as shorthand for $A \leq CB$, where C is an unspecified constant.

From this we see that we actually get an estimate of the form (3) whenever $q \leq p'$ and $1 \leq p \leq 2$. This turns out to be all the estimates there are here, so that ends that story.

The situation becomes much more interesting when we restrict the Fourier transform $\hat{f}$ not to an open set like B, but to a hypersurface like the unit sphere S^{n-1}. In other words, we now ask for estimates of the form

$$\|\hat{f}\|_{L^q(S^{n-1})} \leq C\|f\|_{L^p(\mathbf{R}^n)}. \tag{5}$$

Of course we give S^{n-1} the usual surface measure $d\sigma$.

If f is just an arbitrary L^p function, it is not even clear that this estimate makes sense. For instance, if f is an arbitrary L^2 function, then $\hat{f}$ is also an arbitrary L^2 function (by Plancherel's theorem), and in general an arbitrary L^2 function cannot be meaningfully restricted to a surface such as S^{n-1}, as this is a measure zero set. So (5) cannot possibly be true (or even make sense) when $p = 2$. On the other hand, if $p = 1$, then f is an L^1 function, and $\hat{f}$ is therefore continuous (and decays at infinity) by the Riemann-Lebesgue lemma. Thus $\hat{f}$ can be meaningfully restricted to S^{n-1}, and we get (5) for $q = \infty$ by (2), and hence for all $1 \leq q \leq \infty$ by (4).

To summarize: when $p = 2$, there are no estimates of the form (5). When $p = 1$, we have (5) for all q.

There is still the question of what happens in the intermediate values of p. (We can't interpolate between an estimate and an absence of an estimate!) This problem is called the restriction problem - it asks when one can meaningfully restrict the Fourier transform of an L^p function to the sphere - and is largely open. Of course one can also phrase this problem for surfaces other than the sphere. Nevertheless, even though we do not have a complete solution to this problem yet, there is a lot of progress so far.

We can transform (5) by duality. Suppose that we found a value of p, q for which (5) holds. In particular we have

$$\sup_{\|f\|_p=1} \|\hat{f}\|_{L^q(S^{n-1})} \lesssim 1.$$

By duality, this implies

$$\sup_{\|f\|_p=1} \sup_{\|g\|_{L^{q'}(S^{n-1})}=1} |\int \hat{f}(\xi)g(\xi)\ d\omega(\xi)| \lesssim 1.$$

Interchanging the sup's and applying Parseval, we get

$$\sup_{\|g\|_{L^{q'}(S^{n-1})}=1} \sup_{\|f\|_p=1} |\int f(x)\widehat{gd\omega}(x)\ dx| \lesssim 1.$$

Undoing the duality, we obtain

$$\sup_{\|g\|_{L^{q'}(S^{n-1})}=1} \|\widehat{gd\omega}\|_{p'} \lesssim 1$$

and therefore

$$\|\widehat{gd\omega}\|_{p'} \lesssim \|g\|_{q'}. \tag{6}$$

All the above steps are reversible, so that (6) and (5) are actually equivalent. The statement (5) is sometimes referred to as a restriction theorem, and (6) is referred to as an extension theorem. (Don't ask me why they're called "theorems" rather than "estimates" or "conjectures". One may as well ask why "Fermat's Last Theorem" used to be called a theorem before Wiles.)

Extension theorems such as (6) are of interest to PDE, because expressions such as $\widehat{gd\omega}$ turn up quite frequently. For instance, a (tempered distributional) solution to the Helmholtz equation

$$\Delta u + 4\pi^2 u = 0$$

is automatically of the form $u = \widehat{gd\omega}$ for some distribution g. (Exercise!) Similarly, a global solution to the free wave equation

$$\partial_t^2 u(t,x) - \Delta u(t,x) = 0$$

is of the form $u(t,x) = \widehat{gd\omega}$, where $d\omega$ is surface measure on the light cone $\{(\tau,\xi) \in \mathbf{R}^{1+n} : |\tau| = |\xi|\}$ and we are taking the space-time Fourier transform. Or, a global solution to the free Schrödinger equation

$$2\pi i\partial_t u(t,x) + \Delta u(t,x) = 0$$

is of the form $u(t,x)\widehat{gd\omega}$, where $d\omega$ is now surface measure on the paraboloid $\{(\tau,\xi) \in \mathbf{R}^{1+n} : \tau = |\xi|^2\}$. And so forth.

