WEAK TYPE ESTIMATES AND COTLAR INEQUALITIES FOR CALDERÓN-ZYGMUND OPERATORS ON NONHOMOGENEOUS SPACES

F. NAZAROV, S. TREIL, AND A. VOLBERG

0. Introduction (historical remarks).

The classical theory of Calderón–Zygmund operators started with the study of convolution operators on the real line with singular kernels (a typical example of such an operator is the so called Hilbert transform, defined by $Hf(t) = \int_{\mathbb{R}} \frac{f(s)\,ds}{t-s}$). Later it has developed into a large branch of analysis covering a quite wide class of singular integral operators on abstract measure spaces (so called "spaces of homogeneous type"). To see how far the theory has evolved during the last 30 years, it is enough to compare the classical textbook [St1] by Stein published in 1970 (which still remains an excellent introduction into the subject) to the modern outline of the theory in [DJ], [St2], [Ch2], and [CW].

The only thing that has remained unchallenged until very recently was the doubling property of the measure, i.e., the assumption that for some constant $C > 0$,

$$\mu(B(x, 2r)) \leqslant C\mu(B(x, r)) \qquad \text{for every } x \in \mathcal{X},\ r > 0,$$

where $\mathcal{X}$ is some metric space endowed with a Borel measure μ, and, as usual, $B(x,r) = \{y \in \mathcal{X} \,:\, \operatorname{dist}(x,y) \leqslant r\}$ is the closed ball of radius r centered at x.

The main result we want to present to the reader can be now described in one short sentence:

The doubling condition is superfluous for the most part of the classical theory.

The reader may ask: "Why should one try to eliminate the doubling condition at all?"

The simplest example where such a necessity arises is just a standard singular integral operator considered in an open domain $\Omega \subset \mathbb{R}^n$ (with the usual n-dimensional Lebesgue measure) instead of the whole space, or on a surface S (say, 2-dimensional surface in $\mathbb{R}^3$) with respect to the usual surface area measure. If the boundary of Ω is nice, or if S is a Lipshitz surface, we get a space of homogeneous type and everything is well understood. But for domains with "wild" boundaries, the doubling property for the Lebesgue measure fails and the results for spaces of homogeneous type can no longer be directly applied to them. Similarly, a few spikes on the two-dimensional surface S often do not spoil the upper bound $\mu(B(x,r)) \leqslant \operatorname{Const} r^2$, but they can easily ruin the doubling condition.

Such singular integral operators are sometimes claimed to "appear naturally in the study of PDE". We will abstain from any comment on this issue, but the problem seems to be very natural indeed, and definitely is of its own interest. As far as we know, in

F. Nazarov is partially supported by NSF grant DMS-9706775.

S. Treil and A. Volberg are partially supported by NSF grant DMS-9622936.

both cases no satisfactory theory of Calderón–Zygmund operators has been previously developed.

Another example, which actually was the main motivation for our work, concerns the action of the Cauchy integral operator on the complex plane. The problem here is the following:

> *Given a finite Borel measure μ on the complex plane $\mathbb{C}$, figure out whether the Cauchy integral operator*
>
> $$\mathcal{C}f(x) = \mathcal{C}_\mu f(x) = \int_{\mathbb{C}} \frac{f(y)\, d\mu(y)}{x-y}$$
>
> *acts in $L^2(\mu)$ (in $L^p(\mu)$, from $L^1(\mu)$ to $L^{1,\infty}(\mu)$, and so on).*

The partial case of this question, when μ is the one-dimensional Hausdorff measure on some strange compact set Γ on the plane, has for a long time been one of the central problems in the study of analytic capacity (see [DM], [D], [MMV], [Ch1] and [Ch2]). If Γ is a Lipschitz curve or something similar, then we have a space of homogeneous type and one can apply most of the classical techniques with suitable modifications (see [Ch2]).

But in general the measure does not satisfy the doubling condition and, until recently, one had to look for an alternative approach. One possible way to go around the difficulty was proposed by Melnikov and Verdera. They noted that the kernel $\frac{1}{x-y}$ of the Cauchy integral operator satisfies the following beautiful identity:

$$\sum_{\sigma \in S_3} \frac{1}{(x_{\sigma(1)} - x_{\sigma(2)})\,\overline{(x_{\sigma(1)} - x_{\sigma(3)})}} = \frac{1}{R(x_1, x_2, x_3)^2},$$

where S_3 is the permutation group of order 3, as usual, and $R(x, y, z)$ is the radius of the circumscribed circumference of the triangle with vertices $x, y, z \in \mathbb{C}$. This observation allowed them to reduce the investigation of the *oscillatory* kernel $\frac{1}{x-y}$ to the study of the *non-negative* kernel $\frac{1}{R(x,y,z)^2}$. Later, Tolsa, developing their ideas, has become the first to construct a satisfactory theory of the Cauchy integral operator on the complex plane with respect to an arbitrary "one-dimensional" (see the definition below) measure μ. He went as far as to prove the existence of the principal value of the improper integral $\mathcal{C}f(x) = \int_{\mathbb{C}} \frac{f(y)\, d\mu(y)}{x-y}$ μ-almost everywhere.

On the other hand, for the general theory of Calderón-Zygmund operators, such an approach was a mere disaster: the Cauchy integral operator, which had always been one of the *most natural and important* examples of Calderón-Zygmund operators, was thus *completely excluded* from the general framework.

The present paper can be considered a complement to [NTV1]. There we dealt with the L^2-part of the theory, and the main result was, roughly speaking, that the Cauchy integral operator is bounded in $L^2(\mu)$ if and only if it is bounded on the characteristic functions of squares, which is equivalent to the celebrated Melnikov-Verdera curvature condition:

$$\iiint\limits_{Q^3} \frac{d\mu(x)\, d\mu(y)\, d\mu(z)}{R(x,y,z)^2} \leqslant \operatorname{Const} \mu(Q) \qquad \text{for each square } Q \subset \mathbb{C}.$$

The main difference between [NTV1] and the corresponding earlier work by Tolsa [T1] was that the proofs in [NTV1] remained valid for a quite wide class of Calderón-Zygmund integral operators. So, in a sense, [NTV1] could be viewed as the first approximation to the *general L^2-theory* of Calderón-Zygmund integral operators in non-homogeneous spaces.

Below we, conversely, are going to deal only with the part of the theory concerning the boundedness of a Calderón–Zygmund operator T and the associated maximal operator $T^\sharp$ (see the definition below) in the L^p-spaces ($1 < p < \infty$) and from $L^1(\mu)$ to $L^{1,\infty}(\mu)$ under the *a priori* assumption that T is bounded in $L^2(\mu)$. The main problem here is that one of the basic and most frequently used tools of the classical theory — the Calderón-Zygmund decomposition — fails to work in the non-doubling setting and should be given up completely. The most important question we are going to answer in the present paper is "What can one replace it with?"

Acknowledgements: Most of the results in this note arose from the visit of the first author to Paris, France in Summer 1997. We are very grateful to the University Cergy-Pontoise and to the entire French Analysis Community for the invitation and for their generous financial support of the visit. Our special thanks go to Francoise Lust-Piquard for her organizing the trip and fruitful discussions and to Guy David for his brilliant ideas and for his willingness to share them with other people.

1. Some definitions and the formulation of the main result.

Fix $n > 0$ (not necessarily an integer). Let $\mathcal{X}$ be a separable metric space endowed with a non-negative "n-dimensional" Borel measure μ, i.e., a measure satisfying

$$\mu(B(x,r)) \leqslant r^n \qquad \text{for all } x \in \mathcal{X},\ r > 0.$$

Let $L^p(\mu)$, $1 \leqslant p \leqslant \infty$ be the usual Lebesgue spaces, and let $L^{1,\infty}(\mu)$ be defined by

$$L^{1,\infty}(\mu) := \big\{ f : \mathcal{X} \to \mathbb{C} \,:\, \|f\|_{L^{1,\infty}(\mu)} := \sup_{t>0} t \cdot \mu\{x \in \mathcal{X} \,:\, |f(x)| > t\} < +\infty \big\}.$$

Note that the "norm" $\|f\|_{L^{1,\infty}(\mu)}$ is not actually a norm in the sense that it does not satisfy the triangle inequality. Still, we have

$$\|cf\|_{L^{1,\infty}(\mu)} = |c| \cdot \|f\|_{L^{1,\infty}(\mu)} \quad \text{ and } \quad \|f+g\|_{L^{1,\infty}(\mu)} \leqslant 2\big(\|f\|_{L^{1,\infty}(\mu)} + \|g\|_{L^{1,\infty}(\mu)}\big)$$

for every $c \in \mathbb{C}$, $f, g \in L^{1,\infty}(\mu)$. (The latter is just the observation that in order to have the sum greater than t, one should have at least one term greater than $t/2$).

Let $M(\mathcal{X})$ be the space of all complex-valued Borel measures on $\mathcal{X}$. We will denote by $\|\nu\|$ the total variation of the measure $\nu \in M(\mathcal{X})$.

