Cambridge University Press
978-0-521-79473-2 - Wavelets: Calderon-Zygmund and Multilinear Operators
Yves Meyer and Ronald Coifman
Excerpt
More information

7

The new Calderón-Zygmund operators

1 Introduction

The Calderón-Zygmund operators we discuss in this chapter differ significantly from those which Calderón and Zygmund considered more than thirty years ago. We shall try to describe and explain how the definition of these operators has evolved.

The theory of Calderón-Zygmund operators began in the 1950s when Calderón and Zygmund systematically studied convolution operators appearing in elliptic partial differential equations.

The best-known example is given by the Riesz transforms

$$R_j = -i\frac{\partial}{\partial x_j}(-\Delta)^{-1/2}$$

for $1 \le j \le n$, where

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}\,.$$

Riesz transforms arise when we study the Neumann problem in the upper half-plane. More precisely, let $\Omega \subset \mathbb{R}^{n+1}$ be the open set defined by $t > 0$ and $x \in \mathbb{R}^n$. We consider harmonic functions in the Sobolev space $H^2(\Omega)$, that is, those functions $u(x,t)$ on Ω such that $\Delta u + \partial^2 u/\partial t^2 = 0$. Then we can define the boundary values of $u(x,t)$ and the gradient of $u(x,t)$ on $t = 0$ (the boundary $\partial\Omega$ of Ω). In this context, the Riesz transforms R_j, $1 \le j \le n$, enable us to pass from the boundary values of the normal derivative $(-\partial u/\partial t)|_{t=0}$ to the boundary values of the tangential derivatives $(\partial u/\partial x_j)|_{t=0}$.

The continuity of the Riesz operators on $L^2(\mathbb{R}^n)$ follows immediately. We can either use Green's formula or conjugate each R_j by the Fourier transform $\mathcal{F}$. In the second case we get $\mathcal{F}R_j f(\xi) = \xi_j|\xi|^{-1}\hat{f}(\xi)$; it is enough to observe that $|\xi_j| \le |\xi|$ in order to conclude that $\|R_j(f)\|_2 \le \|f\|_2$. More precisely, we have $\sum_1^n \|R_j(f)\|_2^2 = \|f\|_2^2$.

On the other hand, it is not at all obvious that the R_j, $1 \le j \le n$, are continuous on $L^p(\mathbb{R}^n)$, $1 < p < \infty$. That result is achieved by the real variable methods of Calderón and Zygmund, which we now describe.

Calderón and Zygmund considered the more general problem of operators on $L^2(\mathbb{R}^n)$ defined by the following algorithm. We start with a function $\Omega(x)$ in $C(\mathbb{R}^n\backslash\{0\})$ which is homogeneous of degree 0 and which satisfies the condition $\int_{S^{n-1}} \Omega(x)\, d\sigma(x) = 0$, where $d\sigma$ is the rotation-invariant probability measure on the unit sphere S^{n-1}.

The first generation Calderón-Zygmund operators are convolution operators. We take the distribution $S = \mathrm{PV}\,\Omega(x)|x|^{-n}$ defined by

$$\langle S, \phi\rangle = \lim_{\varepsilon\downarrow 0} \int_{|x|\ge\varepsilon} \frac{\Omega(x)}{|x|^n}\phi(x)\, dx\,, \qquad \phi \in \mathcal{D}(\mathbb{R}^n)\,.$$

Then the Calderón-Zygmund operator T is given by $T(f) = S \star f$, for each $f \in \mathcal{D}(\mathbb{R}^n)$. In other words,

$$(T(f))(x) = \lim_{\varepsilon\downarrow 0} \int_{|y|\ge\varepsilon} f(x-y)\frac{\Omega(y)}{|y|^n}\, dy\,. \tag{1.1}$$

The limit exists if f is continuous, of Hölder exponent $\gamma > 0$, and square-summable.

Then $\mathcal{F}S = m(\xi)$, where $m(\xi) \in C^\infty(\mathbb{R}^n \setminus \{0\})$, where $m(\lambda\xi) = m(\xi)$, for $\lambda > 0$, and where the mean of $m(\xi)$ on the unit sphere is zero. For the Riesz transforms, $m(\xi) = \xi_j/|\xi|$.

