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

$$\frac{1}{q} = \frac{1}{p} + \frac{1}{r}; \tag{1.1}$$

---

* Research supported by N.S.F. Grant GP-8959.
** Research supported by U.S. Army Grant DA-ARO-D-31-124-72-G143.

by C we denote any expression which depends only on p, r and the dimension n; f is a function in $L_p(\mathbf{R}^n)$, a—a function on $\mathbf{R}^n$ such that $\operatorname{grad} a \in L^r(\mathbf{R}^n)$. By $\Lambda_p^p(f)$ we denote the Hardy–Littlewood maximal function of $|f|^p$, i.e.

$$\Lambda_p(f)(x) = \sup_\delta \left(\frac{1}{|S(\delta)|} \int_{S(x,\delta)} |f(t)|^p\,dt\right)^{1/p}$$

similar expression defined $\Lambda_r(\operatorname{grad} a)$ if $r < \infty$, we shall use the convention $\Lambda_\infty(\operatorname{grad} a)(x) = \|\operatorname{grad} a\|_\infty$ for every x. We write Λ instead of Λ_1.

The function Ω defined on $\mathbf{R}^n - \{0\}$ is assumed to be homogenous of degree zero, to satisfy

$$\int_{|x|=1} x^{(i)}\,\Omega(x)\,d\sigma(x) = 0, \quad i = 1, 2, \ldots, n,$$

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

$$|\Omega(x) - \Omega(y)| \leqslant |x-y| \quad \text{for} \quad |x| = |y| = 1 \tag{1.2}$$

and

$$|\Omega(x)| \leqslant 1 \quad \text{for} \quad |x| = 1. \tag{1.3}$$

The operator $T_\varepsilon(a,f)$, which we occasionally write $T_\varepsilon(f)$, is defined by

$$T_\varepsilon(a,f)(x) = \int_{|y-x|>\varepsilon} \frac{a(x)-a(y)}{|x-y|^{n+1}}\,\Omega(x-y) f(y)\,dy.$$

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

$$M(a,f)(x) = \sup_\varepsilon |T_\varepsilon(a,f)(x)|.$$

With the notation just introduced we can state our result.

Theorem. *The following inequalities hold:*

(P) $$M(a,f)(x) \leqslant \Lambda\big(T(a,f)\big)(x) + C\Lambda_r(\operatorname{grad} a)(x)\,\Lambda_p(f)(x)$$

almost everywhere if $q > 1$;

(W) $$\big|\{x \mid M(a,f)(x) > \lambda\}\big| \leqslant C\left(\frac{\|\operatorname{grad} a\|_r \|f\|_p}{\lambda}\right)^q$$

if $q \geqslant 1$.

The letters P, W stand for pointwise, respectively weak-type.

The inequality (W) for the case $q > 1$, and even the stronger inequality

$$\|M(a,f)\|_q \leqslant C\|\operatorname{grad} a\|_r \|f\|_p \quad \text{if} \quad q > 1, \tag{1.4}$$

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 $T_\varepsilon(a,f)$ tends to a limit $T(a,f)$ in $L^q(\mathbf{R}^n)$ as $\varepsilon \to 0$, and the commutator singular integral $T(a,f)$ satisfies

$$\|T(a,f)\|_q \leqslant C\|\operatorname{grad} a\|_r \|f\|_p. \tag{1.5}$$

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

$$D(a)(x) = \sup_{y\in\mathbf{R}^n} \frac{|a(x)-a(y)|}{|x-y|},$$

is known to satisfy $\|D(a)\|_r \leqslant C\|\operatorname{grad} a\|_r$ only for $r > n$. We circumvent this difficulty by using the operators

$$\sup_\delta \left[\frac{1}{S(\delta)} \int_{S(x,\delta)} \left|\frac{a(x)-a(t)}{x-t}\right|^s dt\right]^{1/s}$$

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 (1.5)

$$\big\|\Lambda\big(T(a,f)\big)\big\|_q \leqslant C\|T(a,f)\|_q \leqslant C\|\operatorname{grad} a\|_r \|f\|_p,$$

which gives the desired bound for $\big|\{x \mid \Lambda\big(T(a,f)\big)(x) > \lambda/2\}\big|$, and, second, that for any positive γ,

$$\{x \mid \Lambda_r(\operatorname{grad} a)(x)\,\Lambda_p(f)(x) > \mu\} \subset \{x \mid \Lambda_r(\operatorname{grad} a)(x) > \gamma\mu^{q/r}\} \cup \left\{x \mid \Lambda_p(f)(x) > \frac{1}{\gamma}\mu^{q/p}\right\}$$

and by Hardy–Littlewood maximal theorem each of the last two sets is easily seen to have measure $\leqslant C\left(\frac{\|f\|_p \|\operatorname{grad} a\|_r}{\lambda}\right)$ if we choose $\gamma = \|\operatorname{grad} a\|_r^{q/p} \|f\|_p^{-q/r}$.