Basically, the estimate (6) measures how quickly the function $\widehat{gd\omega}$ decays at infinity. The lower the exponent p', the more decay we obtain.

These decay estimates are especially important in semilinear equations such as

$$\partial_t^2 u(t,x) - \Delta u(t,x) = u^3(t,x).$$

The non-linearity $u^3(t,x)$ is trying to make the solution blow up in finite time (note that the ODE $u''(t) = u^3(t)$ does indeed blow up in finite time, as can be witnessed by the solutions $u(t) = 1/(C-t)$ for any constant C), but the linear part of the equation is trying to make the solution decay. So it's a contest between the two terms - and the way to predict the winner is via estimates such as (6).

Finally, we'll see that these estimates are also related to a purely geometric conjecture known as the Kakeya conjecture. We can state it as follows. Define a Besicovitch set to be any set which contains a line segment in every direction. (Such sets tend to resemble hedgehogs). It is known that such a set can have measure zero; however it is unknown whether their Hausdorff dimension can ever be strictly less than n. The Kakeya conjecture asserts that this is false, that Besicovitch sets always have Hausdorff dimension equal to n.

It turns out that the restriction problem is very closely related to the Kakeya problem. One way to see this is via the PDE interpretation of the restriction problem. All the linear equations mentioned earlier have "wave packet" solutions, which resemble a pulse or train of waves moving in tandem for a short while before dispersing. These can be thought of as oscillatory approximations to straight lines (just as the wave equation is as oscillatory approximation of geometric optics). A general solution to such a PDE tends to consist of a superposition of several such wave packets, which one can think of as an approximation to a Besicovitch set. The smaller one can make a Besicovitch set, the worse the L^p norm of the solution becomes, and the more likely (6) is to fail. We'll have more precise versions of these heuristics later.

2. Why Hausdorff-Young is best possible

We now show that the only estimates of the form (1) are those given by the Hausdorff-Young inequality. More precisely:

Proposition 2.1. *Let $1 \leq p, q \leq \infty$ be exponents such that* (1) *holds for all test functions f. Then we have $q = p'$ and $p \leq 2$.*

Proof Assume that (1) holds. We first show the condition $q = p'$.

The quickest proof is by dimensional analysis. If f is dimensionless, and space has the units of length L, then the L^p norm $\|f\|_p$ has the units of $L^{n/p}$. Also, the Fourier transform $\hat{f}$ has units of L^n, being a spatial integral of a dimensionless quantity. Since the frequency variable ξ has units of L^{-1}, we see that the L^q norm of $\hat{f}$ thus has units of $L^n L^{-n/q}$. Since there is no preferred unit of length in (1), the only way this estimate can be valid is if the units match up, which yields the condition $q = p'$.

A more constructive proof runs by choosing f to be a dilated test function

$$f(x) = \psi(x/\lambda)$$

where $\lambda > 0$ is a parameter we can vary, and ψ is some Schwarz function. Note that $\|f\|_p \sim \lambda^{n/p}$. The Fourier transform of this is

$$\hat{f}(\xi) = \lambda^n \hat{\psi}(\lambda \xi);$$

(Exercise!) thus $\|\hat{f}\|_q \sim \lambda^n \lambda^{-n/q}$. Thus, if (1) holds for all f, we must have

$$\lambda^n \lambda^{-n/q} \lesssim \lambda^{n/p}.$$

Since we can make λ arbitrarily large or arbitrarily small, the only way this can occur is if $n - n/q = n/p$, which is just the condition $q = p'$.