For $f \in L^p(\mu)$, we will denote by $\operatorname{supp} f$ the essential support of the function f, i.e., the smallest closed set $F \subset \mathcal{X}$ for which f vanishes μ-almost everywhere outside F. Also, for $\nu \in M(\mathcal{X})$, we will denote by $\operatorname{supp}\nu$ the smallest closed set $F \subset \mathcal{X}$ for which ν vanishes on $\mathcal{X} \setminus F$ (i.e., $\nu(E) = 0$ for every Borel set $E \subset \mathcal{X} \setminus F$).

Since $\mathcal{X}$ is a separable metric space, such smallest closed set always exists. If $\{\mathcal{B}_j\}_{j=1}^\infty$ is some countable base of topology in $\mathcal{X}$, then for $\nu \in M(\mathcal{X})$, the support $\operatorname{supp}\nu$ is just the complement of the union of those $\mathcal{B}_j$, on which the measure ν vanishes. For a function $f \in L^p(\mu)$, we obviously have $\operatorname{supp} f = \operatorname{supp}\nu$ where $d\nu = |f|^p\, d\mu$.

Let $K : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ be a classical "n-dimensional" Calderón–Zygmund kernel on $\mathcal{X}$, i.e., for some $A > 0$, $\varepsilon \in (0, 1]$,

1. $$|K(x,y)| \leqslant \frac{A}{\operatorname{dist}(x,y)^n},$$

2. $$|K(x,y)-K(x',y)|,\ |K(y,x)-K(y,x')| \leqslant A\cdot\frac{\operatorname{dist}(x,x')^\varepsilon}{\operatorname{dist}(x,y)^{n+\varepsilon}}$$

   whenever $x,x',y\in\mathcal{X}$ and $\operatorname{dist}(x,x')\leqslant\frac{1}{2}\operatorname{dist}(x,y)$.

Remark. We want to attract the attention of the reader to the fact that, though we call the number n "dimension" all the time, it is the dimension of the *measure* μ and of the *kernel* $K(x,y)$, but *by no means* is it the topological (or metric, or whatever else) dimension of the *space* $\mathcal{X}$. For instance, for the case of the Cauchy integral operator on the complex plane, $n=1$, not 2! Actually, the topological dimension of the space $\mathcal{X}$ may be even infinite — we do not care.

Definition. A bounded linear operator T in $L^2(\mu)$ is called a Calderón-Zygmund (integral) operator with the Calderón–Zygmund kernel K if for every $f\in L^2(\mu)$,

$$Tf(x)=\int_{\mathcal{X}}K(x,y)f(y)d\mu(y)$$

for μ-almost every $x\in\mathcal{X}\setminus\operatorname{supp}f$.

Obviously, the adjoint operator T^* is also bounded in $L^2(\mu)$ and has the kernel $K^*(x,y)=\overline{K(y,x)}$, which is a Calderón–Zygmund kernel as well.

Let $\nu\in M(\mathcal{X})$ and $x\in\mathcal{X}\setminus\operatorname{supp}\nu$. For technical reasons it will be convenient to put *by definition*

$$T\nu(x):=\int_{\mathcal{X}}K(x,y)\,d\nu(y).$$

Note that we *do not* attempt here to define the values $T\nu(x)$ for $x\in\operatorname{supp}\nu$.

The maximal operator $T^\sharp$ associated with the Calderón–Zygmund operator T is defined as follows. For every $r>0$, put

$$T_rf(x):=\int_{\mathcal{X}\setminus B(x,r)}K(x,y)f(y)\,d\mu(y)$$

for $f\in L^p(\mu)$, and

$$T_r\nu(x):=\int_{\mathcal{X}\setminus B(x,r)}K(x,y)\,d\nu(y)$$

for $\nu\in M(\mathcal{X})$.

Define

$$T^\sharp f(x):=\sup_{r>0}|T_rf(x)|$$

for $f\in L^p(\mu)$, and

$$T^\sharp\nu(x):=\sup_{r>0}|T_r\nu(x)|$$

for $\nu\in M(\mathcal{X})$. Now we are able to formulate the main result of this paper:

Theorem 1.1. *For any (bounded in $L^2(\mu)$) Calderón-Zygmund operator T, the following statements hold:*

1. **L^p-action:** *For every $p\in(1,+\infty)$, the operator T is bounded in $L^p(\mu)$ in the sense that for every $f\in L^p(\mu)\cap L^2(\mu)$,*

$$\|Tf\|_{L^p(\mu)}\leqslant C\|f\|_{L^p(\mu)}$$

with some constant $C > 0$ not depending on f.

2. **Weak type 1-1 estimate:** *The operator T is bounded from $L^1(\mu)$ to $L^{1,\infty}(\mu)$ in the sense that for every $f \in L^1(\mu) \cap L^2(\mu)$,*

$$\|Tf\|_{L^{1,\infty}(\mu)} \leqslant C\|f\|_{L^1(\mu)}$$

with some constant $C > 0$ not depending on f.

3. **Action of the maximal operator in $L^p(\mu)$:** *For every $p \in (1, +\infty)$, the operator $T^\sharp$ is bounded in $L^p(\mu)$ in the sense that for every $f \in L^p(\mu)$,*

$$\|T^\sharp f\|_{L^p(\mu)} \leqslant C\|f\|_{L^p(\mu)}$$

with some constant $C > 0$ not depending on f.

4. **Weak type 1-1 estimate for the maximal operator:** *The operator $T^\sharp$ is bounded from $M(\mathcal{X})$ to $L^{1,\infty}(\mu)$ in the sense that for every $\nu \in M(\mathcal{X})$,*

$$\|T^\sharp \nu\|_{L^{1,\infty}(\mu)} \leqslant C\|\nu\|$$

with some constant $C > 0$ not depending on ν.

Remark. The above theorem will remain true if we replace the a priori assumption $\|T\|_{L^2(\mu)\to L^2(\mu)} < +\infty$ in the definition of the Calderon-Zygmund operator T by the assumption that T is bounded in some other $L^{p_0}(\mu)$ with $1 < p_0 < \infty$ (all the proofs below do not use any special properties of the number 2). It is even possible (though much less trivial) to show that the assumption $\|T\|_{L^1(\mu)\to L^{1,\infty}(\mu)} < +\infty$ would suffice.

On the other hand, we should confess that we do not know of any example of a Calderon-Zygmund operator T for which to check its, say, $L^4(\mu)$-boundedness would be easier than to check its $L^2(\mu)$-boundedness.

2. The plan of the paper

For notational simplicity, we will restrict ourselves to the case of *real-valued* functions, measures, and kernels (to obtain the result for the complex-valued case, it is enough to consider the real and the imaginary parts separately). In Section 3 we shall outline some preliminary lemmas (all of them well-known) that will be used throughout the rest of the paper (sometimes even without an explicit reference). In Sections 4–5 we shall prove the weak type $1-1$ estimate for $T\nu$ where ν is a finite linear combination of unit point masses with non-negative coefficients (these two sections constitute the core of the whole article). In Section 6 we shall present a simple approximation scheme allowing to switch in the weak type $1-1$ estimate from such "elementary" measures to functions $f \in L^1(\mu) \cap L^2(\mu)$. The L^p-boundedness of T will then follow immediately via the standard interpolation and duality tricks. In Section 7 we shall prove a Cotlar type inequality for the maximal operator $T^\sharp$, which will allow us to establish its L^p-boundedness for $1 < p < +\infty$. At last, in Sections 8–9 we shall prove the boundedness of $T^\sharp$ from $M(\mathcal{X})$ to $L^{1,\infty}(\mu)$, thus finishing the story.

Our aim was to make the paper completely self-contained (up to a few well-known facts that could be found in any standard textbook), so we apologize in advance if the reader finds some sections too boring (this especially concerns Sections 3, 6 and 9).

3. Preliminary observations.

Recall that the Hardy-Littlewood maximal function $Mf(x)$ is defined (for Borel measurable functions f) by

$$Mf(x) := \sup_{r>0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f|\, d\mu.$$

Note that if $x \in \operatorname{supp}\mu$, then $\mu(B(x,r)) > 0$ for every $r > 0$ (otherwise one could drop a small open ball centered at x from the support of μ), so the definition makes sense μ-almost everywhere.

If the measure μ satisfies the doubling property, or if $\mathcal{X}$ has nice geometric structure (similar to that of $\mathbb{R}^N$), the Hardy-Littlewood maximal function operator is well-known to be bounded in all $L^p(\mu)$ with $1 < p \leqslant +\infty$ and from $L^1(\mu)$ to $L^{1,\infty}(\mu)$. But, unfortunately, for arbitrary separable metric space $\mathcal{X}$ and measure μ, the best one can say is that M is bounded in $L^\infty(\mu)$ (which is just the obvious observation that the integral does not exceed the essential supremum of the integrand times the measure of the domain of integration). Fortunately, to save the game is not too hard: all one needs is to replace the measure of the ball $B(x,r)$ in the denominator by the measure of the *three times larger* ball, i.e., to define

$$\widetilde{M}f(x) := \sup_{r>0} \frac{1}{\mu(B(x,3r))} \int_{B(x,r)} |f|\, d\mu.$$

Note that always $\widetilde{M}f(x) \leqslant Mf(x)$ and, if the measure μ satisfies the doubling condition, $Mf(x) \leqslant C \cdot \widetilde{M}f(x)$ for some constant $C > 0$ (the square of the constant in the doubling condition).