This theory gave a unified treatment of earlier work of J. Marcinkiewicz (1938) and G. Giraud (1936). Calderón and Zygmund showed that the Marcinkiewicz multipliers associated with elliptic partial differential equations were the Fourier coefficients of the (periodified) distributions $\mathrm{PV}\,\Omega(x)|x|^{-n}$ used by Giraud. An essential tool in that unification was the notion of the Fourier transform of a tempered distribution (introduced by L. Schwartz between 1943 and 1945).

By these means, Calderón and Zygmund rediscovered, in a very natural way, the rules for composing two Calderón-Zygmund operators, T_1 and T_2: it is enough to employ the usual product of the corresponding multipliers $m_1(\xi)$ and $m_2(\xi)$. However, the symbol $m_3(\xi) = m_1(\xi)m_2(\xi)$ does not always have zero integral on the unit sphere, so it is necessary to consider the algebra of operators $cI + T$, where c is a constant and T is defined by (1.1).

After doing this, Calderón and Zygmund tried to extend the continuity properties of Calderón-Zygmund operators to $L^p(\mathbb{R}^n)$, $1 < p < \infty$. The complex variable method, introduced by Littlewood and Paley, and used by Marcinkiewicz to prove his multiplier theorem, did not work in the n-dimensional case. It was this obstacle which led Calderón and Zygmund to invent the real variable method. The essential ingredient is the Calderón-Zygmund decomposition which, for any parameter $\lambda > 0$, lets us write any function f in $L^1(\mathbb{R}^n)$ as a sum $g+h$, where $g \in L^2(\mathbb{R}^n)$ satisfies $\|g\|_2 \leq C\sqrt{\lambda}$ and where h is the sum of a series of oscillatory terms, each with support in a certain cube Q_j, such that the sum of the measures of the Q_j does not exceed C/λ. These oscillatory terms foreshadow the "atoms" of Coifman and Weiss, as well as wavelets.

The L^2 estimate for an operator T, together with quite a weak hypothesis on the kernel $K(x,y)$ of T ($\int_{|x-y|\geq 2|y'-y|} |K(x,y')-K(x,y)|\, dx \leq C$ is enough) gives the L^1-weak-L^1 estimate which we shall describe explicitly in this chapter. This, together with the L^2 estimate, allows us to derive the L^p estimates, for $1 < p \leq 2$, by interpolation. We get the L^q estimates, for $2 \leq q < \infty$, by duality, provided that we impose corresponding conditions where the roles of x and y in $K(x,y)$ are interchanged.

A second preoccupation of Calderón and Zygmund was to give a pointwise meaning to (1.1), when $f(x)$ was an arbitrary function in $L^2(\mathbb{R}^n)$, by proving the existence, for almost every $x \in \mathbb{R}^n$, of

$$\lim_{\varepsilon\downarrow 0} \int_{|y|\geq\varepsilon} f(x-y)\frac{\Omega(y)}{|y|^n}\, dy\,.$$

This problem leads to the *maximal operator* T^*, defined by

$$T^*f(x) = \sup_{\varepsilon>0}\left|\int_{|y|\geq\varepsilon} f(x-y)\frac{\Omega(y)}{|y|^n}\, dy\right|, \tag{1.2}$$

and to the proof of the *maximal inequality*

$$\|T^*f\|_2 \leq C\|f\|_2, \qquad f \in L^2(\mathbb{R}^n)\,. \tag{1.3}$$

The convolution operators which we have introduced correspond to elliptic problems with constant coefficients. To deal with partial differential equations with variable, but extremely regular, coefficients, Calderón and Zygmund invented what we shall call "second generation" Calderón-Zygmund operators.

