STUDIA MATHEMATICA
Managing Editors: Z. Ciesielski, W. Orlicz (Editor-in-Chief), A. Pełczyński, W. Żelazko

The journal prints original papers in English, French, German and Russian, mainly on functional analysis, abstract methods of mathematical analysis and on the theory of probabilities. Usually 3 issues constitute a volume.

The papers submitted should be typed on one side only and accompanied by abstracts, normally not exceeding 300 words. The authors are requested to send two copies, one of them being the typed, not Xerox copy. Authors are advised to retain a copy of the paper submitted for publication.

Manuscripts and the correspondence concerning editorial work should be addressed to:

STUDIA MATHEMATICA
ul. Śniadeckich 8, 00-950 Warszawa, Poland

Correspondence concerning exchange should be addressed to:
INSTITUTE OF MATHEMATICS, POLISH ACADEMY OF SCIENCES
ul. Śniadeckich 8
00-950 Warszawa, Poland

The journal is available at your bookseller or at

ARPS POLONA
Kraków, ul. Przedmieście 7
00-068 Warszawa, Poland

© Copyright by Państwowe Wydawnictwo Naukowe, Warszawa, 1978

PRINTED IN POLAND

WROCŁAWSKA DRUKARZNA NAUKOWA

STUDIA MATHEMATICA, T. LXII. (1978)

Pointwise estimates for commutator singular integrals

by

B. M. BAISHANSKI* and R.R. COIFMAN** (St. Louis, Mo.)

Abstract. Certain weak type estimates are proved for the commutator integral of Calderón in \( R^n \). These extend previous results obtained by the rotation method.

1. In this note we prove pointwise and weak-type estimates for the maximal operator of the commutator singular integral of Calderón [2]. The characteristic \( \Omega \) of the singular integral is assumed here to satisfy Lipschitz condition. This enables us to obtain weak-type estimates for the limiting case (case \( q = 1 \) of the theorem). These estimates cannot be obtained by the rotation method as used in [1].

The results of this paper and some methods used here can be applied to obtain refinements and extensions of the original result of Calderón [2], that will be published in a forthcoming paper.

We consistently use the following notation. Points in \( R^n \) are denoted by \( x, y, t \); the coordinates of the point \( x \) are \( x^{(i)} = 1, 2, \ldots, n; \delta, \varepsilon, \lambda \) are arbitrary positive numbers. The ball with center \( x \) and radius \( \delta \) is denoted by \( B(x, \delta) \); \( \chi_a \) or \( \chi_\delta \) is the characteristic function of that ball; the Lebesgue measure of a set \( E \) in \( R^n \) is denoted by \( |E| \); in particular, \( |S(\delta)| \) is the \( n \)-dimensional volume of a ball of radius \( \delta \). The element of the surface area is denoted \( da(a) \). A cube in \( R^n \) will always mean a cube all the edges of which are parallel to the coordinate axes. If \( Q \) is a cube, \( Q \) denotes the cube concentric with \( Q \) and with diameter twice the diameter of \( Q \). By \( p, q, r \) we denote real numbers satisfying \( 1 \leq p < \infty \) and \( 1 \leq q < \infty, 1 < r < \infty \) and

\[
\int_1^q \frac{1}{q} = \frac{1}{p} + \frac{1}{r},
\]

* Research supported by N.S.F. Grant GP-8909.
** Research supported by U.S. Army Grant DA-ARO-D-31-122-12-G143.
by \( C \) we denote any expression which depends only on \( p, r \) and the dimension \( n \); \( f \) is a function in \( L^p(\mathbb{R}^n) \), \( a \) a function on \( \mathbb{R}^n \) such that \( \text{grad} \ a \in L^r(\mathbb{R}^n) \). By \( A_r(f) \) we denote the Hardy–Littlewood maximal function of \( |f|^r \), i.e.,

\[
A_r(f)(x) = \sup \left\{ \frac{1}{|B(0)|} \int_{B(x,r)} |f(t)|^r \, dt \right\}^{1/r}
\]

similar expression defined \( A_r(\text{grad} \ a) \) if \( r < \infty \), we shall use the convention \( A_r(\text{grad} \ a)(x) = |\text{grad} \ a(x)|_\infty \) for every \( x \). We write \( A \) instead of \( A_1 \).

