Continuity of Calderón-Zygmund Operators on the Space BMO

Qixiang YANG

Departement of Mathematics
Wuhan University
430072 Hubei — China
qxyang@whu.edu.cn

ABSTRACT

David and Journé discovered a criterion for the continuity on L^2 of Calderón-Zygmund operators defined by singular integrals. In their approach the distributional kernel of the given operator is locally Hölder continuous outside the diagonal. The aim of this paper is to prove a David-Journé theorem where this smoothness assumption is replaced by a weaker one. Our approach strongly relies on an algorithm developed by Beylkin, Coifman, and Rokhlin.

Key words: $T(1)$ theorem for BMO, Hörmander condition, Daubechies wavelets, Bony paraproducts, BCR algorithm, pseudo-annular decomposition.

2000 Mathematics Subject Classification: primary 42B20, 42C15; secondary 47A66.

1. The main theorem

Boundedness of Calderón-Zygmund operators under optimally weak regularity assumptions is an intriguing problem. The ultimate goal is to replace the standard regularity assumptions on the distributional kernel of the operator (see (2)) by Hörmander condition (see (6)). For the reader's convenience, the definition of the distributional kernel $K(x, y)$ of an operator T will be given now. Let T be a linear operator which is defined on the space $S(\mathbb{R}^n)$ of testing functions with values in the dual space $S'(\mathbb{R}^n)$ of tempered distributions. The distributional kernel $K(x, y)$ of $T : S(\mathbb{R}^n) \to S'(\mathbb{R}^n)$ is a distribution in $2n$ variables which is defined by the following condition: for every

pair (f, g) of two testing functions, we have $\langle T(f), g \rangle = \iint K(x, y) g(x) f(y) \, dx \, dy$. In the standard theory of Calderón-Zygmund operators the distributional kernel $K(x, y)$ of T is locally Hölder continuous in the open set $\Omega = \{(x, y) \in \mathbb{R}^{2n}; x \neq y\}$ and satisfies the following properties:

$$|K(x, y)| \leq \frac{C}{|x-y|^n} \text{ for every } x \text{ and every } y \neq x. \tag{1}$$

There exists an exponent γ belonging to $(0, 1)$ and a constant C such that $|K(x, y) - K(x, y')| + |K(y, x) - K(y', x)| \leq \frac{C|y-y'|^\gamma}{|x-y|^{n+\gamma}}$ when $|x - y| \geq$ $2|y - y'| > 0$. $\tag{2}$

These two properties are describing the restriction to $\Omega = \{(x, y) \in \mathbb{R}^{2n}; x \neq y\}$ of the distributional kernel of T. This kernel is not locally integrable and the integral $Tf(x) = \int K(x, y) f(y) \, dy$ is a singular integral. That is why T is often referred to as a singular integral operator. In [3] David and Journé proved a theorem which immediately became famous under the name of the "$T(1)$ theorem." It says the following. Let T be an operator whose kernel satisfies (1) and (2). Then T is bounded on L^2 if and only if the following two conditions are satisfied:

$|\iint K(x, y) f(x) g(y) \, dx \, dy| \leq C|Q|(\|f\|_\infty + |Q|^{\frac{1}{n}} \|f'\|_\infty)(\|g\|_\infty + |Q|^{\frac{1}{n}} \|g'\|_\infty)$
for every cube Q with volume $|Q|$ and every pair (f, g) of two continuously $\tag{3}$
differentiable functions supported by Q.

$$\alpha = T(1) \in \text{BMO}, \quad \beta = T^*(1) \in \text{BMO}. \tag{4}$$

Let us comment on (3) and (4). Property (3) is usually referred as the "weak boundedness property." It is a scale invariant version of the hypothesis that T is defined on the space of compactly supported testing functions with values in the space of tempered distributions. The continuity on L^2 obviously implies (3). In (4) 1 stands for the constant function 1. The meaning of $T(1)$ is not clear. It is a tempered distribution modulo a constant function. Indeed if ψ denotes a test function in the Schwartz class with a vanishing integral then $\langle T(1), \psi \rangle = \langle 1, T^*(\psi) \rangle$ which is well defined since (2) implies that $T^*(\psi)$ is $O(|x|^{-n-\gamma})$ at infinity. The space BMO of functions with bounded mean oscillations was defined and studied by John and Nirenberg in [7]. This space consists of all functions $f(x) \in L^2_{\text{loc}}$ for which a constant C exists with the following property: for every ball $B \subset \mathbb{R}^n$ we have $(\frac{1}{|B|} \int_B |f(x) - m_B f|^2 dx)^{\frac{1}{2}} \leq C$ where $m_B f = \frac{1}{|B|} \int_B f dx$. The optimal C is the norm of f in BMO which implies that constant functions have a zero norm. Therefore a function in BMO is only defined modulo a constant. The space BMO contains L^∞ but enjoys some important properties which are not shared by L^∞. Calderón-Zygmund operators are not bounded on L^∞ but are bounded on BMO. More precisely Jack Peetre proved that any operator T which is bounded on L^2 and satisfies (2) maps L^∞ into the space BMO. Furthermore T maps BMO into itself if and only if $T(1) = 0$. The argument used by Peetre gives more as it will be told below. To conclude (3) and (4) are necessary to the continuity of T on L^2.

