Preface

This book could have been entitled "Analysis and Geometry." The authors are addressing the following issue: Is it possible to perform some harmonic analysis on a set? Harmonic analysis on groups has a long tradition. Here we are given a metric set X with a (positive) Borel measure μ and we would like to construct some algorithms which in the classical setting rely on the Fourier transformation. Needless to say, the Fourier transformation does not exist on an arbitrary metric set.

This endeavor is not a revolution. It is a continuation of a line of research which was initiated, a century ago, with two fundamental papers that I would like to discuss briefly.

The first paper is the doctoral dissertation of Alfred Haar, which was submitted at to University of Göttingen in July 1907. At that time it was known that the Fourier series expansion of a continuous function may diverge at a given point. Haar wanted to know if this phenomenon happens for every orthonormal basis of $L^2[0,1]$. He answered this question by constructing an orthonormal basis (today known as the Haar basis) with the property that the expansion (in this basis) of any continuous function uniformly converges to that function.

Today we know that Haar was the grandfather of wavelets and we also know that wavelet bases offer a powerful and flexible alternative to Fourier analysis. Indeed wavelet bases are unconditional bases of most of the functional spaces we are using in analysis. In other words wavelet expansions offer an improved numerical stability, as compared with Fourier series expansions. One of the goals of this book is to construct wavelets on any metric set equipped with a positive measure which is compatible with the given metric. In this setting we do not have Fourier analysis at our disposal.

The second paper which preluded the authors' endeavor was written in French by Marcel Riesz in 1926. It is entitled "Sur les fonctions conjuguées." The author proves that the Hilbert transform is bounded on $L^p(\mathbb{R})$ when $1 < p < \infty$. The Hilbert transform H is the convolution with $\frac{1}{\pi} p.v. \frac{1}{x}$, which is a distribution. In other words $H(f)(x) = \frac{1}{\pi} p.v. \int \frac{f(y)}{x-y} dy$. The Fourier transform of $H(f)$ is $-i\,\mathrm{sign}(\xi)\widehat{f}(\xi)$ when $\widehat{f}(\xi)$ is the Fourier transform of f. Therefore, H is isometric on $L^2(\mathbb{R})$.

The proof given by Riesz relies on the properties of holomorphic functions F in the unit disc $\mathbb{D}$ of the complex plane. The boundary Γ of $\mathbb{D}$ is the unit circle identified to $[0, 2\pi]$ and functions on Γ can be written as Fourier series. If a holomorphic function F in $\mathbb{D}$ extends to the boundary Γ, then the Fourier series of F on Γ coincides with its Taylor series. Moreover if u is the real part of a holomorphic function F and v is the imaginary part, then v is the Hilbert transform of u on Γ.

To prove his claim, Riesz used the Cauchy formula and the fact that F^p (F raised to the power p) is still holomorphic when p is an integer or when F has no zero in $\mathbb{D}$. This attack was named "complex methods" by Antoni Zygmund.

In the 1950s Alberto Calderón and Zygmund discovered a new strategy for proving L^p estimates. They could not use complex methods anymore since they were interested in operators acting on $L^2(\mathbb{R}^n)$. The operators constructed by Calderón and Zygmund are the famous pseudo-differential operators and soon became one of the most powerful tools in partial differential equations.

Let us sketch the proof of L^p estimates discovered by Calderón and Zygmund. It begins with a lemma which is known as the "Calderón–Zygmund decomposition." It says the following. Let f be any function in $L^1(\mathbb{R}^n)$ and let $\lambda > 0$ be a given threshold. Then f can be split into a sum $u + v$ where $|u|$ is bounded by λ and belongs to $L^2(\mathbb{R}^n)$, while v is oscillating and supported by a set of measure not exceeding $\frac{C}{\lambda}$. As noticed by Joseph Doob, the proof of this lemma is indeed a stopping time argument applied to a dyadic martingale. On the other hand, the Haar basis yields a martingale expansion. Calderón and Zygmund argued as follows. They assumed that the distributional kernel $K(x, y)$ of an operator T satisfies the following conditions: There exists a constant C such that for every $x \in \mathbb{R}^n$ and every $x' \neq x$ one has

$$\int\limits_{|y-x|\geq 2|x'-x|} |K(x', y) - K(x, y)|dy \leq C$$

and there exists a constant C' such that for every $y \in \mathbb{R}^n$ and every $y' \neq y$ one has

$$\int\limits_{|x-y'|\geq 2|y-y'|} |K(x, y') - K(x, y)|dx \leq C'. \tag{†}$$

Calderón and Zygmund proved a remarkable result. If T is bounded on $L^2(\mathbb{R}^n)$ and if the distributional kernel $K(x, y)$ of T satisfies (†), then for every f in $L^1(\mathbb{R}^n), T(f)$ belongs to weak L^1. There exists a constant C such that for every positive λ the measure of the set of points x for which $|T(f)(x)| > \lambda$ does not exceed $C\frac{\|f\|_1}{\lambda}$. This is optimal, since $f = \delta_{x_0}$(Dirac mass at x_0) yields $T(f)(x) = K(x, x_0)$ which belongs to weak L^1 and not to L^1. This theorem follows from the Calderón–Zygmund decomposition. Then the Marcinkiewicz interpolation theorem implies the required L^p estimates for $1 < p \le 2$. Applying the same argument to the adjoint operator T^*, we obtain the L^p estimates for $2 \le p < \infty$.

The arguments which were used in these two steps do not rely on Fourier methods; therefore, this scheme easily extends to geometrical settings where the Fourier transformation does not exist. Such generalizations were achieved by Ronald Coifman and Guido Weiss. They discovered that the "spaces of homogeneous type" are the metric spaces to which the Calderón–Zygmund theory extends naturally. A space of homogeneous type is a metric space X endowed with a positive measure μ which is compatible with the given metric in a sense which will be detailed in this book. Roughly speaking, the measure $\mu(B(x, r))$ of a ball centered at x with radius r scales as a power of r.