The function \( D \) defined on \( \mathbb{R}^n - \{0\} \) is assumed to be homogenous of degree zero, to satisfy

\[
\int |x|^{i-1} \frac{\partial^i \Omega(x)}{\partial x_i} \, dx = 0, \quad i = 1, 2, \ldots, n,
\]

and the Lipschitz condition. So, without restricting generality, we assume

\[
|\Omega(x) - \Omega(y)| \leq |x - y| \quad \text{for} \quad |x| = |y| = 1
\]

and

\[
|\Omega(x)| \leq 1 \quad \text{for} \quad |x| = 1.
\]

The operator \( T(a, f) \), which we occasionally write \( T(f) \), is defined by

\[
T(a, f)(x) = \int_{\mathbb{R}^n} \frac{a(x) - a(y)}{|x|^{n+1}} \Omega(x - y)f(y) \, dy.
\]

It is not difficult to verify that, for almost every \( x \) in \( \mathbb{R}^n \), \( T(a, f)(x) \) is defined and finite for every \( \varepsilon \). (For a proof, see Section 4 of [1].) By \( T(a, f)(x) \) we denote \( \lim_{\varepsilon \to 0} T(a, f)(x) \) which exists a.e. for \( q > 1 \). Finally, we set

\[
M(a, f)(x) = \sup_{\varepsilon > 0} |T(a, f)(x)|.
\]

With the notation just introduced we can state our result.

Theorem. The following inequalities hold:

\[ \text{P)} \quad M(a, f) < A(T(a, f)) + C|\text{grad} \ a|A_r(f)(x) \]

almost everywhere if \( q > 1 \);

\[ \text{W)} \quad |\Omega| M(a, f) > \lambda \leq C\left(\frac{|\text{grad} \ a|}{\lambda}\right)^q \]

if \( q > 1 \).

The letters \( \text{P}, \text{W} \) stand for pointwise, respectively weak-type. The inequality (W) for the case \( q > 1 \), and even the stronger inequality

\[ \text{(1.4) } \quad M(a, f) \leq C|\text{grad} \ a|_{L^q} f_{L^q}, \quad \text{if} \quad q > 1, \]

has been proved by us earlier [1], and will be used in the proof of the inequality (W) for the case \( q = 1 \). We shall also appeal repeatedly to Calderón’s theorem [2]: if \( q > 1 \), then \( T(a, f) \) tends to a limit \( T(a, f) \) in \( L^q(\mathbb{R}^n) \) as \( \varepsilon \to 0 \), and the commutator singular integral \( T(a, f) \) satisfies

\[ |T(a, f)|_{L^q} \leq C|\text{grad} \ a|_{L^q} f_{L^q}. \]

The auxiliary results needed for the proof of the theorem are in Sections 2, 3, 4. In Section 5 we prove (P). There we adapt the method introduced by Cotlar [3] in the study of the maximal operator for the Hilbert transform. The main difficulty there was lying in the fact that the operator \( D \), defined by

\[ D(a, f)(x) = \sup_{y \in \mathbb{R}^n} \frac{|a(x) - a(y)|}{|x - y|}, \]

is known to satisfy \( |D(a, f)|_{L^q} \leq C|\text{grad} \ a| \), only for \( r > n \). We circumvent this difficulty by using the operators

\[ \sup_{\varepsilon > 0} \left( \frac{1}{|B(x, \varepsilon)|} \int_{B(x, \varepsilon)} |a(x) - a(t)|^q \, dt \right)^{1/q} \]

of Lemma 1 instead of the operator \( D \).

Finally, the proof of (W) for the remaining case \( q = 1 \) is given in Section 6.

Remark 1. It is easy to see that, in the case \( q > 1 \), (W) follows from (P). To prove that, we observe first that by Hardy–Littlewood maximal theorem and by Calderón’s theorem (1.6)

\[ |\Omega| M(T(a, f)) \leq C|T(a, f)|_{L^q} \leq C|\text{grad} \ a|_{L^q} f_{L^q}, \]

which gives the desired bound for \( |\Omega| M(T(a, f)) > \lambda, |\Omega| \text{grad} \ a|_{L^q} f_{L^q} \), and, second, that for any positive \( \gamma \),

\[ \text{if} \quad |\Omega| M(T(a, f)) > \lambda \]

and by Hardy–Littlewood maximal theorem each of the last two sets is easily seen to have measure \( \leq C \left( \frac{1}{\lambda} \right)^{1/q} f_{L^q} \).

Remark 2. In this paper we appeal twice to the main result of [1], a result which has rather involved proof: in the proof of (P) we need

\[ T(a, f)(x) \text{ converges a.e. as } \varepsilon \to 0 \text{ if } q > 1, \]

and in the proof of (W) for the case \( q = 1 \) we need (W) for the case \( q > 1 \).
It is possible to remove this dependence on [1] if we assume that $\Omega$ is continuously differentiable. In that case it is relatively easy to show that, for functions $f$ with compact support and satisfying a Lipschitz condition, (1.6) holds. (A proof is given in Section 5 of [1].) From Remark 1 it follows that once we have a (P)-inequality for a function $f$, the (W)-inequality for the same $f$ holds. The restrictions imposed on $f$ are then removed by a standard approximation argument. Moreover, (1.4) follows by the Marcinkiewicz interpolation theorem.

