LECTURE NOTES 7 FOR 247B

TERENCE TAO

1. The $T(1)$ theorem

We now return to the study of Calderón-Zygmund operators T, using our recently established paraproduct estimates to establish one of the most celebrated theorems in the subject - the $T(1)$ theorem of David and Journé.

Recall that Calderón-Zygmund operators T have two properties. Firstly, they have a singular kernel $K(x, y)$, thus $K : \mathbf{R}^d \times \mathbf{R}^d \to \mathbf{C}$ obeys the bounds

$$K(x, y) = O_d(|x - y|^{-d}) \quad (1)$$

and

$$K(x, y) - K(x', y) = O_{d,\theta}(|x - x'|^\theta |x - y|^{-d-\theta}) \quad (2)$$

$$K(x, y) - K(x, y') = O_{d,\theta}(|y - y'|^\theta |x - y|^{-d-\theta}) \quad (3)$$

whenever $x \neq y$ and $|x - x'|, |y - y'| \leq \frac{1}{2}|x - y|$ and some $0 < \theta \leq 1$. We assume K is a kernel for T in the sense that

$$Tf(x) = \int_{\mathbf{R}^d} K(x, y) f(y)\, dy \quad (4)$$

whenever f is bounded with compact support, and x lies outside the support of f. The second property is that T is bounded on $L^2(\mathbf{R}^d)$.

Once one has these two properties, one obtains several additional properties for free, such as boundedness on $L^p(\mathbf{R}^d)$ for all $1 < p < \infty$; also T and T^* both map L^∞ to BMO. However, it is not always easy to verify these properties in practice. The singular kernel properties are usually not too difficult, as the kernel can often be expressed directly, and then estimated pointwise in a fairly straightforward fashion. The L^2 boundedness however can often be tricky - we are asking to bound Tf for *all* functions $f \in L^2(\mathbf{R}^d)$, which is a large class to exhaust over. With assumptions such as translation invariance one can diagonalise T, which makes verification of L^2 boundedness much easier (it is equivalent to the Fourier multiplier symbol being bounded), but in general such invariance is not available. (In the case of pseudodifferential operators there is an almost orthogonality property which is sort of a weak version of translation invariance and serves to almost block-diagonalise T; again, this is not always available in general.)

There is however a remarkable theorem - the $T(1)$ theorem - which says that once one has the kernel bounds, verification of the L^2 bounds is equivalent to verifying bounds on Tf (or T^*f) on just a handful of functions, such as the constant function 1. This is particularly useful in a number of PDE applications in which $T1$ and T^*1 can be easily computed. The function 1 can be replaced by more general functions b - leading to a family of $T(b)$ theorems - but we will not pursue these generalisations here.

To avoid some minor technicalities let us make the qualitative assumptions that the singular kernel K is bounded and compactly supported (but our final bounds will not depend on exactly how bounded or compactly supported K is). Also, we assume that the formula (4) is valid for all $f \in L^2(\mathbf{R}^d)$ and $x \in \mathbf{R}^d$, not just compactly supported f and for x outside the support of f; note from Schur's test and our qualitative hypotheses that the integral operator in (4) is absolutely convergent and bounded on $L^2(\mathbf{R}^d)$ (but with a qualitative bound rather than a quantitative one). In particular the adjoint of T is given by

$$T^*g(y) = \int_{\mathbf{R}^d} \overline{K(x,y)} g(x)\ dx.$$

It also allows us to define the functions $T1$ and T^*1 without any technical difficulty.

Theorem 1.1 ($T(1)$ theorem)**.** *Let K be a singular kernel which is also bounded and compactly supported, and let T be the integral operator* (4). *Suppose also we have the following three properties:*

- $\|T(1)\|_{\mathrm{BMO}(\mathbf{R}^d)} \lesssim_d 1$.
- $\|T^*(1)\|_{\mathrm{BMO}(\mathbf{R}^d)} \lesssim_d 1$.
- *(Weak boundedness) For any ball B, we have $\langle T1_B, 1_B \rangle = O_d(|B|)$.*

Then we have $\|T\|_{L^2(\mathbf{R}^d) \to L^2(\mathbf{R}^d)} \lesssim_{d,\theta} 1$.

In the converse direction, if T is bounded on L^2, then it is a CZO, and the boundedness of $T(1)$ and $T^*(1)$ follows (with a norm of $O_{d,\theta}(1)$) since T and T^* map L^∞ to BMO. The weak boundedness follows from the strong boundedness by Cauchy-Schwarz. Thus the above three properties give a necessary and sufficient condition for L^2 boudnedness of operators with singular kernels, at least assuming some qualitative assumptions. It turns out that none of the three hypotheses are redundant; they deal with three separate components of T, namely the "low-to-high", "high-to-low", and "high-to-high" frequency behaviours (more on this later; the similarity with the paraproduct decomposition is not coincidental).