Second generation Calderón-Zygmund operators are not convolution operators, but are given by slightly more elaborate distribution-kernels. More precisely, for a test function f (square-summable and sufficiently regular),

$$(1.4)\qquad Tf(x) = \mathrm{PV}\int K(x,y)f(y)\,dy = \lim_{\varepsilon\downarrow 0}\int_{|x-y|\geq\varepsilon} K(x,y)f(y)\,dy\,,$$

where $K(x,y)$ satisfies the following properties:–

$$(1.5)\qquad K(x,y) = L(x,x-y)\,;$$

$$(1.6)\qquad L(x,\lambda u) = \lambda^{-n}L(x,u)\,,$$

for every $x\in\mathbb{R}^n$, every $u\in\mathbb{R}^n\setminus\{0\}$ and every $\lambda>0$;

$$(1.7)\qquad |\partial_u^\alpha\partial_x^\beta L(x,u)|\leq C\,,\ \text{when } |u|=1,\ |\alpha|\leq N \text{ and } |\beta|\leq m;$$

$$(1.8)\qquad \int_{S^{n-1}} L(x,u)\,d\sigma(u) = 0 \qquad \text{identically in } x\in\mathbb{R}^n.$$

By separation of variables (involving expansion in spherical harmonics with respect to the variable $u\in S^{n-1}$), when N was sufficiently large compared to n, Calderón and Zygmund reduced the operators of the second generation to absolutely convergent series $\sum_0^\infty M_jT_j$, where the T_j were first generation operators and the M_j were operators of pointwise multiplication by bounded functions.

The integer m is irrelevant, and may even be zero, as long as we restrict ourselves to studying the L^2 or L^p continuity of second generation operators. On the other hand, m plays an essential part as soon as we seek to compose the operators in question, in order to obtain an algebra with a precise symbolic calculus. To get this, we need to work modulo the ideal $\mathcal{I}$ of regularizing operators of order 1, that is, the operators S which are bounded on $L^2(\mathbb{R}^n)$ and are such that $(\partial/\partial x_j)S$ and $S\partial/\partial x_j$ are also bounded, for $1\leq j\leq n$. Calderón and Zygmund proved that, modulo $\mathcal{I}$, the second generation Calderón-Zygmund operators formed a commutative algebra, as long as the integers m and N were large enough. Here again, it was necessary to include the operators of pointwise multiplication by those bounded functions $b(x)$ which had all their derivatives bounded.

Several years later (1965), Calderón took another look at the problem of the symbolic calculus when he sought conditions of minimal regularity, with respect to x, on kernels $L_1(x,u)$ and $L_2(x,u)$, corresponding to operators T_1 and T_2, which would imply the existence of a symbolic calculus modulo the regularizing operators of order 1. In applications to partial differential equations, the regularity with respect to x is given by that of the coefficients $c_\alpha(x)$ of the differential operators $\sum c_\alpha(x)\partial^\alpha$ considered, whereas the kernels themselves are infinitely differentiable with respect to the variable $u\in\mathbb{R}^n\setminus\{0\}$.

The symbolic calculus problem that we have just raised reduces to the study of the commutators $[A,T]$, where A is the operator of pointwise multiplication by a function $a(x)$, whose regularity is measured by the integer m of (1.7), and where $T = (\partial/\partial x_1)T_1 + \cdots + (\partial/\partial x_n)T_n$, the T_j being first generation Calderón-Zygmund operators.

In dimension 1, T_1 is replaced by the Hilbert transform H and T is replaced by the "Calderón operator" $\Lambda = DH$, where $D = -i(d/dx)$.

In 1965, Calderón showed that the commutator $[A,\Lambda]$ was bounded on $L^2(\mathbb{R})$ if and only if the function $a(x)$ was Lipschitz, that is, if there was a constant C such that $|a(x) - a(y)| \le C|x-y|$, for all $x, y \in \mathbb{R}$. It is easy to see that this condition is necessary. The reverse implication is deep, and Calderón's proof relies on the characterization, established by Calderón for this purpose, of the complex $\mathbb{H}^1$ space by the integrability of Lusin's area function. This characterization opened the way to the study of the operators on $\mathbb{H}^1$ and of the dual of the real version of $\mathbb{H}^1$ ([109]).

To pass to dimension n, Calderón used the method of rotations (invented a few years earlier by Calderón and Zygmund) and showed that the set of operators defined by (1.4) to (1.8), with $m = 1$ and N large enough, becomes a commutative Banach algebra, once it is reduced modulo the regularizing operators of order 1.

This remarkable theorem will be proved in Chapter 9. It leads to a substantial improvement of most of the results obtained by applying the classical pseudo-differential calculus (where the symbols $\sigma(x,\xi)$ are infinitely differentiable with respect to x).