**Remark 3.** A careful analysis of the proofs will show that the smoothness condition (1.3) is not necessary, the only smoothness condition we need is

$$\frac{1}{\delta^a} \int_{|x|<\delta} |Q(x) - Q(x+u)| \, du \leq C \frac{\delta}{|x|} \quad \text{for } |x| > \delta. \quad (1.7)$$

**2.** We introduce two different maximal operators, closely related to the Hardy–Littlewood maximal operator. If $\grad a \in L^p(\mathbb{R}^n)$, $1 < r < \infty$, these operators are defined by

$$\tilde{A}_p(x)(x) = \sup_{Q \in \mathcal{Q}} \left( \frac{1}{|Q|} \int_{Q} |a(x) - a(t)|^p \, dt \right)^{1/p},$$

where $1 < p < \infty$ and supremum is taken over all the cubes $Q$ which contain the point $x$, and

$$A(x)(x) = \sup_{0 < \delta < |x|} \int_{|y| < \delta} \frac{|a(x) - a(y)|}{|x-y|^{n+2}} \, dy.$$

**Lemma 1.** There exist constants $C_p$, and $C_\delta$ such that

$$\tilde{A}_p(x)(x) \leq C_{p,\delta} \tilde{A}_p(\grad a)(x), \quad (2.1)$$

$$A(x)(x) \leq C_{\delta} A(x)(x). \quad (2.2)$$

**Proof.** Let $Q$ be a cube containing the point $x$ and $\delta$ the radius of the smallest ball with center at $x$ containing $Q$. Then

$$\frac{1}{|Q|} \int_{Q} |a(x) - a(t)|^p \, dt \leq \frac{C_p}{|x|} \int_{|y|<\delta} \frac{|a(x) - a(y)|^p}{|x-y|^{n+2}} \, dy.$$

and observe that

$$\frac{1}{|x|} \int_{|y|<\delta} \frac{|a(x) - a(y)|^p}{|x-y|^{n+2}} \, dy \leq \frac{C_p}{|x|} \int_{|y|<\delta} \frac{|a(x) - a(t)|^p}{|x-t|^{n+2}} \, dt.$$

We shall prove (2.1) by showing that

$$\frac{1}{|Q|} \int_{Q} |a(x) - a(t)|^p \, dt \leq C_{p,\delta} \tilde{A}_p(\grad a)(x). \quad (2.3)$$

Since

$$\frac{|a(x) - a(x+u)|}{|x|} \leq \frac{\int |\grad a(x+\lambda u)| \, d\lambda}{|x|},$$

we have

$$\frac{1}{|x|} \int_{|x|<\delta} \frac{|a(x) - a(x+u)|}{|x|} \, du \leq \frac{1}{|x|} \int_{|x|<\delta} \frac{|\grad a(x+y)|}{|y|^{n+2}} \, dy.$$

Writting

$$\frac{1}{r} \int_{|x|<\delta} \frac{1}{r} \int_{|x|<\delta} \frac{|a(x) - a(x+ru)|}{|x|} \, du \leq \frac{1}{r} \int_{|x|<\delta} \frac{|\grad a(x+y)|}{|y|^{n+2}} \, dy,$$

and observing that

$$\frac{1}{r} \int_{|x|<\delta} \frac{|\grad a(x+y)|}{|y|^{n+2}} \, dy \leq \frac{1}{r} \int_{|x|<\delta} \frac{|\grad a(x+y)|}{|y|^{n+2}} \, dy,$$

we deduce from (2.4) that (2.3) holds.

To prove (2.2) we write

$$\int_{|y|<\delta} \frac{|a(x) - a(y)|}{|x-y|^{n+2}} \, dy = \sum_{k=0}^{\infty} \int_{|y|<\delta} \frac{|a(x) - a(y)|}{|x-y|^{n+2}} \, dy,$$

and observe that

$$\int_{|y|<\delta} \frac{|a(x) - a(y)|}{|x-y|^{n+2}} \, dy \leq \sum_{k=0}^{\infty} \int_{|y|<\delta} \frac{|a(x) - a(y)|}{|x-y|^{n+2}} \, dy.$$

3. **Lemma 2.** Let $a$ be a function on $\mathbb{R}^n$, such that $\grad a \in L^p(\mathbb{R}^n)$, $1 < r < \infty$, let $a$ be a point in $\mathbb{R}^n$, and $\delta$ a positive number. Then there exists a func-
Riesz $A$ with support in $S(a, 2\delta)$ such that

(i) $\mathcal{A}(u) - \mathcal{A}(t) = a(u) - a(t)$ for $u, t \in S(a, \delta)$

and

(ii) $\|\text{grad} \mathcal{A}\|_L^p \leq C^{\omega_1} \mathcal{A}(\text{grad} a)(u)$.

