LECTURE NOTES 2 FOR 254B

TERENCE TAO

1. The Tomas-Stein restriction theorem

Let $R_S(p \to q)$ denote the estimate

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

When S is the unit sphere $S = S^{n-1}$, we have seen that this estimate is only possible if $p < 2n/(n-1)$ and $p' \geq (n+1)/(n-1)q$. The restriction conjecture asserts that these necessary conditions are also sufficient. However, as far as positive results go, we have only the trivial observation that $R_S(1 \to q)$ holds for all q.

We now show the restriction theorem of Tomas and Stein, which shows that the restriction conjecture is true for $q = 2$. In other words:

Theorem 1.1. *If $1 \leq p \leq 2(n+1)/(n+3)$, then $R_S(p \to 2)$ holds.*

As we shall see, the proof relies very heavily on the fact that $q = 2$, and it has been very difficult (though not completely impossible) to push this argument beyond to $q < 2$. In the next set of notes we shall apply this estimate to Bochner-Riesz summation and semilinear Schrödinger equations.

We apply a standard trick known as the TT^* method. In general, an operator T is bounded from L^p to L^2 if and only if its square TT^* is bounded from $L^{p'} to L^p$. Let's see how this principle works in our concrete context. We square (1); we now have to prove

$$\int |\hat{f}(\xi)|^2 d\omega(\xi) \leq C\|f\|_p^2.$$

We rewrite the left-hand side as

$$\langle \hat{f}, \hat{f} d\omega \rangle \leq C\|f\|_p^2.$$

Since the Fourier transform intertwines multiplication and convolution, this is

$$\langle \hat{f}, \widehat{f * \widehat{d\omega}} \rangle \lesssim \|f\|_p^2.$$

Note that $d\omega$ is real and symmetric, so that there is no distinction between its Fourier transform and inverse Fourier transform. Since the Fourier transform is a unitary operator, this becomes

$$\langle f, f * \widehat{d\omega} \rangle \lesssim \|f\|_p^2.$$

From Hölder's inequality, it thus suffices to show that

$$\|f * \widehat{d\omega}\|_{p'} \lesssim \|f\|_p. \tag{2}$$

We have thus transformed our problem into an $L^p \to L^{p'}$ operation on a certain convolution operator. This trick is very special to the $q = 2$ case; it is remarkable how little light it sheds on the other cases.

2. First attempt: fractional integration.

The above observations were first made in 1970 by Fefferman and Stein. To attack (2), the most obvious tool to use was the fractional integration lemma of Hardy-Littlewood-Sobolev:

Lemma 2.1. *If $0 < \alpha < n$, $1 < p, q < \infty$, and*

$$\frac{1}{q} + 1 = \frac{1}{p} + \frac{\alpha}{n}, \tag{3}$$

then we have

$$\|f * \frac{1}{|x|^\alpha}\|_q \lesssim \|f\|_p.$$

This can be thought of as an endpoint version of Young's inequality; the point is that $1/|x|^\alpha$ is "almost" in $L^{n/\alpha}$. (Actually it's in the slightly weaker space $L^{n/\alpha,\infty}$).

We shall spend some time to prove this estimate, because I want to introduce some techniques we need to use later. First, let's introduce the notion of a *restricted weak-type estimate*. Let X and Y be measure spaces. Recall that if T is a linear operator from functions of X to functions of Y, we say that T is of strong-type (p, q) if

$$\|Tf\|_q \lesssim \|f\|_p \text{ for all } f.$$

We say that T is of weak-type (p, q) if

$$|\{|Tf| \geq \lambda\}| \lesssim \|f\|_p^q \lambda^{-q} \text{ for all } f, \lambda > 0;$$

Clearly weak-type is weaker than strong-type. An even weaker estimate obtains if we restrict f to be a characteristic function:

$$|\{|T\chi_E| \geq \lambda\}| \lesssim |E|^{q/p}\lambda^{-q} \text{ for all } E, \lambda > 0; \tag{4}$$

in this case we say T is of restricted weak-type (p, q).

It is convenient to rephrase (4) in a more symmetric form.

Lemma 2.2. *Suppose* $1 < p, q < \infty$. *Then* (4) *holds if and only if one has*

$$|\langle T\chi_E, \chi_F\rangle| \lesssim |E|^{1/p}|F|^{1/q'} \tag{5}$$

for all sets E, F.