Lemma 3.1. *The modified maximal function operator $\widetilde{M}$ is bounded in $L^p(\mu)$ for each $p \in (1, +\infty]$ and acts from $L^1(\mu)$ to $L^{1,\infty}(\mu)$.*

Proof. The boundedness in $L^\infty(\mu)$ is obvious. To prove the weak type $1-1$ estimate, we will use the celebrated

Vitali covering theorem. *Fix some $R > 0$. Let $E \subset \mathcal{X}$ be any set and let $\{B(x,r_x)\}_{x\in E}$ be a family of balls of radii $0 < r_x < R$. Then there exists a countable subfamily $\{B(x_j,r_j)\}_{j=1}^\infty$ (where $x_j \in E$ and $r_j := r_{x_j}$) of disjoint balls such that $E \subset \cup_j B(x_j, 3r_j)$.*

We refer the reader for the proof of the Vitali covering theorem to his favorite textbook in geometric measure theory.

Now, to prove the lemma, fix some $t > 0$. Pick $R > 0$ and consider the set E of the points $x \in \operatorname{supp}\mu$ for which

$$\widetilde{M}^{(R)}f(x) := \sup_{0<r<R} \frac{1}{\mu(B(x,3r))} \int_{B(x,r)} |f|\, d\mu > t.$$

For every such x, there exists some radius $r_x \in (0,R)$ such that

$$\int_{B(x,r_x)} |f|\, d\mu > t\mu(B(x,3r_x)).$$

Choose the corresponding collection of pairwise disjoint balls $B(x_j, r_j)$. We have

$$\mu(E) \leqslant \sum_j \mu(B(x_j, 3r_j)) \leqslant \frac{1}{t} \sum_j \int_{B(x_j,r_j)} |f|\, d\mu \leqslant \frac{\|f\|_{L^1(\mu)}}{t}.$$

It remains only to note that $\widetilde{M}^{(R)} f \nearrow \widetilde{M} f$ as $R \to +\infty$.

The boundedness in $L^p(\mu)$ for $1 < p < +\infty$ follows now from the Marcinkiewicz interpolation theorem. □

Remark. Exactly the same proof allows to show that for every finite non-negative measure ν on $\mathcal{X}$, the function

$$\widetilde{M}\nu(x) := \sup_{r>0} \frac{\nu(B(x,r))}{\mu(B(x,3r))}$$

belongs to $L^{1,\infty}(\mu)$ and satisfies the estimate

$$\|\widetilde{M}\nu\|_{L^{1,\infty}(\mu)} \leqslant \nu(\mathcal{X}).$$

We shall also need the modification of the maximal function $\widetilde{M}f$, in which the averaging of $|f|$ over balls is done with some power $\beta \neq 1$. Namely, for each $\beta > 0$, put

$$\widetilde{M}_\beta f(x) := \big[\widetilde{M}(|f|^\beta)(x)\big]^{\frac{1}{\beta}} = \sup_{r>0} \Big[\frac{1}{\mu(B(x,3r))} \int_{B(x,r)} |f|^\beta\, d\mu\Big]^{\frac{1}{\beta}}.$$

Note that the greater β is, the greater $\widetilde{M}_\beta f(x)$ is (the Holder inequality). Note also that $\widetilde{M}_\beta$ is bounded in $L^p(\mu)$ for every $p \in (\beta, +\infty]$ (to say that $\widetilde{M}_\beta$ is bounded in $L^p(\mu)$ is exactly the same as to say that $\widetilde{M}$ is bounded in $L^{p/\beta}(\mu)$).

We shall however need one less trivial (though no less standard) observation:

Lemma 3.2. *For any $\beta \in (0,1)$, the maximal operator $\widetilde{M}_\beta$ is bounded in $L^{1,\infty}(\mu)$, i.e.,*

$$\|\widetilde{M}_\beta f\|_{L^{1,\infty}(\mu)} \leqslant C\|f\|_{L^{1,\infty}(\mu)}$$

with some constant $C > 0$ not depending on f.

Proof. Let $f \in L^{1,\infty}(\mu)$. Write $f = f_t + f^t$ where

$$f_t(x) = \begin{cases} f(x), & \text{if } |f(x)| \leqslant t; \\ 0, & \text{if } |f(x)| > t; \end{cases} \qquad \text{and} \qquad f^t(x) = \begin{cases} 0, & \text{if } |f(x)| \leqslant t; \\ f(x), & \text{if } |f(x)| > t. \end{cases}$$

Since $\|\widetilde{M}_\beta f_t\|_{L^\infty(\mu)} \leqslant \|f_t\|_{L^\infty(\mu)} \leqslant t$ and $[\widetilde{M}_\beta f]^\beta \leqslant [\widetilde{M}_\beta f_t]^\beta + [\widetilde{M}_\beta f^t]^\beta$ (additivity of integral), we have

$$\begin{aligned}\mu\{x \in \mathcal{X} : \widetilde{M}_\beta f > 2^{\frac{1}{\beta}} t\} &\leqslant \mu\{x \in \mathcal{X} : \widetilde{M}_\beta f^t > t\} \\ &= \mu\{x \in \mathcal{X} : \widetilde{M}|f^t|^\beta > t^\beta\} \leqslant t^{-\beta} \int_{\mathcal{X}} |f^t|^\beta\, d\mu\end{aligned}$$

according to the weak type $1-1$ estimate for $\widetilde{M}$.

On the other hand, we have

$$\int_{\mathcal{X}} |f^t|^\beta \, d\mu = t^\beta \mu\{|f| > t\} + \int_t^{+\infty} \beta s^{\beta-1} \mu\{|f| > s\} ds$$
$$\leqslant t^\beta \frac{1}{t} \|f\|_{L^{1,\infty}(\mu)} + \|f\|_{L^{1,\infty}(\mu)} \int_t^{+\infty} \beta s^{\beta-2} ds = \frac{1}{1-\beta} \frac{1}{t} t^\beta \|f\|_{L^{1,\infty}(\mu)}.$$

So, finally we get

$$\mu\{x \in \mathcal{X} : \widetilde{M}_\beta f > 2^{\frac{1}{\beta}} t\} \leqslant \frac{1}{1-\beta} \frac{1}{t} \|f\|_{L^{1,\infty}(\mu)},$$

i.e.,

$$\|\widetilde{M}_\beta f\|_{L^{1,\infty}} \leqslant \frac{2^{1/\beta}}{1-\beta} \|f\|_{L^{1,\infty}},$$

proving the lemma. □

3.1. **Comparison lemma.**

Lemma 3.3. *Let $U : (0, +\infty) \to [0, +\infty)$ be a continuous non-negative decreasing function. Let ν be any non-negative Borel measure on $\mathcal{X}$. Then for every $x \in \mathcal{X}$ and $R > 0$,*

$$\int_{\mathcal{X} \setminus B(x,R)} U(\operatorname{dist}(x,y)) \, d\nu(y) \leqslant 3^n \widetilde{M}\nu(x) \Big[R^n U(R) + n \int_R^{+\infty} t^{n-1} U(t) \, dt \Big].$$

Proof. Consider first the case when U is a "step-function", i.e., $U(t) = \chi_{(0,T]}$ for some $T > 0$ (as usual, by χ_E we denote the characteristic function of the set E). If $T \leqslant R$, the inequality is obvious because the left hand part is 0. For $T > R$, it is equivalent to the estimate

$$\nu(B(x,T) \setminus B(x,R)) \leqslant 3^n \widetilde{M}\nu(x) \cdot T^n,$$

which easily follows from the definition of $\widetilde{M}\nu(x)$ and the inequality $\mu(B(x,3T)) \leqslant 3^n \cdot T^n$.

Now to obtain the lemma, it is enough to recall that every non-negative continuous decreasing function $U(t)$ can be represented as the limit of an increasing sequence of linear combinations of step-functions with non-negative coefficients. □

3.2. **Hörmander inequality.**

We shall need one more standard observation about Calderón–Zygmund kernels.