We begin proving the $T(1)$ theorem. The first step is to show that the weak boundedness property, when combined with the BMO bounds, can be bootstrapped to a stronger estimate.

Lemma 1.2. *Let T be as in Theorem 1.1. Then for any ball B, we have*

$$\int_B |T1_B|^2\ dx \lesssim_{d,\theta} |B|$$

and

$$\int_B |T^*1_B|^2\ dx \lesssim_{d,\theta} |B|$$

The point here is that the absolute values are *inside* the integral, whereas with weak boundedness they are outside.

Proof As the hypotheses of Theorem 1.1 are invariant under replacing T with T^*, it suffices to prove the first claim.

Since $T1$ lies in BMO, we have

$$\int_B |T1(x) - c_B|^2 \, dx \lesssim_{d,\theta} |B|$$

for some c_B. Also, using (1), (2), (4) we have the pointwise bounds

$$|T(1 - 1_B)(x) - T(1 - 1_B)(x_B)| \lesssim_{d,\theta} 1 + \log \frac{r_B}{\operatorname{dist}(x, \partial B)}$$

for all $x \in B$, where r_B is the radius of the ball and ∂B is the boundary (one should deal with the contribution of $1_{2B} - 1_B$ and $1 - 1_{2B}$ separately). Thus

$$\int_B |T(1 - 1_B)(x) - T(1 - 1_B)(x_B)|^2 \, dx \lesssim_{d,\theta} |B|$$

and thus by the triangle inequality

$$\int_B |T1_B(x) - c'_B|^2 \, dx \lesssim_{d,\theta} |B|$$

for some c'_B. By Cauchy-Schwarz we see in particular that

$$|\int_B (T1_B(x) - c'_B) \, dx| \lesssim_{d,\theta} |B|.$$

On the other hand, from weak boudnedness we have

$$|\int_B T1_B(x) \, dx| \lesssim_{d,\theta} |B|$$

and so by the triangle inequality we have $c'_B = O_{d,\theta}(1)$. The claim then follows from one final application of the triangle inequality. ∎

We now allow 1_B to be replaced by other functions adapted to B.

Lemma 1.3. *Let T be as before. Let ϕ_B be a bump function adapted to a ball B. Then*

$$\int_B |T\phi_B|^2 \, dx \lesssim_{d,\theta} |B|$$

and

$$\int_B |T^*\phi_B|^2 \, dx \lesssim_{d,\theta} |B|$$

Proof For $x \in B$, let us examine the commutator

$$T\phi_B(x) - \phi_B(x)T1_B(x).$$

We can expand this expression as

$$\int_B K(x,y)(\phi_B(y) - \phi_B(x)) \, dy.$$

Since ϕ_B is a bump function, we can bound $\phi_B(y) - \phi_B(x) = O_d(|y-x|/r_B)$. Applying (1) we then conclude that

$$T\phi_B(x) - \phi_B(x)T1_B(x) = \int_B K(x,y)(\phi_B(y) - \phi_B(x))\ dy = O_d(1)$$

for all $x \in B$. The first claim then follows from the previous lemma and the triangle inequality; the second claim is proven similarly. ■

For technical reasons we shall need to replace bump functions by Schwartz functions of mean zero.

Lemma 1.4. *Let T be as before. Let ϕ_B be a Schwartz function of height 1 adapted to a ball B with radius r_B, such that $\int_{\mathbf{R}^d} \phi_B = 0$. Then*

$$\int_{B'} |T\phi_B|^2 \lesssim_{d,\theta} |B| \langle \mathrm{dist}(B, B')/r_B \rangle^{-d-\theta}$$

for all balls B' of radius equal to that of B.

Proof We can split up into two cases, when ϕ_B is supported in $5B'$ and when it is supported outside of $2B'$. (A standard smooth decomposition doesn't quite work because one needs to preserve the mean zero condition. Instead, remove the portion of ϕ_B supported on and around $2B'$, and move it, say, $2r_B$ units in an arbitrary direction, to create the component which vanishes outside of $2B'$ and which continues to have mean zero; the other term is then the difference between the new function and the old one.). If ϕ_B is supported in $5B'$, then (by the Schwartz property) it is equal to $\langle \mathrm{dist}(B, B')/r_B \rangle^{-100d}$ (say) times a bump function adapted to $5B'$, then the claim follows from the previous lemma. If instead ϕ_B vanishes outside of $5B'$, we can use the mean zero property of ϕ_B to write

$$T\phi_B(x) = \int_{\mathbf{R}^d} (K(x,y) - K(x_{B'}, y))\phi_B(y)\ dy.$$

By (2) and the decay of ϕ_B we see that $T\phi_B(x) = O_{d,\theta}(\langle \mathrm{dist}(B, B')/r_B \rangle^{-d-\theta}|B|)$ for all $x \in B'$, and the claim follows. ■

Until now we have only bounded T on very special types of functions. Now we take our first step toward bounding T on all functions. We use Littlewood-Paley decomposition to write

$$T = \sum_{j,k} \psi_j(D)\psi_j(D)T\psi_k(D)\psi_k(D) \tag{5}$$

for appropriate Littlewood-Paley multipliers ψ_j.