(A general note: if f has size roughly A on a set of volume V, and is rapidly decreasing outside of this set, then the L^p norm of f is about $AV^{1/p}$.)

Now for the proof of $p \leq 2$. We again choose a Schwarz function ψ; let's say it's supported on the unit cube $[0,1]^n$. We shall choose f to be the random function

$$f(x) = \sum_{k=1}^{N} \varepsilon_k \psi(x - ke_1)$$

where e_1 is one of the basis vectors of $\mathbf{R}^n$, N is a large integer that we may vary, and ε_k are a collection of independent identically distributed (iid) signs $\varepsilon_k = \pm 1$ with an equal probability of each.

Clearly we have $\|f\|_p \sim N^{1/p}$ regardless of the choice of signs. What about the Fourier transform? It's easy to show that

$$\hat{f}(\xi) = \sum_{k=1}^{N} \varepsilon_k \hat{\psi}(\xi) e^{2\pi i k \xi_1}$$

(Exercise!), so that our Fourier transform is a randomized sum of highly oscillatory functions. How does one deal with this? The key tool - and the reason why randomization is actually a useful technique - is Khinchin's inequality:

Lemma 2.2 (Khinchin's inequality)**.** *If $f_1(x), \ldots, f_N(x)$ are a collection of functions, and $\varepsilon_1, \ldots, \varepsilon_n$ are randomized signs, then for any $1 < p < \infty$ we have*

$$\mathbf{E}(\| \sum_{k=1}^{N} \varepsilon_k f_k \|_p^p) \sim \| (\sum_{k=1}^{N} |f_k|^2)^{1/2} \|_p^p,$$

where the constants in the $\sim$ symbol are independent of N and the f_k (although they do depend on p), and $\mathbf{E}$ denotes the expectation.

In other words, the function $\sum_{k=1}^{N} \varepsilon_k f_k$ behaves in magnitude like the square function $(\sum_{k=1}^{N} |f_k|^2)^{1/2}$ on the average. Assuming this lemma for the moment, we see that

$$\mathbf{E}(\|\hat{f}\|_q^q) \sim \| (\sum_{k=1}^{N} |\hat{\psi}(\xi) e^{2\pi i k \xi_1}|^2)^{1/2} \|_q^q.$$

The oscillating factors $e^{2\pi i k \xi_1}$ are harmless, and we see that the right-hand side is comparable to $N^{q/2}$. Thus, there must exist some choice of signs for which

$\|\hat{f}\|_q \gtrsim N^{1/2}$. Combining this with our bound on $\|f\|_p$, we see from (1) that $N^{1/2} \lesssim N^{1/p}$. Letting $N \to \infty$ we obtain a contradiction unless $p \leq 2$.

Now let's prove Khinchin's inequality. It suffices to show

$$\mathbf{E}(|\sum_{k=1}^N \varepsilon_k a_k|^p)^{1/p} \sim (\sum_{k=1}^N |a_k|^2)^{1/2}, \tag{7}$$

since one can then apply this inequality with $a_k = f_k(x)$ for each x, raise this to the p^{th} power, and integrate in x. (Note that the expectation operator $\mathbf{E}$ is linear).

It suffices to prove (7) assuming that

$$(\sum_{k=1}^N |a_k|^2)^{1/2} = 1.$$

This is because (7) is unaffected by the operation of multiplying a_k by a constant.

We first prove this for $p = 2$, in which case we have equality. Indeed:

$$\mathbf{E}(|\sum_{k=1}^N \varepsilon_k a_k|^2) = \sum_{k=1}^N \sum_{k'=1}^N \mathbf{E}(\varepsilon_k \varepsilon_{k'} a_k a_{k'})$$

by linearity of expectation. By independence, the expectation vanishes unless $k = k'$, so our sum becomes

$$\sum_{k=1}^N \mathbf{E}(\varepsilon_k^2 a_k^2) = \sum_{k=1}^N a_k^2 = 1$$

as desired.