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

$$T_\varepsilon(a,f)(x) \text{ converges a.e. as } \varepsilon\to 0 \quad \text{if} \quad q > 1, \tag{1.6}$$

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 Ω 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 (1.2) is not necessary, the only smoothness condition we need is

$$\frac{1}{\delta^n}\int_{|u|<\frac{\delta}{2}}|\Omega(z)-\Omega(z+u)|\,du \leqslant C\frac{\delta}{|z|} \quad \text{for} \quad |z|>\delta. \tag{1.7}$$

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

$$\tilde{\Lambda}_p(a)(x)=\sup_Q\left[\frac{1}{|Q|}\int_Q\left|\frac{a(x)-a(t)}{x-t}\right|^p dt\right]^{1/p},$$

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

$$\tilde{\Lambda}(a)(x)=\sup_\delta \delta\int_{|y-x|>\delta}\frac{|a(x)-a(y)|}{|x-y|^{n+2}}\,dy.$$

LEMMA 1. *There exist constants $C_{p,n}$ and C_n such that*

$$\tilde{\Lambda}_p(a)(x)\leqslant C_{p,n}\Lambda_p(\operatorname{grad} a)(x), \tag{2.1}$$

$$\tilde{\Lambda}(a)(x)\leqslant C_n\tilde{\Lambda}_1(a)(x). \tag{2.2}$$

Proof. Let Q be a cube containing the point x, and δ the radius of the smallest ball with center at x containing Q. Then

$$\frac{1}{|Q|}\int_Q\left|\frac{a(x)-a(t)}{x-t}\right|^p dt\leqslant\frac{C_n}{\delta^n}\int_{|t-x|<\delta}\left|\frac{a(x)-a(t)}{x-t}\right|^p dt$$
$$\leqslant\frac{C^n}{\delta^n}\int_0^\delta r^{n-1}\,dr\int_{S_{n-1}}\left|\frac{a(x)-a(x+ru)}{r}\right|^p d\sigma(u).$$

We shall prove (2.1) by showing that

$$\int_{S_{n-1}}\left|\frac{a(x)-a(x+ru)}{r}\right|^p d\sigma(u)\leqslant C_n\Lambda_p^p(\operatorname{grad} a)(x). \tag{2.3}$$

Since

$$|a(x)-a(x+ru)|\leqslant\int_0^r|\operatorname{grad} a(x+\lambda u)|\,d\lambda$$
$$\leqslant r^{1-1/p}\Big(\int_0^r|\operatorname{grad} a(x+\lambda u)|^p d\lambda\Big)^{1/p},$$

we have

$$\int_{S_{n-1}}\left|\frac{a(x)-a(x+ru)}{r}\right|^p d\sigma(u)\leqslant\frac{1}{r}\int_{|y|\leqslant r}|\operatorname{grad} a(x+y)|^p\frac{dy}{|y|^{n-1}}. \tag{2.4}$$

Writting

$$\frac{1}{r}\int_{|y|\leqslant r}=\sum_{i=1}^\infty\frac{1}{r}\int_{2^{-i}r<|y|\leqslant 2^{-i+1}r}$$

and observing that

$$\frac{1}{r}\int_{2^{-i}r<|y|\leqslant 2^{-i+1}r}|\operatorname{grad} a(x+y)|^p\frac{dy}{|y|^{n-1}}$$
$$\leqslant\frac{2^{n-i}}{(2^{-i+1}r)^n}\int_{|y|\leqslant 2^{-i+1}r}|\operatorname{grad} a(x+y)|^p\,dy$$
$$\leqslant 2^{n-i}\sup_\varepsilon\frac{1}{\varepsilon^n}\int_{|y|\leqslant\varepsilon}|\operatorname{grad} a(x+y)|^p\,dy$$
$$\leqslant C_n 2^{-i}\Lambda_p^p(\operatorname{grad} a)(x),$$