Coifman and Weiss observed that any bounded operator $T : L^2(X, d\mu) \to L^2(X, d\mu)$ whose distributional kernel satisfies (†)—with $|x - y'| \ge 2|y - y'|$ replaced by $d(x, y') \ge 2d(y, y')$—maps L^1 into weak L^1. That implies L^p estimates for $1 < p \le 2$. This can be found in the remarkable book *Analyse Harmonique Non- commutative sur Certains Espaces Homogènes* which was published in 1971.

But this does not tell us how to prove the fundamental L^2 estimate. We will return to this issue after a detour.

In the 1960s Calderón launched an ambitious program. He wanted to free the pseudo-differential calculus from the unnecessary smoothness assumptions which were usually required to obtain commutator estimates. The first issue he addressed was the following problem. Let A be the pointwise multiplication by a function $A(x)$ and let T be any pseudo-differential operator of order 1. Can we find a necessary and sufficient condition on A implying that all commutators $[A, T]$ are bounded on $L^2(\mathbb{R}^n)$? This is required for every pseudo-differential operator of order 1 and the particular choices $T_j = \frac{\partial}{\partial x_j}, 1 \le j \le n$, show that A must be a Lipschitz function. The other way around is much more difficult and was proved by Calderón in 1965. The proof relies on new estimates on the Hardy space $\mathcal{H}^1(\mathbb{R})$. Calderón proved that the $\mathcal{H}^1$ norm of a holomorphic function F is controlled by the L^1 norm of the Lusin area function of F. This connection between an L^2 estimate and the Hardy space $\mathcal{H}^1$ is the most surprising. An explanation will be given by the $T(1)$ theorem of David and Journé.

This spectacular achievement gave a second life to the theory of Hardy spaces and Charles Fefferman, in collaboration with Elias Stein, proved that the dual of $H^1(\mathbb{R}^n)$ is $BMO(\mathbb{R}^n)$. Here $H^1(\mathbb{R}^n)$ is the real variable version of the Hardy space $H^1(\mathbb{R})$. In other words, H^1 is the subspace of L^1 which is defined by $n{+}1$ conditions $f \in L^1$ and $R_j f \in L^1$, where $R_j, 1 \le j \le n$, are the Riesz transforms.

Calderón conjectured that the Cauchy kernel on a Lipschitz curve Γ is bounded on $L^2(\mathbb{R})$. A Lipschitz curve Γ is the graph of a (real-valued)

Lipschitz function A. The curve Γ admits a parameterization given by $z(x) = x + iA(x), -\infty < x < \infty$, and the Cauchy operator can be written as

$$C(f)(x) = p.v.\frac{1}{\pi i}\int_{-\infty}^{\infty}(z(x)-z(y))^{-1}f(y)dy.$$

If $\|A'\|_\infty < 1$, the Cauchy operator is given by a Taylor expansion $\sum_0^\infty C_n(f)$, where C_n are the iterated commutators between A (the pointwise multiplication with $A(x)$) and D^nH. Here, as above, H is the Hilbert transform and $D = -i\frac{d}{dx}$.

In 1977 Calderón used a refinement of the method which was successful for the first commutator and could prove the boundedness of the Cauchy kernel under the frustrating condition $\|A'\|_\infty < \beta$, where β is a small positive number. Guy David combined this result with new real variable methods and got rid of the limitation in Calderón's theorem.

But the main breakthrough came when David and Jean-Lin Journé attacked a much more general problem. They moved to $\mathbb{R}^n$ and studied singular integral operators which are defined by

$$T(f)(x) = p.v.\int K(x,y)f(y)dy,$$

where $K(x,y) = -K(y,x)$, $|K(x,y)| \le C|x-y|^{-n}$, and $|\nabla_x K(x,y)| \le C'|x-y|^{-n-1}$.

They discovered that T is bounded on $L^2(\mathbb{R}^n)$ if and only if $T(1) \in BMO(\mathbb{R}^n)$. Here $T(1)(x) = p.v.\int K(x,y)dy$ and in many situations this calculation is trivial. For instance, when $K_n(x,y) = \frac{(A(x)-A(y))^n}{(x-y)^{n+1}}$ is the n-th commutator,

$$p.v.\int K_n(x,y)dy = -\frac{1}{n}p.v.\int K_{n-1}(x,y)A'(y)dy,$$

which immediately yields Calderón's theorem. Complex methods are beaten by real variable methods and the surprising connection between Hardy spaces and L^2 estimates is explained. Indeed BMO is the dual of H^1.

A spectacular discovery by David, Journé, and S. Semmes is the generalization of the $T(1)$ theorem to spaces of homogeneous type.

This version of the $T(1)$ theorem will receive a careful exposition in this book. It paves the road to a broader program which is the extension to spaces of homogeneous type of the Littlewood–Paley theory. The Littlewood–Paley theory began with the fundamental achievements of J. E. Littlewood and R. E. A. C. Paley.

Let me say a few words on this discovery. We consider the Fourier series $\sum_{-\infty}^{\infty} c_k\, exp(ikx)$ of a 2π-periodic function $f(x)$ and we define the dyadic blocks $D_j(f)(x), j \in \mathbb{N}$, by

$$D_j f(x) = \sum_{2^j \le |k| < 2^{j+1}} c_k \, exp(ikx).$$

Then the square function $S(f)$ of Littlewood and Paley is defined by

$$S(f)(x) = \big(\sum_0^\infty |D_j(f)(x)|^2 \big)^{\frac{1}{2}}.$$

Littlewood and Paley proved that we have

$$c_p \|f\|_p \le |c_0| + \|S(f)\|_p \le C_p \|f\|_p$$

when $1 < p < \infty$.

The definition of the square function $S(f)$ was generalized by Elias Stein. Then $L^p[0, 2\pi]$ can be replaced by $L^p(\mathbb{R}^n)$. Jean-Michel Bony used Stein's version of the Littlewood–Paley theory to construct his famous paraproducts. Such paraproducts play a pivotal role in the proof of the $T(1)$ theorem.

The authors of this book show us how to extend the Littlewood–Paley theory to spaces of homogeneous type. This is a key achievement since most of the usual functional spaces admit simple characterizations using the Littlewood–Paley theory.