Lemma 3.4. *Let $\eta \in M(\mathcal{X})$, $\eta(\mathcal{X}) = 0$, and $\operatorname{supp} \eta \subset B(x,\rho)$ for some $\rho > 0$. Then for every non-negative Borel measure ν on $\mathcal{X}$, we have*

$$\int_{\mathcal{X} \setminus B(x,2\rho)} |T\eta| d\nu \leqslant A_1 \, \widetilde{M}\nu(x) \, \|\eta\|$$

where $A_1 > 0$ depends only on the dimension n and the constants A and ε in the definition of the Calderón–Zygmund kernel K. In particular,

$$\int_{\mathcal{X} \setminus B(x,2\rho)} |T\eta| \cdot |f| d\mu \leqslant A_1 \, \widetilde{M}f(x) \, \|\eta\|$$

for every Borel measurable function f on $\mathcal{X}$, and

$$\int_{\mathcal{X}\setminus B(x,2\rho)} |T\eta| d\mu \leqslant A_1 \|\eta\|$$

Proof. For any $y \in \mathcal{X} \setminus B(x, 2\rho)$, we have

$$|T\eta(y)| = \Big|\int_{B(x,\rho)} K(y,x')d\eta(x')\Big| = \Big|\int_{B(x,\rho)} [K(y,x') - K(y,x)]d\eta(x')\Big| \leqslant$$
$$\leqslant \|\eta\| \sup_{x' \in B(x,\rho)} |K(y,x') - K(y,x)| \leqslant \|\eta\| \frac{A\rho^\varepsilon}{\operatorname{dist}(x,y)^{n+\varepsilon}}.$$

It remains only to notice that, according to the Comparison Lemma (Lemma 3.3) applied to $R = 2\rho$ and $U(t) = \dfrac{\rho^\varepsilon}{t^{n+\varepsilon}}$,

$$\int_{\mathcal{X}\setminus B(x,2\rho)} \frac{\rho^\varepsilon \, d\nu(y)}{\operatorname{dist}(x,y)^{1+\varepsilon}}$$
$$\leqslant 3^n \widetilde{M}\nu(x) \Big[(2\rho)^n \frac{\rho^\varepsilon}{(2\rho)^{n+\varepsilon}} + n\int_{2\rho}^{+\infty} t^{n-1}\frac{\rho^\varepsilon}{t^{n+\varepsilon}}\Big] dt = 3^n 2^{-\varepsilon}(1 + \tfrac{n}{\varepsilon})\widetilde{M}\nu(x).$$

□

4. The Guy David lemma.

The following lemma is implicitly contained in [D].

Lemma 4.1. *For any Borel set $F \in \mathcal{X}$ of finite measure and for any point $x \in \operatorname{supp}\mu$,*

$$T^\sharp \chi_F(x) \leqslant 2\cdot 3^n \, \widetilde{M}T\chi_F(x) + A_2$$

where $A_2 > 0$ depends only on the dimension n, the constants A and ε in the definition of the Calderón–Zygmund kernel K, and the norm $\|T\|_{L^2(\mu)\to L^2(\mu)}$.

Proof. Let $r > 0$. Consider the sequence of balls $B(x, r_j)$ where $r_j := 3^j r$, and the corresponding sequence of measures $\mu_j := \mu(B(x, r_j))$ $(j = 0, 1, \dots)$.

Note that we cannot have $\mu_j > 2 \cdot 3^n \mu_{j-1}$ for every $j \geqslant 1$. Indeed, otherwise we would have for every $j = 1, 2, \dots,$

$$\mu(B(x,r)) = \mu_0 \leqslant [2\cdot 3^n]^{-j}\mu_j \leqslant [2\cdot 3^n]^{-j} r_j^n = 2^{-j} r^n.$$

Since the right hand part tends to 0 as $j \to +\infty$, we could conclude from here that $\mu(B(x, r)) = 0$, which is impossible.

Therefore there exists the smallest positive integer k for which $\mu_k \leqslant 2 \cdot 3^n \mu_{k-1}$. Put $R := r_{k-1} = 3^{k-1} r$. Observe that

$$|T_r\chi_F(x) - T_{3R}\chi_F(x)| \leqslant \int_{B(x,3R)\setminus B(x,r)} |K(x,y)|\, d\mu(y)$$
$$= \sum_{j=1}^{k} \int_{B(x,r_j)\setminus B(x,r_{j-1})} |K(x,y)|\, d\mu(y) =: \sum_{j=1}^{k} \mathcal{I}_j.$$

Now recall that $|K(x,y)| \leqslant \frac{A}{\mathrm{dist}(x,y)^n}$ and therefore $\mathcal{I}_j \leqslant A\frac{\mu_j}{r_{j-1}^n}$ for every $j = 1, \dots, k$. Note that $\mu_j \leqslant [2\cdot 3^n]^{(j+1-k)}\mu_{k-1}$ and $r_{j-1} = 3^{j-k}r_{k-1}$ for $j = 1, \dots, k$. Hence

$$\sum_{j=1}^{k} \mathcal{I}_j \leqslant A\sum_{j=1}^{k}\frac{\mu_j}{r_{j-1}^n} \leqslant A\cdot 2\cdot 3^n\,\frac{\mu_{k-1}}{r_{k-1}^n}\sum_{j=1}^{k} 2^{j-k} \leqslant 4\cdot 3^n A$$

(for $\mu_{k-1} = \mu(B(x, r_{k-1})) \leqslant r_{k-1}^n$).

And that is basically the main part of the reasoning, because now it is enough to pick any standard proof based on the doubling condition to get the desired estimate for $T_{3R}\chi_F(x)$ (recall that $\mu(B(x,3R)) \leqslant 2\cdot 3^n\mu(B(x,R))\,!$).

One of such standard ways is to compare $T_{3R}\chi_F(x)$ to the average

$$V_R(x) := \frac{1}{\mu(B(x,R))}\int_{B(x,R)} T\chi_F\, d\mu$$

(the quantity, which is clearly bounded by $\frac{\mu(B(x,3R))}{\mu(B(x,R))}\widetilde{M}T\chi_F(x) \leqslant 3\cdot 2^n\widetilde{M}T\chi_F(x)$).

We have (here and below δ_x is the unit point mass at the point $x \in \mathcal{X}$):

$$T_{3R}\chi_F(x) - V_R(x) =$$
$$\int_{F\setminus B(x,3R)} T^*[\delta_x - \tfrac{1}{\mu(B(x,R))}\chi_{B(x,R)}d\mu]d\mu - \frac{1}{\mu(B(x,R))}\int_{\mathcal{X}}\chi_{B(x,R)}\cdot T\chi_{F\cap B(x,3R)}d\mu.$$

The first term does not exceed $2A_1$ according to Lemma 3.4 (applied to the adjoint operator T^* instead of T), while the second can be estimated by

$$\frac{1}{\mu(B(x,R))}\|T\|_{L^2(\mu)\to L^2(\mu)}\cdot\|\chi_{B(x,R)}\|_{L^2(\mu)}\cdot\|\chi_{F\cap B(x,3R)}\|_{L^2(\mu)} \leqslant$$
$$\frac{1}{\mu(B(x,R))}\|T\|_{L^2(\mu)\to L^2(\mu)}\sqrt{\mu(B(x,R))}\sqrt{\mu(B(x,3R))} \leqslant \sqrt{2\cdot 3^n}\,\|T\|_{L^2(\mu)\to L^2(\mu)}.$$

Combining all the above inequalities, we see that one can take $A_2 = 4\cdot 3^n\, A + 2A_1 + \sqrt{2\cdot 3^n}\|T\|_{L^2(\mu)\to L^2(\mu)}$.

□

The lemma we just outlined is crucial for our proof of the weak type $1-1$ estimate, but, unfortunately, not sufficient alone. In the next section we will present a construction that, however simple and natural, seems to have been completely overlooked (at any rate we do not know about any other paper in which it is used).

5. An alternative to the Calderón–Zygmund decomposition

Let $\nu \in M(\mathcal{X})$ be a finite linear combination of unit point masses with positive coefficients, i.e.,

$$\nu = \sum_{i=1}^{N}\alpha_i\delta_{x_i}.$$

Theorem 5.1.

$$\|T\nu\|_{L^{1,\infty}(\mu)} \leqslant A_4\|\nu\|$$

with some $A_4 > 0$ depending only on the dimension n, the constants A and ε in the definition of the Calderón–Zygmund kernel K, and the norm $\|T\|_{L^2(\mu)\to L^2(\mu)}$.

Here there is no problem with the definition of $T\nu$: it is just the finite sum $\sum_{i=1}^N \alpha_i K(x,x_i)$, which makes sense everywhere except finitely many points.

Proof. Without loss of generality, we may assume that $\|\nu\| = \sum_i \alpha_i = 1$ (this is just a matter of normalization). Thus we have to prove that $\|T\nu\|_{L^{1,\infty}(\mu)} \leqslant A_4$.

Fix some $t > 0$ and suppose first that $\mu(\mathcal{X}) > \frac{1}{t}$. Let $B(x_1,\rho_1)$ be the smallest (closed) ball such that $\mu(B(x_1,\rho_1)) \geqslant \dfrac{\alpha_1}{t}$ (since the function $\rho \to \mu(B(x_1,\rho))$ is increasing and continuous from the right, tends to 0 as $\rho \to 0$, and is greater than $\dfrac{1}{t} \geqslant \dfrac{\alpha_1}{t}$ for sufficiently large $\rho > 0$, such ρ_1 exists and is strictly positive).