We can first deal with a single term:

Lemma 1.5. *Let the notations and hypotheses be as above. For any j, k, we have*

$$\|\psi_j(D)T\psi_k(D)f\|_{L^2(\mathbf{R}^d)} \lesssim_{d,\theta} \|f\|_{L^2(\mathbf{R}^d)}.$$

for all $f \in L^2(\mathbf{R}^d)$.

Proof By duality we may assume that $k \geq j$. Observe that

$$\psi_j(D)T\psi_k(D)f(x) = \int_{\mathbf{R}^d} K_{j,k}(x,y)f(y)\ dy$$

where

$$K_{j,k}(x,y) = \int_{\mathbf{R}^d} \check{\psi}_j(x-z)T\mathrm{Trans}_y\check{\psi}_k(z)\ dz.$$

Applying the previous lemma, noting that $\check{\psi}_k$ is a Schwartz functions of height 2^{dk} adapted to $B(0,2^{-k})$, we have

$$\int_{B(w,2^{-k})} |T\mathrm{Trans}_y\check{\psi}_k(z)|\ dz \lesssim_{d,\theta} \langle 2^k|w-y|\rangle^{-d-\theta}$$

for all balls $B(w,2^{-k})$ of radius 2^{-k}. On the other hand, $\check{\psi}_j$ is a Schwartz function of height 2^{dj} adapted to the ball $B(0,2^{-j})$, which one can cover by $O(2^{d(k-j)})$ balls of radius 2^k. From this we conclude that

$$K_{j,k}(x,y) = O_{d,\theta}(2^{dj}\langle 2^j|x-y|\rangle^{-d-\theta})$$

and the claim then follows from Schur's test. ∎

The above lemma does not quite let us sum in j,k; it turns out we need some additional decay away from the diagonal $j=k$. Let us first consider the case $k < j-10$, so that the input is much lower frequency than the output. The idea here is to approximate a low frequency function by a constant. Roughly speaking, we want to exploit the heuristic

$$T\phi \approx \phi T1$$

when ϕ is low frequency (and thus slowly varying). The precise way of formulating this requires the following variant of Lemma 1.4.

Lemma 1.6. *Let T be as before. Let ϕ_B be a Schwartz function of height 1 adapted to a ball B of radius r_B. Let B' be another ball with smaller radius $r_{B'} < r_B$, and let $\phi_{B'}$ be a Schwartz function of height 1 adapted to B' such that*

$$\int_{\mathbf{R}^d} \phi_{B'} = \int_{\mathbf{R}^d} \phi_B\phi_{B'} = 0.$$

Then

$$|\int_{\mathbf{R}^d} \phi_{B'}T\phi_B - \phi_{B'}\phi_B T1| \lesssim_{d,\theta} \frac{r_{B'}}{r_B})^{\theta/2}|B'|\langle \mathrm{dist}(B,B')/r_B\rangle^{-d-\theta/2} \qquad (6)$$

for all balls B' of radius equal to that of B.

Proof We can write the integral in (6) explicitly as

$$\int_{\mathbf{R}^d}\int_{\mathbf{R}^d} \phi_{B'}(x)K(x,y)[\phi_B(y)-\phi_B(x)]\ dxdy.$$

To compute this integral we shall split K into two parts. Let $r := \sqrt{r_B r_{B'}}$, and let χ_r be a bump function adapted to $B(0,r)$ which equals 1 on $B(0,r)$, and split $K = K_1+K_2$ where $K_1(x,y) := K(x,y)\chi_r(x-y)$ and $K_2(x,y) := K(x,y)(1-\chi_r(x-y))$. Observe that K_1 and K_2 are also singular kernels.