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

To prove (2.2) we write

$$\delta\int_{|y-x|>\delta}\frac{|a(x)-a(y)|}{|x-y|^{n+2}}\,dy=\sum_{i=0}^\infty\delta\int_{2^i\delta<|y-x|\leqslant 2^{i+1}\delta}\frac{|a(x)-a(y)|}{|x-y|^{n+2}}\,dy$$

and observe that

$$\delta\int_{2^i\delta<|y-x|\leqslant 2^{i+1}\delta}\frac{|a(x)-a(y)|}{|x-y|^{n+2}}\,dy\leqslant\frac{2^{n-i}}{(2^{i+1}\delta)^n}\int_{|y-x|\leqslant 2^{i+1}\delta}\left|\frac{a(x)-a(y)}{x-y}\right|dy$$
$$\leqslant 2^{n-i}\sup_\varepsilon\frac{1}{\varepsilon^n}\int_{|y-x|\leqslant\varepsilon}\left|\frac{a(x)-a(y)}{x-y}\right|dy$$
$$\leqslant C_n 2^{-i}\tilde{\Lambda}_1(a)(x).$$

3. LEMMA 2. *Let a be a function on $\boldsymbol{R}^n$, such that $\operatorname{grad} a\in L^r(\boldsymbol{R}^n)$, $1<r\leqslant+\infty$, let x be a point in $\boldsymbol{R}^n$, and δ a positive number. Then there exists a func-*

tion A with support in $S(x, 2\delta)$ such that

(i) $A(u)-A(t) = a(u)-a(t)$ *for* $u, t \in S(x, \delta)$

and

(ii) $\|\operatorname{grad} A\|_r \leqslant C\delta^{n/r}\Lambda_r(\operatorname{grad} a)(x)$.

Proof. Let φ be an infinitely differentiable function on $\boldsymbol{R}^n$ satisfying the following conditions: $0 \leqslant \varphi(u) \leqslant 1$ for every u; $\varphi(u) = 1$ for $u \in S(x, \delta)$; $\varphi(u) = 0$ for $u \notin S(x, 2\delta)$; $|\operatorname{grad}\varphi(u)| \leqslant 2/\delta$ for every u. Set

$$A(u) = \left[a(u) - \frac{1}{|S(\delta)|}\int_{S(x,\delta)} a(t)\,dt\right]\varphi(u).$$

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

Since $\|\operatorname{grad}\varphi\| \leqslant \frac{2}{\delta}\chi_{2\delta}$, we have

$$|\operatorname{grad} A(u)| \leqslant |\operatorname{grad} a(u)|\,|\varphi(u)| + \left|a(u) - \frac{1}{S(\delta)}\int_{S(x,\delta)} a(t)\,dt\right| |\operatorname{grad}\varphi(u)|$$

$$\leqslant |\operatorname{grad} a(u)|\,\chi_{2\delta}(u) + |a(u)-a(x)|\,\frac{2}{\delta}\chi_{2\delta}(u) +$$

$$+\frac{2}{\delta|S(\delta)|}\int_{S(x,\delta)} |a(t)-a(x)|\,dt\cdot\chi_{2\delta}(u).$$

To prove (ii) it is sufficient to show that the rth norm of each of the three summands in the last expression is $\leqslant C\delta^{n/r}\Lambda_r(\operatorname{grad} a)(x)$.

For the first term this is obvious since

$$\|\operatorname{grad} a\cdot\chi_{2\delta}\|_r = \left(\int_{|u-x|\leqslant 2\delta} |\operatorname{grad} a(u)|^r\,du\right)^{1/r}$$

$$= C\delta^{n/r}\left(\frac{1}{S(2\delta)}\int_{S(x,2\delta)} |\operatorname{grad} a(u)|^r\right)^{1/r}.$$

For the second term we have by Lemma 1

$$\frac{1}{\delta}\big\|\big(a(\cdot)-a(x)\big)\cdot\chi_{2\delta}\big\|_r = \frac{1}{\delta}\left(\int_{|u-x|\leqslant 2\delta} |a(u)-a(x)|^r\,du\right)^{1/r}$$

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

$$\leqslant C\delta^{n/r}\Lambda\,(\operatorname{grad} a)(x).$$

For the last term which we denote by $l(u)$, we obtain from Hölder's

inequality and Lemma 1

$$|l(u)| \leqslant \frac{2}{\delta}\left(\frac{1}{S(\delta)}\int_{S(x,\delta)} |a(t)-a(x)|^r\,dt\right)^{1/r}\chi_{2\delta}(u)$$

$$\leqslant 2\left(\frac{1}{S(\delta)}\int_{S(x,\delta)}\left|\frac{a(t)-a(x)}{t-x}\right|^r dt\right)^{1/r}\chi_{2\delta}(u)$$

$$\leqslant 2\Lambda_r(\operatorname{grad} a)(x)\,\chi_{2\delta}(u),$$

so

$$\|l\|_r \leqslant C\Lambda_r(\operatorname{grad} a)(x)\,\|\chi_{2\delta}\|_r \leqslant C\delta^{n/r}\Lambda_r(\operatorname{grad} a)(x).$$