Note that for the corresponding *open* ball $B'(x_1,\rho_1) := \{y \in \mathcal{X} : \operatorname{dist}(x_1,y) < \rho_1\}$, we have $\mu(B'(x_1,\rho_1)) = \lim_{\rho\to\rho_1-0} \mu(B(x_1,\rho)) \leqslant \dfrac{\alpha_1}{t}$. Since the measure μ is σ-finite and non-atomic, one can choose a Borel set E_1 satisfying

$$B'(x_1,\rho_1) \subset E_1 \subset B(x_1,\rho_1) \qquad \text{and} \qquad \mu(E_1) = \frac{\alpha_1}{t}.$$

Let $B(x_2,\rho_2)$ be the smallest ball such that $\mu(B(x_2,\rho_2) \setminus E_1) \geqslant \dfrac{\alpha_2}{t}$ (since $\mu(\mathcal{X}) > \frac{1}{t}$, the measure of the remaining part $\mathcal{X} \setminus E_1$ is still greater than $\dfrac{1-\alpha_1}{t} \geqslant \dfrac{\alpha_2}{t}$). Again for the corresponding open ball $B'(x_2,\rho_2)$, we have $\mu(B'(x_2,\rho_2)\setminus E_1) \leqslant \dfrac{\alpha_2}{t}$, and therefore there exists a Borel set E_2 satisfying

$$B'(x_2,\rho_2) \setminus E_1 \subset E_2 \subset B(x_1,\rho_1) \setminus E_1 \qquad \text{and} \qquad \mu(E_2) = \frac{\alpha_2}{t}.$$

In general, for $i = 3, 4, \dots, N$, let $B(x_i,\rho_i)$ be the smallest ball such that

$$\mu\Big(B(x_i,\rho_i) \setminus \bigcup_{\ell=1}^{i-1} E_\ell\Big) \geqslant \frac{\alpha_i}{t},$$

and let E_i be a Borel set satisfying

$$B'(x_i,\rho_i) \setminus \bigcup_{\ell=1}^{i-1} E_\ell \subset E_i \subset B(x_i,\rho_i) \setminus \bigcup_{\ell=1}^{i-1} E_\ell \quad \text{and} \quad \mu(E_i) = \frac{\alpha_i}{t}.$$

Put $E := \bigcup_i E_i$. Clearly

$$\bigcup_i B'(x_i,\rho_i) \subset E \subset \bigcup_i B(x_i,\rho_i) \qquad \text{and} \qquad \mu(E) = \frac{1}{t}.$$

Now let us compare $T\nu$ to $t\sum_i \chi_{\mathcal{X}\setminus B(x_i,2\rho_i)} \cdot T\chi_{E_i} =: t\sigma$ outside E. We have

$$T\nu - t\sigma = \sum_i \varphi_i$$

where

$$\varphi_i = \alpha_i T\delta_{x_i} - t\,\chi_{\mathcal{X}\setminus B(x_i,2\rho_i)} \cdot T\chi_{E_i}.$$

Note now that

$$\int_{\mathcal{X}\setminus E} |\varphi_i|\,d\mu \leqslant \int_{\mathcal{X}\setminus B(x_i,2\rho_i)} \big|T[\alpha_i\delta_{x_i} - t\chi_{E_i}\,d\mu]\big|\,d\mu + \int_{B(x_i,2\rho_i)\setminus B'(x_i,\rho_i)} \alpha_i |T\delta_{x_i}|\,d\mu.$$

But, according to Lemma 3.4, the first integral does not exceed

$$A_1\|\alpha_i\delta_{x_i} - t\,\chi_{E_i} d\mu\| = 2A_1\alpha_i,$$

while $|T\delta_{x_i}| \leqslant A\rho_i^{-n}$ outside $B'(x_i,\rho_i)$ and therefore the second integral is not greater than $\alpha_i A\rho_i^{-n}\mu(B(x_i,2\rho_i)) \leqslant 2^n A\alpha_i$. Finally we conclude that

$$\int_{\mathcal{X}\setminus E} |T\nu - t\sigma| d\mu \leqslant (2A_1 + 2^n A)\sum_i \alpha_i = 2A_1 + 2^n A,$$

and thereby $|T\nu - t\sigma| \leqslant (2A_1 + 2^n A)t$ everywhere on $\mathcal{X}\setminus E$, except, maybe, a set of measure $\frac{1}{t}$. To accomplish the proof of the theorem, we will show that for sufficiently large A_3,

$$\mu\{|\sigma| > A_3\} \leqslant \frac{2}{t}.$$

Then, combining all the above estimates, we shall get

$$\mu\{x \in \mathcal{X} \,:\, |T\nu(x)| > (A_3 + 2A_1 + 2^n A)t\} \leqslant \frac{4}{t}.$$

Since the same inequality is obviously true in the case when $\mu(\mathcal{X}) \leqslant \frac{1}{t}$, one may take $A_4 = 4(A_3 + 2A_1 + 2^n A)$.

We will apply the standard Stein-Weiss duality trick. Assume that the inverse inequality $\mu\{|\sigma| > A_3\} > \frac{2}{t}$ holds. Then either $\mu\{\sigma > A_3\} > \frac{1}{t}$, or $\mu\{\sigma < -A_3\} > \frac{1}{t}$. Assume for definiteness that the first case takes place and choose some set $F \subset \mathcal{X}$ of measure exactly $\frac{1}{t}$ such that $\sigma > A_3$ everywhere on F. Then, clearly,

$$\int_{\mathcal{X}} \sigma\chi_F d\mu > \frac{A_3}{t}.$$

On the other hand, this integral can be computed as

$$\sum_i \int_{\mathcal{X}} [T\chi_{E_i}]\cdot\chi_{F\setminus B(x_i,2\rho_i)}\,d\mu = \sum_i \int_{\mathcal{X}} \chi_{E_i}\cdot[T^*\chi_{F\setminus B(x_i,2\rho_i)}]\,d\mu.$$

Note that for every point $x \in E_i \subset B(x_i,\rho_i)$,

$$|T^*\chi_{F\setminus B(x_i,2\rho_i)}(x) - T^*\chi_{F\setminus B(x,\rho_i)}(x)| \leqslant \int_{B(x_i,2\rho_i)\setminus B(x,\rho_i)} |K(y,x)|\,d\mu(y) \leqslant$$
$$\leqslant A\rho_i^{-n}\mu(B(x_i,2\rho_i)) \leqslant 2^n A$$

and thereby for every $x \in E_i \cap \operatorname{supp}\mu$,

$$|T^*\chi_{F\setminus B(x_i,2\rho_i)}(x)| \leqslant (T^*)^\sharp\chi_F(x) + 2^n A \leqslant 2\cdot 3^n\,\widetilde{M}T^*\chi_F(x) + A_2 + 2^n A$$

according to the Guy David lemma (Lemma 4.1). Hence

$$\int_{\mathcal{X}} \sigma\chi_F d\mu \leqslant (A_2 + 2^n A)\mu(E) + 2\cdot 3^n \int_{\mathcal{X}} \chi_E\cdot\widetilde{M}T^*\chi_F d\mu.$$

But the first term equals $\dfrac{A_2 + 2^n A}{t}$ while the second one does not exceed

$$2\cdot 3^n\,\|\chi_E\|_{L^2(\mu)}\|\widetilde{M}T^*\chi_F\|_{L^2(\mu)} \leqslant \frac{2\cdot 3^n}{t}\|\widetilde{M}\|_{L^2(\mu)\to L^2(\mu)}\|T^*\|_{L^2(\mu)\to L^2(\mu)}.$$

Recalling that $\|T^*\|_{L^2(\mu)\to L^2(\mu)} = \|T\|_{L^2(\mu)\to L^2(\mu)}$, we see that one can take

$$A_3 = A_2 + 2^n A + 2\cdot 3^n \,\|\widetilde{M}\|_{L^2(\mu)\to L^2(\mu)}\|T\|_{L^2(\mu)\to L^2(\mu)}$$

to get a contradiction. Since the norm $\|\widetilde{M}\|_{L^2(\mu)\to L^2(\mu)}$ is bounded by some absolute constant (the constant in the Marcinkiewicz interpolation theorem), we are done. ☐

6. From finite linear combinations of point masses to $L^1(\mu)$-functions

Note first of all that Theorem 5.1 remains valid (with twice larger constant) for finite linear combinations of point masses with *arbitrary* real coefficients. Indeed, every such measure ν can be represented as $\nu_+ - \nu_-$ where $\nu_\pm$ are finite linear combinations of point masses with *positive* coefficients and $\|\nu\| = \|\nu_+\| + \|\nu_-\|$. Hence

$$\|T\nu\|_{L^{1,\infty}(\mu)} \leqslant 2\big(\|T\nu_+\|_{L^{1,\infty}(\mu)} + \|T\nu_-\|_{L^{1,\infty}(\mu)}\big) \leqslant 2A_4(\|\nu_+\| + \|\nu_-\|) = 2A_4\|\nu\|.$$

Now we are ready to prove

Theorem 6.1. *Let $f \in L^1(\mu)\cap L^2(\mu)$. Then*

$$\|Tf\|_{L^{1,\infty}(\mu)} \leqslant A_5\|f\|_{L^1(\mu)}$$

with some $A_5 > 0$ depending only on the dimension n, the constants A and ε in the definition of the Calderón–Zygmund kernel K, and the norm $\|T\|_{L^2(\mu)\to L^2(\mu)}$.