Let's first deal with the contribution of the local kernel K_1, where we have $x-y=O(r)$. In this regime we have the Lipschitz bound $\phi_B(y)-\phi_B(x)=O_d(\frac{|x-y|}{r_B}\langle \mathrm{dist}(x,B)/r_B\rangle^{-100d})$. Inserting this bound, and the bounds $K_1(x,y)=O_d(|x-y|^{-d}1_{|x-y|=O(r)})$, we see that

$$|\int_{\mathbf{R}^d} K_1(x,y)[\phi_B(y)-\phi_B(x)]\ dy| \ll_d \frac{r}{r_B}\langle \mathrm{dist}(x,B)/r_B\rangle^{-100d}$$

and from this we quickly see that the contribution of K_1 to (6) is acceptable.

Now we turn to K_2. Here we use duality to rewrite the contribution to (6) as

$$\int_{\mathbf{R}^d} \phi_B T_2^t \phi_{B'} - T_2^t(\phi_B\phi_{B'})\ dx$$

where T_2 is the integral operator associated to K_2, and T_2^t is the transpose operator. Let's look at $T_2^t\phi_{B'}(x)$. Using the fact that $\phi_{B'}$ has mean zero, we can write this as

$$T_2^t\phi_{B'}(x) = \int_{\mathbf{R}^d}(K_2(x,y)-K_2(x,x_{B'}))\phi_{B'}(y)dy$$

where $x_{B'}$ is the centre of B'. But we can verify (by some tedious case checking) that

$$|K_2(x,y)-K_2(x,x_{B'})| \lesssim_{d,\theta} \langle y-x_{B'}/r\rangle^{100d}\frac{|x_{B'}-y|^\theta}{(r+|x-x_{B'}|)^{d+\theta}}$$

and thus

$$|T_2^t\phi_{B'}(x)| \lesssim_{d,\theta} \frac{r_{B'}^\theta}{(r+|x-x_{B'}|)^{d+\theta}}|B'|$$

which shows (with the usual decay bounds on ϕ_B) that the contribution of $\int_{\mathbf{R}^d}\phi_B T_2^t\phi_{B'}$ is acceptable. Similarly one can show that

$$|T_2^t(\phi_B\phi_{B'})(x)| \lesssim_{d,\theta} \frac{r_{B'}^\theta}{(r+|x-x_{B'}|)^{d+\theta}}|B'|\langle \mathrm{dist}(B,B')/r_B\rangle^{-100d}$$

and so the second integral $\int_{\mathbf{R}^d} T_2^t(\phi_B\phi_{B'})$ is also acceptable. ∎

Corollary 1.7. *Let the notations and hypotheses be as above. If $k<j-10$, then*

$$\|\psi_j(D)T\psi_k(D)f - \psi_j(D)((T1)(\psi_k(D)f))\|_{L^2(\mathbf{R}^d)} \lesssim_{d,\theta} 2^{-\theta(j-k)/2}\|f\|_{L^2(\mathbf{R}^d)}$$

for all $f\in L^2(\mathbf{R}^d)$.

Proof We can write

$$\psi_j(D)T\psi_k(D)f - \psi_j(D)((T1)(\psi_k(D)f))(x) = \int_{\mathbf{R}^d} K'_{jk}(x,y)f(y)\ dy$$

where

$$K'_{jk}(x,y) := \int_{\mathbf{R}^d}\int_{\mathbf{R}^d} \check\psi_j(x-z)T\mathrm{Trans}_y\check\psi_k(z) - \check\psi_j(x-z)\mathrm{Trans}_y\check\psi_k(z)T1(z)\ dz.$$

Applying the previous lemma we obtain

$$|K'_{jk}(x,y)| \lesssim 2^{-\theta(j-k)/2}2^{-dk}\langle 2^k|x-y|\rangle^{-d-\theta/2}$$

and the claim follows from Schur's test. ∎

Now we can prove the $T(1)$ theorem. By duality it suffices to show that

$$|\langle Tf, g\rangle| \lesssim_{d,\theta} \|f\|_{L^2(\mathbf{R}^d)}\|g\|_{L^2(\mathbf{R}^d)}$$

for all Schwartz f, g. Using (5) and the triangle inequality, it suffices to show that

$$|\sum_j \sum_k \langle \psi_j(D)\psi_j(D)T\psi_k(D)\psi_k(D)f, g\rangle| \lesssim_{d,\theta} \|f\|_{L^2(\mathbf{R}^d)}\|g\|_{L^2(\mathbf{R}^d)}.$$

Let us first deal with the diagonal case $j = k + O(1)$. Here we use Lemma 1.5 and Cauchy-Schwarz (throwing one factor of $\psi_j(D)$ to the other side) to bound this contribution by

$$\lesssim_{d,\theta} \sum_{j,k: j=k+O(1)} \|\psi_j(D)g\|_{L^2(\mathbf{R}^d)}\|\psi_k(D)f\|_{L^2(\mathbf{R}^d)}.$$

By Schur's test this is

$$\lesssim_{d,\theta} (\sum_j \|\psi_j(D)g\|_{L^2(\mathbf{R}^d)}^2)^{1/2}(\sum_k \|\psi_k(D)f\|_{L^2(\mathbf{R}^d)}^2)^{1/2}$$

which is acceptable by the Littlewood-Paley inequality (or orthogonality).