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

LEMMA 3. *If S is a sublinear operator of weak-type (p_0, q_0), a sufficient condition that S be also of weak-type (p, q), where $1/p - 1/q = 1/p_0 - 1/q_0$, $p_0 > p \geqslant 1$, is that for every sequence of pairwise disjoint cubes Q_i and every function h in $L^p(\boldsymbol{R}^n)$ 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

$$\left|\{x|\ x\in\boldsymbol{R}^n - \bigcup\overline{Q}_i,\ S(h)(x) > \lambda\}\right| \leqslant C(\|h\|_p/\lambda)^q,$$

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

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

$$\left|\{x|\ |S(f)(x)| > \lambda\}\right| \leqslant C(\|f\|_p/\lambda)^q, \tag{4.1}$$

it is sufficient in view of sublinearity of S to prove (4.1) for f in L^p such that $f \geqslant 0$ and $\|f\|_p = 1$. 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

$$\lambda^q \leqslant \frac{1}{|Q_i|}\int_{Q_i} f^p\,dx \leqslant 2^n\lambda^q \tag{4.2}$$

and

$$f^p(x) \leqslant \lambda^q \quad \text{for almost all } x \text{ in } \boldsymbol{R}^n - \bigcup Q_i. \tag{4.3}$$

From (4.2) we obtain

$$\sum |Q_i| \leqslant \frac{1}{\lambda^q}\sum\int_{Q_i} f^p(x)\,dx \leqslant \frac{1}{\lambda^q}\|f\|_p^p = \frac{1}{\lambda^q}, \tag{4.4}$$

and

(4.5) $$\frac{1}{|Q_i|}\int_{Q_i} f(x)\,dx \leqslant \left(\frac{1}{|Q_i|}\int_{Q_i} f^p(x)\,dx\right)^{1/p} \leqslant (2^n\lambda^q)^{1/p}.$$

Let

$$g(x) = \begin{cases} \dfrac{1}{|Q_i|}\displaystyle\int_{Q_i} f(x)\,dx, & x \in Q_i, \\ f(x), & x \in \boldsymbol{R}^n - \bigcup Q_i, \end{cases}$$

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

From the first inequality in (4.5) we have

$$\int_{Q_i} g^p(x)\,dx \leqslant \int_{Q_i} f^p(x)\,dx,$$

hence

$$\|g\|_p \leqslant \|f\|_p = 1$$

and thus

$$\|h\|_p \leqslant 2.$$

Obviously, the support of h is contained in $\bigcup Q_i$ and $\int_{Q_i} h(x)\,dx = 0$ for every i; so that using the assumption of Lemma 3 we have

(4.7) $$\left|\{x \mid x \in \boldsymbol{R}^n - \bigcup \overline{Q_i} \text{ and } |S(h)(x)| > \lambda\}\right| \leqslant C(\|h\|_p/\lambda)^q \leqslant 2^q C/\lambda^q.$$

On the other hand, (4.4) implies

$$|\bigcup \overline{Q}_i| \leqslant 2^n/\lambda^q,$$

which, together with (4.7), gives

$$\left|\{x| \ |S(h)(x)| > \lambda\}\right| \leqslant C/\lambda^q.$$

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

(4.8) $$\left|\{x| \ |S(g)(x)| > \lambda\}\right| \leqslant C/\lambda^q.$$

By the assumption of Lemma 3, S is of weak-type (p_0, q_0) which means that

(4.9) $$\left|\{x| \ |S(g)(x)| > \lambda\}\right| \leqslant C(\|g\|_{p_0}/\lambda)^{q_0}$$

Using (4.3) and (4.5) we obtain

$$g(x) \leqslant 2^n \lambda^{q/p} \text{ a.e.,}$$

which together with the assumptions $p_0 > p \geqslant 1$, $1/p - 1/q = 1/p_0 - 1/q_0$ and the fact that $\|g\|_p \leqslant 1$ gives

$$\begin{aligned} \|g\|_{p_0} &= 2^n\lambda^{q/p}\left(\int \left(\frac{g(x)}{2^n\lambda^{q/p}}\right)^{p_0} dx\right)^{1/p_0} \\ &\leqslant 2^n\lambda^{q/p}\left(\int \left(\frac{g(x)}{2^n\lambda^{q/p}}\right)^{p} dx\right)^{1/p_0} \\ &\leqslant C\lambda^{1-q/q_0}. \end{aligned}$$

From this and (4.9) the desired estimate (4.8) follows.