The transition from the L^2 continuity to the L^p continuity ($1 < p < \infty$) of the commutator $[A,\Lambda]$ is achieved by repeating the proof used for the case of the Riesz transforms.

From 1966, Calderón proposed studying commutators of higher order $T_k = [A,[A,\ldots,[A,D^kH]\ldots]]$ and proving that the T_k were continuous on $L^2(\mathbb{R})$ whenever A was an operator of pointwise multiplication by a Lipschitz function $a(x)$. The generating series $\sum_0^\infty \xi^k T_k$ is the operator defined by the Cauchy integral PV $\int_\Gamma (z-w)^{-1} f(w)\,dw$ (the reader will recall that PV means "principal value"), where Γ is the graph of a Lipschitz function and where f lies in $L^2(\Gamma, ds)$. Calderón posed the problem of the continuity of this latter operator on the Hilbert space $L^2(\Gamma, ds)$.

A further operator belonging to "Calderón's programme" is related to the classical method of using a double layer potential to solve the Dirichlet or the Neumann problem in a Lipschitz domain. The operator involved is then given in local coordinates by

$$Tf(x) = \frac{1}{\omega_n} \mathrm{PV} \int_{\mathbb{R}^n} K(x,y) f(y)\, dy\,,$$

where

$$K(x,y) = \frac{a(x) - a(y) - (x-y)\cdot \nabla a(y)}{(|x-y|^2 + (a(x)-a(y))^2)^{(n+1)/2}}\,,$$

and where $f \in L^2(\mathbb{R}^n)$. The Lipschitz domain is defined (locally) by $t > a(x)$, where $x \in \mathbb{R}^n$, $t \in \mathbb{R}$, and the function $a(x)$ is Lipschitz.

If $n = 1$, this kernel is precisely the real part of the Cauchy kernel and in dimension n it can be studied by the method of rotations, once we know that the Cauchy kernel is bounded on $L^2(\mathbb{R})$.

In 1977, Calderón proved the continuity of the operator defined by the Cauchy integral for every Lipschitz graph $y = a(x)$, where $\|a'\|_\infty < \varepsilon$; the method used did not give the value of the mysterious constant ε.

The third generation Calderón-Zygmund operators generalize the examples we have just described. In 1976, Calderón asked the following question. Let $K(x,y)$ be a function of two real variables, defined for $x \neq y$, which satisfies the estimates $|K(x,y)| \leq C_0|x-y|^{-1}$ and $|(\partial/\partial x)K(x,y)| \leq C_1(x-y)^{-2}$ and the condition $K(y,x) = -K(x,y)$. Is it true that the operator $T : \mathcal{D} \to \mathcal{D}'$ whose kernel is PV $K(x,y)$ can be extended to a continuous operator on $L^2(\mathbb{R})$? Such a result would have given "for free" the continuity of the Cauchy kernel on all Lipschitz curves, etc. It is unfortunately not the case: J.L. Journé discovered in 1982 that it was necessary to assume that $T(1)$ belonged to BMO, this latter condition being necessary and sufficient for the L^2 continuity of the operator T. To define $T(1)$, we let g_ε denote the Gaussian $e^{-\varepsilon x^2}$ and we prove that there are renormalization constants $c(\varepsilon)$, such that $T(g_\varepsilon) - c(\varepsilon)$ converges, in the sense of distributions, to what we shall call $T(1)$. This distribution is, therefore, only defined modulo the constant functions, which is, after all, the case for elements of BMO.

The "$T(1)$ theorem" was finally proved by G. David and Journé in 1983. It is a remarkable result because it reduces all the previous results of Calderón, and the authors of this book, to simple integrations by parts. In particular, it immediately gives the continuity of the Cauchy kernel on Lipschitz graphs of small slope.

The extension to general Lipschitz curves is obtained by new real methods discovered by David and then simplified by T. Murai.

To define the "third generation Calderón-Zygmund operators", we replace $\mathbb{R}$ by $\mathbb{R}^n$ and drop the antisymmetry hypothesis $K(y,x) = -K(x,y)$. It is then no longer possible to define the operator $T : \mathcal{D} \to \mathcal{D}'$

just by the function $K(x,y)$ (which is defined on $\mathbb{R}^n \times \mathbb{R}^n \setminus \Delta$, where Δ is the diagonal).