The last but not the least contribution of the authors is the construction of wavelet bases on spaces of homogeneous type. Once again, wavelets offer an alternative to Fourier analysis. As we know, wavelet analysis can be traced back to a fundamental identity discovered by Calderón. If ψ is a radial function in the Schwartz class with a vanishing integral and if, for $t > 0, \psi_t(x) = t^{-n}\psi(\frac{x}{t})$, then for $f \in L^2(\mathbb{R}^n)$ we have

$$f = c \int_0^\infty f * \widetilde{\psi}_t * \psi_t \frac{dt}{t},$$

where $c > 0$ is a normalizing factor and $\widetilde{\psi}(x) = \overline{\psi}(-x)$. In other words, one computes the wavelet coefficients by

$$W(y,t) = \int f(x)\overline{\psi}_t(x-y)dx$$

and one recovers f through

$$f(x) = c \int_0^\infty \int_{\mathbb{R}^n} W(y,t)\psi_t(x-y)dy\frac{dt}{t}.$$

Everything works as if the wavelets $\psi_{t,y}(x) = t^{-n/2}\psi(\frac{x-y}{t})$ were an orthonormal basis of $L^2(\mathbb{R}^n)$. Indeed, orthonormal wavelet bases exist. There exist

$2^n - 1$ functions $\psi_\epsilon \in \mathscr{S}(\mathbb{R}^n), \epsilon \in F, \#F = 2^n - 1$, such that the functions $\psi_\epsilon(x) = 2^{\frac{nj}{2}}\psi_\epsilon(2^j x - k), j \in \mathbb{Z}, k \in \mathbb{Z}^n, \epsilon \in F$, are an orthonormal basis of $L^2(\mathbb{R}^n)$.

The authors succeeded in generalizing the construction of wavelet bases to spaces of homogeneous type; however, wavelet bases are replaced by frames, which in many applications offer the same service.

One is amazed by the dramatic changes that occurred in analysis during the twentieth century. In the 1930s complex methods and Fourier series played a seminal role. After many improvements, mostly achieved by the Calderón–Zygmund school, the action takes place today on spaces of homogeneous type. No group structure is available, the Fourier transform is missing, but a version of harmonic analysis is still present. Indeed the geometry is conducting the analysis.

Donggao Deng passed away after completing a preliminary version of this book. In his last moments he knew his efforts were not in vain and that his collaboration with Yongsheng Han would eventually lead to this remarkable treatise.

China 2007 *Yves Meyer*

Chapter 2
The Boundedness of Calderón-Zygmund Operators on Wavelet Spaces

We first define test functions and wavelet spaces on spaces of homogeneous type. Then we prove the main result of this chapter, namely that Calderón-Zygmund operators whose kernels satisfy an additional smoothness condition are bounded on *wavelet spaces*. This result will be a crucial tool to provide wavelet expansions of functions and distributions on spaces of homogeneous type in the next chapter.

We first introduce test functions on spaces of homogeneous type.

Definition 2.1. Fix $0 < \gamma, \beta < \theta$. A function f defined on X is said to be a test function of type (x_0, r, β, γ), $x_0 \in X$, and $r > 0$, if f satisfies the following conditions:

(i) $|f(x)| \leq C \frac{r^\gamma}{(r+\rho(x,x_0))^{1+\gamma}}$;

(ii) $|f(x) - f(y)| \leq C \Big(\frac{\rho(x,y)}{r+\rho(x,x_0)} \Big)^\beta \frac{r^\gamma}{(r+\rho(x,x_0))^{1+\gamma}}$ for all $x, y \in X$ with $\rho(x,y) \leq \frac{1}{2A}(r + \rho(x, x_0))$.

Such functions exist and the reader will find a recipe two lines after Definition 1.2. If f is a test function of type (x_0, r, β, γ), we write $f \in \mathcal{M}(x_0, r, \beta, \gamma)$, and the norm of f in $\mathcal{M}(x_0, r, \beta, \gamma)$ is defined by

$$\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} = \inf\{C : (i) \text{ and } (ii) \text{ hold}\}.$$

One should observe that if $f \in \mathcal{M}(x_0, r, \beta, \gamma)$, then

$$\|f\|_1 \approx \|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)}.$$

We say that a function f is a *scaling function* if $f \in \mathcal{M}(x_0, r, \beta, \gamma)$ and $\int f(x) d\mu(x) = 1$.

Now fix $x_0 \in X$ and denote $\mathcal{M}(\beta, \gamma) = \mathcal{M}(x_0, 1, \beta, \gamma)$. It is easy to see that $\mathcal{M}(x_1, r, \beta, \gamma) = \mathcal{M}(\beta, \gamma)$ with equivalent norms for all $x_1 \in X$ and $r > 0$. Furthermore, it is also easy to check that $\mathcal{M}(\beta, \gamma)$ is a Banach space with respect to the norm in $\mathcal{M}(\beta, \gamma)$.

D. Deng and Y. Han, *Harmonic Analysis on Spaces of Homogeneous Type*, Lecture Notes in Mathematics 1966,

Definition 2.2. A function f defined on X is said to be a wavelet of type (x_0, r, β, γ) if $f \in \mathcal{M}(x_0, r, \beta, \gamma)$ and $\int f(x)d\mu(x) = 0$. We denote this by $f \in \mathcal{M}_0(x_0, r, \beta, \gamma)$.

These wavelets are named *molecules* by Guido Weiss. A compactly supported molecule is an *atom*. Atomic decompositions preluded wavelet analysis, as indicated in the Introduction. Moreover Caderón-Zygmund operators T satisfying $T(1) = T^*(1) = 0$ have the remarkable property map a molecule into a molecule. We use the notation $\mathcal{M}_0(\beta, \gamma)$, when the dependence in x_0 and r can be forgotten, as a space of wavelets with regularity (β, γ).

To study the boundedness of Calderón-Zygmund singular integral operators on a *wavelet space*, we define the following "strong" weak boundedness property.

Definition 2.3. An operator T defined by a distributional kernel K, is said to have the "strong weak boundedness property" if there exist $\eta > 0$ and $C < \infty$ such that

$$|\langle K, f\rangle| \le Cr \tag{2.1}$$

for all $f \in C_0^\eta(X \times X)$ with $\text{supp}(f) \subseteq B(x_1, r) \times B(y_1, r), x_1$ and $y_1 \in X, \|f\|_\infty \le 1, \|f(\cdot, y)\|_\eta \le r^{-\eta}$, and $\|f(x, \cdot)\|_\eta \le r^{-\eta}$ for all x and $y \in X$.