5. In this section we prove the pointwise inequality (P). We fix $\delta > 0$ and x, $x \in \boldsymbol{R}^n$ and such that $T_\varepsilon(a, f)(x)$ is defined and finite for every $\varepsilon > 0$. (As was mentioned earlier, almost every point in $\boldsymbol{R}^n$ has this property.) We denote by χ_δ the characteristic function of the ball $S(x, \delta)$. Then, by the theorem we have proved in [1], for almost every point t in $S(x, \delta/2)$ both $T(f)(t)$ and $T(\chi_\delta f)(t)$ are defined. For such a point t we have

(5.1) $$\begin{aligned} &T_\delta(f)(x) - T(f)(t) + T(\chi_\delta f)(t) \\ &= \int_{|x-y|>\delta} f(y)\left[\frac{a(x)-a(y)}{|x-y|^{n+1}}\,\Omega(x-y) - \frac{a(t)-a(y)}{|t-y|^{n+1}}\,\Omega(t-y)\right]dy \\ &= \int_{|x-y|>\delta} f(y) \sum_{i=1}^{3} \Delta_i(x, y, t)\,dy, \end{aligned}$$

where

$$\Delta_1 = \frac{a(x)-a(y)}{|x-y|^{n+1}}\,[\Omega(x-y) - \Omega(t-y)],$$

$$\Delta_2 = [a(x)-a(y)]\left[\frac{1}{|x-y|^{n+1}} - \frac{1}{|y-t|^{n+1}}\right]\Omega(t-y),$$

$$\Delta_3 = \frac{a(x)-a(t)}{|t-y|^{n+1}}\,\Omega(t-y).$$

For $|x-y| > \delta$, $|x-t| < \delta/2$ we have

$$\left|\frac{1}{|x-y|^{n+1}} - \frac{1}{|t-y|^{n+1}}\right| \leqslant C\frac{\delta}{|x-y|^{n+2}},$$

so that

$$|\Delta_2| \leqslant C\frac{\delta}{|x-y|^{n+2}}\,|a(x)-a(y)|.$$

It is also easy to verify

$$|\Delta_3| \leqslant C \frac{\delta}{|x-y|^{n+1}} \left| \frac{a(x)-a(t)}{x-t} \right|.$$

From the last two inequalities and (5.1) we obtain

$$|T_\delta(f)(x)| \leqslant |T(f)(t)| + |T(\chi_\delta f)(t)| +$$
$$+ \int_{|x-y|>\delta} |f(y)| \frac{|a(x)-a(y)|}{|x-y|^{n+1}} |\Omega(x-y)-\Omega(t-y)|\,dy +$$
$$+ C \int_{|x-y|>\delta} \frac{\delta}{|x-y|^{n+2}} |a(x)-a(y)|\ |f(y)|\,dy +$$
$$+ C \int_{|x-y|>\delta} \frac{\delta}{|x-y|^{n+1}} \left| \frac{a(x)-a(t)}{x-t} \right| |f(y)|\,dy.$$

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)

$$\frac{1}{S(\delta/2)} \int_{S(x,\delta/2)} |\Omega(x-y)-\Omega(t-y)|\,dt \leqslant C \frac{\delta}{|x-y|},$$

that by Lemma 1

$$\frac{1}{S(\delta/2)} \int_{S(x,\delta/2)} \left| \frac{a(x)-a(t)}{x-t} \right| dt \leqslant C\Lambda(\operatorname{grad} a)(x),$$

and that

$$\frac{1}{S(\delta/2)} \int_{S(x,\delta/2)} |T(f)(t)|\,dt \leqslant \Lambda\big(T(f)\big)(x),$$

we have

$$(5.2)\qquad |T_\delta(f)(x)| \leqslant \Lambda\big(T(f)\big)(x) + \frac{C}{\delta^n} \int_{|t-x|>\delta/2} |T(\chi_\delta f)(t)|\,dt +$$
$$+ C \int_{|x-y|>\delta} \frac{\delta}{|x-y|^{n+2}} |a(x)-a(y)|\ |f(y)|\,dy +$$
$$+ C\Lambda(\operatorname{grad} a)(x) \int_{|x-y|>\delta} \frac{\delta}{|x-y|^{n+1}} |f(y)|\,dy.$$

We shall show that each of the last three summands on the right-hand side of (5.2) is $\leqslant C\Lambda_r(\operatorname{grad} a)(x)\Lambda_p(f)(x)$, which will imply that $\Lambda\big(T(a,f)\big)(x) + C\Lambda_r(\operatorname{grad} a)(x)\Lambda_p(f)(x)$, an expression not depending on δ, is a majorant of $M(a,f)(x)$. This will end the proof of the pointwise inequality.