We proceed in the reverse direction and start with a linear (and continuous) operator $T : \mathcal{D} \to \mathcal{D}'$. We say that this operator corresponds to a singular integral if the distribution-kernel of T, restricted to $\mathbb{R}^n \times \mathbb{R}^n \setminus \Delta$, is a function $K(x,y)$ satisfying the following hypotheses:

(1.9) $$|K(x,y)| \le C_0 |x-y|^{-n},$$

(1.10) $$\left|\frac{\partial K(x,y)}{\partial x_j}\right| \le C_1 |x-y|^{-n-1}, \qquad 1 \le j \le n,$$

(1.11) $$\left|\frac{\partial K(x,y)}{\partial y_j}\right| \le C_1 |x-y|^{-n-1}, \qquad 1 \le j \le n.$$

In particular, if f is a test function and if x does not lie in the (compact) support of f, then

(1.12) $$Tf(x) = \int K(x,y) f(y)\, dy\,.$$

The fundamental problem (which will be solved in Chapter 8) is to find a necessary and sufficient condition for the continuity of T on the space $L^2(\mathbb{R}^n)$. We then say that T is a Calderón-Zygmund operator. We shall prove the existence of a bounded measurable function $m(x)$ and of a sequence ε_j of measurable functions on $\mathbb{R}^n$, taking positive values $\varepsilon_j(x) > 0$, such that, for every function $f(x)$ in $L^2(\mathbb{R}^n)$, we have

(1.13) $$Tf(x) = m(x) f(x) + \lim_{j\to\infty} \int_{|x-y|\ge \varepsilon_j(x)} K(x,y) f(y)\, dy\,,$$

for almost all $x \in \mathbb{R}^n$.

In the case of a convolution operator, $\varepsilon_j(x)$ does not depend on x; if ε_j can also be taken to be an arbitrary sequence tending to 0, then we are back to the classical idea of the principal value of a singular integral. The representation (1.13) explains the sense in which the operator T "corresponds to a singular integral".

Besides the examples we have already mentioned (the Calderón commutators, the Cauchy integral on Lipschitz curves and the double-layer potential), third generation Calderón-Zygmund operators arise in a completely different context. To show that a wavelet basis ψ_λ, $\lambda \in \Lambda$, constructed by the algorithms of Chapter 3 of *Wavelets and Operators*, is an unconditional basis of a classical space of functions or distributions B, we need to show that the operators $T : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ which are diagonal in our wavelet basis are also continuous on the space B. Now such operators are automatically Calderón-Zygmund operators, whose continuity on B will be established by the real variable methods developed in this chapter and in Chapter 10.

2 Definition of Calderón-Zygmund operators corresponding to singular integrals

As we stated in the introduction, we do not want to define a Calderón-Zygmund operator as the principal value of a singular integral.

More precisely, we also want to use kernels $K(x,y)$ for which the limit

$$\lim_{\varepsilon\downarrow 0}\int_{|x-y|\geq\varepsilon} K(x,y)f(y)\,dy$$

might not exist in the usual sense, even when $f(y)$ is a test function or $K(x,y)$ is antisymmetric. Now, in the antisymmetric case, that limit exists in the sense of distributions. This leads us to base the theory on the bilinear form $J(f,g)=\langle T(f),g\rangle=\langle S, g\otimes f\rangle$, where f and g are test functions, T is a linear operator from $\mathcal{D}$ into $\mathcal{D}'$ and S is the distribution-kernel of T. Then K is derived from S by the condition that K is the restriction of S to the complement of the diagonal.

Here is an example to clarify these distinctions. Let $\theta(x)$ be an odd function in $\mathcal{D}(\mathbb{R})$ such that $\int_{-\infty}^{\infty}\theta(x)\sin x\,dx=1$. Further, let the support of θ be the union of the intervals $[-4/3,-2/3]$ and $[2/3,4/3]$. Then we define the kernel $K(x,y)$ by

$$K(x,y)=\sum_0^{\infty}2^k\theta(2^k(x-y))e^{i2^k(x+y)}\,.$$

We immediately get $K(y,x)=-K(x,y)$, $|K(x,y)|\leq C_0|x-y|^{-1}$ and $|\partial k/\partial x|\leq C_1|x-y|^{-2}$. If f is in $\mathcal{D}(\mathbb{R})$ and equals 1 on $[-10,10]$, the existence of the limit $\lim_{\varepsilon\downarrow 0}\int_{|x-y|\geq\varepsilon}K(x,y)f(y)\,dy$, for $-1\leq x\leq 1$, would imply that the series $\sum_0^{\infty}e^{i2^k x}$ converged. This series diverges everywhere (pointwise), but converges in the sense of distributions.