Now that we have proven (7) for $p = 2$, it suffices to prove the upper bound of (7) for all $1 < p < \infty$. This will automatically imply the lower bound in (7), by applying Hölder's inequaity

$$\mathbf{E}(X \cdot X) \leq \mathbf{E}(|X|^p)^{1/p} \mathbf{E}(|X|^{p'})^{1/p'}$$

with $X = \sum_{k=1}^N \varepsilon_k a_k$.

To show the upper bound in (7) we first consider the related expression

$$\mathbf{E}(e^{\lambda \sum_{k=1}^N \varepsilon_k a_k})$$

where $\lambda > 0$ is a parameter. This is of course equal to

$$\mathbf{E}(\prod_{k=1}^N e^{\lambda \varepsilon_k a_k}).$$

By independence one can take the product outside of the expectation:

$$\prod_{k=1}^N \mathbf{E}(e^{\lambda \varepsilon_k a_k}) = \prod_{k=1}^N \cosh(\lambda a_k).$$

By comparing Taylor series, we see that $\cosh(x) \leq e^{x^2/2}$ for all x, so we have

$$\mathbf{E}(e^{\lambda \sum_{k=1}^N \varepsilon_k a_k}) \leq \prod_{k=1}^N e^{\lambda^2 a_k^2/2} = e^{\lambda^2/2}.$$

From the Chebyshev inequality

$$\mathbf{P}(X > \alpha) \leq \frac{1}{\alpha}\mathbf{E}(X)$$

where $\mathbf{P}(E)$ is the probability of an event E, we obtain

$$\mathbf{P}(e^{\lambda \sum_{k=1}^N \varepsilon_k a_k} > \alpha) \lesssim \alpha^{-1} e^{\lambda^2/2}.$$

This works for every α; we choose $\alpha = e^{\lambda^2}$. The estimate now becomes

$$\mathbf{P}(\sum_{k=1}^N \varepsilon_k a_k > \lambda) \lesssim e^{-\lambda^2/2}.$$

Since the random variable $\sum_{k=1}^N \varepsilon_k a_k$ is clearly symmetric around the origin, we thus have

$$\mathbf{P}(|\sum_{k=1}^N \varepsilon_k a_k| > \lambda) \lesssim e^{-\lambda^2/2}.$$

In particular, we have

$$\mathbf{P}(|\sum_{k=1}^N \varepsilon_k a_k|^p > \lambda) \lesssim e^{-\lambda^{2/p}/2}.$$

If we now use the identity

$$\mathbf{E}(X) = \int_0^\infty \mathbf{P}(X > \lambda)\ d\lambda$$

for any random variable X with finite expectation (exercise!), we thus obtain

$$\mathbf{E}(X) \lesssim 1$$

for all $1 < p < \infty$, as desired. ∎

We will need this randomization trick a bit later in this course.

3. Hausdorff-Young on compact sets

As observed in the introduction, when one restricts the Fourier transform of f to a compact set such as the unit ball $B(0,1)$, one gets a larger set of exponents. Indeed, we have already seen that (3) holds whenever $q \leq p'$ and $1 \leq p \leq 2$. Again, this is best possible:

Proposition 3.1. *Let $1 \leq p, q \leq \infty$ be exponents such that (3) holds for all test functions f. Then we have $q \leq p'$ and $p \leq 2$.*

The proof is identical to that of Proposition 2.1. The proof of $p \leq 2$ is unaffected; one just has to ensure that $\hat{\psi}$ does not vanish completely on B. The second argument which showed that $q = p'$ no longer words as stated, because one can make λ arbitrarily small but not arbitrarily large, which is why one gets the condition $q \leq p'$ instead. (The dimensional analysis argument breaks down because the set B now introduces a preferred length scale).