We observe first that the expression

$$\frac{1}{\delta^n} \int_{|t-x|<\delta/2} |T(\chi_\delta f)(t)|\,dt$$
$$= \frac{1}{\delta^n} \int_{|t-x|<\delta/2} \left| \int_{|y-x|<\delta} \frac{a(t)-a(y)}{|t-y|^{n+1}} \Omega(t-y) f(y)\,dy \right| dt$$

does not change its value if the function a is replaced by the function A of Lemma 2. Using Hölder's inequality, Calderón's theorem (1.5) and Lemma 2, we have

$$\frac{1}{\delta^n} \int_{|t-x|<\delta/2} |T(a, \chi_\delta f)(t)|\,dt = \frac{1}{\delta^n} \int_{|t-x|<\delta/2} |T(A, \chi_\delta f)(t)|\,dt$$
$$\leqslant \frac{C}{\delta^n} \delta^{n/q'} \|T(A, \chi_\delta f)\|_q \leqslant C\delta^{-n/q} \|\operatorname{grad} A\|_r \|\chi_\delta f\|_p$$
$$\leqslant C\delta^{n/r-n/q} \Lambda_r(\operatorname{grad} a)(x) \Big(\int_{|u-x|<\delta} |f(u)|^p\,du \Big)^{1/p}$$
$$\leqslant C\Lambda_r(\operatorname{grad} a)(x) \Big(\frac{1}{\delta^n} \int_{|u-x|<\delta} |f(u)|^p\,du \Big)^{1/p}$$
$$\leqslant C\Lambda_r(\operatorname{grad} a)(x)\Lambda_p(f)(x),$$

which is the desired estimate for the second term on the right-hand side of (5.2).

Since $q > 1$, we have $p' < r$ and so $\Lambda_{p'}(\operatorname{grad} a)(x) \leqslant \Lambda_r(\operatorname{grad} a)(x)$. From this and from Lemma 1 we obtain for $i = 0, 1, \ldots$

$$K_i(x) = \int_{2^i\delta<|x-y|\leqslant 2^{i+1}\delta} \delta \frac{|a(x)-a(y)|}{|x-y|^{n+2}} |f(y)|\,dy$$
$$\leqslant \frac{\delta}{(2^i\delta)^{n+1}} \int_{|x-y|\leqslant 2^{i+1}\delta} \left| \frac{a(x)-a(y)}{x-y} \right| |f(y)|\,dy$$
$$\leqslant \frac{2^{n-i}}{(2^{i+1}\delta)^n} \Big(\int_{|x-y|\leqslant 2^{i+1}\delta} \left| \frac{a(x)-a(y)}{x-y} \right|^{p'} dy \Big)^{1/p'} \Big(\int_{|x-y|\leqslant 2^{i+1}\delta} |f(y)|^p\,dy \Big)^{1/p}$$
$$\leqslant 2^{n-i} \Big(\frac{1}{(2^{i+1}\delta)^n} \int_{|x-y|\leqslant 2^{i+1}\delta} \left| \frac{a(x)-a(y)}{x-y} \right|^{p'} dy \Big)^{1/p'} \times$$

$$\times\left(\frac{1}{(2^{i+1}\delta)^n}\int_{|x-y|\leqslant 2^{i+1}\delta}|f(y)|^p\,dy\right)^{1/p}$$
$$\leqslant 2^{-i}C'\Lambda_{p'}(\operatorname{grad}a)(x)\Lambda_p(f)(x)$$
$$\leqslant 2^{-i}C\Lambda_r(\operatorname{grad}a)(x)\Lambda_p(f)(x).$$

Since

$$\int_{|x-y|>\delta}\delta\frac{|a(x)-a(y)|}{|x-y|^{n+2}}|f(y)|\,dy=\sum_i K_i(x)\leqslant C\Lambda_r(\operatorname{grad}a)(x)\Lambda_p(f)(x),$$

the desired estimate is proved for the third term on the right-hand side of (5.2).

Since

$$\delta\int_{|y-x|>\delta}\frac{|f(y)|}{|x-y|^{n+1}}\,dy=\sum_{i=0}^{\infty}\delta\int_{2^i\delta<|y-x|\leqslant 2^{i+1}\delta}\frac{|f(y)|}{|x-y|^{n+1}}\,dy$$
$$\leqslant\sum_{i=0}^{\infty}\frac{2^{n-i}}{(2^{i+1}\delta)^n}\int_{|y-x|\leqslant 2^{i+1}\delta}|f(y)|\,dy$$
$$\leqslant C\sum_{i=0}^{\infty}2^{-i}\Lambda(f)(x)$$
$$\leqslant C\Lambda(f)(x),$$

we obtain that the last term on the right-hand side of (5.2) is $\leqslant C\Lambda(\operatorname{grad}a)(x)\Lambda(f)(x)$. Since $\Lambda(\operatorname{grad}a)(x)\leqslant\Lambda_r(\operatorname{grad}a)(x)$, $\Lambda(f)(x)\leqslant\Lambda_p(f)(x)$, the desired estimate holds for the last term too.