If T has the "strong weak boundedness property", we write $T \in SWBP$.

Note that if ψ and ϕ are functions satisfying the conditions in Definition 1.15, then $f(x, y) = \psi(x) \times \phi(y)$ satisfies the conditions in Definition 2.3, and hence $|\langle T\psi, \phi\rangle| = |\langle K, f\rangle| \le Cr$ if T has the "strong weak boundedness property". This means that the strong weak boundedness property implies the weak boundedness property. However, in the standard situation of $\mathbb{R}^n$, the weak boundedness property implies the strong one. Indeed any smooth function $f(x, y), x \in B, y \in B$, supported by $B \times B$ can be written, by a double Fourier series expansion, as $\sum \alpha_j f_j(x) g_j(y)$ with $\sum |\alpha_j| < \infty, \|f_j\|_{C_0^\beta} \le 1, \|g_j\|_{C_0^\beta} \le 1$.

If $T \in CZK(\epsilon)$, we say that $T^*(1) = 0$ if $\int T(f)(x)dx = 0$ for all $f \in \mathcal{M}_0(\beta, \gamma)$. Similarly, $T(1) = 0$ if $\int T^*(f)(x)dx = 0$ for all $f \in \mathcal{M}_0(\beta, \gamma)$.

The main result in this chapter is the following theorem.

Theorem 2.4. *Suppose that $T \in CZK(\epsilon) \cap SWBP$, and $T(1) = T^*(1) = 0$. Suppose further that $K(x, y)$, the kernel of T, satisfies the following condition:*

$$\begin{aligned} |K(x, y) - K(x', y) - K(x, y') + K(x', y')| \\ \le C\rho(x, x')^\epsilon \rho(y, y')^\epsilon \rho(x, y)^{-(1+2\epsilon)} \end{aligned} \tag{2.2}$$

for $\rho(x, x'), \rho(y, y') \le \frac{1}{2A}\rho(x, y)$. Then there exists a constant C such that for each wavelet $f \in \mathcal{M}_0(x_0, r, \beta, \gamma)$ with $x_0 \in X, r > 0$ and $0 < \beta, \gamma < \epsilon$, $Tf \in \mathcal{M}_0(x_0, r, \beta, \gamma)$. Moreover

$$\|T(f)\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \le C\|T\|\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \tag{2.3}$$

where $\|T\|$ denote the smallest constant in the "strong weak boundedness property" and in the estimates of the kernel of T.

Before proving Theorem 2.4, we observe that this theorem will provide wavelet expansions which, as in the standard case of $\mathbb{R}^n$, will be the building blocks of most functional spaces.

To prove Theorem 2.4, we first need the following lemma.

Lemma 2.5. *Suppose that T is a continuous linear operator from $\dot{C}_0^\eta$ to $(\dot{C}_0^\eta)'$ satisfying $T \in CZK(\epsilon) \cap SWBP$ with $\eta < \epsilon$, and $T(1) = 0$. Then there exists a constant C such that*

$$\|T\phi\|_\infty \le C \tag{2.4}$$

whenever there exist $x_0 \in X$ and $r > 0$ such that $\mathrm{supp}\phi \subseteq B(x_0, r)$ with $\|\phi\|_\infty \le 1$ and $\|\phi\|_\eta \le r^{-\eta}$.

Proof. We follow the idea of the proof in [M1]. Fix a function $\theta \in C^\infty(\mathbb{R})$ with the following properties: $\theta(x) = 1$ for $|x| \le 1$ and $\theta(x) = 0$ for $|x| > 2$. Let $\chi_0(x) = \theta(\frac{\rho(x,x_0)}{2r})$ and $\chi_1 = 1-\chi_0$. Then $\phi = \phi\chi_0$ and for all $\psi \in C_0^\eta(X)$,

$$\begin{aligned}\langle T\phi, \psi\rangle &= \langle K(x,y), \phi(y)\psi(x)\rangle = \langle K(x,y), \chi_0(y)\phi(y)\psi(x)\rangle \\ &= \langle K(x,y), \chi_0(y)[\phi(y) - \phi(x)]\psi(x)\rangle + \langle K(x,y), \chi_0(y)\phi(x)\psi(x)\rangle \\ &:= p + q\end{aligned}$$

where $K(x,y)$ is the distribution kernel of T.

To estimate p, let $\lambda_\delta(x,y) = \theta(\frac{\rho(x,y)}{\delta})$. Then

$$\begin{aligned}p &= \langle K(x,y), (1-\lambda_\delta(x,y))\chi_0(y)[\phi(y) - \phi(x)]\psi(x)\rangle \\ &\quad + \langle K(x,y), \lambda_\delta(x,y)\chi_0(y)[\phi(y) - \phi(x)]\psi(x)\rangle \\ &:= p_{1,\delta} + p_{2,\delta}.\end{aligned} \tag{2.5}$$

Since K is locally integrable on $\Omega = \{(x,y) \in X \times X : x \neq y\}$, the first term on the right hand side of (2.5) satisfies

$$\begin{aligned}|p_{1,\delta}| &= \left|\int_\Omega K(x,y)(1-\lambda_\delta(x,y))\chi_0(y)[\phi(y) - \phi(x)]\psi(x)d\mu(x)d\mu(y)\right| \\ &\le C\int_X\int_X |K(x,y)\chi_0(y)[\phi(y) - \phi(x)]\psi(x)|d\mu(x)d\mu(y) \\ &\le C\int_X |\psi(x)|d\mu(x) = C\|\psi\|_1.\end{aligned}$$

Thus it remains to show that $\lim_{\delta\to 0} p_{2,\delta} = 0$, i.e.,

$$\lim_{\delta\to 0}\langle K(x,y),\lambda_\delta(x,y)\chi_0(y)[\phi(y)-\phi(x)]\psi(x)\rangle = 0, \tag{2.6}$$

and it is here that we use the "strong" weak boundedness property of T :

$$|\langle K,f\rangle| \le Cr \tag{2.7}$$

for all $f \in C_0^\eta(X\times X)$ satisfying supp$f \subseteq B(x_0,r)\times B(y_0,r), \|f\|_\infty \le 1, \|f(\cdot,y)\|_\eta \le r^{-\eta}$ and $\|f(x,\cdot)\|_\eta \le r^{-\eta}$ for all $x,y\in X$.

