§0. Introduction (historical remarks).

The classical theory of Calderon-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 [] by Stein written in ...(which, by the way, still remains an excellent introduction into the subject for newcomers) to the modern outline of the theory in [], [], and [].

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 $C > 0$,

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

where X is some metric space endowed with a Borel measure μ, and, as usual, $B(x, r) = \{y \in X : \operatorname{dist}(x, y) \leqslant r\}$ 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 first example of the situation when such a necessity arises concerns the action of the Cauchy integral operator on the complex plane. The problem here is as the following:

Given a finite Borel measure μ on the complex plane $\mathbb{C}$, figure out when the Cauchy integral operator

$$Cf(x) = C_\mu f(x) = \int_{\mathbb{C}} \frac{f(y)\,d\mu(y)}{x-y}$$

acts in $L^2(\mu)$ (in $L^p(\mu)$, from $L^1(\mu)$ to $L^{1,\infty}(\mu)$ and so on).

The leading example of such a measure (coming from study of analytic capacity) is the one-dimensional Hausdorf measure on some strange compact set K on the plane. If K is a Lipshitz curve or something similar, then we have a space of homogeneous type and almost all standard techniques do apply (see [Christ]). But in general the measure does not satisfy the doubling condition and one had to look for an alternative approach (see for instance [Mel'nikov] and [Tolsa]). This was a real pity, because the Cauchy integral operator is one of the *most natural and important* examples of Calderon-Zygmund operators, and to have it excluded from the general framework would be an unforgivable drawback of the theory.

Another example is just some standard singular integral operator considered in an open domain $\Omega \subset \mathbb{R}^n$ (with the usual n-dimensional Lebesgue measure) instead

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

of the whole space. Such singular integral operators are claimed in several papers 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. Again, if the boundary of Ω is nice, we get a space of homogeneous type and everything is well understood. But, if we don’t mistake, for domains with “wild” boundaries no satisfactory theory of Calderon-Zygmund operators has been previously known.

One more good question is ”Why then hasn’t it all been done long ago?” We do not know any really good answer to that. Nevertheless let us attract the attention of the reader to the fact that in order to drop the doubling condition from the proofs, it is impossible to “build on the top of the existing theory”: one should go to the very roots and to rebuild the whole thing from a scratch (for the experts in the field it will be enough to say that one has to give up the Calderon-Zygmund decomposition completely and to severely restrict himself in the usage of the Hardy-Littlewood maximal function). Such a task requires a lot of courage *just to start*.

We plan to outline the “basics” of the theory of Calderon-Zygmund operators in non-homogeneous spaces in one large paper to appear in “Algebra and Analysis” (known in the West as “St.-Petersburg Math. J.”). This note is a kind of advertisement for that forthcoming paper. It can also be considered as a complement to [], where the L^2-theory was (at least partially) constructed for the special case of the Cauchy integral operator on the complex plane. Here we are going to deal only with the part of the theory concerning the boundedness of a Calderon-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 question we are going to answer below is ”What can one replace the Calderon-Zygmund decomposition 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 are 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

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

For $1 \leqslant p < +\infty$, define, as usual,

$$L^p(\mu) := \Big\{ f : \mathcal{X} \to \mathbb{C} \,:\, \|f\|_{L^p(\mu)} := \Big[\int_{\mathcal{X}} |f|^p \, d\mu\Big]^{\frac{1}{p}} < +\infty \Big\};$$

$$L^\infty(\mu) := \big\{ f : \mathcal{X} \to \mathbb{C} \,:\, \|f\|_{L^\infty(\mu)} := \operatorname*{ess\,sup}_{x \in \mathcal{X}} |f(x)| < +\infty \big\};$$

and

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

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

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

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 $\frac{t}{2}$).

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" Calderon-Zygmund kernel on $\mathcal{X}$, i.e., for some $A > 0$, $\varepsilon \in (0, 1]$,

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

and

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

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

Assume that T is a bounded linear operator in $L^2(\mu)$ with the Calderon-Zygmund kernel K. As usual, it means that

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

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 Calderon-Zygmund kernel as well.

For $\nu \in M(\mathcal{X})$ and $x \in \mathcal{X} \setminus \operatorname{supp} \nu$, we will also put $T\nu(x) := \int_{\mathcal{X}} K(x, y) \, d\nu(y)$.