Now we need to deal with off-diagonal cases. By symmetry it suffices to deal with the contributions when $k < j - 10$. We split this contribution into

$$|\sum_{j,k:k<j-10} \langle \psi_j(D)\psi_j(D)[(\psi_k(D)\psi_k(D)f)T1], g\rangle|$$
$$+ |\sum_{j,k:k<j-10} \langle \psi_j(D)(\psi_j(D)T\psi_k(D)\psi_k(D)f - \psi_j(D)[(\psi_k(D)\psi_k(D)f)T1]), g\rangle|.$$

Let's first deal with the error term. Applying Corollary 1.7 and Cauchy-Schwarz, we can bound this by

$$\lesssim_{d,\theta} \sum_{j,k:k<j-10} 2^{-\theta(j-k)/2}\|\psi_k(D)f\|_{L^2(\mathbf{R}^d)}\|\psi_j(D)f\|_{L^2(\mathbf{R}^d)}$$

which is acceptable by Schur's test as before. As for the main term, we can rewrite this as

$$|\langle \pi_{lh}(f, T1), g\rangle|$$

for some low-high paraproduct π_{lh}. But since $\|T1\|_{\mathrm{BMO}} = O_{d,\theta}(1)$, we see that

$$\|\pi_{lh}(f, T1)\|_{L^2} \lesssim_{d,\theta} \|f\|_{L^2}$$

and the claim follows by Cauchy-Schwarz. (Note that the hypothesis that T^*1 is bounded in BMO will be similarly used to deal with the opposite case $k > j + 10$.)

2. Sample application: the Cauchy integral

Just to illustrate the $T(1)$ theorem, we use it to analyse the Cauchy integral for Lipschitz curves with small Lipschitz constant. (One can also handle the case of large Lipschitz constant with substantially more effort; but in that case it is actually better to replace the $T(1)$ theorem by a more general theorem, the $T(b)$ theorem, which we will not discuss here. Suffice to say that sometimes 1 is not the most convenient function to test T or T^* on.)

We begin with some (rather informal) motivation from complex analysis; further discussion can be obtained for instance in Michael Christ's lectures on singular integral operators.

Let Γ be a compact subset of the complex plane $\mathbf{C}$ (typically Γ will be a segment of a curve, or something similar), and let μ be a positive Radon measure on Γ. Given any $f \in L^2(\mu)$, we can define the *Cauchy integral* $C_\mu(f)$ on $\mathbf{C}\backslash\Gamma$ by

$$C_\mu(f)(z) := \frac{1}{2\pi i}\int_\Gamma \frac{f(w)}{z-w}\, d\mu(w).$$

This is complex analytic outside of Γ. One can also (under reasonable conditions on Γ, f, μ) define the Cauchy integral on Γ itself by a subtly different formula:

$$C_\mu(f)(z) := \frac{1}{2\pi i} p.v. \int_\Gamma \frac{f(w)}{z-w}\, d\mu(w) := \lim_{\varepsilon\to 0}\frac{1}{2\pi i}\int_{\Gamma; |w-z|>\varepsilon} \frac{f(w)}{z-w}\, d\mu(w).$$

It is usually not the case that the value of $C_\mu(f)$ outside of Γ converges pointwise to the value of $C_\mu(f)$ on Γ; recall for instance the Plemelj formulae from previous notes. However, they are certainly related. For instance, if f and $C_\mu(f)$ are both bounded on Γ, one can show (again assuming reasonable hypotheses on Γ) that $C_\mu(f)$ is then also bounded on $\mathbf{C}\backslash\Gamma$.

In this section we shall consider an important special case, when Γ is a Lipschitz curve

$$\Gamma := \{x + iA(x) : x \in I\}$$

where I is a compact interval and $A : I \to \mathbf{R}$ is a Lipschitz function, and μ is the pushforward of Lebesgue measure dx by the map $x \mapsto x + iA(x)$. There are many reasons why it is natural to impose the Lipschitz condition, but one of them is that the Cauchy integral C_μ on Γ is conjugate (using the map $x \mapsto x + iA(x)$, which is invertible on L^2) to the operator

$$\tilde{C}_\mu f(x) := p.v. \frac{1}{2\pi i}\int_I \frac{f(y)}{x-y+i(A(x)-A(y))}\, dy$$

on $L^2(I)$. Notice that the Lipschitz condition on A ensures that kernel $\frac{1}{2\pi i}\frac{1}{x-y+i(A(x)-A(y))}$ is a singular kernel (ignoring for now the technicality that the domain is only I instead of a Euclidean space such as $\mathbf{R}$; one can introduce Calderón-Zygmund theory on much more general domains than $\mathbf{R}^d$, but we will not pursue this topic here.). A fundamental theorem here, first conjectured by Calderón and then proven in full generality by Coifman, McIntosh, and Meyer, is