As a comparison, note that the strong-type estimate is equivalent to

$$|\langle Tf, g\rangle| \leq \|f\|_p \|g\|_{q'}.$$

Proof For our purposes, we only need the implication (4) $\implies$ (5); I leave the opposite implication as an exercise. (Hint: apply (5) to the set $F = \{Re(T\chi_E) > \lambda\}$).

We have

$$|\langle T\chi_E, \chi_F\rangle| \leq \int_F |T\chi_E(x)| \, dx.$$

We rewrite this as

$$\int_F \int_0^\infty \chi_{|T\chi_E|>\lambda}(x) \, d\lambda \, dx.$$

Interchanging the integrals, and then doing the x integration, this becomes

$$\int_0^\infty |\{x \in F : |T\chi_E(x)| > \lambda\}| \, d\lambda.$$

We have two estimates for the integrand. The first is just $|F|$. The second is $O(|E|^{q/p}\lambda^{-p})$. Thus we can estimate this integral by

$$O(\int_0^\infty \min(|F|, |E|^{q/p}\lambda^{-q}) \, d\lambda).$$

The first term in the minimum is smaller when $\lambda \leq |E|^{1/p}|F|^{-1/q}$, and the second term is smaller otherwise. Breaking up the integral accordingly, and doing some freshman calculus, we eventually get (5). ■

Restricted weak-type estimates are clearly weaker than strong type estimates; however for interpolation purposes they are equally good.

Lemma 2.3 (Marcinkeiwicz interpolation). *Suppose* $1 < p_0 < q_0 < \infty$, $1 < p_1 < q_1 < \infty$, $p_0 < p_1$, $q_0 < q_1$, *and and* T *is of restricted weak-type* (p_0, q_0) *and* (p_1, q_1). *Then* T *is of strong-type* (p_θ, q_θ) *for any* $0 < \theta < 1$, *where* $1/p_\theta = (1-\theta)/p_0 + \theta/p_1$ *and similarly for the qs.*

Proof Fix θ, and write p, q for p_θ, q_θ. We have to show that

$$|\langle Tf, g\rangle| \lesssim \|f\|_p \|g\|_{q'}.$$

We may assume that f, g are non-negative, because any complex function is the linear combination of at four non-negative functions.

By re-arranging X and Y, we may assume that X and Y are the positive real line $(0, \infty)$ with the standard measure, and f, g are monotone decreasing functions. By a limiting argument we can assume that f, g are test functions.

We now write f as a linear combination of characteristic functions:

$$f = \int_0^\infty f(\lambda)\delta_\lambda \, d\lambda = -\int_0^\infty f'(\lambda)\chi_{[0,\lambda]} \, d\lambda$$

and similarly for g. By linearity, we thus have

$$\langle Tf, g\rangle = \int_0^\infty \int_0^\infty f'(\lambda) g'(\beta) \langle T\chi_{[0,\lambda]}, \chi_{[0,\beta]}\rangle \, d\lambda d\beta.$$

Fromm the previous lemma and our assumptions we have

$$|\langle T\chi_{[0,\lambda]}, \chi_{[0,\beta]}\rangle| \lesssim \lambda^{1/p_i} \beta^{1/q'_i}$$

for $i = 1, 2$. Thus we have

$$|\langle Tf, g\rangle| \lesssim \int_0^\infty \int_0^\infty f'(\lambda) g'(\beta) \min_{i=1,2}(\lambda^{1/p_i}\beta^{1/q'_i}) \, d\lambda d\beta.$$

Integrating by parts, this is[1]

$$C\int_0^\infty \int_0^\infty f(\lambda) g(\beta) \min_{i=1,2}(\lambda^{1/p_i - 1}\beta^{1/q'_i - 1}) \, d\lambda d\beta.$$

Let a, b be such that

$$1/q'_i = -a/p_i + ab \text{ for } i = 1, 2;$$

note that this implies that

$$1/q = -a/p + ab. \tag{6}$$

The integral now becomes

$$C\int_0^\infty \int_0^\infty f(\lambda) g(\beta) \min_{i=1,2}(\lambda/\beta^a)^{1/p_i} \beta^{ab} \frac{d\lambda}{\lambda}\frac{d\beta}{\beta}.$$

Making the change of variables $\eta = \beta^a$, this becomes

$$C\int_0^\infty \int_0^\infty f(\lambda) g(\eta^{1/a}) \min_{i=1,2}(\lambda/\eta)^{1/p_i} \eta^b \frac{d\lambda}{\lambda}\frac{d\eta}{\eta}.$$

[1]Note that the discontinuity caused by the minimum does not cause a difficulty.