To show (2.6), let $\{y_j\}_{j\in\mathbb{Z}} \in X$ be a maximal collection of points satisfying

$$\frac{1}{2}\delta < \inf_{j\neq k}\rho(y_j,y_k) \le \delta. \tag{2.8}$$

By the maximality of $\{y_j\}_{j\in\mathbb{Z}}$, we have that for each $x\in X$ there exists a point y_j such that $\rho(x,y_j)\le\delta$. Let $\eta_j(y)=\theta(\frac{\rho(y,y_j)}{\delta})$ and $\bar\eta_j(y) = \big[\sum_i \eta_i(y)\big]^{-1}\eta_j(y)$. To see that $\bar\eta_j$ is well defined, it suffices to show that for any $y\in X$, there are only finitely many η_j with $\eta_j(y)\neq 0$. This follows from the following fact: $\eta_j(y)\neq 0$ if and only if $\rho(y,y_j)\le 2\delta$ and hence this implies that $B(y_j,\delta)\subseteq B(y,4A\delta)$. Inequalities (2.8) show $B(y_j,\frac{\delta}{4A})\cap B(y_k,\frac{\delta}{4A})=\phi$, and thus there are at most CA points $y_j\in X$ such that $B(y_j,\frac{\delta}{4A})\subseteq B(y,4A\delta)$. Now let $\Gamma=\{j:\bar\eta_j(y)\chi_0(y)\neq 0\}$. Note that $\#\Gamma\le Cr\delta$ since $\mu(\text{supp}\chi_0)\sim r$ and $\mu(\text{supp}\bar\eta_j)\sim\delta$. We write

$$\lambda_\delta(x,y)\chi_0(y)[\phi(y)-\phi(x)]\psi(x) = \sum_{j\in\Gamma}\lambda_\delta(x,y)\bar\eta_j(y)\chi_0(y)[\phi(y)-\phi(x)]\psi(x),$$

and we obtain

$$\begin{aligned}&\langle K(x,y),\lambda_\delta(x,y)\chi_0(y)[\phi(y)-\phi(x)]\psi(x)\rangle\\ &\quad=\sum_{j\in\Gamma}\langle K(x,y),\lambda_\delta(x,y)\bar\eta_j(y)\chi_0(y)[\phi(y)-\phi(x)]\psi(x)\rangle.\end{aligned}$$

It is then easy to check that supp$\{\lambda_\delta(x,y)\bar\eta_j(y)\chi_0(y)[\phi(y)-\phi(x)]\psi(x)\}\subseteq B(y_j,3A\delta)\times B(y_j,2\delta)$ and

$$\|\lambda_\delta(x,y)\bar\eta_j(y)\chi_0(y)[\phi(y)-\phi(x)]\psi(x)\|_\infty \le C\delta^\eta$$

where C is a constant depending only on θ,ϕ,ψ,x_0, and r but not on δ and j.

We claim that

$$\|\lambda_\delta(.,y)\bar\eta_j(y)\chi_0(y)[\phi(y)-\phi(.)]\psi(.)\|_\eta \le C, \tag{2.9}$$

and

$$\|\lambda_\delta(x,.)\bar\eta_j(.)\chi_0(.)[\phi(.)-\phi(x)]\psi(x)\|_\eta \le C. \tag{2.10}$$

We accept (2.9) and (2.10) for the moment. Then, since T satisfies the "strong" weak boundedness property, we have

$$\begin{aligned}
&|\langle K(x,y), \lambda_\delta(x,y)\chi_0(y)[\phi(y)-\phi(x)]\psi(x)\rangle| \\
&\quad \le \sum_{j\in\Gamma} |\langle K(x,y), \lambda_\delta(x,y)\bar{\eta}_j(y)\chi_0(y)[\phi(y)-\phi(x)]\psi(x)\rangle| \\
&\quad \le \sum_{j\in\Gamma} C\mu(B(y_j, 3A\delta))\delta^\eta \le C\frac{r}{\delta}CA\delta\delta^\eta = CAr\delta^\eta
\end{aligned}$$

which yields (2.6).

It remains to show (2.9) and (2.10). We prove only (2.9) since the proof of (2.10) is similar. To show (2.9) it suffices to show that for $x, x_1 \in X$ and $\rho(x, x_1) \le \delta$,

$$\begin{aligned}
&|\bar{\eta}_j(y)\chi_0(y)|\big|\lambda_\delta(x,y)[\phi(y)-\phi(x)]\psi(x) - \lambda_\delta(x_1,y)[\phi(y)-\phi(x_1)]\psi(x_1)\big| \\
&\quad \le C\rho(x,x_1)^\eta,
\end{aligned}$$

since if $\rho(x, x_1) \ge \delta$, then the expansion on the left above is clearly bounded by

$$\begin{aligned}
&|\bar{\eta}_j(y)\chi_0(y)|\big\{|\lambda_\delta(x,y)[\phi(y)-\phi(x)]\psi(x)| + |\lambda_\delta(x_1,y)[\phi(y)-\phi(x_1)]\psi(x_1)|\big\} \\
&\quad \le C\delta^\eta \le C\rho(x,x_1)^\eta.
\end{aligned}$$

By the construction of $\bar{\eta}_j$, it follows that

$$|\bar{\eta}_j(y)\chi_0(y)| \le C$$

for all $y \in X$. Thus

$$\begin{aligned}
&|\bar{\eta}_j(y)\chi_0(y)||\lambda_\delta(x,y)[\phi(y)-\phi(x)]\psi(x) - \lambda_\delta(x_1,y)[\phi(y)-\phi(x_1)]\psi(x_1)| \\
&\quad \le C|\lambda_\delta(x,y)[\phi(y)-\phi(x)]\psi(x) - \lambda_\delta(x_1,y)[\phi(y)-\phi(x_1)]\psi(x_1)| \\
&\quad \le C|[\lambda_\delta(x,y)-\lambda_\delta(x_1,y)][\phi(y)-\phi(x)]\psi(x)| \\
&\quad\quad +|\lambda_\delta(x_1,y)[\phi(x)-\phi(x_1)]\psi(x)| \\
&\quad\quad +|\lambda_\delta(x_1,y)[\phi(y)-\phi(x_1)][\psi(x)-\psi(x_1)]| \\
&\quad := I + II + III.
\end{aligned}$$