The maximal operator $T^\sharp$ associated with the Calderon-Zygmund operator T is defined as the following. For every $r > 0$, put

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

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 $T^\sharp f(x)$ by

$$T_r f(x) := \sup_{r>0} |T_r f(x)|$$

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

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

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

Theorem:

1) 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.

T is a bounded operator from $L^1(\mu)$ to $L^{1,\infty}(\mu)$ in the sense that for every $f \in L^1(\mu) \cap L^2(\mu)$ and for every $t > 0$,

$$\mu\{x \in \mathcal{X} : |Tf(x)| > t\} \leqslant C\frac{\|f\|_{L^1(\mu)}}{t}$$

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.

$T^\sharp$ is a bounded operator from $M(\mathcal{X})$ to $L^{1,\infty}(\mu)$ in the sense that for every $\nu \in M(\mathcal{X})$ and for every $t > 0$,

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

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

§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 §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 §§4-5 we shall prove the weak type $1-1$ estimate for $T\nu$ where ν is a finite linear combination of point masses (these two sections constitute the core of the whole article). In §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 §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 §§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 §3 and §6).

§3. Preliminary observations.

Recall that the Hardy-Littlewood maximal function $Mf(x)$ is defined (for locally integrable 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 $L^\infty(\mu)$. But 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). So one needs to switch to some *"restricted"* versions of the Hardy-Littlewood maximal function. We will use two of them:

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

and

$$\widehat{M}f(x) := \sup\Big\{ \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f|\, d\mu \,:\, r>0,\ \mu(B(x,3r)) \leqslant 9^n \mu(B(x,r)) \Big\}.$$

Thus, in the definition of $\widetilde{M}f$ we just replaced the denominator $\mu(B(x,r))$ with the larger number r^n, while in the definition of $\widehat{M}f$ we restricted ourselves to the balls that do satisfy the doubling condition. Note, by the way, that we always have a lot of such balls for every $x \in \operatorname{supp}\mu$. Indeed, if for some $r_0 > 0$ the inverse inequality $\mu(B(x,3r)) > 9^n \mu(B(x,r))$ holds for each $r \geqslant r_0$, then for every $k \in \mathbb{N}$,

$$\mu(B(x,r_0)) \leqslant 9^{-nk}\mu(B(x,3^k r_0)) \leqslant 3^{-nk} r_0^n.$$

Since the right hand part tends to 0 as k tends to $+\infty$, we obtain $\mu(B(x,r_0)) = 0$, which is impossible.

Lemma 3.1.

Both the restricted maximal function operators $\widetilde{M}$ and $\widehat{M}$ are bounded in $L^p(\mu)$ for each $p \in (1,+\infty]$ and act from $L^1(\mu)$ to $L^{1,\infty}(\mu)$.

Proof:

Note first of all that for every $x \in \mathcal{X}$ both $\widetilde{M}(x)$ and $\widehat{M}(x)$ do not exceed $Mf(x) \leqslant \|f\|_{L^\infty(\mu)}$, so 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 $\{B(x,r_x)\}_{x\in E}$ 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 $E \subset \cup_j B(x,3r_j)$.

The weak type $1-1$ estimate for $\widetilde{M}$.

Fix some $t > 0$. Pick $R > 0$ and consider the set E of the points $x \in \mathcal{X}$ for which

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

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

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

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

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

and therefore

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

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

The weak type $1-1$ estimate for $\widehat{M}$.

Again fix $t>0$, pick $R>0$ and consider the set E of the points $x\in\mathcal{X}$ for which

$$\widehat{M}^{(R)}f(x):=\sup\Big\{\frac{1}{\mu(B(x,r))}\int_{B(x,r)}|f|\,d\mu\,:\,0<r<R,\ \mu(B(x,3r))\leqslant 9^n\mu(B(x,r))\Big\}>t$$

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

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

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

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

and therefore

$$\mu(E)\leqslant\sum_j\mu(B(x_j,3r_j))\leqslant 9^n\sum_j\mu(B(x_j,r_j))\frac{\|f\|_{L^1(\mu)}}{t}.$$

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

Remark:

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

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

and

$$\widehat{M}f(x):=\sup\Big\{\frac{\nu(B(x,r))}{\mu(B(x,r))}\ :\ r>0,\ \ \mu(B(x,3r))\leqslant 9^n\mu(B(x,r))\Big\}$$

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

$$\|\widetilde{M}\nu\|_{L^{1,\infty}},\|\widehat{M}\nu\|_{L^{1,\infty}}\leqslant 9^n\nu(\mathcal{X}).$$

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

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

and

$$\widehat{M}_\beta f(x):=\big[\widehat{M}(|f|^\beta)(x)\big]^{\frac{1}{\beta}}$$
$$=\sup\Big\{\Big[\frac{1}{\mu(B(x,r))}\int_{B(x,r)}|f|^\beta\,d\mu\Big]^{\frac{1}{\beta}}\ :\ r>0,\ \ \mu(B(x,3r))\leqslant 9^n\mu(B(x,r))\Big\}.$$

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

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

Lemma 3.2.

For any $\beta \in (0,1)$, the maximal operator $\widehat{M}_\beta$ 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.

The same, of course, is true for $\widetilde{M}_\beta$, but we shall not use that version of the lemma anywhere below.

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

$$\begin{aligned} \mu\{x \in \mathcal{X} : \widehat{M}_\beta f > 2^{\frac{1}{\beta}} t\} &\leqslant \mu\{x \in \mathcal{X} : \widehat{M}_\beta f^t > t\} \\ &= \mu\{x \in \mathcal{X} : \widehat{M}|f^t|^\beta > t^\beta\} \leqslant 9^n t^{-\beta} \int_{\mathcal{X}} |f^t|^\beta \, d\mu \end{aligned}$$

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

On the other hand, we have

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

So, finally we get

$$\mu\{x \in \mathcal{X} : \widehat{M}_\beta f > 2^{\frac{1}{\beta}} t\} \leqslant 9^n \tfrac{1}{1-\beta} \tfrac{1}{t} \|f\|_{L^{1,\infty}(\mu)},$$

i.e.,

$$\|\widehat{M}_\beta f\|_{L^{1,\infty}} \leqslant \frac{9^n 2^{1/\beta}}{1-\beta} \|f\|_{L^{1,\infty}},$$

proving the lemma.

Comparison lemma.

Lemma 3.3.

Let $U : (0, +\infty) \to [0, +\infty)$ be a continuous nonnegative decreasing function. Let f be any nonnegative Borel measurable function on $\mathcal{X}$. Then for every $x \in \mathcal{X}$ and $R > 0$,

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

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 inequality

$$\int_{B(x,T)\setminus B(x,R)} f\,d\mu \leqslant \widetilde{M}f(x)\cdot T^n,$$

which immediately follows from the definition of $\widetilde{M}f$.

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

The Hoermander condition.

We will need one more standard observation about Calderon-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

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

where A_1 depends only on the dimension n and the constants A and ε in the definition of the Calderon-Zygmund kernel K.

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 applied to $R = 2\rho$, $f \equiv 1$ and $U(t) = \dfrac{\rho^\varepsilon}{t^{n+\varepsilon}}$,

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

§4. The Guy David lemma.

The following lemma is implicitly contained in [David].

Lemma 4.1.

For any Borel set $F \in \mathcal{X}$ and for any point $x \in \operatorname{supp}\mu$,

$$T^\sharp \chi_F(x) \leqslant \widehat{M} T \chi_F(x) + A_2$$

where A_2 depends only on the dimension n, the constants A and ε in the definition of the Calderon-Zygmund kernel K, and the norm $\|T\|_{L^2(\mu)\to L^2(\mu)}$ (the condition $x \in \operatorname{supp}\mu$ is imposed only to have the maximal function $\widehat{M} T\chi_F(x)$ well-defined at the point x).