Now making the change of variables $\gamma = \lambda/\eta$, this becomes

$$C \int_0^\infty \int_0^\infty f(\eta\gamma) g(\eta^{1/a}) \min_{i=1,2} \gamma^{1/p_i} \eta^b \frac{d\gamma}{\gamma} \frac{d\eta}{\eta}.$$

Interchanging the integrals, this is

$$C \int_0^\infty \min_{i=1,2} \gamma^{1/p_i} \int_0^\infty f(\eta\gamma) g(\eta^{1/a}) \eta^{b-1} \, d\eta \frac{d\gamma}{\gamma}.$$

By Hölder, the inner integral is bounded by

$$(\int_0^\infty f(\eta\gamma)^p \, d\eta)^{1/p} (\int_0^\infty (g(\eta^{1/a})\eta^{b-1})^{p'} \, d\eta)^{1/p'}.$$

A tedious computation using (6) shows that this simplifies to

$$C\gamma^{-1/p} \|f\|_p \|g\|_{q'}.$$

Putting all this together, we get

$$|\langle Tf, g\rangle| \lesssim \|f\|_p \|g\|_{q'} \int_0^\infty \gamma^{-1/p} \min_{i=1,2} \gamma^{1/p_i} \frac{d\gamma}{\gamma}.$$

Since $p_0 < p < p_1$, we see that the integral converges, and we are done. ■

Now we can prove the Hardy-Littlewood-Sobolev inequality. From the above considerations it suffices to prove the restricted weak-type estimate

$$|\{\chi_E * \frac{1}{|x|^\alpha} > \lambda\} \lesssim |E|^{q/p} \lambda^{-q}$$

for the same range of exponents, since one can automatically upgrade this to the strong type estimate by Marcinkeiwicz interpolation. By Lemma 2.2, this is equivalent to

$$\langle \chi_E * \frac{1}{|x|^\alpha}, \chi_F \rangle \lesssim |E|^{1/p} |F|^{1/q'}. \tag{7}$$

To estimate this, we first observe that

$$\int_X \frac{1}{|x|^\alpha} \, dx \lesssim |X|^{1-\frac{\alpha}{n}}$$

for any set $X \subset \mathbf{R}^n$. Indeed, it is clear that for a fixed size $|X|$, the left-hand side is maximized when X is a ball centered at the origin (since this is where $\frac{1}{|x|^\alpha}$ is largest), and in this case the estimate is clear. In particular, we have

$$\|\chi_E * \frac{1}{|x|^\alpha}\|_\infty \lesssim |E|^{1-\frac{\alpha}{n}}.$$

Thus the left-hand side of (7) is bounded by

$$|E|^{1-\frac{\alpha}{n}} |F|.$$

On the other hand, we can pass the convolution kernel onto the other side of the inner product:

$$\langle \chi_E * \frac{1}{|x|^\alpha}, \chi_F \rangle = \langle \chi_E, \chi_F * \frac{1}{|x|^\alpha} \rangle$$

and argue symmetrically to obtain the bound of

$$|E||F|^{1-\frac{\alpha}{n}}.$$

The estimate (7) then follows by taking a suitable average of the two estimates, using (3). This completes the proof of the Hardy-Littlewood-Sobolev inequality.

We now return to (2). To exploit fractional integration, we simply use the decay estimate

$$|\widehat{d\omega}(x)| \lesssim |x|^{-(n-1)/2} \tag{8}$$

to crudely estimate

$$\|f * \widehat{d\omega}\|_{p'} \leq \||f| * |x|^{-(n-1)/2}\|_{p'}.$$

In order for this to be controlled by $\|f\|_p$, we need

$$\frac{1}{p'} + 1 = \frac{1}{p} + \frac{(n-1)/2}{n},$$

which works out to

$$p = 4n/(3n+1).$$

In other words, we have just proved that $R_S(4n/(3n+1) \to 2)$ holds. By interpolation with the trivial estimate $R_S(1 \to 2)$, we thus get $R_S(p \to 2)$ for all $1 \leq p \leq 4n/(3n+1)$. This is a non-trivial statement, however, it is not best possible; the restriction conjecture asserts that $R_S(p \to 2)$ for all $1 \leq p \leq 2(n+1)/(n+3)$, which is a significantly larger range.