Proof. Let $C_0(\mathcal{X})$ be the space of bounded continuous functions on $\mathcal{X}$ with bounded support (a function is said to have bounded support if it vanishes outside some (large) ball of finite radius). Clearly, $C_0(\mathcal{X}) \subset L^1(\mu)\cap L^2(\mu)$, and it is a standard fact from measure theory that $C_0(\mathcal{X})$ is dense in $L^1(\mu)\cap L^2(\mu)$ with respect to the norm $\|\cdot\|_{L^1(\mu)} + \|\cdot\|_{L^2(\mu)}$. Therefore it is enough to prove the desired inequality for $f \in C_0(\mathcal{X})$.

Fix $t > 0$ and put $G := \{x \in \mathcal{X} : |f(x)| > t\}$, $f^t := f\cdot\chi_G$ and $f_t = f\cdot\chi_{\mathcal{X}\setminus G}$. We have $Tf = Tf^t + Tf_t$. Now observe, as usual, that

$$\int_{\mathcal{X}} |f_t|^2\,d\mu \leqslant t\int_{\mathcal{X}} |f_t|\,d\mu \leqslant t\|f\|_{L^1(\mu)}.$$

Therefore $\int_{\mathcal{X}} |Tf_t|^2\,d\mu \leqslant \|T\|^2_{L^2(\mu)\to L^2(\mu)} t\|f\|_{L^1(\mu)}$, and

$$\mu\big\{x\in\mathcal{X} : |Tf_t(x)| > t\cdot\|T\|_{L^2(\mu)\to L^2(\mu)}\big\} \leqslant \frac{\|f\|_{L^1(\mu)}}{t}.$$

Note now that G is an *open* set (this is the only place where we use the continuity of f) and that $\mu(G) \leqslant \frac{1}{t}\|f\|_{L^1(\mu)}$. Recall that every open set G in a separable metric space allows a "Whitney decomposition", i.e., it can be represented as a union of countably many pairwise disjoint Borel sets G_i $(i = 1, 2, \dots)$ satisfying

$$\operatorname{diam} G_i \leqslant \tfrac{1}{2}\operatorname{dist}(G_i, \mathcal{X}\setminus G).$$

Put $f_i := f\cdot\chi_{G_i}$. Then $f^t = \sum_{i=1}^\infty f_i$ where the series converges at least in $L^2(\mu)$. Let $f^{(N)}$ be the N-th partial sum of this series. Define

$$\alpha_i := \int_{\mathcal{X}} f_i\,d\mu = \int_{G_i} f\,d\mu.$$

Obviously, $\sum_{i=1}^{\infty}|\alpha_i| \leqslant \|f\|_{L^1(\mu)}$. Choose one point x_i in every set G_i and put $\nu_N = \sum_{i=1}^{N}\alpha_i\delta_{x_i}$. Consider the difference $Tf^{(N)} - T\nu_N$ outside G. We have

$$\int_{\mathcal{X}\setminus G}\left|Tf^{(N)} - T\nu_N\right| d\mu \leqslant \sum_{i=1}^{N}\int_{\mathcal{X}\setminus G}\left|T[f_i d\mu - \alpha_i\delta_{x_i}]\right| d\mu \leqslant 2A_1\sum_{i=1}^{N}|\alpha_i| \leqslant 2A_1\|f\|_{L^1(\mu)}$$

according to Lemma 3.4. Thus $|Tf^{(N)} - T\nu_N| \leqslant 2A_1 t$ everywhere outside G save, maybe, some exceptional set of measure at most $\frac{1}{t}\|f\|_{L^1(\mu)}$. As we have seen above,

$$\mu\{x\in\mathcal{X} \,:\, |T\nu_N(x)| > 2A_4 t\} \leqslant \frac{1}{t}\|\nu_N\| \leqslant \frac{1}{t}\|f\|_{L^1(\mu)}.$$

Hence

$$\mu\big\{x\in\mathcal{X}\setminus G \,:\, |Tf^{(N)}(x)| > 2(A_1+A_4)t\big\} \leqslant \frac{2}{t}\|f\|_{L^1(\mu)},$$

and

$$\mu\big\{x\in\mathcal{X} \,:\, |Tf^{(N)}(x)| > 2(A_1+A_4)t\big\} \leqslant \frac{3}{t}\|f\|_{L^1(\mu)}.$$

Since $f^{(N)} \to f^t$ in $L^2(\mu)$ as $N\to+\infty$, we have $Tf^{(N)} \to Tf^t$ in $L^2(\mu)$ as $N\to+\infty$, which is more than enough to pass to the limit and to conclude that

$$\mu\big\{x\in\mathcal{X} \,:\, |Tf^{t}(x)| > 2(A_1+A_4)t\big\} \leqslant \frac{3}{t}\|f\|_{L^1(\mu)}.$$

Thus, we can take $A_5 = 4\big[\,\|T\|_{L^2(\mu)\to L^2(\mu)} + 2(A_1+A_4)\,\big]$.

As usual, by the Marcinkiewicz interpolation theorem, we obtain that the operator T is bounded in $L^p(\mu)$ for every $1 < p \leqslant 2$. By duality, this result automatically extends to all $p\in(1,+\infty)$. □

7. Cotlar type inequalities and boundedness of $T^\sharp$ in $L^p(\mu)$.

Now we are ready to prove the boundedness of the maximal operator $T^\sharp$ in all spaces $L^p(\mu)$ with $1 < p < +\infty$. It follows immediately from

Theorem 7.1. *Let $f\in L^2(\mu)$. For any $\beta > 1$ and $x\in\operatorname{supp}\mu$,*

$$T^\sharp f(x) \leqslant 4\cdot 9^n\widetilde{M}Tf(x) + B(\beta)\,\widetilde{M}_\beta f(x)$$

where the constant $B(\beta) > 0$ depends on the parameter $\beta > 1$, the dimension n, the constants ε and A in the definition of the Calderón–Zygmund kernel K, and the norm $\|T\|_{L^2(\mu)\to L^2(\mu)}$ only.

Proof. It is just a minor modification of the proof of the Guy David lemma. Let again $r > 0$. Put $r_j = 3^j r$ and $\mu_j = \mu(B(x,r_j))$ as before, but let now k be the smallest positive integer for which $\mu_{k+1} \leqslant 4\cdot 9^n\mu_{k-1}$ (i.e., we look now *two steps forward* when checking for the doubling). Note that such an integer k exists, because otherwise for every *even* j,

$$\mu(B(x,r)) \leqslant 2^{-j}3^{-nj}\mu(B(x,r_j)) \leqslant 2^{-j}r^n,$$

and thereby $\mu(B(x,r)) = 0$, which is impossible. Put $R = r_{k-1} = 3^{k-1}r$ exactly as before. We have

$$\begin{aligned}|T_r f(x) - T_{3R} f(x)| &\leqslant \int_{B(x,3R)\setminus B(x,r)} |K(x,y)|\cdot|f(y)|\,d\mu(y)\\ &= \sum_{j=1}^{k}\int_{B(x,r_j)\setminus B(x,r_{j-1})} |K(x,y)|\cdot|f(y)|\,d\mu(y) =: \sum_{j=1}^{k}\mathcal{I}_j.\end{aligned}$$

Note that

$$\mathcal{I}_j \leqslant A r_{j-1}^{-n}\int_{B(x,r_j)} |f|\,d\mu \leqslant A r_{j-1}^{-n}\mu_{j+1}\widetilde{M}f(x).$$

Observe now that $r_{j-1} = 3^{j-1-k}r_k$ and $\mu_{j+1} \leqslant [4\cdot 9^n]^{\frac{j+2-k}{2}}\mu_k$ for $1 \leqslant j \leqslant k$ (it is enough to check this inequality for $j = k, k-1$ and $k-2$). Hence

$$\sum_{j=1}^{k}\mathcal{I}_j \leqslant 4\cdot 27^n\, A\,\widetilde{M}f(x)\,\frac{\mu_k}{r_k^n}\sum_{j=1}^{k} 2^{j-k} \leqslant 8\cdot 27^n\, A\,\widetilde{M}f(x) \leqslant 8\cdot 27^n\, A\,\widetilde{M}_\beta f(x).$$

So, again, we need only to estimate $T_{3R}f(x)$. As before, consider the average

$$V_R(x) := \frac{1}{\mu(B(x,R))}\int_{B(x,R)} Tf\,d\mu,$$

which is bounded by $\frac{\mu(B(x,3R))}{\mu(B(x,R))}\widetilde{M}Tf(x) \leqslant 4\cdot 9^n\widetilde{M}Tf(x)$ according to our choice of k, and write

$$T_{3R}f(x) - V_R(x) =$$
$$\int_{\mathcal{X}\setminus B(x,3R)} T^*[\delta_x - \tfrac{1}{\mu(B(x,R))}\chi_{B(x,R)}\,d\mu]\,f\,d\mu - \frac{1}{\mu(B(x,R))}\int_{\mathcal{X}}\chi_{B(x,R)}\cdot T[f\chi_{B(x,3R)}]d\mu.$$