Proof:

Let $x \in \mathcal{X}, r > 0$. Consider the sequence of balls $B(x, r_j)$ where $r_j := 3^j r$ and the corresponding sequence of measures $\mu_j := \mu(B(x, r_j))$ $(j = 0, 1, \ldots)$. As we have seen above, it is impossible to have $\mu_j > 9^n \mu_{j-1}$ for every $j \geqslant 1$. Therefore there exists the smallest positive integer k for which $\mu_k \leqslant 9^n \mu_{k-1}$. Put $R := r_{k-1} = 3^{k-1} r$. Note that

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

Now recall that $|K(x,y)| \leqslant \frac{A}{\operatorname{dist}(x,y)^n}$ and therefore $\mathcal{I}_j \leqslant A \frac{\mu_j}{r_{j-1}^n}$ for every $j = 1, \ldots, k$. Note that $\mu_j \leqslant 9^{n(j+1-k)}\mu_k$ and $r_j = 3^{j-1-k} r_k$ for $j = 1, \ldots, k$. Hence

$$\sum_{j=1}^{k} \mathcal{I}_j \leqslant A \sum_{j=1}^{k} \frac{\mu_j}{r_{j-1}^n} \leqslant A \cdot 27^n \frac{\mu_k}{r_k^n} \sum_{j=1}^{k} 3^{n(j-k)} \leqslant \frac{A \cdot 27^n}{1 - 3^{-n}}$$

(because $\mu_k = \mu(B(x, r_k) \leqslant r_k^n)$.

And that is basically the main part of the reasoning, because now it is enough to pick up any standard proof based on the doubling condition to get the desired estimate for $T_{3R}\chi_F(x)$ (recall that $\mu(B(x,3R) \leqslant 9^n \mu(B(x,R))\,!)$.

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 $\widehat{M} T \chi_F(x)$).

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

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

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

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

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 don't know of any other paper in which it is used).

§5. An alternative to the Calderon-Zygmund decomposition.

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

$$\nu=\sum_{j=1}^N\alpha_j\delta_{x_j}.$$

Theorem 5.1.

$$\mu\{x\in\mathcal{X}\,:\,|T\nu(x)|>t\}\leqslant\frac{A_4\|\nu\|}{t}$$

with some A_4 depending only on the dimension n, the constants A and ε in the definition of the Calderon-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 $\sum_{j=1}^N\alpha_jK(x,x_j)$, which makes sense everywhere except finitely many points.

Proof:

Without loss of generality, we may assume that $||\nu||=\sum_j\alpha_j=1$ (this is just a matter of normalization) Thus we have to prove that

$$\mu\{|T\nu|>t\}\leqslant\frac{A_4}{t}.$$

If $\mu(\mathcal{X})\leqslant\frac{1}{t}$, there is nothing to do. Otherwise let $B(x_1,r_1)$ be the smallest (closed) ball such that $\mu(B(x_1,r_1))\geqslant\dfrac{\alpha_1}{t}$ (since the function $r\to\mu(B(x,r))$ is increasing and continuous from the right, tends to 0 as $r\to 0$, and is greater than $\dfrac{1}{t}\geqslant\dfrac{\alpha_1}{t}$ for sufficiently large $r>0$, such r_1 exists and is strictly positive).

Note that for the corresponding *open* ball $B'(x_1,r_1):=\{y\in\mathcal{X}\,:\,\operatorname{dist}(x,y)<r_1\}$, we have $\mu(B'(x_1,r_1))=\lim_{r\to r_1-0}\mu(B(x,r))\leqslant\dfrac{\alpha_1}{t}$. Since the measure μ is non-atomic, one can choose a Borel set E_1 satisfying

$$B'(x_1,r_1)\subset E_1\subset B(x_1,r_1)\qquad\text{and}\qquad\mu(E_1)=\frac{\alpha_1}{t}.$$

Let $B(x_2, r_2)$ be the smallest ball such that $\mu(B(x_2, r_2)\setminus E_1) \geqslant \dfrac{\alpha_2}{t}$ (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 $B'(x_2, r_2)$, we have $\mu(B'(x_2, r_2) \setminus E_1) \leqslant \dfrac{\alpha_2}{t}$, and therefore there exists a Borel set E_2, satisfying

$$B'(x_2, r_2) \setminus E_1 \subset E_2 \subset B(x_1, r_1) \setminus E_1 \qquad \text{and} \qquad \mu(E_2) = \frac{\alpha_2}{t}.$$