6. In this section we prove (W) for the case $q=1$. We fix r, $1<r\leqslant\infty$, and a, $\operatorname{grad}a\in L^r(\boldsymbol{R}^n)$, and apply Lemma 3 to the sublinear operator $M(f)=M(a,f)$, which is known to be of weak-type (p_0,q_0) if $q_0>1$, $1/q_0-1/p_0=1/r$. We need only show that the condition of Lemma 3 is satisfied.

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

(6.1)
$$I(x,\varepsilon)=\{i|\ Q_i\cap S(x,\varepsilon)=\emptyset\},$$
$$J(x,\varepsilon)=\{i|\ Q_i\cap S(x,\varepsilon)\neq\emptyset\ \text{ and }\ Q_i-S(x,\varepsilon)\neq\emptyset\}.$$

Then

(6.2)
$$T_\varepsilon(a,h)(x)=\sum_i\int_{Q_i-S(x,\varepsilon)}\ldots dy$$
$$=\sum_{i\in I(x,\varepsilon)}\int_{Q_i}\ldots dy+\sum_{i\in J(x,\varepsilon)}\int_{Q_i-S(x,\varepsilon)}\ldots dy,$$

where each of the integrands is $\dfrac{a(x)-a(y)}{|x-y|^{n+1}}\Omega(x-y)h(y)$.

We show first that

(6.3)
$$\Big|\sum_{i\in J(x,\varepsilon)}\int_{Q_i-S(x,\varepsilon)}\ldots dy\Big|\leqslant C\Lambda_r(\operatorname{grad}a)(x)\Lambda_p(h)(x).$$

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

$$Q_i\subset\{y|\ \alpha\varepsilon<|y-x|<\beta\varepsilon\}\quad\text{for every } i\in J(x,\varepsilon).$$

Since Q_i's are pairwise disjoint, it follows that

$$\sum_{i\in J(x,\varepsilon)}\int_{Q_i-S(x,\varepsilon)}\frac{|a(x)-a(y)|}{|x-y|^{n+1}}|\Omega(x-y)|\ |h(y)|\,dy$$
$$\leqslant\int_{\alpha\varepsilon\leqslant|y-x|\leqslant\beta\varepsilon}\frac{|a(x)-a(y)|}{|x-y|^{n+1}}|\Omega(x-y)|\ |h(y)|\,dy$$
$$\leqslant\frac{C}{\varepsilon^n}\int_{|y-x|\leqslant\beta\varepsilon}\left|\frac{a(x)-a(y)}{x-y}\right||h(y)|\,dy$$
$$\leqslant C\left(\frac{1}{\varepsilon^n}\int_{|y-x|\leqslant\beta\varepsilon}\left|\frac{a(x)-a(y)}{x-y}\right|^{p'}dy\right)^{1/p'}\left(\frac{1}{\varepsilon^n}\int_{|y-x|\leqslant\beta\varepsilon}|h(y)|^p\,dy\right)^{1/p}$$
$$\leqslant C\Lambda_{p'}(\operatorname{grad}a)(x)\Lambda_p(h)(x)$$

which proves (6.3).

We show now that for $i\in I(x,\varepsilon)$,

(6.4)
$$\Big|\int_{Q_i}\ldots dy\Big|\leqslant C\delta_i\int_{Q_i}\frac{|a(x)-a(y)|}{|x-y|^{n+2}}|h(y)|\,dy+$$
$$+C\frac{\delta_i}{|x-y_i|^{n+1}}\int_{Q_i}\Lambda(\operatorname{grad}a)(y)|h(y)|\,dy,$$

where y_i is the center and δ_i the diameter of the cube Q_i. The right-hand side of the last inequality we denote by $A_i(x)$.