Using Lemma 3.4, we can now estimate the absolute value of the minuend by $2A_1\widetilde{M}f(x) \leqslant 2A_1\widetilde{M}_\beta f(x)$. As to the subtrahend, at this stage we know that T is bounded in $L^\beta(\mu)$, and therefore the absolute value of the subtrahend does not exceed

$$\frac{1}{\mu(B(x,R))}\|T\|_{L^\beta(\mu)\to L^\beta(\mu)}\|\chi_{B(x,R)}\|_{L^{\beta'}(\mu)}\cdot\|f\chi_{B(x,3R)}\|_{L^\beta(\mu)}$$

where $\beta' := \frac{\beta}{\beta-1}$ is the conjugate exponent to β. Clearly

$$\|\chi_{B(x,R)}\|_{L^{\beta'}(\mu)} = \{\mu(B(x,R))\}^{1/\beta'}.$$

The point is that now, according to our choice of k, we have $\mu(B(x,9R)) \leqslant 4\cdot 9^n\mu(B(x,R))$, and therefore

$$\|f\chi_{B(x,3R)}\|_{L^\beta(\mu)} \leqslant \widetilde{M}_\beta f(x)\{\mu(B(x,9R))\}^{1/\beta} \leqslant \widetilde{M}_\beta f(x)\{4\cdot 9^n\mu(B(x,R))\}^{1/\beta}.$$

This allows us to conclude finally that the subtrahend is bounded by

$$[4\cdot 9^n]^{1/\beta}\|T\|_{L^\beta(\mu)\to L^\beta(\mu)}\widetilde{M}_\beta f(x),$$

proving the theorem with $B(\beta) = 8\cdot 27^n A + 2A_1 + [4\cdot 9^n]^{1/\beta}\|T\|_{L^\beta(\mu)\to L^\beta(\mu)}$. □

8. Weak type 1-1 estimate for the maximal operator $T^\sharp$

Now, to complete the "classical L^p-theory", it remains to prove that the maximal operator $T^\sharp$ is bounded from $M(\mathcal{X})$ to $L^{1,\infty}(\mu)$, i.e., that for every signed measure $\nu \in M(\mathcal{X})$,

$$\|T^\sharp \nu\|_{L^{1,\infty}(\mu)} \leqslant C\|\nu\|$$

with some constant $C > 0$, not depending on ν.

We will start again with "elementary" measures $\nu \in M(\mathcal{X})$, i.e., with the measures of the kind $\nu = \sum_{i=1}^N \alpha_i \delta_{x_i}$ where $x_i \in \mathcal{X}$, $\alpha_i > 0$ $(i = 1, \dots, N)$.

Theorem 8.1. *Let $\beta \in (0, 1)$. For every elementary measure $\nu \in M(\mathcal{X})$ and for every $x \in \operatorname{supp} \mu$,*

$$[T^\sharp \nu(x)]^\beta \leqslant 4 \cdot 9^n [\widetilde{M}_\beta T\nu(x)]^\beta + B(\beta)\,[\widetilde{M}\nu(x)]^\beta$$

with some $B(\beta) > 0$ depending only on the parameter $\beta < 1$, dimension n, the constants A and ε in the definition of the Calderón–Zygmund kernel K, and the norm $\|T\|_{L^2(\mu)\to L^2(\mu)}$.

Note again that $T\nu$ is well-defined everywhere except finitely many points, so the first term on the right does make sense.

Corollary 8.2. *For every elementary measure $\nu \in M(\mathcal{X})$,*

$$\|T^\sharp \nu\|_{L^{1,\infty}(\mu)} \leqslant A_6\|\nu\|$$

with $A_6 > 0$ depending only on the dimension n, the constants A and ε in the definition of the Calderón–Zygmund kernel K, and the norm $\|T\|_{L^2(\mu)\to L^2(\mu)}$.

Proof of Theorem 8.1. Take some $r > 0$. Put $r_j = 3^j r$ and $\mu_j = \mu(B(x, r_j))$ as usual, and let again (like in Section 7) k be the smallest positive integer for which $\mu_{k+1} \leqslant 4 \cdot 9^n \mu_{k-1}$. Put $R = r_{k-1} = 3^{k-1} r$.

The same reasoning as in the proof of Theorem 7.1 yields

$$|T_r\nu(x) - T_{3R}\nu(x)| \leqslant 8 \cdot 27^n\, A\,\widetilde{M}\nu(x)$$

Now represent the measure ν as $\nu_1 + \nu_2$, where

$$\nu_1 := \sum_{i:\, x_i \in B(x,3R)} \alpha_i \delta_{x_i} \quad \text{and} \quad \nu_2 := \sum_{i:\, x_i \notin B(x,3R)} \alpha_i \delta_{x_i}.$$

For any $x' \in B(x, R)$, we have

$$\begin{aligned} |T_{3R}\nu(x) - T\nu_2(x')| &= |T\nu_2(x) - T\nu_2(x')| = \left|\int_{\mathcal{X}} T^*[\delta_x - \delta_{x'}]\, d\nu_2\right| \\ &\leqslant \int_{\mathcal{X}} |T^*[\delta_x - \delta_{x'}]|\, d\nu_2 = \int_{\mathcal{X}\setminus B(x,3R)} |T^*[\delta_x - \delta_{x'}]|\, d\nu \leqslant 2A_1\widetilde{M}\nu(x) \end{aligned}$$

(see Lemma 3.4). Hence

$$\frac{1}{\mu(B(x,R))}\int_{B(x,R)} |T_{3R}\nu(x) - T\nu_2(x')|^\beta\, d\mu(x') \leqslant \left[2A_1\widetilde{M}\nu(x)\right]^\beta.$$

On the other hand,

$$\frac{1}{\mu(B(x,R))}\int_{B(x,R)}|T\nu_2(x')-T\nu(x')|^\beta\,d\mu(x')$$
$$=\frac{1}{\mu(B(x,R))}\int_{B(x,R)}|T\nu_1(x')|^\beta\,d\mu(x')$$
$$=\frac{1}{\mu(B(x,R))}\int_0^{+\infty}\beta s^{\beta-1}\mu\{x'\in B(x,R)\,:\,|T\nu_1(x')|>s\}\,ds.$$

Note now that for every $s>0$,

$$\mu\{x'\in B(x,R)\,:\,|T\nu_1(x')|>s\}\leqslant\min\Big(\mu(B(x,R)),\frac{A_4\,\|\nu_1\|}{s}\Big)$$
$$\leqslant\mu(B(x,R))\min\big(1,\tfrac{\mu(B(x,9R))}{\mu(B(x,R))}\,\tfrac{A_4\,\widetilde{M}\nu(x)}{s}\big)\leqslant\mu(B(x,R))\min\big(1,\tfrac{4\cdot 9^n\,A_4\,\widetilde{M}\nu(x)}{s}\big).$$

Therefore

$$\frac{1}{\mu(B(x,R))}\int_0^{+\infty}\beta s^{\beta-1}\mu\{x'\in B(x,R)\,:\,|T\nu_1(x')|>s\}\,ds$$
$$\leqslant\int_0^{+\infty}\beta s^{\beta-1}\min\Big(1,\frac{4\cdot 9^n\,A_4\,\widetilde{M}\nu(x)}{s}\Big)\,ds$$
$$=\big[4\cdot 9^n\,A_4\,\widetilde{M}\nu(x)\big]^\beta\int_0^{+\infty}\beta s^{\beta-1}\min(1,\tfrac{1}{s})\,ds=\tfrac{1}{1-\beta}\big[4\cdot 9^n\,A_4\,\widetilde{M}\nu(x)\big]^\beta.$$

Using the elementary inequality $|a+b|^\beta\leqslant|a|^\beta+|b|^\beta$ $(a,b\in\mathbb{R};\ \beta\in(0,1)\,)$, we obtain

$$\frac{1}{\mu(B(x,R))}\int_{B(x,R)}|T_{3R}\nu(x)-T\nu(x')|^\beta\,d\mu(x')\leqslant\big([2A_1]^\beta+\tfrac{1}{1-\beta}[4\cdot 9^n\,A_4]^\beta\big)\,[\widetilde{M}\nu(x)]^\beta.$$

Using it twice more, we finally get

$$|T_r\nu(x)|^\beta\leqslant\frac{1}{\mu(B(x,R))}\int_{B(x,R)}|T\nu|^\beta\,d\mu$$
$$+\big([8\cdot 27^n A]^\beta+[2A_1]^\beta+\tfrac{1}{1-\beta}[4\cdot 9^n\,A_4]^\beta\big)\,[\widetilde{M}\nu(x)]^\beta.$$

To prove the theorem, it remains only to note that

$$\frac{1}{\mu(B(x,R))}\int_{B(x,R)}|T\nu|^\beta\,d\mu\leqslant\frac{\mu(B(x,3R))}{\mu(B(x,R))}[\widetilde{M}_\beta T\nu]^\beta\leqslant 4\cdot 9^n[\widetilde{M}_\beta T\nu]^\beta.$$

□

To prove Corollary 8.2, it is enough to recall that $\widetilde{M}_\beta$ is bounded in $L^{1,\infty}(\mu)$ for any $\beta\in(0,1)$, and that $\|T\nu\|_{L^{1,\infty}(\mu)}\leqslant A_4\|\nu\|$ and $\|\widetilde{M}\nu\|_{L^{1,\infty}(\mu)}\leqslant\|\nu\|$.