This leads us to consider

$$J_\varepsilon(f,g)=\iint_{|x-y|\geq\varepsilon}K(x,y)f(y)g(x)\,dy\,dx\,,$$

where $K(x,y)$ is an antisymmetric kernel with $|K(x,y)|\leq C_0|x-y|^{-1}$ and where f and g are functions in $\mathcal{D}(\mathbb{R})$. Now

$$J_\varepsilon(f,g)=\frac{1}{2}\iint_{|x-y|\geq\varepsilon}K(x,y)(f(y)g(x)-f(x)g(y))\,dy\,dx$$

and we can define $J(f,g)$ by the absolutely converging integral

$$J(f,g)=\frac{1}{2}\iint K(x,y)(f(y)g(x)-f(x)g(y))\,dy\,dx\,.$$

Then $J(f,g)=\lim_{\varepsilon\downarrow 0}J_\varepsilon(f,g)$, which enables us to complete the construction of $T:\mathcal{D}\to\mathcal{D}'$ by the formula $\langle T(f),g\rangle=J(f,g)$.

The case of non-antisymmetric kernels is more subtle: we can no longer use a kernel to define the operator. Reversing rôles, we start with an operator $T : \mathcal{D} \to \mathcal{D}'$. Let $S \in \mathcal{D}'(\mathbb{R}^n \times \mathbb{R}^n)$ denote the kernel of T. The existence of S is guaranteed by Schwartz's kernel theorem and S is related to T by the identity $\langle T(f), g\rangle = \langle S, g \otimes f\rangle$. The left-hand side involves the duality between $\mathcal{D}(\mathbb{R}^n)$ and $\mathcal{D}'(\mathbb{R}^n)$, whereas the duality on the right-hand side is that between $\mathcal{D}(\mathbb{R}^n \times \mathbb{R}^n)$ and $\mathcal{D}'(\mathbb{R}^n \times \mathbb{R}^n)$.

We let $\Omega = \{(x,y) \in \mathbb{R}^n : x \neq y\}$ and we consider the restriction $K(x,y)$ of S to Ω.

Definition 1. *Let $T : \mathcal{D}(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^n)$ be a continuous linear operator. We say that T is a Calderón-Zygmund operator if there are constants C_0, C_1, C_2, and an exponent $\gamma \in (0,1]$ such that the following four conditions are satisfied:*

(2.1) *$K(x,y)$ is a locally integrable function on Ω and satisfies*

$$|K(x,y)| \le C_0|x-y|^{-n}\,;$$

(2.2) *if $(x,y) \in \Omega$ and $|x'-x| \le |x-y|/2$, then*

$$|K(x',y) - K(x,y)| \le C_1|x'-x|^{\gamma}|x-y|^{-n-\gamma}\,;$$

(2.3) *if $(x,y) \in \Omega$ and $|y'-y| \le |x-y|/2$, then*

$$|K(x,y') - K(x,y)| \le C_1|y'-y|^{\gamma}|x-y|^{-n-\gamma}\,;$$

(2.4) *T extends to a continuous linear operator on $L^2(\mathbb{R}^n)$ with*

$$\|T\| \le C_2\,.$$

If conditions (2.2) and (2.3) are satisfied for an exponent $\gamma > 0$, they are satisfied, a fortiori, for every exponent $\gamma' \in (0,\gamma)$.