Proof. Let $\phi$ be an infinitely differentiable function on $\mathbb{R}^n$ satisfying the following conditions: $0 \leq \phi(u) \leq 1$ for every $u$; $\phi(u) = 1$ for $u \in S(a, \delta)$; $\phi(u) = 0$ for $u \notin S(a, 2\delta)$; $|\text{grad} \phi(u)| \leq 2/\delta$ for every $u$. Set

$$
\mathcal{A}(u) = \left[ a(u) - \frac{1}{\delta^p} \int_{S(\delta)} a(t)dt \right] \phi(u).
$$

Then (i) is obviously satisfied, and so is (ii) in the case $r = \infty$.

Since $\|\text{grad} \phi\| \leq \frac{2}{\delta} 2\pi$, we have

$$
\|\text{grad} \mathcal{A}\|_{L^p} \leq \|\text{grad} a(u)\|_{L^p} |\phi(u)| + \left| a(u) - \frac{1}{\delta^p} \int_{S(\delta)} a(t)dt \right| \|\text{grad} \phi(u)\|,
$$

$$
\leq \|\text{grad} a(u)\|_{L^p} \|\phi(u)|_{L^p} + |a(u) - a(0)| \frac{2}{\delta} \chi_{2\delta}(u) +
$$

$$
+ \frac{2}{\delta^p} \int_{S(\delta)} \left| a(t) - a(0) \right| dt \cdot \chi_{2\delta}(u).
$$

To prove (ii) it is sufficient to show that the $r$th norm of each of the three summands in the last expression is $\leq C^{\omega_1} \mathcal{A}(\text{grad} a)(u)$.

For the first term this is obvious since

$$
\|\text{grad} \mathcal{A}\|_{L^p} \leq \int_{|x| \leq 2\delta} |\text{grad} a(u)|^p du \leq C^{\omega_1} \mathcal{A}(\text{grad} a)(u)
$$

For the second term we have by Lemma 1

$$
\frac{1}{\delta} \left\| a(t) - a(0) \right\|_{L^p} \leq \frac{1}{\delta} \left( \int_{|x| \leq 2\delta} \left| a(u) - a(0) \right|^p du \right)^{1/p}
$$

$$
\leq \frac{1}{\delta} \left( \frac{2\delta}{\delta^p} \int_{|x| \leq 2\delta} \left| a(u) - a(0) \right|^p du \right)^{1/p}
$$

$$
\leq C^{\omega_1} \mathcal{A}(\text{grad} a)(u).
$$

For the last term which we denote by $l(u)$, we obtain from Hölder's inequality and Lemma 1:

$$
|l(u)| \leq \frac{2}{\delta} \left( \frac{1}{\delta^p} \int_{S(\delta)} \left| a(t) - a(0) \right|^p dt \right)^{1/p} \chi_{2\delta}(u)
$$

$$
\leq \frac{2}{\delta} \left( \frac{1}{\delta^p} \int_{S(\delta)} \left| a(t) - a(0) \right|^p \right)^{1/p} \chi_{2\delta}(u)
$$

$$
\leq 2\mathcal{A}(\text{grad} a)(u) \chi_{2\delta}(u),
$$

so

$$
\|l\|_r \leq C\mathcal{A}(\text{grad} a)(u) \|\chi_{2\delta}\|_r \leq C^{\omega_1} \mathcal{A}(\text{grad} a)(u).
$$

4. The proof of (W) will be based on the following lemma, which is an extension of a well-known result of Calderón and Zygmund.

**Lemma.** If $S$ is a sublinear operator of weak-type $(p, q)$, a sufficient condition that $S$ be also of weak-type $(p, q)$, where $1/p - 1/q = 1/p_0 - 1/q_0$, $p_0 > p > 1$, is that for every sequence of pairwise disjoint cubes $Q_i$, and every function $h$ in $L^p(\mathbb{R}^n)$ having support in $\bigcup Q_i$ and such that

$$
\int_{Q_i} h(x) dx = 0 \quad \text{for every } i,
$$

the following estimate holds

$$
|\{x| \text{ in } \mathbb{R}^n \setminus \bigcup Q_i, S(h)(x) > \lambda\}| \leq C(\lambda h_{p_0}/\lambda)^p,
$$

where $Q_i$ is the cube concentric with $Q_i$ and such that $\text{diam}(Q_i) = 2 \text{diam}(Q_i)$.

Proof. To show that $S$ is of weak-type $(p, q)$, i.e. that for every $f$ in $L^p$

(4.1) $|\{x| \text{ in } S(f)(x) > \lambda\}| \leq C(\lambda h_{p_0}/\lambda)^p$, it is sufficient in view of sublinearity of $S$ to prove (4.1) for $f$ in $L^p$ such that $f \geq 0$ and $\int f = 1$. Applying a well-known lemma of Calderón and Zygmund to the function $f^p$ we obtain, for each $\lambda > 0$, a sequence of pairwise disjoint cubes $Q_i$ such that

(4.2) $\lambda \leq \frac{1}{|Q_i|} \int_{Q_i} f^p dx \leq \lambda^{2^{n/p}}$

and

(4.3) $f^p(x) \leq \lambda^{2^{n/p}}$ for almost all $x$ in $\mathbb{R}^n \setminus \bigcup Q_i$.