The converse implication in the "$T(1)$ theorem" is deeper and says that (2), (3), and (4) imply the continuity on L^2. To prove it David and Journé introduced two auxiliary operators T_α and T_β which are defined by the following conditions: $T_\alpha(f)$ is the paraproduct between f and $\alpha \in$ BMO and similarly the adjoint of T_β is the paraproduct between f and $\beta \in$ BMO. We remind the reader that $T_\alpha(f) = \sum_{-\infty}^{\infty} \Delta_j(\alpha) S_{j-3}(f)$ where the notations refer to the classical Littlewood-Paley analysis. We then obviously have $T_\alpha(1) = \alpha$ and $T_\alpha^*(1) = 0$. Nowadays the boundedness of these two operators on L^2 is an easy exercise using the famous characterization of BMO by Carleson measures (see Theorem 2.2). The continuity of T on L^2 is then reduced to the boundedness of $R = T - T_\alpha - T_\beta$. In other words (4) can be replaced by

$$T(1) = T^*(1) = 0. \tag{5}$$

If (2), (3), and (5) are satisfied the continuity of T can be proved using a beautiful lemma devised by Mischa Cotlar and improved by Cotlar and Stein. We do not say more since Cotlar's lemma does not apply to our framework and the reader is referred to [11, 12].

This decomposition of T as a sum $T = T_\alpha + T_\beta + R$ can be used for proving estimates on most of the functional spaces. When T satisfies (1), (2), (3), and (5) one writes $T \in \mathrm{OpE}_\gamma$. For $T \in \mathrm{OpE}_\gamma$ and $0 \leq s < \gamma$, P. G. Lemarié-Rieusset proved that T is continuous on the homogeneous Besov spaces $\dot{B}_p^{s,q}$ [9]. Frazier, Jawerth, Han, and Weiss extended this theorem to the Triebel-Lizorkin spaces $\dot{F}_p^{s,q}$. Indeed, they proved in [5] that an atom is mapped into a molecule by T.

More difficult problems arise if one tries to replace (2) by a weaker condition.

Definition 1.1. One writes $T \in \mathrm{OpH}$ if (3) and (5) are satisfied together with the following Hörmander condition:

> There exists a constant C such that for every $y \in \mathbb{R}^n$ and every $y' \neq y$, we have $\int_{|x-y| \geq 2|y-y'|} \{|K(x,y) - K(x,y')| + |K(y,x) - K(y',x)|\}\, dx \leq C.$ (6)

The Hörmander condition is weaker than (2) and is needed to give a meaning to $T(1)$ and $T^*(1)$ in (5). At the present time we do not know whether any operator $T \in \mathrm{OpH}$ is bounded on L^2 or not. More generally the continuity of $T \in \mathrm{OpH}$ on Besov or Triebel-Lizorkin spaces is raising many interesting problems. Let us begin with two observations.

Lemma 1.2. *If $T \in \mathrm{OpH}$, then T maps the homogeneous Besov space $\dot{B}_1^{0,1}$ into L^1.*

This is proved in [10] and it does not require the boundedness of T on L^2. We obviously have $\dot{B}_1^{0,1} \subset H^1$ where H^1 denotes the Hardy space. Then a remarkable result by Jack Peetre says more.

Lemma 1.3. *Any $T \in \mathrm{OpH}$ which is bounded on L^2 maps the Hardy space H^1 into L^1.*

We say that T is a convolution operator if $T(f) = f * S$ for some generalized function S. If $T \in \mathrm{OpH}$ is a convolution operator then T is bounded on L^2 and on the Hardy space H^1. Let us sketch the proof of these results. Let $\psi_j(x) = 2^{nj}\psi(2^j x)$ where ψ is a function in the Schwartz class with a vanishing integral. This sequence ψ_j is bounded in $\dot{B}_1^{0,1}$. Therefore $\|S * \psi_j\|_1$ is bounded by Lemma 1.2. Moving to the Fourier transforms $\hat{S}\hat{\psi}_j$ is a bounded sequence in L^∞ and this implies $\hat{S} \in L^\infty$. Therefore T is bounded on L^2. We now treat the Hardy space case when $T \in \mathrm{OpH}$ is a convolution operator. If R_j, $1 \le j \le n$, denote the Riesz transforms then $f \in H^1$ means that f and its n Riesz transforms $R_j(f)$ altogether belong to L^1. Moreover the Riesz transforms are bounded on H^1. These properties imply that every convolution operator T which maps H^1 into L^1 is mapping H^1 into H^1. Indeed we have $R_j T = T R_j$ which ends the proof. Arguing by duality it implies that T is bounded on the space BMO. One would like to extend this result to every $T \in \mathrm{OpH}$. Theorem 1.6 below yields a partial answer. The kernel $K(x,y)$ is assumed to be locally integrable in $\Omega = \{(x,y) \in \mathbb{R}^{2n}; x \ne y\}$ and new moduli of continuity are now defined as in [10]:

Definition 1.4. For $(u,v,y) \in \mathbb{R}^{3n}$, for $r > 0$ and for each integer $R \in \mathbb{N}$, one considers the integrals

$$I(R,r,u,v,y) = \int\limits_{2^R r \le |x-y| < 2^{R+1} r} \{|K(x+u,y+v) - K(x,y)| + |K(y+u,x+v) - K(y,x)|\}\, dx.$$

One then defines $\varepsilon(K,R) = \varepsilon(R)$ as being the supremum of $I(R,r,u,v,y)$ over all $r > 0$, all $(u,v) \in \mathbb{R}^{2n}$ fulfilling $|u| + |v| \le r$, and all $y \in \mathbb{R}^n$.

For $a > 0$ and $u \in \mathbb{R}^n$, the kernels $K(x,y)$ and $\tilde{K}(x,y) = a^n K(ax + u, ay + u)$ are defining two operators T and $\tilde{T}$ which have the same operator norm acting on L^p. One immediately notices that $\epsilon(\tilde{K}, R) = \epsilon(K, R)$.

One could also consider the double integral

$$J(R,r,u,v,z) = r^{-n} \iint\limits_{\substack{2^R r \le |x-y| < 2^{R+1} r \\ |y-z| \le r}} \{|K(x+u,y+v) - K(x,y)| + |K(y+u,x+v) - K(y,x)|\}\, dx\, dy$$

and define $\eta(R)$ as being the supremum of $J(R,r,u,v,z)$ over (r,u,v,z) as above. All the theorems which will be proved in this paper will remain valid if $\varepsilon(R)$ is being replaced by $\eta(R)$.

Let $\mu \ge 0$ be an exponent. A new class of Calderón-Zygmund operators is now defined:

Definition 1.5. One writes $T \in \mathrm{OpM}_\mu$ if (3), (5) are satisfied together with

$$\sum_{R \ge 1} R^\mu \varepsilon(R) < +\infty.$$

We have $\mathrm{OpM}_\mu \subset \mathrm{OpM}_\nu$ if $\nu \leq \mu$. If $T \in \mathrm{OpM}_\mu$ and $\mu \geq 0$, then $T \in \mathrm{OpH}$. The property $T \in \mathrm{OpM}_0$ is slightly more precise that Hörmander's condition (6). If $T \in \mathrm{OpM}_\mu$ so does its adjoint T^*.

Theorem 2.6 in the next section says that $T \in \mathrm{OpE}_\gamma$ implies $\varepsilon(R) < C2^{-R\gamma}$. But $\varepsilon(R) \leq CR^{-\beta}$ and $\beta > 1 + \mu$ imply $T \in \mathrm{OpM}_\mu$. Obviously $\mathrm{OpE}_\gamma \subset \mathrm{OpM}_\mu$.

Meyer proved the following: if $T \in \mathrm{OpM}_1$ then T and its adjoint T^* are bounded on $\dot{B}_1^{0,1}$. Therefore T is bounded on L^2. The hypothesis $T \in \mathrm{OpM}_1$ was at that time the weakest regularity assumption implying L^2 estimates. The proof uses the action of T on the so-called weak molecules [10]. Afterwards Han and Hofmann proved that if T belongs to a space which is slightly different from OpM_1 then for $1 \leq p, q \leq \infty$, T is bounded on $\dot{B}_p^{0,q}$. Moreover for $1 < p, q < \infty$, T is bounded on $\dot{F}_p^{0,q}$ [8]. But these methods do not extend to the continuity on $\mathrm{BMO} = (\dot{F}_1^{0,2})^*$. Deng, Yan, and Yang constructed an operator $T \in \mathrm{OpH}$ which is neither bounded on $\dot{B}_1^{0,1}$ nor bounded on $\dot{F}_1^{0,2}$ [5]. This operator is not a convolution operator. This construction raises the following problem. What is the smallest exponent μ such that every $T \in \mathrm{OpM}_\mu$ is bounded on BMO? The answer given by Theorem 1.6 says that this exponent belongs to the interval $[1, 3/2]$. Concerning the continuity on L^2 the author proved that the minimal exponent ν with the property that every $T \in \mathrm{OpM}_\nu$ is bounded on L^2 belongs to $[0, 1/2]$. This is striking since it says that for $T \in \mathrm{OpH}$ the continuity on L^2 does not imply the continuity on BMO. This sharply contrasts with the case of a convolution operator.