9. The weak type $1-1$ estimate for arbitrary measures $\nu\in M(X)$

Theorem 9.1. *For any finite non-negative measure $\nu\in M(\mathcal{X})$, one has*

$$\|T^\sharp\nu\|_{L^{1,\infty}(\mu)}\leqslant A_6\|\nu\|,$$

where A_6 is the same constant as in the corollary 8.2.

Theorem 9.1 essentially says that elementary measures are "weakly dense" in the set of all finite non-negative measures. Though by no means surprising, it is not completely obvious (or, maybe, it is, but we just do not see how), because we work with a space that is not locally compact and with a kernel that is not everywhere continuous. That is why we decided to include a formal proof.

Corollary 9.2. *For every $\nu \in M(\mathcal{X})$,*

$$\|T^\sharp \nu\|_{L^{1,\infty}(\mu)} \leqslant 2A_6\|\nu\|.$$

Proof of Theorem 9.1. Fix $t > 0$. Our aim is to show that

$$\mu\{x \in \mathcal{X} \,:\, T^\sharp\nu(x) > t\} \leqslant \frac{A_6\|\nu\|}{t}.$$

Take $R > 0$ and consider the truncated maximal operator

$$T^\sharp_R\nu(x) := \sup_{r>R} |T_r\nu(x)|.$$

Since $T^\sharp_R\nu \nearrow T^\sharp\nu$ pointwise on $\mathcal{X}$ as $R \to 0$, it is enough to check that

$$\mu\{x \in \mathcal{X} \,:\, T^\sharp_R\nu(x) > t\} \leqslant \frac{A_6\|\nu\|}{t}$$

for every $R > 0$.

For every $N \in \mathbb{N}$, consider the random elementary measure

$$\nu_N := \frac{\|\nu\|}{N}\sum_{i=1}^{N}\delta_{x_i}$$

where the random points $x_i \in \mathcal{X}$ are independent and $\mathcal{P}\{x_i \in E\} = \frac{\nu(E)}{\|\nu\|}$ for every Borel set $E \subset \mathcal{X}$. (Here and below we denote by $\mathcal{P}\{X\}$ the probability of the event X, by $\mathcal{E}\xi$ the mathematical expectation of a random variable ξ, and by $\mathcal{D}\xi := \mathcal{E}|\xi - \mathcal{E}\xi|^2 = \mathcal{E}|\xi|^2 - |\mathcal{E}\xi|^2$ the dispersion of the random variable ξ).

Note that for every fixed $x \in \mathcal{X}$ and $r > R$,

$$\mathcal{E}\, T_r\delta_{x_i}(x) = T_r(\mathcal{E}\,\delta_{x_i})(x) = \frac{1}{\|\nu\|}T_r\nu(x)$$

and

$$\mathcal{D}\, T_r\delta_{x_i}(x) \leqslant \mathcal{E}\,|T_r\delta_{x_i}(x)|^2 \leqslant \frac{A^2}{r^{2n}} \leqslant \frac{A^2}{R^{2n}}.$$

Hence

$$\mathcal{E}\, T_r\nu_N(x) = T_r\nu(x) \qquad \text{and} \qquad \mathcal{D}\, T_r\nu_N(x) \leqslant \frac{1}{N}\,\frac{A^2\|\nu\|^2}{R^{2n}}.$$

Fix a very small number $\gamma > 0$ and note that for every point $x \in \mathcal{X}$ satisfying $|T_r\nu(x)| > t$, we have

$$\mathcal{P}\{|T_r\nu_N(x)| \leqslant (1-\gamma)t\} \leqslant \mathcal{P}\{|T_r\nu_N(x) - T_r\nu(x)| > \gamma t\} \leqslant \frac{\mathcal{D}\, T_r\nu_N(x)}{\gamma^2 t^2} \leqslant \frac{1}{N}\,\frac{A^2\|\nu\|^2}{R^{2n}\gamma^2 t^2} \leqslant \gamma,$$

provided that $N \in \mathbb{N}$ is large enough. From here we incur that for every point $x \in \mathcal{X}$ satisfying $|T^\sharp_R \nu(x)| > t$, we have

$$\mathcal{P}\{|T^\sharp \nu_N(x)| \leqslant (1-\gamma)t\} \leqslant \gamma.$$

Let now E be any Borel set *of finite measure* such that $T^\sharp_R \nu(x) > t$ for every $x \in E$. We have

$$\mathcal{E}\,\mu\{x \in E : |T^\sharp \nu_N(x)| \leqslant (1-\gamma)t\} = \int_E \mathcal{P}\{|T^\sharp \nu_N(x)| \leqslant (1-\gamma)t\}\, d\mu(x) \leqslant \gamma\mu(E).$$

Thus there exists at least one choice of points x_i $(i = 1, \dots, N)$ for which $\mu\{x \in E : |T^\sharp \nu_N(x)| \leqslant (1-\gamma)t\} \leqslant \gamma\mu(E)$ and therefore

$$\mu\{x \in E : |T^\sharp \nu_N(x)| > (1-\gamma)t\} \geqslant (1-\gamma)\mu(E).$$

According to the weak type $1-1$ estimate for elementary measures, this implies

$$\mu(E) \leqslant \frac{A_6\|\nu_N\|}{(1-\gamma)^2 t} = \frac{A_6\|\nu\|}{(1-\gamma)^2 t}.$$

Since $\gamma > 0$ was arbitrary, we get $\mu(E) \leqslant \frac{A_6\|\nu\|}{t}$. At last, since μ is σ-finite and E was an arbitrary subset of finite measure of the set of the points $x \in \mathcal{X}$ for which $T^\sharp_R \nu(x) > t$, we conclude that

$$\mu\{x \in \mathcal{X} : T^\sharp_R \nu(x) > t\} \leqslant \frac{A_6\|\nu\|}{t},$$

proving the theorem. □

To prove Corollary 9.2, it is enough to recall that every signed measure $\nu \in M(\mathcal{X})$ can be represented as $\nu_+ - \nu_-$, where $\nu_\pm$ are *finite non-negative* measures and $\|\nu_+\| + \|\nu_-\| = \|\nu\|$.

References

[Ch1] M. Christ, *A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral*, Colloq. Math. **60/61** (1990), 601-628.

[Ch2] M. Christ, Lectures on singular integral operators. Regional Conference Series in Mathematics,**77**, Amer. Math. Soc., 1990.

[CW] R. R. Coifman, G. Weiss, *Analyse harmonique non-commutative sur certaines espaces homogènes*, Lect. Notes in Math.**242**, Springer-Verlag, 1971.

[D] G. David, *Completely unrectifiable 1-sets on the plane have vanishing analytic capacity*, Prépublications Mathématiques d'ORSAY, No. 61, 1997, 1-94.

[DJ] G. David, J. L. Journé, *A boundedness criterion for generalized Calderón–Zygmund operators*, Ann. of Math. (2), **120** (1984), 157-189.

[DM] G. David, P. Mattila, *Removable sets for Lipschitz harmonic functions in the plane*, Prépublications Mathématiques d'ORSAY, No. 31, 1997, 1-59.

[MMV] P. Mattila, M. Melnikov, J. Verdera, *The Cauchy integral, analytic capacity, and uniform rectifiability*, Ann. of Math. (2) **144** (1996), 127-136.

[NTV1] F. Nazarov, S. Treil, A. Volberg, *Cauchy integral and Calderón–Zygmund operators on non-homogeneous spaces*, IMRN (International Math. Res. Notices), 1997, **No. 15**, pp. 703-726.

[St1] E. M. Stein, Singular integrals and differentiability properties of functions. Princeton Mathematical Series, No. 30 Princeton University Press, Princeton, N.J. 1970 xiv+290 pp

[St2] E. M. Stein, Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals. With the assistance of Timothy S. Murphy. Princeton Math. Series, **43**, Monographs in Harmonic Analysis, Princeton Univ. press, Princeton, NJ, 1993.

[T1] X. Tolsa, *L^2-boundedness of the Cauchy integral operator for continuous measures*, Preprint, 1997, 1-24.

[T2] X. Tolsa, *Cotlar's inequality and existence of principal values for the Cauchy integral without doubling conditions*, Preprint, 1997, 1-31.

All authors: Dept of Math, Michigan State University, East Lansing, MI 48824
E-mail address: fedja@math.msu.edu

E-mail address: treil@math.msu.edu

E-mail address: volberg@math.msu.edu