4. Are there restriction theorems for the plane?

In the previous section we have seen what happens to the Fourier transform $\hat{f}$ of an L^p function on a ball. Now let's try to restrict $\hat{f}$ to hypersurfaces.

The easiest case to consider is when the surface is a hyperplane; let's say it's the plane $\{x_n = 0\}$. In fact, let's only look at a compact subset of this plane, e.g.

$$S = \{x : x_n = 0, |x| \leq 1\}.$$

We give this plane the obvious surface measure $d\sigma$.

By the remarks in the introduction, we have the estimate

$$\|\hat{f}\|_{L^q(S)} \lesssim \|f\|_p \tag{8}$$

if $p = 1$ and q is arbitrary. However, these are the only estimates available.

Proposition 4.1. *Suppose one has* (8) *for all test functions f. Then one must have $p = 1$.*

Proof The idea is to construct a function f whose Fourier transform is strongly concentrated on and near S.

Let ψ be a Schwarz function such that $\hat{\psi} \sim 1$ near the origin. For instance, we could take ψ to be a Gaussian. Let f be the function

$$f(x_1, \dots, x_{n-1}, x_n) = \psi(x_1, \dots, x_{n-1}, x_n/\lambda)$$

where λ is a parameter we may vary. Observe that $\|f\|_p \sim \lambda^{1/p}$. The Fourier transform of f is

$$\hat{f}(\xi_1, \dots, \xi_{n-1}, \xi_n) = \lambda\psi(\xi_1, \dots, \xi_{n-1}, \lambda\xi_n),$$

In other words, $\hat{f}$ has size λ and is concentrated in a squashed ball of dimensions $1 \times \ldots \times 1 \times 1/\lambda$. When restricted to S, $\hat{f}$ therefore has size λ and is concentrated in a unit $n-1$-ball. Thus $\|\hat{f}\|_{L^q(S)} \sim \lambda$. Thus in order for (8) to hold for all f, we must have $\lambda \lesssim \lambda^{1/p}$. Letting $\lambda \to \infty$ we see that we must have $p = 1$. ∎

In summary, there are no non-trivial restriction estimates for the plane, even if we only look at a compact piece. The problem is that the plane is so flat that one can easily manipulate $\hat{f}$ to be large on and near the surface.

5. Curved surfaces

Let's now look at more general hypersurfaces. It's time for some notation.

Definition 5.1. If $1 \leq p, q \leq \infty$ and $S = (S, d\sigma)$ is a hypersurface with boundary with measure $d\sigma$, we say that we have a *(L^p, L^q) restriction theorem for S* if (8) holds for all test functions f. We denote this theorem by $R_S(p \to q)$.

Here are some trivial observations, which I leave as exercises.

- $R_S(1 \to q)$ holds for all q.
- $R_S(2 \to q)$ fails for all q.
- The statement $R_S(p \to q)$ is unaffected by any rotation or reflection of S, or indeed by any invertible linear transformation L as long as L and L^{-1} both have bounded co-efficients.
- The statement $R_S(p \to q)$ is unaffected by any translation of S, even if this translation is arbitrarily large.
- If S is the union of two smaller surfaces S_1 and S_2, then $R_S(p \to q)$ holds if and only if $R_{S_1}(p \to q)$ and $R_{S_2}(p \to q)$ both hold.
- If S is compact and $R_S(p \to q)$ holds for some q, it automatically holds for any lower value of q.
- If S is compact and $R_S(p \to q)$ holds for some p, it automatically holds for any lower value of p. (Hint: interpolate with the case $p = 1$).

In other words, smaller values of p and q are easier to prove; for higher values of p and q the restriction theorem is more likely to fail.