If $\gamma = 1$, conditions (2.2) and (2.3) can be written more simply as

$$\left|\frac{\partial K}{\partial x_j}\right| + \left|\frac{\partial K}{\partial y_j}\right| \le C_1|x-y|^{-n-1}, \qquad 1 \le j \le n.$$

Let us consider the "second generation" Calderón-Zygmund operators of the introduction. Then the kernel $K(x,y)$ is $L(x, x-y)$, where $L(x,u)$ satisfies (1.5), (1.6) and (1.7). Such a kernel never satisfies the estimate (2.2), unless $L(x,u)$ is independent of x, in which case the operator is a convolution operator, in other words, a "first generation" Calderón-Zygmund operator.

To justify our remark, we simply observe that the kernels $K(x,y)$ and $\lambda^n K(\lambda x \lambda y)$ satisfy (2.2) with the same constant: C_1. Using the homogeneity of $L(x,u)$ in u, we find that, if (2.2) holds, then, for every $\lambda \ge 1$ and for every u of length 1, we have

$$|L(\lambda x, u) - L(\lambda x', u)| \le C_1|x'-x|^{\gamma}\,,$$

when $|x'-x| \le 1/2$. Letting λ tend to infinity concludes the argument.

Thus condition (2.2) fails when $|x-y|$ is large.

A subset $\mathcal{B} \subset \mathcal{L}(L^2(\mathbb{R}^n), L^2(\mathbb{R}^n))$ is called a *bounded* set of Calderón-Zygmund operators if the operators $T \in \mathcal{B}$ satisfy (2.1), (2.2), (2.3) and (2.4) with the same exponent γ and the same constants C_0, C_1 and C_2.

Let $\mathcal{G}$ denote the group of unitary isomorphisms of $L^2(\mathbb{R}^n)$ of the form $Uf(x) = \delta^{-n/2} f(\delta^{-1}(x - x_0))$ for $x_0 \in \mathbb{R}^n$ and $\delta > 0$. Then, if T is a Calderón-Zygmund operator, the collection UTU^{-1}, $U \in \mathcal{G}$, is a bounded set of Calderón-Zygmund operators. The kernel of UTU^{-1} is $\delta^{-n} K(\delta^{-1}(x - x_0))$ and we note again that conditions (2.1) to (2.4) are invariant under change of scale.

We shall use the following "weak compactness theorem".

Proposition 1. *Let T_j, $j \in \mathbb{N}$, be a bounded sequence of Calderón-Zygmund operators. Then there exist a Calderón-Zygmund operator T and a subsequence $T_{j(m)}$ such that*

(2.5) $$\langle T_{j(m)}(f), g\rangle \to \langle T(f), g\rangle ,$$

for $f \in L^2(\mathbb{R}^n)$ and $g \in L^2(\mathbb{R}^n)$.

We shall write $T_{j(m)} \rightharpoonup T$ in the above situation.

Since the operator norms of the $T_j : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ form a bounded sequence, there is a subsequence $T_{j(m)}$ which converges weakly to a linear operator $T : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$.

It remains to prove that T is a Calderón-Zygmund operator. We let S_m and S denote the kernels of $T_{j(m)}$ and T. Then S is the limit of the S_m in the sense of convergence of distributions in $\mathcal{D}'(\mathbb{R}^n \times \mathbb{R}^n)$. The restrictions of the S_m to each compact subset of Ω are continuous functions satisfying Hölder conditions uniformly in m. Ascoli's theorem gives the uniform convergence of S_m to S on those compact subsets. Conditions (2.1), (2.2) and (2.3) on S are obtained by passing to the limit in the inequalities.

Conversely, every Calderón-Zygmund operator T can be written, as in (2.5), as the limit of a bounded sequence of Calderón-Zygmund operators T_m, whose kernels S_m are not only distributions, but also infinitely differentiable, bounded functions. More precisely, we have

Proposition 2. *Let $T : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ be a Calderón-Zygmund operator with kernel $K(x, y)$. Then there is a sequence $K_m(x, y)$ in $L^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ such that $\partial K_m/\partial x_j$ and $\partial K_m/\partial y_j$ are also in $L^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ and the following properties hold:*

(2.6) *the operators T_m defined by $T_m(x) = \int K_m(x, y) f(y)\, dy$ form a bounded sequence of Calderón-Zygmund operators;*

(2.7) *for every function $f \in L^2(\mathbb{R}^n)$,*

$$\lim_{m \to \infty} \|T(f) - T_m(f)\|_2 = 0.$$