Theorem 2.1 (Cauchy integral is bounded on Lipschitz curves)**.** *Let $A : I \to \mathbf{R}$ be Lipschitz. Then C_μ is bounded on $L^2(\Gamma)$ (or equivalently, $\tilde{C}_\mu$ is bounded on $L^2(I)$), with constant depending only on the Lipschitz constant of A.*

We will not fully prove this celebrated theorem here, but will show a special case of it, at least, can be obtained from the $T(1)$ theorem.

The set Γ is said to have *non-zero analytic capacity* if it is possible to have a non-constant bounded analytic function on $\mathbf{C}\backslash\Gamma$. The full study of analytic capacity is well beyond the scope of this course. But from the preceding discussion, we see that one necessary condition for having non-zero analytic capacity (again assuming some reasonable hypotheses on Γ) is that there exists a measure μ supported on Γ non-trivial $f \in L^\infty(\mu)$ for which $C_\mu(f)$ also lies in $L^\infty(\mu)$. (Actually, one does not even need f to be bounded; it is enough for f to be in $L^1(\mu)$, as long as all truncated versions of $C_\mu(f)$ are uniformly in $L^\infty(\mu)$.)

At first glance, this appears difficult to achieve, since the Cauchy integral is a singular integral operator and thus highly unlikely to map $L^\infty(\mu)$ to itself; dually, it is unlikely to map $L^1(\mu)$ to itself. However, we only need to find *one* non-trivial $L^\infty(\mu)$ function whose image is also in $L^\infty(\mu)$. Curiously, such a fact is implied by a dual statement, namely that of being of *weak-type* $(1,1)$. For technical reasons (having to do with the bad duality properties of $L^\infty(\mu)$) we will have to work instead with the space of continuous functions $C(\Gamma)$, together with its dual $C(\Gamma)^*$, the space of finite Radon measures. More specifically:

Lemma 2.2 (Weak $(1,1)$ and duality)**.** *Let Γ be a locally compact Hausdorff space with a non-negative finite Radon measure μ which is not identically zero, and let $T : C(\Gamma)^* \to C(\Gamma)$ be a linear operator whose adjoint $T^* : C(\Gamma)^* \to C(\Gamma)$ obeys the weak-type $(1,1)$ bound*

$$\|T^*\nu\|_{L^{1,\infty}(\mu)} < K\|\nu\|_{C(\Gamma)^*}$$

for some $K > 0$. Then there exists $f \in C(\Gamma)$ with the pointwise bound $0 \leq f \leq 1$ and $\int_\Gamma f\ d\mu \geq \mu(\Gamma)/2$ (so in particular f is non-trivial) such that

$$\|Tf\|_{C(\Gamma)} \leq 2K.$$

Proof We normalise $K = 1/2$ and $\mu(\Gamma) = 1$. Suppose for contradiction that the claim failed, then we see that the convex set

$$\{Tf : 0 \leq f \leq 1; \int_\Gamma f\ d\mu \geq 1/2\}$$

avoids the unit ball of $C(\Gamma)$. Applying the Hahn-Banach theorem, there thus exists ν with $\|\nu\|_{C(\Gamma)^*} = 1$ such that

$$|\int Tf\ d\nu| \geq 1$$

for all $0 \leq f \leq 1$ with $\int_\Gamma f\ d\mu \geq 1/2$. On the other hand, from the weak-type bound we know that

$$\mu(\{|T^*\nu| \geq 1\}) < \frac{1}{2}$$

so if one sets $f := 1_{|T^*\nu|<1}$ then $\int_\Gamma f\ d\mu \geq 1/2$ and

$$|\int Tf\ d\mu| = \|\int T^*\nu \overline{f}| < 1,$$

a contradiction. ■

Ignoring some moderate technicalities (such as distinguishing the Cauchy integral on Γ with the Cauchy integral away from Γ, and regularising the Cauchy integral so that it does indeed map measures to continuous functions), we thus see that to establish non-zero analytic capacity on a Lipschitz curve Γ it suffices to show that the Cauchy integral operator is of weak-type $(1,1)$. Since this operator has (after conjugating by $x \mapsto x+iA(x)$) a singular kernel, Calderón-Zygmund theory reduces matters to establishing L^2 boundedness. Thus, assuming Theorem 2.1, we obtain

Corollary 2.3. *Any set Γ which contains a Lipschitz curve has non-zero analytic capacity.*

In fact a slight refinement of the argument shows that any set which contains a rectifiable set of positive length has non-zero analytic capacity (settling an old conjecture of Denjoy). The theory of analytic capacity has since been developed substantially, but this is well beyond the scope of these notes.