Let's see what happens if we bend the plane a little bit. Let S be a surface of the form

$$S = \{(\underline{x}, \Phi(\underline{x})) : \underline{x} \in \mathbf{R}^{n-1}, |\underline{x}| \lesssim 1\}$$

where $\Phi : \mathbf{R}^{n-1} \to \mathbf{R}$ is a smooth function such that $\Phi(0) = \nabla\Phi(0) = 0$. If we can handle this case, then by the above observations we can handle any compact hypersurface with boundary. Of course, the case $\Phi \equiv 0$ is that of the plane, which has no non-trivial estimates.

If Φ is curved, then the situation is not quite as bad; we do have the possibility of having some non-trivial restriction theorems. But it all depends on how curved Φ is.

Proposition 5.2. *Suppose Φ vanishes to order k at 0 for some $k \geq 2$, so that $\Phi(x) = O(|x|^k)$. Then the restriction estimate $R_S(p \to q)$ is only possible when*

$$p' \geq \frac{n+k-1}{n-1} q.$$

Thus, the flatter S is, the fewer restriction estimates we expect.

Proof This is a modification of Proposition 4.1. Basically we tweak the dilations used in that proposition in order to better fit the new surface S.

Let ψ be as in Proposition 4.1, and let

$$f(x_1, \dots, x_{n-1}, x_n) = \psi(x_1/\lambda^{1/k}, \dots, x_{n-1}/\lambda^{1/k}, x_n/\lambda)$$

for some $\lambda \gg 1$, which we will later send to infinity. In other words, f is a bump of height about 1 concentrated on a $\lambda^{1/k} \times \dots \times \lambda^{1/k} \times \lambda$ tube. In particular we see that $\|f\|_p \approx \lambda^{(n+k-1)/kp}$.

The Fourier transform[2] of this is

$$\hat{f}(\xi_1, \dots, \xi_{n-1}, \xi_n) = \lambda^{(n+k-1)/k} \hat{\psi}(\lambda^{1/k}\xi_1, \dots, \lambda^{1/k}\xi_{n-1}, \lambda x_n).$$

If we restrict this to S, we see that $\hat{f}$ has size about $\lambda^{(n+k-1)/k}$ on a cap of radius about $\lambda^{-1/k}$ on S. Since this cap has measure about $\lambda^{-(n-1)/k}$, we see that

$$\|\hat{f}\|_q \gtrsim \lambda^{(n+k-1)/k} \lambda^{-(n-1)/kq}.$$

Comparing this with our bound for $\|f\|_p$, we see that $R_S(p \to q)$ is only possible when

$$\lambda^{(n+k-1)/k} \lambda^{-(n-1)/kq} \lesssim \lambda^{(n+k-1)/kp}$$

for all λ. Letting $\lambda \to \infty$ we obtain the result. ■

In particular, since we always have Φ vanishing to at least second order if S is smooth, we see that we have

$$p' \geq \frac{n+1}{n-1} q \tag{9}$$

as a necessary condition regardless of the exact choice of S. With a flatter surface, we expect even stricter necessary conditions.

This is not the only necessary condition for $R_S(p \to q)$ to hold, just as $p' \geq q$ was not the only condition for (3) to hold. One could repeat the argument in Proposition 2.1 and get the condition $p \leq 2$; however we can do better than this. For instance, if S is the unit sphere, we have the improved necessary condition

$$p < 2n/(n+1). \tag{10}$$

To see this, we use duality, and recall that $R_S(p \to q)$ is equivalent to the estimate

$$\|\widehat{g d\sigma}\|_{p'} \lesssim \|g\|_{q'}$$

for all $g \in L^{q'}(S)$. In particular, if we choose g to equal the constant function 1, we have

$$\|\widehat{d\sigma}\|_{p'} \lesssim 1.$$

We now use a fundamental fact about the surface measure of the sphere, which we prove in the next section.

[2]The general relationship between linear changes of variables and the Fourier transform is this: if $f(x) = \psi(L^{-1}x)$ for some linear transformation L, then $\hat{f}(\xi) = \det(L)\hat{\psi}(L^*\xi)$.