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

$$(6.5)\qquad \Big|\int_{Q_i} \frac{a(x)-a(y)}{|x-y|^{n+1}}\,\Omega(x-y)\,h(y)\,dy\Big|$$

$$= \Big|\int_{Q_i}\Big[\frac{a(x)-a(y)}{|x-y|^{n+1}}\,\Omega(x-y) - \frac{a(x)-a(t)}{|x-t|^{n+1}}\,\Omega(x-t)\Big]h(y)\,dy\Big|$$

$$\leqslant \int_{Q_i} \frac{|a(x)-a(y)|}{|x-y|^{n+1}}\,|\Omega(x-y)-\Omega(x-t)|\ |h(y)|\,dy +$$

$$+\int_{Q_i} |a(x)-a(y)|\Big|\frac{1}{|x-y|^{n+1}} - \frac{1}{|x-t|^{n+1}}\Big|\,|h(y)|\,dy +$$

$$+\int_{Q_i} |a(t)-a(y)|\ \frac{1}{|x-t|^{n+1}}\,|h(y)|\ |\Omega(x-t)|\,dy.$$

Observing that, for $y,\ t \in Q_i$, $x \in \boldsymbol{R}^n - \overline{Q}_i$,

$$\frac{1}{|x-t|^{n+1}} \leqslant \frac{\delta_i}{|t-y|}\cdot\frac{C}{|x-y_i|^{n+1}},$$

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

$$\frac{1}{|Q_i|}\int_{Q_i} |\Omega(x-y)-\Omega(x-t)|\,dt \leqslant C\,\frac{\delta_i}{|x-y|}$$

and that by Lemma 1

$$\frac{1}{|Q_i|}\int_{Q_i}\Big|\frac{|a(y)-a(t)|}{y-t}\Big|\,dt \leqslant \Lambda(\operatorname{grad} a)(y),$$

we obtain (6.4).

It follows from (6.2), (6.3) and (6.4) that

$$|T_\varepsilon(a,h)(x)| \leqslant C\Lambda_r(\operatorname{grad} a)(x)\,\Lambda_p(h)(x) + \sum_{i\in I(x,\varepsilon)} A_i(x)$$

$$\leqslant C\Lambda_r(\operatorname{grad} a)(x)\,\Lambda_p(h)(x) + \sum_{i=1}^{\infty} A_i(x).$$

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

$$(6.6)\qquad \big|\{x|\ x\in \boldsymbol{R}^n,\ \Lambda_r(\operatorname{grad} a)(x)\,\Lambda_p(h)(x) > \lambda\}\big| \leqslant C\,\frac{\|\operatorname{grad} a\|_r\,\|h\|_p}{\lambda}$$

and

$$(6.7)\qquad \big|\{x|\ x\in \boldsymbol{R}^n - \bigcup \overline{Q}_i,\ \textstyle\sum A_i(x) > \lambda\}\big| \leqslant C\,\frac{\|\operatorname{grad} a\|_r\,\|h\|_p}{\lambda}.$$

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

$$\sum_{i=1}^{\infty}\int_{\boldsymbol{R}^n-\bigcup\overline{Q}_i} A_i(x)\,dx \leqslant \sum_{i=1}^{\infty}\int_{\boldsymbol{R}^n-\overline{Q}_i} A_i(x)\,dx.$$

There exists a constant γ, dependent only on the dimension n, such that if $x\notin \overline{Q}_i$, $y\in Q_i$, then $|x-y| > \gamma\delta_i$. Thus, using (2.2),

$$\int_{\boldsymbol{R}^n-\overline{Q}_i} A_i(x)\,dx \leqslant C\int_{Q_i}\delta_i\Big(\int_{|x-y|>\gamma\delta_i}\frac{|a(x)-a(y)|}{|x-y|^{n+2}}\,dx\Big)|h(y)|\,dy +$$

$$+C\delta_i\int_{|x-y_i|>\gamma\delta_i}\frac{dx}{|x-y_i|^{n+1}}\int_{Q_i}\Lambda(\operatorname{grad} a)(y)\,|h(y)|\,dy$$

$$\leqslant C\int_{Q_i}\Lambda(\operatorname{grad} a)(y)\ |h(y)|\,dy.$$

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

$$\int_{\boldsymbol{R}^n-\bigcup\overline{Q}_i}\sum A_i(x)\,dx \leqslant C\int_{\boldsymbol{R}^n}\Lambda(\operatorname{grad} a)(y)\,|h(y)|\,dy$$

$$\leqslant C\|\Lambda(\operatorname{grad} a)\|_r\,\|h\|_p$$

$$\leqslant C\,\|\operatorname{grad} a\|_r\,\|h\|_p,$$

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 L^p and involve different methods complementing our own.

References

[1] B. Baishanski and R. Coifman, *On singular integrals*, Proc. Symp. in Pure Math. 10 (1967), pp. 1–17.

[2] A. P. Calderón, *Commutators of singular integral operators*, Proc. Nat. Acad. Sci. U.S.A. 53 (1965), pp. 1092–1099.

[3] M. Cotlar, *Condiciones De Continuidad De Operatores Potentiales Y De Hilbert*, Cursos Y Seminarios De Mathematica, Fasc. 2, Universidad Nacional de Buenos Aires, 1959.

Received August 8, 1974
New version May 12, 1976 (877)