From (4.3) we obtain

(4.4) $\sum_i |Q_i| \leq \frac{1}{\lambda^{2^{n/p}}} \sum_i \int_{Q_i} f^p(x) dx \leq \frac{1}{\lambda^{2^{n/p}}} \int f^p dx = \frac{1}{\lambda^{2^{n/p}}}$.


and
\[ \frac{1}{|Q_i|} \int_Q f(s) \, ds \leq \left( \frac{1}{|Q_i|} \int_Q |f(s)|^p \, ds \right)^{\frac{1}{p}} \leq (2^n R^n)^{\frac{1}{p}}. \]

Let
\[ g(x) = \begin{cases} \frac{1}{|Q_i|} \int_Q f(s) \, ds, & x \in Q_i, \\ f(s), & x \in \mathbb{R}^n - \bigcup Q_i, \end{cases} \]

and let \( h(x) = f(x) - g(x) \).

From the first inequality in (4.5) we have
\[ \int_{Q_i} g^p(s) \, ds \leq \int_{Q_i} f^p(s) \, ds, \]

hence
\[ \|g\|_{L^p} \leq \|f\|_{L^p}, \]

and thus
\[ \|h\|_{L^p} \leq 2. \]

Obviously, the support of \( h \) is contained in \( \bigcup Q_i \) and \( \int h(s) \, ds = 0 \) for every \( i \); so that using the assumption of Lemma 3 we have
\[ \|s \cap \mathbb{R}^n - \bigcup Q_i \| \leq C \|h\|_{L^p} \|s\|_{L^p} \leq 2C \|s\|_{L^p}. \]

On the other hand, (4.4) implies
\[ \|\bigcup Q_i\| \leq 2^n R^n, \]

which, together with (4.7), gives
\[ \|s \cap \mathbb{R}^n - \bigcup Q_i \| \leq C |s^1| \]

In view of the sublinearity of \( S \) and since \( f = g + h \), (4.1) will follow from the last inequality if we show that also
\[ \|s \cap \mathbb{R}^n - \bigcup Q_i \| \leq C |s^1| \]

By the assumption of Lemma 3, \( S \) is of weak-type \((p_1, q_1)\) which means that
\[ \|s \cap \mathbb{R}^n - \bigcup Q_i \| \leq C \|f\|_{L^p} \|s\|_{L^p}^{q_1}, \]

Using (4.3) and (4.5) we obtain
\[ g(x) \leq 2^n R^n \text{ a.e.} \]

which together with the assumptions \( p_1 > p \geq 1, 1/p - 1/q = 1/p_2 - 1/q_2 \) and the fact that \( \|g\|_{L^p} \leq 1 \) gives
\[ \|g\|_{L^p} \leq 2^n R^n \left( \int \left( \frac{g(s)}{|x-y|^{p-1}} \right)^p \, ds \right)^{\frac{1}{p}} \leq 2^n R^n \left( \int \left( \frac{g(s)}{|x-y|^{p-1}} \right)^p \, ds \right)^{\frac{1}{p}} \leq C |s^1| \]

From this and (4.9) the desired estimate (4.3) follows.

5. In this section we prove the pointwise inequality \((P)\). We fix \( \delta > 0 \) and \( x, y \in \mathbb{R}^n \) and such that \( S(x, f)(x) \) is defined and finite for every \( \delta > 0 \). (As was mentioned earlier, almost every point in \( \mathbb{R}^n \) has this property.) We denote by \( \chi \) the characteristic function of the ball \( B(x, \delta) \). Then, by the theorem we have proved in [1], for almost every point \( t \) in \( B(x, \delta) \), both \( T(f)(t) \) and \( T(g)(t) \) are defined. For such a point \( t \) we have
\[ T(f)(x) - T(f)(t) + T(g)(t) \]

\[ = \int_{|x-y|>\delta} f(y) \left[ \frac{a(x) - a(y)}{|x-y|^{n+1}} - O(x-y) \right] \Omega(x-y) \left[ \frac{1}{|x-y|^{n+1}} - \frac{1}{|t-y|^{n+1}} \right] \Omega(t-y) \right] \]

\[ = \int_{|x-y|>\delta} f(y) \sum_{i=1}^4 A_i(x, y, t) \, dy, \]

where
\[ A_1 = \frac{a(x) - a(y)}{|x-y|^{n+1}} - O(x-y), \]
\[ A_2 = a(x) - a(y) \left[ \frac{1}{|x-y|^{n+1}} - \frac{1}{|t-y|^{n+1}} \right] \Omega(t-y), \]
\[ A_3 = a(x) - a(t) \frac{1}{|t-y|^{n+1}} \Omega(t-y), \]