The reason why we did not get such a good estimate here is because we only exploited the decay of the convolution kernel $\widehat{d\sigma}$ in (8). However, $\widehat{d\sigma}$ also has some oscillation, and this should also be exploited to get better estimates. For instance, when $n = 3$ we actually have the exact formula

$$\widehat{d\sigma}(x) = 2\frac{\sin(2\pi|x|)}{|x|},$$

as can be easily verified. Crudely estimating this by $O(1/|x|)$ is very inefficient.

3. Second attempt: Real interpolation.

In 1975 Tomas introduced a very simple argument (his paper was only two pages long!) which allowed one to use both the decay and the oscillation of the kernel $\widehat{d\sigma}$ to get within an epsilon of the sharp result. The idea was to decompose $\widehat{d\sigma}$ dyadically. (This is a very effective technique in harmonic analysis - break up your functions and kernels into many pieces. This works quite well, except when one has to recombine all the pieces together; one can often lose an epsilon this way, but rarely does one lose more than this).

Let ϕ be a radial bump function which equals 1 near 0 and is compactly supported, and define

$$\psi_k(x) = \phi(2^{-k}x) - \phi(2^{1-k}x).$$

Thus ψ_k has size roughly 1 and is supported on the annulus $|x| \sim 2^k$. Also the ψ_k are all related to each other by

$$\psi_k(x) = \psi_0(2^{-k}x). \tag{9}$$

We have the telescoping identity

$$1 = \phi(x) + \sum_{k>0} \psi_k(x)$$

and so we can break up $f * \widehat{d\sigma}$ as

$$f * \widehat{d\sigma} = f * (\phi\widehat{d\sigma}) + \sum_{k>0} f * (\psi_k \widehat{d\sigma}).$$

We now just use the triangle inequality

$$\|f * \widehat{d\sigma}\|_{p'} \leq \|f * (\phi\widehat{d\sigma})\|_{p'} + \sum_{k>0} \|f * (\psi_k \widehat{d\sigma})\|_{p'}$$

and estimate each term separately. (One could try to be a bit more sophisticated than the triangle inequality; we'll see this in the next section).

The first term is really easy. Since $d\sigma$ is a compactly supported finite measure, $\widehat{d\sigma}$ is a C^∞ function (in fact, it is complex analytic), and so $\phi\widehat{d\sigma}$ is just a C_0^∞ bump function. In which case one can easily control this quantity in terms of $\|f\|_p$, just from Young's inequality (for example).

Now let's look at the other terms. We would really like an estimate such as

$$\|f * (\psi_k \widehat{d\sigma})\|_{p'} \lesssim 2^{-\varepsilon k} \|f\|_p \tag{10}$$

for some $\varepsilon > 0$, since this would sum up quite nicely. Now, in general, good $(L^p, L^{p'})$ estimates can be hard to come by; but there are two extremes, namely (L^1, L^∞) and (L^2, L^2) estimates, which are really easy.

First, let's observe that

$$\|f * (\psi_k \widehat{d\sigma})\|_\infty \lesssim 2^{-(n-1)k/2} \|f\|_1 \tag{11}$$

Why is this? Well, from (8) we see that

$$\|\psi_k \widehat{d\sigma}\|_\infty \lesssim 2^{-(n-1)k/2}.$$

So this is clear from (a very easy case of) Young's inequality.

That's good - we have an $(L^p, L^{p'})$ estimate with a really strong decay in k. However, it's just for $p = 1$, and we already know that everything works fine for $p = 1$. To get off this exponent, we need another estimate, and it turns out that an L^2 estimate is perfect for this:

$$\|f * (\psi_k \widehat{d\sigma})\|_2 \lesssim 2^k \|f\|_2 \tag{12}$$

Now why is this true? Well, there is an important fact about convolution operators, which is that their L^2 behaviour is extremely easy to characterize. Indeed, for any kernel K we have

$$\|f * K\|_2 \leq \|\hat{K}\|_\infty \|f\|_2;$$

this is just Plancherel and Hölder. So to show (12) we just need to show that

$$\|\widehat{\psi_k \widehat{d\sigma}}\|_\infty \lesssim 2^k,$$

which simplifies to showing that

$$|\widehat{\psi_k} * d\sigma(x)| \lesssim 2^k \tag{13}$$

for all x.