It remains to establish Theorem 2.1. We first consider the perturbative case when the Lipschitz constant $\|A\|_{\mathrm{Lip}}$ is small; this case was established by Calderón. Here the idea is to simply perform a Taylor expansion of the kernel:

$$\frac{1}{2\pi i}\frac{1}{x-y+i(A(x)-A(y))} = \sum_{j=0}^{\infty}\frac{1}{2\pi i}\frac{(-i(A(x)-A(y)))^j}{(x-y)^{j+1}}.$$

This leads to the study of the *Calderón commutators*

$$T_j f(x) := p.v. \int_I \frac{(A(x)-A(y))^j}{(x-y)^{j+1}} f(y)\ dy.$$

Note that the kernels here are all singular kernels (with constants $O(C^j\|A\|_{\mathrm{Lip}}^j)$ for some absolute constant j). These commutators are so named because they are connected to commutators of the Hilbert transform H with the operation of multiplication by A; for instance, up to constants, T_1 is (formally) equal to $\frac{d}{dx}[H,A]+A'H$, and so the L^2 boundedness of T_1 is equivalent to the fact that $[H,A]$ maps $L^2(\mathbf{R})$ to $W^{1,2}(\mathbf{R})$.

In order to prove Theorem 2.1 in the small Lipschitz constant case, it will (formally) suffice by the triangle inequality to establish an exponential bound of the form

$$\|T_j f\|_{L^2(I)} \lesssim C^{j+1}\|A\|_{\mathrm{Lip}}^j\|f\|_{L^2(I)}$$

for all Schwartz f, all $j \geq 0$ and for some absolute constant C independent of A, I, f. (One can make the summation rigorous by various qualitative tricks such as regularisation; we won't dwell on that here.) Note that by homogeneity we can now normalise $\|A\|_{\mathrm{Lip}} = 1$.

The operator T_0 is (up to constants) just the Hilbert transform and is thus bounded on L^2 (with operator norm $O(1)$). As for T_1, we can use the $T(1)$ theorem. Firstly we can extend the interval I to all of $\mathbf{R}$ (by extending A in some arbitrary Lipschitz manner). Strictly speaking, T_1 is not given by an integral kernel, but this can be achieved by a standard regularisation and limiting argument which we omit. Since the kernel is a singular kernel, it suffices to show that $T_1 1$ and $T_1^* 1$ lie in $\mathrm{BMO}(\mathbf{R})$ with a norm of $O(1)$. Since T_1 is an anti-symmetric operator it just suffices to check

$T_1 1$. But we (formally) have[1]

$$\begin{aligned} T_1 1(x) &= \int_{\mathbf{R}} \frac{A(x)-A(y)}{(x-y)^2}\ dx \\ &= \int_{\mathbf{R}} \int_{x<z<y} A'(z)(x-y)^{-2}\ dzdx \\ &= \int_{\mathbf{R}} A'(z)/(z-y)\ dz \\ &= CHA'(x) \end{aligned}$$

for some absolute constant C. (One needs to take a bit more care with the principal value integration to make this rigorous.) Since A is Lipschitz, A' is bounded and so HA' is in BMO as desired.

The higher commutators T_j can be handled inductively. Again, the main task is to show that $T_j 1$ lies in BMO. This can be achieved by a formal identity of the form

$$T_j 1 = T_{j-1} A' \tag{7}$$

which generalises the previous identity. If T_{j-1} was already shown to be a CZO, then it maps the bounded function A' to BMO, and hence $T_j 1$ is in BMO. From this and induction one can then show that all the T_j are CZOs (with bounds growing exponentially in j.)