Proposition 5.3. *If $d\sigma$ is the surface measure of the unit sphere, then for $|x| \gg 1$ we have*

$$\widehat{d\sigma}(x) = C\frac{e^{2\pi i|x|}}{|x|^{(n-1)/2}} + C\frac{e^{-2\pi i|x|}}{|x|^{(n-1)/2}} + O(|x|^{-n/2}).$$

Assuming this for the moment, we see that $|\widehat{d\sigma}(x)|$ is roughly comparable to $|x|^{-(n-1)/2}$ on the average for large x. This implies that $\widehat{d\sigma}$ is in $L^{p'}$ only if $p \leq 2n/(n-1)$, which explains (10).

We have thus obtained two necessary conditions, (9) and (10), for the restriction problem for the sphere. It is conjectured that these conditions are also sufficient. That is,

Conjecture 5.4 (Restriction conjecture for the sphere)**.** *If $S = S^{n-1}$ is the unit sphere and $1 \leq p, q \leq \infty$, then $R_S(p \to q)$ holds if and only if* (9) *and* (10) *both hold.*

Although this conjecture is stated for the sphere, it is believed that this conjecture should be true for any compact hypersurface S with boundary whose Gaussian curvature is always non-vanishing. (This includes compact subsets of the paraboloid, but not the cone).

In the next series of lectures we give some positive progress on this conjecture (namely, a theorem of Tomas and Stein), and summarize what is known. Then we shall apply the Tomas-Stein theorem to the problems of Bocher-Riesz summation, and global existence for semilinear PDE. After that, we shall return to the restriction problem and study it more deeply.

6. Proof of Proposition 5.3

The proof of this estimate is based around a body of techniques known collectively as "the method of stationary phase". To go into the proofs of these techniques would take us somewhat far afield, and we will delegate this responsibility to "Harmonic Analysis", E.M. Stein, Princeton University Press, Chapter VIII.

The main tools which we will need (and return to later) are as follows.

Lemma 6.1 (Principle of non-stationary phase)**.** *Let M be a smooth manifold, ψ a compactly supported bump function on M, and ϕ be a smooth real function on the support of ψ such that $|\nabla\phi|$ is bounded away from zero. (i.e. ϕ is "non-stationary"). Then we have*

$$\int_M e^{i\lambda\phi(x)}\psi(x)\, dx = O(\lambda^{-N})$$

as $\lambda \to +\infty$, where N is an arbitrary integer.

The most familiar instance of this lemma is when $M = \mathbf{R}$ and $\phi(x) = x$; this lemma now says that the Fourier transform of a bump function is rapidly decreasing. The general case is proven in a similar fashion to this special case, i.e. by repeated integration by parts. The intuitive explanation for this lemma is that if ϕ is non-stationary and λ is large, then the phase function $e^{i\lambda\phi(x)}$ oscillates so rapidly that the integral cancels itself out almost completely.

Essentially, the above lemma states that the integral $\int_M e^{i\lambda\phi(x)}\psi(x)\, dx$ is negligible except when ϕ is stationary (i.e. when $\nabla\phi = 0$). At these stationary points the integral is not negligible, but its asymptotics can often be computed quite accurately. For instance, we have

Lemma 6.2 (Principle of stationary phase for non-degenerate stationary points)**.** *Let x_0 be a point in $\mathbf{R}^n$. Suppose ϕ is a smooth real function on a neighbourhood of x_0 which has a non-degenerate stationary point at x_0. In other words, we suppose that $\nabla\phi(x_0) = 0$, but*

$$\det(\partial_i\partial_j\phi(x_0)) \neq 0,$$

where $\partial_i\partial_j\phi(x_0)$ is the Hessian matrix of ϕ at x_0. Then, if ψ is a bump function supported on a sufficiently small neighbourhood of x_0, then we have

$$\int_{\mathbf{R}^n} e^{i\lambda\phi(x)}\psi(x)\, dx = C\psi(x_0)e^{i\lambda\phi(x_0)}\lambda^{-n/2} + O(\lambda^{-(n+1)/2})$$

as $\lambda \to +\infty$, where C is a constant depending on ϕ.