In general, for $j = 3, 4, \ldots, N$, let $B(x_j, r_j)$ be the smallest ball such that

$$\mu\Big(B(x_j, r_j) \setminus \bigcup_{\ell=1}^{j-1} E_\ell\Big) \geqslant \frac{\alpha_j}{t},$$

and let E_j be a Borel set satisfying

$$B'(x_j, r_j) \setminus \bigcup_{\ell=1}^{j-1} E_\ell \subset E_j \subset B(x_j, r_j) \setminus \bigcup_{\ell=1}^{j-1} E_\ell \quad \text{and} \quad \mu(E_j) = \frac{\alpha_j}{t}.$$

Let $E = \bigcup_j E_j$. Clearly

$$\bigcup_j B'(x_j, r_j) \subset E \subset \bigcup_j B(x_j, r_j) \qquad \text{and} \qquad \mu(E) = \frac{1}{t}.$$

Now let us compare $T\nu$ to $t\sum_j \chi_{\mathcal{X}\setminus B(x_j,2r_j)} \cdot T\chi_{E_j} =: t\sigma$ outside E. We have

$$T\nu - t\sigma = \sum_j \varphi_j$$

where

$$\varphi_j = \alpha_j T\delta_{x_j} - t\chi_{\mathcal{X}\setminus B(x_j,2r_j)} \cdot T\chi_{E_j}.$$

Note now that

$$\int_{\mathcal{X}\setminus E} |\varphi_j| d\mu \leqslant \int_{\mathcal{X}\setminus B(x_j,2r_j)} |T[\alpha_j\delta_{x_j} - t\chi_{E_j} d\mu]| d\mu + \int_{B(x_j,2r_j)\setminus B'(x_j,r_j)} \alpha_j |T\delta_{x_j}| d\mu.$$

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

$$A_1 || \alpha_j \delta_{x_j} - t\chi_{E_j} d\mu || = 2A_1\alpha_j,$$

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

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

and thereby the absolute value of the difference $T\nu - t\sigma$ does not exceed $(2A_1+2^nA)t$ everywhere on $\mathcal{X}\setminus E$, except, maybe, a set of measure $\frac{1}{t}$. To accomplish the proof of the theorem, it is enough to show that for sufficiently large A_3

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

(then we may take $A_4=4(A_3+2A_1+2^nA)$).

We will do it by the standard Stein-Weiss duality trick. Assume that the inverse inequality 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

$$\sum_j\int_{\mathcal{X}}[T\chi_{E_j}]\cdot\chi_{F\setminus B(x_j,2r_j)}\,d\mu=\sum_j\int_{\mathcal{X}}\chi_{E_j}\cdot[T^*\chi_{F\setminus B(x_j,2r_j)}]\,d\mu$$

Note that for every point $x\in E_j\subset B(x_j,r_j)$

$$\begin{aligned}|T^*\chi_{F\setminus B(x_j,2r_j)}(x)-T^*\chi_{F\setminus B(x,r_j)}(x)|&=\Big|\int_{B(x_j,2r_j)\setminus B(x,r_j)}K(y,x)d\mu(y)\Big|\leqslant\\&\leqslant Ar_j^{-n}\mu(B(x_j,2r_j))\leqslant 2^nA\end{aligned}$$

and thereby

$$|T^*\chi_{F\setminus B(x_j,2r_j)}(x)|\leqslant(T^*)^\sharp\chi_F(x)+2^nA\leqslant\widehat{M}T^*\chi_F(x)+A_2+2^nA$$

according to the Guy David lemma. Hence

$$\int_{\mathcal{X}}\sigma\chi_F\,d\mu\leqslant(A_2+2^nA)\mu(E)+\int_{\mathcal{X}}\chi_E\cdot\widehat{M}T^*\chi_F\,d\mu.$$

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

$$\|\chi_E\|_{L^2(\mu)}\|\widehat{M}T^*\chi_F\|_{L^2(\mu)}\leqslant\frac{1}{t}\|\widehat{M}\|_{L^2(\mu)\to L^2(\mu)}\|T^*\|_{L^2(\mu)\to L^2(\mu)}.$$

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

$$A_3=A_2+2^nA+\|\widehat{M}\|_{L^2(\mu)\to L^2(\mu)}\|T\|_{L^2(\mu)\to L^2(\mu)}$$

to get a contradiction. Since the norm $\|\widehat{M}\|_{L^2(\mu)\to L^2(\mu)}$ is finite and depends on the dimension n only, we are done.

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