We will construct an operator T belonging to OpM_μ for every $\mu \in (0, 1)$ with the following two properties: for $1 \leq p < \infty$, $1 \leq q \leq \infty$, T is bounded on $\dot{F}_p^{0,q}$ but T is not bounded on BMO. On the other hand, we will prove that every $T \in \mathrm{OpM}_{3/2}$ is bounded on BMO. Our main result reads as follows:

Theorem 1.6.

(i) *Every $T \in \mathrm{OpM}_{3/2}$ is bounded on* BMO.

(ii) *There exists an operator T with the following properties:*

    (a) *For all $\mu \in (0, 1)$ we have $T \in \mathrm{OpM}_\mu$.*

    (b) *For all $1 \leq p < \infty$, $1 \leq q \leq \infty$, T is bounded on $\dot{F}_p^{0,q}$ and on $\dot{B}_p^{0,q}$.*

    (c) *However T is not bounded on* BMO.

Theorem 5.2 in section 5 slightly improves (i). The proof of (i) begins with writing the BCR analysis of T. The BCR algorithm of Beylkin, Coifman, and Rokhlin will be described in section 2. Next an improved analysis of T will be introduced and the building blocks T_R, $R \geq 0$, provided by this analysis will be named "band operators." Then the norm of T_R acting on BMO will be estimated. These norm estimates will imply Theorem 1.6. The key ingredient in this estimation is the characterization

of BMO by size estimates on wavelet coefficients. A variation of such method can also be applied to the approximation of symbol operators, see [15].

This paper is organized as follows. The BCR algorithm will be detailed in section 2. Using this algorithm, any operator T can be analyzed through $2^{2n} - 1$ sequences M_j^ε, $\varepsilon \in \{0, 1\}^{2n} \setminus \{0\}$, $j \in \mathbb{Z}$, of matrices. These matrices are named the "non-standard representation" of the given operator and the corresponding entries $a_{j,k,l}^\varepsilon$, $j \in \mathbb{Z}$, $k, l \in \mathbb{Z}^n$, are the "non-standard entries" of T. The action of T on a function $f \in L^2$ is not given by a standard algorithm (see Lemma 2.5 below). Every operator $T \in \mathrm{OpE}_\gamma$ will then be characterized by simple size estimates on these non-standard entries. An improved version of the BCR algorithm yields $T = \sum_{\varepsilon_1, \varepsilon_2} \sum_{R \geq 0} T_R^{\varepsilon_1, \varepsilon_2}$. These building blocks $T_R^{\varepsilon_1, \varepsilon_2}$ are named "band operators" and this series is the "pseudo-annular decomposition." The modulus of continuity $\varepsilon(R)$ which is described in Definition 1.1 is now playing a pivotal role. A key result relates the non-standard entries of $T_R^{\varepsilon_1, \varepsilon_2}$ to $\varepsilon(R)$. New norms $\|A_R^{\varepsilon_1, \varepsilon_2}\|_p$ will be defined. Finally $\|T_R^{\varepsilon_1, \varepsilon_2}\|_{L^2 \to L^2}$ is estimated by $\|A_R^{\varepsilon_1, \varepsilon_2}\|_p$. A function $f \in \mathrm{BMO}$ will be identified with its wavelet coefficients $\{a_{j,k}^\varepsilon\}_{(\varepsilon, j, k) \in \Lambda}$. A characterization of BMO using the famous "Carleson measures" is given by weighted l^2 estimates of the wavelet coefficients $\{a_{j,k}^\varepsilon\}_{(\varepsilon, j, k)}$. We are led to computing the wavelet coefficients of $T_R^{\varepsilon_1, \varepsilon_2} f$ for each Carleson box Q. The function $f(x)$ is then split into three pieces $f_Q^i(x)$, $i = 1, 2, 3$ which are adapted to the cube Q. The L^2 norm of the main term $T_R^{\varepsilon_1, \varepsilon_2} f_Q^1$ is estimated by $\|T_R^{(\varepsilon_1, \varepsilon_2)}\|_{L^2 \to L^2} \|f_Q^1\|_2$. The other terms will be treated as error terms and their bounds only depend on the norm of f in the Besov space $\dot{B}_\infty^{0, \infty}$. In the last section some counter-examples are detailed.

2. Wavelets and generalized functions

For the reader's convenience some classical results will be listed.