For \( |x-y| > \delta, |x-t| < \delta/2 \) we have
\[ \left| \frac{1}{|x-y|^{n+1}} - \frac{1}{|t-y|^{n+1}} \right| \leq C \frac{\delta}{|x-y|^{n+1}} \]

so that
\[ |A_1| \leq C \frac{\delta}{|x-y|^{n+1}} |a(x) - a(y)|. \]
It is also easy to verify
\[ |\mathcal{A}| \leq C \frac{\delta}{|x-y|^{n+1}} |a(x)-a(t)|. \]

From the last two inequalities and (5.1) we obtain
\[
|\mathcal{T}_2(f)(x)| \leq |\mathcal{T}(f)(t)| + |\mathcal{T}(x_0 f)(t)| + 
\begin{align*}
&+ \int_{|x-y| > \delta} |f(y)| \frac{|a(x)-a(y)|}{|x-y|^{n+1}} |\Omega(x-y) - \Omega(t-y)| dy + \\
&+ C \int_{|x-y| > \delta} \frac{\delta}{|x-y|^{n+1}} |a(x)-a(y)| f(y) dy + \\
&+ C \int_{|x-y| > \delta} \frac{\delta}{|x-y|^{n+1}} |a(x)-a(t)| f(y) dy.
\end{align*}
\]

We integrate both sides of the last inequality in $t$ over $S(x, \delta/2)$ and divide by $|S(\delta/2)|$. Observing that by (1.7)
\[
\frac{1}{|S(\delta/2)|} \int_{S(\delta/2)} |\Omega(x-y) - \Omega(t-y)| dt \leq C \frac{\delta}{|x-y|},
\]
that by Lemma 1
\[
\frac{1}{|S_{e/2, \delta}||x-y|} \frac{1}{|x-y|} dt \leq C \frac{1}{|x-y|},
\]
and that
\[
\frac{1}{|S(\delta/2)|} \int_{S(\delta/2)} |\mathcal{T}_2(f)(t)| dt \leq \mathcal{A}(\text{grad } a)(x),
\]
we have
\[
(5.2) \quad |\mathcal{T}_2(f)(x)| \leq C \left( \frac{1}{|x-y|^{n+1}} \right) |f(y)| dy.
\]

side of (5.3) is $\leq C \mathcal{A}(\text{grad } a)(x) A_p(f)(x)$, which will imply that $\mathcal{A}(\mathcal{T}(f)(x)) + C \mathcal{A}(|a(x)-a(y)| |f(y)| dy$, an expression not depending on $\delta$, is a majorant of $\mathcal{M}(a,f)(x)$. This will end the proof of the pointwise inequality.

We observe first that the expression
\[
\frac{1}{\delta^n} \int_{|x-y| < \delta/2} |\mathcal{T}_3(f)(t)| dt \leq 
\]
\[
- \frac{1}{\delta^n} \int_{|x-y| < \delta/2} \int_{|y-y'| < \delta} a(t)-a(y) \frac{|x-y'|}{|x-y|^{n+1}} \Omega(t-y)f(y) dy dt
\]
does not change its value if the function $a$ is replaced by the function $\mathcal{A}$ of Lemma 2. Using Hörmander's inequality, Calderón's theorem (1.5) and Lemma 2, we have
\[
\frac{1}{\delta^n} \int_{|x-y| < \delta/2} |\mathcal{T}_3(f)(t)| dt = \frac{1}{\delta^n} \int_{|x-y| < \delta/2} |\mathcal{T}(A, x_0 f)(t)| dt
\]
\[
\leq \frac{C}{\delta^n} \left( \int_{|x-y| < \delta/2} |\mathcal{T}(A, x_0 f)(t)| dt \right) \leq C \mathcal{A}(\text{grad } a)(x) \int_{|x-y| < \delta/2} |f(y)| dy
\]
\[
\leq C \mathcal{A}(\text{grad } a)(x) \left( \int_{|x-y| < \delta/2} |f(y)| dy \right)^p
\]
\[
\leq C \mathcal{A}(\text{grad } a)(x) A_p(f)(x),
\]
which is the desired estimate for the second term on the right-hand side of (5.5).