Recall that $\rho(x, x_1) \le \delta$. If $\rho(x, y) > C\delta$, where C is a constant depending on A but not on δ, then $\lambda_\delta(x, y) = \lambda_\delta(x_1, y) = 0$, so $I = 0$. Thus we may assume that $\rho(x, y) \le C\delta$ and with θ in (1.7),

$$\begin{aligned}
I &\le C\Big|\frac{\rho(x,y)}{\delta} - \frac{\rho(x_1,y)}{\delta}\Big|\rho(x,y)^\eta \le C\delta^{\eta-1}\rho(x,x_1)^\theta[\rho(x,y)+\rho(x_1,y)]^{1-\theta} \\
&\le C\delta^{\eta-\theta}\rho(x,x_1)^\theta \le C\rho(x,x_1)^\eta
\end{aligned}$$

since we may assume $\eta \leq \theta$. Terms II and III are easy to estimate:

$$II \leq C\rho(x, x_1)^\eta,$$

$$III \leq C\rho(x, x_1)^\eta,$$

since we can assume that $\delta < 1$. This completes the proof of (2.9) and implies

$$|p| \leq C\|\psi\|_1.$$

To finish the proof of Lemma 2.5, we now estimate q. It suffices to show that for $x \in B(x_0, r)$,

$$|T\chi_0(x)| \leq C. \tag{2.11}$$

To see this, it is easy to check that $q = \langle T\chi_0, \phi\psi\rangle$, and hence (2.10) implies

$$|q| \leq \|T\chi_0\|_{L^\infty(B(x_0,r))}\|\phi\psi\|_{L^1(B(x_0,r))} \leq C\|\psi\|_1.$$

To show (2.11), we use Meyer's idea again ([M1]). Let $\psi \in C^\eta(X)$ with supp$\psi \subseteq B(x_0, r)$ and $\int \psi(x)d\mu(x) = 0$. By the facts that $T(1) = 0, \int \psi(x)d\mu(x) = 0$, and the conditions on K, we obtain

$$\begin{aligned}|\langle T\chi_0, \psi\rangle| = |-\langle T\chi_1, \psi\rangle| &= \left|\iint [K(x, y) - K(x_0, y)]\chi_1(y)\psi(x)d\mu(x)d\mu(y)\right| \\ &\leq C\|\psi\|_1.\end{aligned}$$

Thus, $T\chi_0(x) = \omega + \gamma(x)$ for $x \in B(x_0, r)$, where ω is a constant and $\|\gamma\|_\infty \leq C$. To estimate ω, choose $\phi_1 \in C_0^\eta(X)$ with supp $\phi_1 \subseteq B(x_0, r), \|\phi_1\|_\infty \leq 1, \|\phi_1\|_\eta \leq r^{-\eta}$ and $\int \phi_1(x)d\mu(x) = Cr$. We then have, by the "strong" weak boundedness property of T,

$$\left|Cr\omega + \int \phi_1(x)\gamma(x)d\mu(x)\right| = |\langle T\chi_0, \phi_1\rangle| \leq Cr$$

which implies $|\omega| \leq C$ and hence Lemma 2.5. ∎

We remark that the calculation above, together with the dominated convergence theorem and $T1 = 0$, yields the following integral representation:

$$\begin{aligned}&\langle T\phi, \psi\rangle \\ &\quad = \int_\Omega K(x, y)\{\chi_0(y)[\phi(y) - \phi(x)] - \chi_1(y)\phi(x)\}\psi(x)d\mu(y)d\mu(x)\end{aligned} \tag{2.12}$$

and

$$\begin{aligned}&\langle K(x, y), [\phi(y) - \phi(x)]\chi_0(y)\rangle \\ &\quad = \lim_{\delta\to 0}\int_{\rho(x,y)\geq\delta} K(x, y)\chi_0(y)[\phi(y) - \phi(x)]d\mu(y)\end{aligned} \tag{2.13}$$

where χ_0, ϕ and ψ are defined as above.

We return to prove the Theorem 2.4. Fix a function $\theta \in C^1(\mathbb{R})$ with supp $\theta \subseteq \{x \in \mathbb{R} : |x| \leq 2\}$ and $\theta = 1$ on $\{x \in \mathbb{R} : |x| \leq 1\}$. Suppose that $f \in \mathcal{M}_0(x_0, r, \beta, \gamma)$ with $x_0 \in X, r > 0$ and $0 < \beta, \gamma < \epsilon$. We first prove that $T(f)(x)$ satisfies the size condition (i) of Definition 2.1. To do this, we first consider the case where $\rho(x, x_0) \leq 5r$. Set $1 = \xi(y) + \eta(y)$ where $\xi(y) = \theta(\frac{\rho(y,x_0)}{10Ar})$. Then we have

$$T(f)(x) = \int K(x,y)\xi(y)[f(y) - f(x)]d\mu(y) + \int K(x,y)\eta(y)f(y)d\mu(y)$$
$$+f(x)\int K(x,y)\xi(y)d\mu(y) := I + II + III.$$

Using (2.13),

$$\begin{aligned}|I| &\leq C\int_{\rho(x,y)\leq 25A^2r} |K(x,y)||f(y) - f(x)|d\mu(y)\\ &\leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)}\int_{\rho(x,y)\leq 25A^2r} \rho(x,y)^{-1}\frac{\rho(x,y)^\beta}{r^{1+\beta}}d\mu(y)\\ &\leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)}r^{-1}.\end{aligned}$$

By Lemma 2.5,

$$|III| \leq C|f(x)| \leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)}r^{-1}.$$

For term II we have

$$\begin{aligned}|II| &\leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)}\int_{\rho(x,y)\geq 10Ar} \rho(x,y)^{-1}\frac{r^\gamma}{\rho(y,x_0)^{1+\gamma}}d\mu(y)\\ &\leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)}r^{-1}\end{aligned}$$

since $\rho(x, x_0) \leq 5r$.