2.1. Daubechies wavelets

For any integer M, Ingrid Daubechies constructed a pair $\varphi(x) = \varphi_M(x)$, $\psi(x) = \psi_M(x)$ of two functions of a real variable x with the following properties. The function $\varphi(x)$ has q_M continuous derivatives (where q_M tends to infinity as M tends to infinity), $\varphi(x)$ is supported by $[-2^M, 2^M]$ and the functions $\varphi(x - k)$, $k \in \mathbb{Z}$, are an orthonormal basis of V_0. The latter space is the first rung of a ladder $V_j, j \in \mathbb{Z}$, which is named a multiresolution analysis of L^2. Let $\psi(x)$ be the "mother wavelet." The wavelet ψ is supported by $[-2^M, 2^M]$ and the sequence $\psi(x - k)$, $k \in \mathbb{Z}$, is an orthonormal basis of W_0 which orthogonally complements V_0 in V_1. Similarly W_j orthogonally complements V_j in V_{j+1}. Moreover the following property can be fixed in the construction:

$$\psi(x) > \frac{1}{4} \quad \text{if } x \in [-2^{-N}, 2^{-N}] \quad \text{where } N \text{ is large enough.}$$

We write $\Phi^{(0)}(x) = \varphi(x), \Phi^{(1)}(x) = \psi(x)$ and $\tilde{\Phi}(x) = \Phi^{(0)}(2x)$. A sequence $\{h_u\}_{u \in \mathbb{Z}}$ exists such that

$$h_u = 0 \quad \text{if} \quad |u| > 2^{M+1} \text{ and } \sum_u h_u = 0,$$

$$\Phi^{(1)}(x) = \sum_u h_u \Phi^{(0)}(2x - u).$$

Let us introduce

$$\tilde{h}_u = \sum_{v=-\infty}^{u} h_v.$$

Then we have

$$\psi(x) = \Phi^{(1)}(x) = \sum_{u \in \mathbb{Z}} \tilde{h}_u \left\{ \tilde{\Phi}\left(x - \frac{u}{2}\right) - \tilde{\Phi}\left(x - \frac{u+1}{2}\right) \right\}. \tag{7}$$

We now turn to the n-dimensional wavelets. For $j \in \mathbb{Z}$, $k \in \mathbb{Z}^n$, $\varepsilon \in \{0,1\}^n$, $x \in \mathbb{R}^n$, the scaling function and the wavelets are

$$\Phi^{(\varepsilon)}(x) = \prod_{i=1}^{n} \Phi^{(\varepsilon_i)}(x_i), \Phi_{j,k}^{(\varepsilon)} = 2^{\frac{nj}{2}} \prod_{i=1}^{n} \Phi^{(\varepsilon_i)}(2^j x_i - k_i).$$

More precisely $\Phi_{j,k}^{(\varepsilon)}(x)$, $\varepsilon \in \{0,1\}^n \setminus \{0\}$, $j \in \mathbb{Z}$, $k \in \mathbb{Z}^n$, are the n-dimensional wavelets. These wavelets provide us with an orthonormal basis of $L^2(\mathbb{R}^n)$. They are labeled by $\Lambda_n = \{ \lambda = (\varepsilon, j, k); \varepsilon \in \{0,1\}^n \setminus \{0\}, j \in \mathbb{Z}, k \in \mathbb{Z}^n \}$. The function $\Phi^{(0)}(x)$ which is excluded is the $n-$dimensional scaling function. We denote by i_ε the smallest subscript i such that $\varepsilon_i \neq 0$ and let $e^\varepsilon \in \{0,1\}^n$ be defined by $e_i^\varepsilon = 1$ if $i = i_\varepsilon$ and $e_i^\varepsilon = 0$ if $i \neq i_\varepsilon$. Finally one sets

$$\tilde{\Phi}^{(\varepsilon)}(x) = \Phi^{(0)}(2x_{i_\varepsilon}) \prod_{i=1}^{-1+i_\varepsilon} \Phi^{(\varepsilon_i)}(x_i) \prod_{i=1+i_\varepsilon}^{n} \Phi^{(\varepsilon_i)}(x_i) \quad \text{and} \quad \tilde{\Phi}_u^{(\varepsilon)}(x) = \tilde{\Phi}^{(\varepsilon)}\left(x - \frac{u}{2}e^\varepsilon\right).$$

Summarizing our introduction to the n-dimensional wavelets we have

Theorem 2.1. *Keeping the preceding notations we have*

(i) $\{\Phi_{j,k}^{(\varepsilon)}\}_{\lambda \in \Lambda_n}$ *is an orthonormal basis of* $L^2(\mathbb{R}^n)$.

(ii) $\Phi^{(\varepsilon)}(2^j x - k) = \sum_u \tilde{h}_u \{\tilde{\Phi}_u^{(\varepsilon)}(2^j x - k) - \tilde{\Phi}_u^{(\varepsilon)}(2^j x - k - \frac{1}{2}e^\varepsilon)\}$.

Each $\tilde{\Phi}_u^{(\varepsilon)}$ *is supported by* $[-2^{M+1}, 2^{M+1}]^n$.

The proof of (i) can be found in [11] and (7) implies (ii).