In the previous discussion we tried to control the Cauchy integral operator C_μ by a power series expansion, which by its nature only is going to work in a perturbative regime. One can instead look for arguments that tackle the operator directly, so that they have a better chance of working non-perturbatively. One starting point is the identity

$$p.v. \frac{2\pi i}{\int}_{\Gamma} \frac{1}{z-w}\ dw = 0$$

which (formally) follows from the Cauchy integral formula (at least if Γ is smooth and flat at infinity). In terms of the conjugated operator $\tilde{C}_\mu$, this is essentially asserting that

$$\tilde{C}_\mu(1+iA') = 0$$

or equivalently that

$$\tilde{C}_\mu(1) = -i\tilde{C}_\mu(A').$$

This identity is closely related to (7), indeed one can view (7) as the j^{th} term of the multilinear expansion of the above identity in A. Now formally, since CZOs map L^∞ to BMO we have

$$\|\tilde{C}_\mu(A')\|_{\mathrm{BMO}} \lesssim \|\tilde{C}_\mu\|_{CZO} \|A\|_{\mathrm{Lip}}$$

while the $T(1)$ theorem (in a quantitative form) shows that

$$\|\tilde{C}_\mu\|_{CZO} \lesssim 1 + \|\tilde{C}_\mu(1)\|_{\mathrm{BMO}}.$$

Combining all these together, we obtain (formally at least) another proof of the boundedness of the Cauchy integral when the Lipschitz constant is small.

---

[1]This is related to the identity $T_1 = C(\frac{d}{dx}[H,A] + A'H)$ mentioned earlier; indeed, on taking adjoints, we get $T_1 = -T_1^* = C([H,A]\frac{d}{dx} + HA')$, since H and $\frac{d}{dx}$ are skew-adjoint and A' is self-adjoint.

To deal with the large Lipschitz constant case there are a number of possible strategies. One strategy, introduced by Peter Jones, is to use quantitative versions of the Radamacher differentiation theorem, which asserts that Lipschitz functions are almost everywhere differentiable; another way of thinking about this is that curves with large Lipschitz constant can usually be approximated by (affine images of) curves with small Lipschitz constant. By introducing a suitable stopping time decomposition to make this quantitative, one can decompose the majority of a Cauchy integral operator for large Lipschitz constant curves into pieces made up of Cauchy integral operators for small Lipschitz constant curves.

The more modern way to incorporate this strategy is to remove the focus on the function 1 that $\tilde{C}_\mu$ is being applied to, and instead replace it with a more general function b; in this specific case, the most natural function to use is $b := 1+iA'$, since we have $\tilde{C}_\mu b = 0$. This led to the development of $T(b)$ *theorems*, in which hypotheses such as $Tb_1 \in \mathrm{BMO}$ and $T^* b_2 \in \mathrm{BMO}$ were used to deduce L^2 boundedness. Of course, some non-degeneracy hypotheses have to be placed on b_1, b_2 in order for this to work (if $b_1 = b_2 = 0$ for instance then the hypotheses are trivial and thus useless). In the case of $b = 1 + iA'$, the relevant property is that the real part of b is bounded away from zero, a property known as *accretivity* for reasons arising from PDE. More recently, more "local" analogues of accretivity (pseudoaccretivity, para-accretivity, etc.) have been used as substitutes. The proof of the $T(b)$ theorem is most easily accomplished by using Haar-type wavelets adapted to the chosen functions b_1, b_2, rather than (or in conjunction with) the Littlewood-Paley approach presented here; but discussing this in detail is beyond the scope of these notes.

3. Exercises

- Q1. Let K be a singular kernel which is also bounded and compactly supported, and let T be the integral operator (4). Suppose that we have the estimates
$$\int_B |T1_B| \lesssim_d |B|$$
and
$$\int_B |T^* 1_B| \lesssim_d |B|$$
for all balls B. Show that $\|T\|_{L^2(\mathbf{R}^d) \to L^2(\mathbf{R}^d)} \lesssim_{d,\theta} 1$.
- Q2. (Calderón commutator estimate) Let $a \in \mathcal{S}(\mathbf{R})$, and let H be the Hilbert transform. Show that the following are equivalent up to changes of constant:
  - $\|a\|_{\mathrm{BMO}(\mathbf{R})} \lesssim 1$.
  - For all $1 < p < \infty$ and $f \in L^p(\mathbf{R})$, we have
$$\|[a,H]f\|_{L^p(\mathbf{R})} \lesssim_p \|f\|_{L^p(\mathbf{R})}$$
    where $[a,H]f = aHf - H(af)$.

  (Hints: to deduce the latter from the former, one can try to rewrite $[a,H]f$ in terms of various paraproducts and Hilbert transforms in such a way that a is always measured as a high frequency, so that BMO type paraproduct estimates can be exploited. To obtain the reverse implication, look at what

$\langle [a, H]f, g\rangle$ behaves like when f, g are supported on nearby intervals of the same size. It may be better for you to try to prove an estimate of the form $\|a\|_{\mathrm{BMO}(\mathbf{R})} \leq C + \frac{1}{2}\|a\|_{\mathrm{BMO}(\mathbf{R})}$.)

- Q3. Working formally (ignoring issues of convergence, principal value integration, etc.) derive (7) for all $j \geq 1$.

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

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