This implies that $T(f)(x)$ satisfies (i) of Definition 2.1 with $\rho(x, x_0) \leq 5r$. Consider now $\rho(x, x_0) = R > 5r$. Following the proof in [M1], set $1 = I(y) + J(y) + L(y)$, where $I(y) = \theta(\frac{4A\rho(y,x)}{R}), J(y) = \theta(\frac{4A\rho(y,x_0)}{R})$, and $f_1(y) = f(y)I(y), f_2(y) = f(y)J(y)$, and $f_3(y) = f(y)L(y)$. Then it is easy to check the following estimates:

$$|f_1(y)| \leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)}\frac{r^\gamma}{R^{1+\gamma}}; \tag{2.14}$$

$$\begin{aligned}|f_1(y) - f_1(y')| &\leq |I(y)||f(y) - f(y')| + |f(y')||I(y) - I(y')| \qquad (2.15)\\ &\leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)}\frac{\rho(y,y')^\beta}{R^\beta}\frac{r^\gamma}{R^{1+\gamma}}\end{aligned}$$

for all y and y';

$$|f_3(y)| \leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \frac{r^\gamma}{\rho(y,x_0)^{1+\gamma}} \chi_{\{y\in X:\rho(y,x_0)>\frac{1}{4A}R\}}; \quad (2.16)$$

$$\int |f_3(y)| d\mu(y) \leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \frac{r^\gamma}{R^\gamma}; \quad (2.17)$$

$$\begin{aligned} \left|\int f_2(y)d\mu(y)\right| &= \left| -\int f_1(y)d\mu(y) - \int f_3(y)d\mu(y)\right| \quad (2.18) \\ &\leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \frac{r^\gamma}{R^\gamma}. \end{aligned}$$

We write

$$\begin{aligned} T(f_1)(x) &= \int K(x,y)u(y)[f_1(y) - f_1(x)]d\mu(y) + f_1(x) \int K(x,y)u(y)d\mu(y) \\ &= \sigma_1(x) + \sigma_2(x) \end{aligned}$$

where $u(y) = \theta(\frac{2A\rho(x,y)}{R})$. Applying the estimate (2.15) and Lemma 2.5, we obtain

$$\begin{aligned} |\sigma_1(x)| &\leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \int_{\rho(x,y)\leq \frac{R}{A}} \rho(x,y)^{-1} \frac{\rho(x,y)^\beta}{R^\beta} \frac{r^\gamma}{R^{1+\gamma}} d\mu(y) \\ &\leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \frac{r^\gamma}{R^{1+\gamma}}; \end{aligned}$$

and

$$|\sigma_2(x)| \leq C|f_1(x)| \leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \frac{r^\gamma}{R^{1+\gamma}}.$$

Notice that x is not in the support of f_2. We can write

$$\begin{aligned} T(f_2)(x) &= \int [K(x,y) - K(x,x_0)]f_2(y)d\mu(y) + K(x,x_0) \int f_2(y)d\mu(y) \\ &= \delta_1(x) + \delta_2(x). \end{aligned}$$

Using the estimates on the kernel of T and on f_2 in (2.18), we then get

$$\begin{aligned} |\delta_1(x)| &\leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \int_{\rho(x_0,y)\leq \frac{R}{2A}} \frac{\rho(x_0,y)^\epsilon}{R^{1+\epsilon}} \frac{r^\gamma}{\rho(x_0,y)^{1+\gamma}} d\mu(y) \\ &\leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \frac{r^\gamma}{R^{1+\gamma}} \end{aligned}$$

since $\gamma < \epsilon$, and

$$|\delta_2(x)| \leq CR^{-1}\left|\int f_2(y)d\mu(y)\right| \leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \frac{r^\gamma}{R^{1+\gamma}}.$$

Finally, since x is not in the support of f_3, (2.16) implies

$$
\begin{aligned}
|T(f_3)(x)| &\le C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \int_{\rho(x,y)\ge \frac{R}{4A},\rho(x_0,y)\ge \frac{R}{4A}} \rho(x,y)^{-1} \frac{r^\gamma}{\rho(x_0,y)^{1+\gamma}} d\mu(y) \\
&\le C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \frac{r^\gamma}{R^{1+\gamma}}.
\end{aligned}
$$

This yields that $T(f)(x)$ satisfies (i) of Definition 2.1 for $\rho(x,x_0) > 5r$ and hence, estimate (i) of Definition 2.1 for all $x \in X$.

Now we prove that $T(f)(x)$ satisfies the smoothness condition (ii) of Definition 2.1. To do this, set $\rho(x,x_0) = R$ and $\rho(x,x') = \delta$. We consider first the case where $R \ge 10r$ and $\delta \le \frac{1}{20A^2}(r+R)$. As in the above, set $1 = I(y) + J(y) + L(y)$, where $I(y) = \theta(\frac{8A\rho(y,x)}{R})$, $J(y) = \theta(\frac{8A\rho(y,x_0)}{R})$, and $f_1(y) = f(y)I(y)$, $f_2(y) = f(y)J(y)$, and $f_3(y) = f(y)L(y)$. We write

$$
\begin{aligned}
T(f_1)(x) = &\int K(x,y)u(y)[f_1(y) - f_1(x)]d\mu(y) \\
&+ \int K(x,y)v(y)f_1(y)d\mu(y) + f_1(x)\int K(x,y)u(y)d\mu(y)
\end{aligned}
$$

where $u(y) = \theta(\frac{\rho(x,y)}{2A\delta})$ and $v(y) = 1-u(y)$. Denote the first term of the above right-hand side by $p(x)$ and the last two terms by $q(x)$. The size condition of K and the smoothness of f_1 in (2.15) yield

$$
\begin{aligned}
|p(x)| &\le C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \int_{\rho(x,y)\le 4A\delta} \rho(x,y)^{-1} \frac{\rho(x,y)^\beta}{R^\beta} \frac{r^\gamma}{R^{1+\gamma}} d\mu(y) \\
&\le C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \frac{\delta^\beta}{R^\beta} \frac{r^\gamma}{R^{1+\gamma}}.
\end{aligned}
$$