2.2. Wavelets and BMO

For each dyadic cube $Q_{s,p} = 2^{-s}p + 2^{-s}[0,1[^n,\ \forall s \in \mathbb{Z},\ p \in \mathbb{Z}^n$ and for $t > 0$, the enlarged cube is defined by

$$Q_{t,s,p} = 2^{-s}p + 2^{-s}[0,1[^n + 2^{-s}[-t,t]^n = 2^{-s}p + 2^{-s}[-t, 1+t[^n.$$

The space BMO is characterized by simple size conditions on wavelet coefficients [11]:

Theorem 2.2. *For every function* $f \in L^2_{loc}(\mathbb{R}^n)$ *the following three properties are equivalent:*

(i) $f(x) = \sum_{\lambda \in \Lambda_n} a_{j,k}^{(\varepsilon)} \Phi_{j,k}^{(\varepsilon)}(x) \in$ BMO.

(ii) *There exists a constant* C *such that for every dyadic cube* $Q_{s,p}$ *and every* $\varepsilon \in \{0,1\}^n \setminus \{0\}$ *one has*

$$\sum_{Q_{j,k} \subset Q_{s,p}} |a_{j,k}^{(\varepsilon)}|^2 \leq C|Q_{s,p}|.$$

(iii) *For every* $t \geq 0$, *there exists a constant* $C = C(t)$ *such that*

$$\sum_{Q_{j,k} \subset Q_{t,s,p}} |a_{j,k}^{(\varepsilon)}|^2 \leq C|Q_{s,p}|.$$

A proof can be found in [11].

The following lemma will be used in sections 5 and 6 to estimate norms in the Hardy space H^1. This Hardy space contains the homogeneous Besov space $\dot{B}_1^{0,1}$ and the norm in the former space does not exceed the norm in the latter. We have

Lemma 2.3. *If* $\phi(x)$ *is continuously differentiable and if the diameter of the support of* ϕ *does not exceed* C *there exists a constant* $C' > 0$ *such that if* $R \geq 1$ *and if* $\sum_{|k| \leq R} a_k = 0$, *we have*

$$\Big\| \sum_{|k| \leq R} a_k \phi(x+k) \Big\|_{\dot{B}_1^{0,1}} \leq C' \log R \sum_{|k| \leq R} |a_k|. \tag{8}$$

Every $f \in H^1$ or in $\dot{B}_1^{0,1}$ has a vanishing integral. Therefore if $\sum_{|k| \leq R} a_k \neq 0$ the function $\sum_{|k| \leq R} a_k \phi(x+k)$ cannot belong to $\dot{B}_1^{0,1}$. Simple examples show that the logarithmic factor in (8) is optimal. Lemma 2.3 says that, up to a logarithmic factor, the norm in H^1 of functions $f \in V_0$ with a vanishing integral and a (large) compact support does not exceed the L^1 norm. Let us provide the reader with a proof of this simple estimate. Since the integral of f vanishes we have $f = \partial_1 g_1 + \cdots + \partial_n g_n$ where the L^1 norms of g_p, $1 \leq p \leq n$, are bounded by $CR\|f\|_1$. This is well known when $R =$

1 and a simple rescaling gives the general case. For computing the Besov norm we consider a Littlewood-Paley analysis where $S_j(f) = f * \phi_j$ and $\phi_j(x) = 2^{nj}\phi(2^j x)$. The function ϕ is smooth, radial, compactly supported and the integral of ϕ is 1. We consider $\Delta_j = S_{j+1} - S_j$ and the homogeneous Besov space $\dot{B}_1^{0,1}$ is then defined by $\|\Delta_j(f)\|_1 \in l^1(\mathbb{Z})$. This l^1 norm is the Besov norm of f. The properties of f which will be used as j tends to $+\infty$ are the trivial estimates of the L^1 norms of f and of its gradient. When $j \geq 0$ the smoothness of f yields $\|\Delta_j(f)\|_1 \leq C2^{-j}\|\nabla f\|_1$. These estimates sum up to $C\|\nabla f\|_1 = O(\|f\|_1)$. Then we define $j_0 = \log R$ and we estimate $\|\Delta_j(f)\|_1$ by $C\|f\|_1$ when $-j_0 \leq j < 0$. This is the trivial estimate. These trivial estimates sum up to $C \log R\|f\|_1$. When $j < -j_0$ we use the fact that f is a sum of derivatives of functions in L^1 and we obtain $\|\Delta_j(f)\|_1 \leq 2^j R\|f\|_1$. These estimates sum up to $C2^{-j_0}R\|f\|_1 = O(\|f\|_1)$ and Lemma 2.3 is proved.