Since $g > 1$, we have $p' < r$ and so $A_p(\text{grad } a)(x) \leq A_r(\text{grad } a)(x)$. From this and from Lemma 1 we obtain for $i = 0, 1, \ldots$
\[
K_i(x) = \int_{|x-y| < \delta/2} |x-y|^{n+1} |f(y)| dy
\]
\[
\leq C \left( \frac{1}{|x-y|^{n+1}} \right) |f(y)| dy
\]
\[
\leq \frac{C}{|x-y|^{n+1}} \left( \int_{|x-y| < \delta/2} |f(y)| dy \right)^{r'}
\]
\[
\leq C \left( \frac{1}{|x-y|^{n+1}} \right) \left( \int_{|x-y| < \delta/2} |f(y)| dy \right)^{r'}
\]
\[
\leq C \mathcal{A}(\text{grad } a)(x) A_p(f)(x),
\]
We shall show that each of the last three summands on the right-hand side
\[
\left( \frac{1}{(2^l + 1)^n} \int_{|y| = 2^l - 0} |f(y)|^p \, dy \right)^{1/p} \leq 2^{-l} C A_p (\text{grad } a)(x) A_p (f)(x) \leq 2^{-l} C A_p (\text{grad } a)(x) A_p (f)(x) - \leq C A_p (\text{grad } a)(x) A_p (f)(x),
\]

Since
\[
\int_{|y| = 2^l - 0} \frac{|a(x) - a(y)|}{|x - y|^{n+1}} |f(y)| \, dy = \sum_i K_i(x) \leq C A_p (\text{grad } a)(x) A_p (f)(x),
\]

the desired estimate is proved for the third term on the right-hand side of (5.3).

Since
\[
\frac{2^{-l} \int_{|y| = 2^l - 0} |f(y)| \, dy}{|x - y|^{n+1}} \leq \sum_{i=0}^\infty \frac{2^{-l} \int_{|y| = 2^l - 0} |f(y)| \, dy}{|x - y|^{n+1}} \leq C \sum_{i=0}^\infty \frac{1}{|x - y|^{n+1}} A_p (f)(x),
\]

we obtain that, the last term on the right-hand side of (5.2) is \( \leq C A_p (\text{grad } a)(x) A_p (f)(x) \).

Since \( A_p (\text{grad } a)(x) \leq A_p (\text{grad } a)(x) \leq A_p (f)(x) \), the desired estimate holds for the last term too.

6. In this section we prove (W) for the case \( q = 1 \). We fix \( r, 1 < r \leq \infty \), and \( a, \text{grad } a \in L^p(\mathbb{R}^n) \), and apply Lemma 3 to the sublinear operator \( M(f) = M(a, f) \), which is known to be of weak-type \((p, q)\) if \( q > 1 \), \( 2q - 1 \leq p \leq 1 + \). We need only show that the condition of Lemma 3 is satisfied.

Let \( h \) be a function in \( L^p(\mathbb{R}^n) \) with the support contained in the union of pairwise disjoint cubes \( Q_i \) and let \( h(y) \, dy = 0 \) for every \( i \). We fix \( a \) in \( \mathbb{R}^n - \bigcup Q_i \), and \( \epsilon > 0 \). Let

\[
Q_i(x, r) = \{ y : Q_i \cap S(x, \epsilon) = \emptyset \},
\]

\[
J(x, \epsilon) = \{ i : Q_i \cap S(x, \epsilon) \neq \emptyset \}
\]

\( \epsilon > 0 \) and \( Q_i - S(x, \epsilon) \neq \emptyset \).

Then
\[
I(a, h)(x) = \sum_{i=1}^\infty \int_{Q_i(x, \epsilon)} \cdots \, dy = \sum_{i=1}^\infty \int_{Q_i(x, \epsilon)} \cdots \, dy
\]

where each of the integrands is \( \frac{|a(x) - a(y)|}{|x - y|^{n+1}} \cdot O(|x - y| h(y)) \).

We show first that
\[
\left( \sum_{i=1}^\infty \int_{Q_i(x, \epsilon)} \cdots \, dy \right)^{1/p} \leq C A_p (\text{grad } a)(x) A_p (h)(x).
\]

From (6.1) and the assumption \( x \in \mathbb{R}^n - \bigcup Q_i \), it is easily seen that there exist positive constants \( a, \beta \), dependent only on the dimension \( n \) such that

\[
Q_i = \{ y : a < |y - x| < \beta \epsilon \} \quad \text{for every} \quad i \in J(x, \epsilon).
\]

Since \( Q_i \)'s are pairwise disjoint, it follows that

\[
\sum_{i=1}^\infty \int_{Q_i(x, \epsilon)} \cdots \, dy \leq \int_{\mathbb{R}^n - \bigcup Q_i} \cdots \, dy
\]

\[
\leq C \int_{\mathbb{R}^n - \bigcup Q_i} \cdots \, dy
\]

which proves (6.3).

We show now that for \( i \in J(x, \epsilon) \),

\[
\left( \int_{Q_i(x, \epsilon)} \cdots \, dy \right)^{1/p} \leq C \int_{Q_i(x, \epsilon)} \cdots \, dy
\]

\[
+ C \left( \frac{1}{\epsilon} \int_{Q_i(x, \epsilon)} \cdots \, dy \right)^{1/p} \int_{Q_i(x, \epsilon)} \cdots \, dy
\]

\[
\leq C \int_{Q_i(x, \epsilon)} \cdots \, dy
\]
where \( y_i \) is the center and \( \delta_i \) the diameter of the cube \( Q_i \). The right-hand side of the last inequality we denote by \( A_4(x) \).