This estimate still holds with x replaced by x' for $\rho(x,x') = \delta$. Thus

$$
|p(x) - p(x')| \le C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \frac{\delta^\beta}{R^\beta} \frac{r^\gamma}{R^{1+\gamma}}.
$$

For $q(x)$, using the condition $T1 = 0$, we obtain

$$
\begin{aligned}
q(x) - q(x') &= \int [K(x,y) - K(x',y)]v(y)[f_1(y) - f_1(x)]d\mu(y) \\
&\quad +[f_1(y) - f_1(x)] \int K(x,y)u(y)d\mu(y) \\
&= I + II.
\end{aligned}
$$

Using Lemma 2.5 and the estimate for f_1 in (2.15),

$$
|II| \le C|f_1(x) - f_1(x')| \le C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \frac{\delta^\beta}{R^\beta} \frac{r^\gamma}{R^{1+\gamma}}.
$$

Observing

$$|f_1(y) - f_1(x)||v(y)| \leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \frac{\rho(x,y)^\beta}{R^\beta} \frac{r^\gamma}{R^{1+\gamma}}$$

for all $y \in X$, we see that I is dominated by

$$\begin{aligned} C \int\limits_{\rho(x,y)\geq 2A\delta} & |K(x,y) - K(x',y)||v(y)||f_1(y) - f_1(x)|d\mu(y) \\ & \leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \int_{\rho(x,y)\geq 2A\delta} \frac{\rho(x,x')^\epsilon}{\rho(x,y)^{1+\epsilon}} \frac{\rho(x,y)^\beta}{R^\beta} \frac{r^\gamma}{R^{1+\gamma}} d\mu(y) \\ & \leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \frac{\delta^\beta}{R^\beta} \frac{r^\gamma}{R^{1+\gamma}} \end{aligned}$$

since $\beta < \epsilon$. This implies

$$|T(f_1)(x) - T(f_1)(x')| \leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \frac{\delta^\beta}{R^\beta} \frac{r^\gamma}{R^{1+\gamma}}.$$

Note that for $\rho(x,x') = \delta \leq \frac{1}{20A^2}(r+R)$ and $R \geq 10r, x$ and x' are not in the supports of f_2 and f_3. Using the condition for K and the estimate for f_2 in (2.18), then

$$\begin{aligned} |T(f_2)(x) - T(f_2)(x')| &= \left| \int [K(x,y) - K(x',y)] f_2(y) d\mu(y) \right| \\ \leq & \int |K(x,y) - K(x',y) - K(x,x_0) - K(x',x_0)||f_2(y)|d\mu(y) \\ & + |K(x,x_0) - K(x',x_0)| \left| \int f_2(y) d\mu(y) \right| \\ \leq & C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \left\{ \int_{\rho(x_0,y)\leq \frac{R}{4A}} \frac{\rho(x,x')^\epsilon \rho(y,x_0)^\epsilon}{R^{2+\epsilon}} \frac{r^\gamma}{\rho(y,x_0)^{1+\gamma}} d\mu(y) \right. \\ & \left. + \frac{\delta^\epsilon}{R^{1+\epsilon}} \frac{r^\gamma}{R^\gamma} \right\} \\ \leq & C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \frac{\delta^\epsilon}{R^\epsilon} \frac{r^\gamma}{R^{1+\gamma}} \end{aligned}$$

since $\gamma < \epsilon$. Finally, we have

$$\begin{aligned} |T(f_3)(x) - T(f_3)(x')| &= \left| \int [K(x,y) - K(x',y)] f_3(y) d\mu(y) \right| \\ \leq C & \int_{\rho(x,y)\geq \frac{R}{8A} \geq 2A\delta} \frac{\rho(x,x')^\epsilon}{\rho(x,y)^{1+\epsilon}} |f_3(y)| d\mu(y) \leq C\|f\|_{\mathcal{M}(x_0,r,\beta,\gamma)} \frac{\delta^\epsilon}{R^\epsilon} \frac{r^\gamma}{R^{1+\gamma}}. \end{aligned}$$

These estimates imply that $T(f)(x)$ satisfies the condition (ii) of Definition 2.1 for the case where $\rho(x,x_0) = R \geq 10r$ and $\rho(x,x') = \delta \leq \frac{1}{20A^2}(r+R)$. We now consider the other cases. Note first that if $\rho(x,x_0) = R$ and $\frac{1}{2A}(r+R) \geq$

$\rho(x,x') = \delta \geq \frac{1}{20A^2}(r+R)$, then the estimate (ii) of Definition 2.1 for $T(f)(x)$ follows from the estimate (i) of Definition 2.1 for $T(f)(x)$. So we only need to consider the case where $R \leq 10r$ and $\delta \leq \frac{1}{20A^2}(r+R)$. This case is similar and easier. In fact, all we need to do is to replace R in the proof above by r. We leave these details to the reader. The proof of Theorem 2.4 is completed.

We remark that the condition in (2.2) is also necessary for the boundedness of Calderón-Zygmund operators on *wavelet spaces*. To be precise, in the next chapter, we will prove all kinds of Calderón's identities and use them to provide all kinds of wavelet expansions of functions and distributions on spaces of homogeneous type. Suppose that T is a Calderón-Zygmund operator and maps the wavelet space $\mathcal{M}_0(x_0, r, \beta, \gamma)$ to itself. By the wavelet expansion given in Theorem 3.25 below, $K(x,y)$, the kernel of T, can be written as $K(x,y) = \sum\limits_{\lambda\in\Lambda} T(\widetilde{\psi}_\lambda)(x)\psi_\lambda(y)$. Since $\widetilde{\psi}_\lambda(x)$ is a wavelet, by the assumption on $T, T(\widetilde{\psi}_\lambda)(x)$ is also a wavelet. Then one can easily check that $K(x,y)$ satisfies the condition (2.2) but the exponent ϵ must be replaced by ϵ' with $0 < \epsilon' < \beta, \gamma$. We leave these details to the reader.