2.3. The Beylkin-Coifman-Rokhlin algorithm

The non-standard representation of operators and the Beylkin-Coifman-Rokhlin algorithm will be defined right now. It is adapted to the analysis of Calderón-Zygmund operators and of some pseudo-differential operators. This analysis was implicit in [9] and advocated by G. Beylkin, R. Coifman, and V. Rokhlin. See [4, 5, 11, 13]. Let us begin with an abstract setting where we are given a Hilbert space H and an increasing sequence V_j, $j \in \mathbb{Z}$, of approximation spaces. Each V_j is closed, the union of the approximation spaces $\bigcup V_j$ is dense in H and their intersection $\bigcap V_j = 0$. Let $P_j : H \mapsto V_j$ be the orthogonal projector and $Q_j = P_{j+1} - P_j$. If we are given a bounded operator $T : H \mapsto H$, we have

$$T = \lim_{j \to +\infty} P_j T P_j = \sum_j (P_j T Q_j + Q_j T P_j + Q_j T Q_j) = \sum_j T_j.$$

It often occurs that each of the three series $\sum_j P_j T Q_j$, $\sum_j Q_j T P_j$, and $\sum_j Q_j T Q_j$ converges. In the proof of the David-Journé theorem Cotlar almost orthogonality lemma is applied to this series $\sum_j T_j$.

Definition 2.4. The non-standard representation of an operator T is defined by the expansion $T = \sum_{j \in \mathbb{Z}}(P_j T Q_j + Q_j T P_j + Q_j T Q_j)$.

From now on we are focusing on a specific example where the V_j are a given by a multiresolution analysis. Then the non-standard representation of T paves the road to the BCR algorithm (Beylkin-Coifman-Rokhlin algorithm). The aim of the BCR algorithm is to decouple the action of T on the dyadic frequency bands provided by the orthogonal complements W_j of V_j in V_{j+1}. This decoupling can be boosted by using an orthogonal wavelet basis. Indeed the building blocks provided by the BCR algorithm can be further written as local operators. We then obtain $T = \sum_\varepsilon \sum_j T_j^\varepsilon$ where the pieces T_j^ε, $\varepsilon = (\varepsilon_1, \varepsilon_2)$, are defined by explicit matrices M_j^ε, $\varepsilon \in \{0,1\}^{2n} \setminus \{0\}$, $j \in \mathbb{Z}$. Here and in what follows 0 denotes either the element $(0,\ldots,0)$ of $\{0,1\}^{2n}$ or of $\{0,1\}^n$.

Let us begin with defining two collections of labels or indices:

$$\Lambda = \Lambda_{2n} = \{\, \lambda = (\varepsilon_1, \varepsilon_2, j, k, l);\ (\varepsilon_1, \varepsilon_2) \in \{0,1\}^{2n} \setminus \{0\},\ \ j \in \mathbb{Z},\ \ k \in \mathbb{Z}^n,\ \ l \in \mathbb{Z}^n \,\}$$

and similarly

$$\Lambda_n = \{\, \lambda = (\varepsilon_1, j, k);\ \varepsilon_1 \in \{0,1\}^n \setminus \{0\},\ \ j \in \mathbb{Z},\ \ k \in \mathbb{Z}^n \,\}.$$

Let T be an operator satisfying the version of (3) where functions in the Schwartz class are now replaced by compactly supported functions with q_M continuous derivatives. We introduce the entries

$$a(\lambda) = a_{j,k,l}^{(\varepsilon_1, \varepsilon_2)} = \langle \Phi_{j,k}^{(\varepsilon_1)}, T\Phi_{j,l}^{(\varepsilon_2)} \rangle, \quad \lambda = (\varepsilon_1, \varepsilon_2, j, k, l) \in \Lambda.$$

These entries $\{a_{j,k,l}^{(\varepsilon_1, \varepsilon_2)}\}_{\lambda \in \Lambda}$ provide us with a full information on T. Indeed the distributional kernel $K(x,y)$ of T is given by the following series expansion:

$$K(x,y) = \sum_{\lambda \in \Lambda} a_{j,k,l}^{(\varepsilon_1, \varepsilon_2)} \Phi_{j,k}^{(\varepsilon_1)}(x) \Phi_{j,l}^{(\varepsilon_2)}(y). \tag{9}$$

One should notice that the right-hand side of (9) is simply the wavelet expansion of $K(x,y)$ viewed as a function of $2n$ variables. This distributional kernel $K(x,y)$ is now written as a sum $K = \sum_{-\infty}^{\infty} K_j$ where $K_j(x,y) = \sum_{\varepsilon_1, \varepsilon_2, k, l} a_{j,k,l}^{(\varepsilon_1, \varepsilon_2)} \Phi_{j,k}^{(\varepsilon_1)}(x) \Phi_{j,l}^{(\varepsilon_2)}(y)$. Therefore the support of $K_j(x,y)$ is contained in the union over k and l of the balls defined by $|x - k2^{-j}| \le 2^{M-j}$, $|y - l2^{-j}| \le 2^{M-j}$. This support is too large and will be partitioned in the next section. If $\tilde{K}_j(x,y) = 2^{-nj} K_j(2^{-j}x, 2^{-j}y)$ then the estimates on $\tilde{K}_j$ will be uniform in j. Meyer says that (9) is the "non-standard analysis" of T and that the coefficients $\{a(\lambda)\}_{\lambda \in \Lambda} = \{a_{j,k,l}^{(\varepsilon_1, \varepsilon_2)}\}_{\lambda \in \Lambda}$ are the coefficients of the non-standard representation of T.