Let \( t \in Q_i \). Since \( \int_{Q_i} \lambda(y) dy = 0 \), we have

\[
\begin{align*}
(6.5) \quad \int_{Q_i} \frac{a(x) - a(y)}{|x - y|^{n+3}} \Omega(x - y) \lambda(y) dy & = \int_{Q_i} \left[ \frac{a(x) - a(y)}{|x - y|^{n+3}} - \frac{a(x) - a(t)}{|x - t|^{n+3}} - \frac{a(x) - a(y)}{|x - t|^{n+3}} \right] \lambda(y) dy \\
& \leq \int_{Q_i} \frac{|a(x) - a(y)|}{|x - y|^{n+3}} |\Omega(x - y) - \Omega(x - t)| |\lambda(y)| dy + \\
& \quad + \int_{Q_i} \frac{|a(x) - a(y)|}{|x - t|^{n+3}} |\lambda(y)| |\Omega(x - t)| dy + \\
& \quad + \int_{Q_i} \frac{|a(t) - a(y)|}{|x - t|^{n+3}} |\lambda(y)| |\Omega(x - t)| dy.
\end{align*}
\]

Observing that, for \( y, t \in Q_i, x \in \mathbb{R}^n - Q_i \),

\[
\frac{1}{|x - t|^{n+3}} \leq C \frac{\delta_i}{|x - y|^{n+3}},
\]

then integrating in \( t \) over \( Q_i \) both sides of (6.5), dividing by \(|Q_i|\), noticing that by (1.7)

\[
\frac{1}{|Q_i|} \int_{Q_i} |\Omega(x - y) - \Omega(x - t)| dt \leq C \frac{\delta_i}{|x - y|}
\]

and that by Lemma 1

\[
\frac{1}{|Q_i|} \int_{Q_i} \frac{|a(x) - a(t)|}{y - t} dt \leq A(\grad a)(y),
\]

we obtain (6.4).

It follows from (6.3), (6.3) and (6.4) that

\[
|T_* (a, \lambda)(x)| \leq C A_4(\grad a)(x) A_4(\lambda)(x) + \sum_{i \in B(x)} A_4(x)
\]

\[
\leq C A(\grad a)(x) A_4(\lambda)(x) + \sum_{i \in B(x)} A_4(x).
\]

The last expression is independent of \( x \) and so it is a majorant for \( M(a, \lambda)(x) \). The condition of Lemma 3 will be satisfied by \( M(a, f) \) if we show

\[
(6.6) \quad \left| \{ x \in \mathbb{R}^n, A_4(\grad a)(x) A_4(\lambda)(x) > \lambda \} \right| \leq C \frac{|\grad a|_p |\lambda|_p}{\lambda}
\]

and

\[
(6.7) \quad |\{ x \in \mathbb{R}^n - \bigcup Q_i, \sum A_4(x) > \lambda \} | \leq C \frac{|\grad a|_p |\lambda|_p}{\lambda}.
\]

Since (6.6) has really been proved in Remark 1, it remains only to prove (6.7). We have

\[
\sum_{i = 1}^{n} \int_{\mathbb{R}^n - Q_i} A_4(x) dx \leq \sum_{i = 1}^{n} \int_{\mathbb{R}^n - Q_i} A_4(x) dx.
\]

There exists a constant \( \gamma_i \), dependent only on the dimension \( n \), such that if \( x \notin Q_i, y \in Q_i \), then \( |x - y| > \gamma_i \). Thus, using (2.2),

\[
\int_{\mathbb{R}^n - Q_i} A_4(x) dx \leq C \int_{Q_i} \delta_i \left( \int_{|x - y| > \gamma_i} \frac{|a(x) - a(y)|}{|x - y|^{n+3}} |\lambda(y)| dy + \\
+ C \delta_i \int_{|x - y| < \gamma_i} \frac{dx}{|x - y|^{n+3}} \int_{Q_i} A(\grad a)(y) |\lambda(y)| dy \right)
\]

\[
\leq C \int_{Q_i} A(\grad a)(y) |\lambda(y)| dy.
\]

Since the cubes \( Q_i \) are disjoint, and \( 1/p + 1/r = 1 \), we obtain

\[
\sum_{i = 1}^{n} \int_{\mathbb{R}^n - Q_i} A_4(x) dx \leq C \int_{\mathbb{R}^n} A(\grad a)(y) |\lambda(y)| dy
\]

\[
\leq C |\grad a|_p |\lambda|_p,
\]

which implies (3.7). This ends the proof of the theorem.

**Added in Proof.** We call the reader’s attention to a recent paper by Calixto Calderón (see Studia Math. 59 (1976), pp. 93–105) in which similar estimates on commutators are obtained. The results there are for a different range of spaces \( D^p \) and involve different methods complementing our own.