From (9) we have

$$\widehat{\psi_k}(x) = 2^{nk} \widehat{\psi_0}(2^k x).$$

Since ψ_0 is a Schwarz function, its Fourier transform is also Schwarz, so we have

$$|\widehat{\psi_k}(x)| \lesssim \frac{2^{nk}}{(1 + 2^k|x|)^N} \tag{14}$$

for any $N > 0$. Thus we just need to show that

$$\left|\frac{2^{nk}}{(1 + 2^k|x|)^N} * d\sigma(x)\right| \lesssim 2^k.$$

But this is just a boring computation. Roughly speaking, the kernel $\frac{2^{nk}}{(1+2^k|x|)^N}$ is an approximation to the identity with thickness about 2^{-k}, so this convolution takes surface measure on the sphere and blurs it out to thickness 2^{-k}. Since blurring preserves L^1 norm, this convolution still has to have L^1 norm comparable to 1. Since it is supported on an annulus of thickness 2^{-k}, this explains why it has size about 2^k.

Anyway, we have an (L^1, L^∞) estimate with a good decay, and an (L^2, L^2) estimate with some growth. If we use real interpolation to interpolate between the two, we end up with (10) with some positive ε provided that $p < 2(n+1)/(n+3)$. (This, of course, is another example of tedious algebra which unfortunately comes up incredibly often in this field. But it's the techniques and ideas which are important, not the numerology - that's really just a way of keeping score).

Thus, by exploiting oscillation (via the Fourier transform-based estimate (12)) in addition to decay, we get $R_S(p \to 2)$ for all $1 \leq p < 2(n+1)/(n+3)$; this is almost, but not quite, the sharp result, as we are still missing the endpoint $p = 2(n+1)/(n+3)$. But it is a very good near miss.

To summarize the basic strategy here: we interpolated between an (L^1, L^∞) estimate, which was based on the decay of the kernel $\widehat{d\sigma}$ and an L^2 estimate, to get the L^p estimates we wanted.

4. Last attempt: complex interpolation

The endpoint estimate $R_S(2(n+1)/(n+3) \to 2)$ was obtained by Stein in 1975 (unpublished), and is the last part we have to prove in the Tomas-Stein theorem.

The basic idea is the same as the Tomas argument, but the key innovation is to replace real interpolation by complex interpolation, and to refuse to give in to the triangle inequality as we did in the previous section. We will also need a special assumption on the cutoff function ψ, but we'll leave you in suspense as to the exact condition required for now.

Let $p = 2(n+1)/(n+3)$. As before, it suffices to show that

$$\| \sum_{k>0} f * (\psi_k \widehat{d\sigma}) \|_{p'} \lesssim \|f\|_p. \tag{15}$$

At this endpoint, (10) only holds with $\varepsilon = 0$, which means that we will lose if we just apply the triangle inequality. Instead, we will show the following enhanced versions of (11) and (12):

$$\| \sum_{k>0} 2^{[\frac{n-1}{2} + it]k} f * (\psi_k \widehat{d\sigma}) \|_\infty \lesssim \|f\|_1 \tag{16}$$

$$\| \sum_{k>0} 2^{[-1+it]k} f * (\psi_k \widehat{d\sigma}) \|_2 \lesssim \|f\|_2 \tag{17}$$

for all $t \in \mathbf{R}$. From these estimates, (15) follows from the Stein complex interpolation theorem (see Stein, Chapter IX). Note that if we were only dealing with a single k rather than the infinite sum, this would just be (11) and (12).

Let's see how (16) works. Rewriting it as

$$\|f * \sum_{k>0} 2^{[\frac{n-1}{2}+it]k} \psi_k \widehat{d\sigma}\|_\infty \lesssim \|f\|_1$$

we see from Young's inequality that we just need to show that

$$\|\sum_{k>0} 2^{[\frac{n-1}{2}+it]k} \psi_k \widehat{d\sigma}\|_\infty \lesssim 1.$$

But this just follows from (8), since

$$\sum_{k>0} 2^{[\frac{n-1}{2}+it]k} \psi_k(x) = O(|x|^{(n-1)/2}).$$

Note how we are being more efficient here than in the proof of (11).

Now let's turn to (17). By arguments of the previous section, we have to show that

$$\|\sum_{k>0} 2^{[-1+it]k} \widehat{\psi_k} * d\sigma\|_\infty \lesssim 1;$$

we will ignore the possible cancellation coming from the 2^{itk} factor, and simply obtain this from

$$\sum_{k>0} 2^{-k} |\widehat{\psi_k} * d\sigma(x)| \lesssim 1. \tag{18}$$

The estimate (13) proven in the previous section is not going to be enough for this task, and we will need the more sophisticated estimate

$$|\widehat{\psi_k} * d\sigma(x)| \lesssim \begin{cases} 2^k (2^k d(x,S))^{-N} & d(x,S) \geq 2^{-k} \\ 1 + 2^k (2^k d(x,S)) & d(x,S) \leq 2^{-k} \end{cases}$$

where $d(x,S) = |1 - |x||$ is the distance from x to the unit sphere. Once we have this estimate, (18) follows by a routine calculation.