A very crude justification of the above lemma can be given as follows. Split the integral into two pieces,

$$\int_{|x-x_0|\ll\lambda^{-1/2}} e^{i\lambda\phi(x)}\psi(x)\, dx \tag{11}$$

and

$$\int_{|x-x_0|\gtrsim\lambda^{-1/2}} e^{i\lambda\phi(x)}\psi(x)\, dx. \tag{12}$$

To handle the first term, we use the Taylor expansion

$$\phi(x) = \phi(x_0) + \nabla\phi(x_0)(x - x_0) + O(|x - x_0|^2) = \phi(x_0) + O(1/\lambda)$$

to conclude that

$$|e^{i\lambda\phi(x)} - e^{i\lambda\phi(x_0)}| \ll 1.$$

Since the integral is over a set of measure about $C\lambda^{-n/2}$, we see that (11) is thus approximately equal to $C\psi(x_0)e^{i\lambda\phi(x_0)}\lambda^{-n/2}$, as desired. As for (12), we see from the non-degeneracy condition that $|\nabla\phi(x)| \sim |x - x_0|$, so that $|\lambda\nabla\phi(x)| \gtrsim 1$ on the domain of integration of (12). Thus the phase in (12) is non-stationary, and one expects the contribution to be more or less negligible from the principle of non-stationary phase.

Complete proofs of these two lemmas can be found in e.g. Stein Chapter VIII. We'll just take these lemmas for granted and prove Proposition 5.3.

By radial symmetry we may assume that $x = \lambda e_n$ for some $\lambda \gg 1$. We thus have

$$\widehat{d\sigma}(\lambda e_n) = \int_{S^{n-1}} e^{-2\pi i\lambda\omega_n} d\omega. \tag{13}$$

The function $\omega \mapsto \omega_n$, which maps S^{n-1} to the interval $[-1, 1]$ is stationary when $\omega = \pm e_n$ and non-stationary otherwise. Thus, we break up (13) into

$$\int_{S^{n-1}} e^{-2\pi i\lambda\omega_n}\psi_+(\omega)d\omega + \int_{S^{n-1}} e^{-2\pi i\lambda\omega_n}\psi_-(\omega)d\omega + \int_{S^{n-1}} e^{-2\pi i\lambda\omega_n}(1-\psi_+-\psi_-)(\omega)d\omega$$

where ψ_+, ψ_- are cutoff functions supported very close to $+e_n$, $-e_n$ respectively.

Let's look at the first integral. We change variables and write this in terms of $\underline{\omega} = (\omega_1, \dots, \omega_{n-1})$. We have

$$\omega_n = (1 - |\underline{\omega}|^2)^{1/2} = \Phi(\underline{\omega}) \text{ (say) } .$$

It's easy to check that Φ has a non-degenerate stationary point at $\underline{\omega} = 0$. Thus by the principle of stationary phase, the contribution of this integral is

$$Ce^{-2\pi i\lambda\Phi(0)}\lambda^{-(n-1)/2} + O(\lambda^{-n/2}),$$

noting that we have an $n-1$-dimensional integral here rather than an n-dimensional integral. Since $\Phi(0) = 1$, the contribution of ψ_+ is thus

$$Ce^{-2\pi i\lambda}\lambda^{-(n-1)/2} + O(\lambda^{-n/2}).$$

Similarly, the contribution of ψ_- is

$$Ce^{2\pi i\lambda}\lambda^{-(n-1)/2} + O(\lambda^{-n/2}).$$

Finally, the contribution of $(1 - \psi_+ - \psi_-)$ is $O(\lambda^{-N})$ for any N by the principle of non-stationary phase. Putting all these estimates together we get the result.

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

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