We have to estimate the quantity

$$|\int_{S^{n-1}} \widehat{\psi_k}(x - \omega)\ d\sigma(\omega)|.$$

We first consider the case when $d(x,S) \geq 2^k$. The claim then follows from (14) and some easy estimates. (For instance, break up the sphere into regions where $d(x,\omega) \sim 2^{k+j}$ for some $j \geq 0$, and then sum in j.)

Now we look at the case when $d(x,S) \leq 2^k$. If we just use (14) we'll end up with a bound of $O(2^k)$, which isn't good enough; the point is that we can also exploit the moment conditions on ψ_k.

We first observe the Lipschitz bound

$$|\nabla(\widehat{\psi_k} * d\sigma(x))| \lesssim 2^{2k}.$$

This just comes from the identity

$$\nabla(\widehat{\psi_k} * d\sigma(x)) = 2^k(2^{-k}\nabla\widehat{\psi_k}) * d\sigma(x)$$

and the fact that $2^{-k}\nabla\widehat{\psi_k}$ also satisfies (14).

Because of this bound, it suffices to prove the estimate when x is on the unit sphere; by rotational symmetry we may take $x = e_1$. In other words, we just need to show

$$|\int_{S^{n-1}} \widehat{\psi_k}(e_1 - \omega)\ d\sigma(\omega)| = O(1).$$

The estimate is trivial in the region $|e_1 - \omega| > 1/10$ from (14), so we may restrict our attention to the region $|e_1 - \omega| < 1/10$. In this case we parameterize ω as

$$\omega = (\underline{\omega}, (1 - |\underline{\omega}|^2)^{1/2})$$

The integrand now becomes

$$\int_{|\underline{\omega}| \ll 1} \widehat{\psi_k}(\underline{\omega}, 1 - (1 - |\underline{\omega}|^2)^{1/2}) J(\underline{\omega}) d\underline{\omega}$$

where J is a Jacobian. We rewrite this as

$$C \int_{\mathbf{R}^{n-1}} \widehat{\psi_k}(\underline{\omega}, O(|\underline{\omega}|^2))(1 + O(|\underline{\omega}|^2))\ d\underline{\omega} \tag{19}$$

modulo extremely tiny errors. We claim that this quantity is

$$C \int_{\mathbf{R}^{n-1}} \widehat{\psi_k}(\underline{\omega}, 0)\ d\omega + O(1). \tag{20}$$

If this were the case, then we simply choose ϕ (and thus ψ_0) so that

$$\int_{\mathbf{R}^{n-1}} \widehat{\psi_0}(\underline{\omega}, 0)\ d\omega = 0;$$

this will achieve the desired goal.

To prove the approximation, we first observe that

$$\widehat{\psi_k}(\underline{\omega}, O(|\underline{\omega}|^2)) = \widehat{\psi_k}(\underline{\omega}, 0) + O(\frac{2^{(n+1)k}|\underline{\omega}|^2}{(1 + 2^k|\underline{\omega}|)^N})$$

for all $N > 0$; this is best shown by considering the cases $|\underline{\omega}| < 2^{-k}$ and $|\underline{\omega}| \geq 2^{-k}$ separately. Thus the error between (19) and (20) is at most

$$\int O(\frac{2^{(n+1)k}|\underline{\omega}|^2}{(1 + 2^k|\underline{\omega}|)^N})\ d\underline{\omega}$$

which is easily calculated to be $O(1)$ (again, treat the case $|\underline{\omega}| < 2^{-k}$ and the cases $|\underline{\omega}| \sim 2^j 2^{-k}$ respectively separately.

Whew! That completes the proof of the Tomas-Stein restriction theorem.

This is a good example of the subtle differences of power between real and complex interpolation. Generally, for non-endpoint results, it doesn't matter which one you use, but at the endpoints it's a toss up which one is better. In this case complex is better. (There are other endpoint cases where real interpolation is definitely superior, but they usually involve technical spaces such as Lorentz and Besov spaces).

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

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