The following result will be used to compute $T(f)$ when T is given by its non-standard coefficients:

Lemma 2.5. *With the above notations, if the wavelet expansion of f is given by $f(x) = \sum_{(\epsilon, j, k) \in \Lambda_n} \alpha^\epsilon(j, k) \Phi_{j,k}^\epsilon(x)$ then one has $T(f) = \sum_{j,k} \sum_{\epsilon \in \{0,1\}^n} b^\varepsilon(j, k) \Phi_{j,k}^\varepsilon$. This is not a wavelet expansion. The coefficients $b^\varepsilon(j, k)$ are given by $b^\varepsilon(j, k) = \beta^\varepsilon(j, k)$ or $\gamma^\varepsilon(j, k)$ where*

$$\beta^\varepsilon(j, k) = \sum_{l, \eta \neq 0} a_{j,k,l}^{(\varepsilon, \eta)} \alpha^\eta(j, l)$$

and

$$\gamma^\varepsilon(j, k) = \sum_{\{j' < j,\, l, l', \eta \neq 0\}} a_{j,k,l}^{(\varepsilon, 0)} \omega^\eta(j, l, j', l') \alpha^\eta(j', l')$$

with $\omega^\eta(j, l, j', l') = \langle \Phi_{j,l}^0, \Phi_{j',l'}^\eta \rangle$.

The proof of this lemma is routine. The expansion of the kernel is used. Indeed

$$T(f)(x) = \int K(x,y)f(y)\,dy = \sum_{\lambda \in \Lambda} a_{j,k,l}^{(\varepsilon_1,\varepsilon_2)} \Phi_{j,k}^{(\varepsilon_1)}(x) \int \Phi_{j,l}^{(\varepsilon_2)}(y)f(y)\,dy.$$

It remains to replace $f(y)$ by its wavelet expansion. Then two cases are occurring. The first one is defined by $\varepsilon_2 \neq 0$. In that case $\alpha(j,l) = \int \Phi_{j,l}^{(\varepsilon_2)}(y)f(y)\,dy$ is a wavelet coefficient of f. The corresponding sum is $\sum_{\lambda \in \Lambda} a_{j,k,l}^{(\varepsilon_1,\varepsilon_2)} \alpha^{\varepsilon_2}(j,l)\Phi_{j,l}^{(\varepsilon_2)}(x)$. The coefficients $\gamma^\varepsilon(j,k)$ appear when $\varepsilon_2 = 0$. Then a scaling function $\Phi_{j,l}^0$ is integrated against all the wavelets which are present in the wavelet expansion of f. We know that any wavelet $\Phi_{j',l'}^{(\varepsilon)}$ is orthogonal to any scaling function $\Phi_{j,l}^0$ such that $j \leq j'$. This explains why we have $j' < j$ in the second series of terms occurring in Lemma 2.5.

Lemma 2.5 says that the non-standard decomposition of an operator yields two types of components. In the first type the dyadic frequency channels are fully decoupled and the coefficients $\beta(j,k)$ are given by a series where j is the same everywhere. Indeed we are staying inside W_j. In the second type couplings between distinct frequency channels are occurring and the summation runs over $j' < j$.

Lemma 2.5 provides an expansion of $T(f)$ where two types of terms appear. The first type which is defined by $\varepsilon \neq 0$ yields a wavelet expansion. The second component is defined by $\varepsilon = 0$. Then the coefficients $\gamma^\varepsilon(j,k)$ do not appear. This component $R(x) = \sum_{j,k,l} \sum_{\eta \neq 0} a_{j,k,l}^{(0,\eta)} \alpha^\eta(j,l)\Phi_{j,k}^0(x)$ is more involved since scaling functions do not have a vanishing integral. If for every j,l we have $\sum_k a_{j,k,l}^{(0,\eta)} = 0$ then $\int R(x)\,dx = 0$ and this remark will play a key role in this paper.

The following theorem will not be used as it stands in this paper since it relies on a stronger hypothesis. It paves the road between non-standard analysis and the Calderón-Zygmund theory.

Theorem 2.6.   (i) *If* (3) *holds, then there exists a unique sequence* $\{a_{j,k,l}^{(\varepsilon_1,\varepsilon_2)}\}_{\lambda \in \Lambda}$
   such that the distributional kernel $K(x,y)$ *of* T *is given by the series* (9) *which converges in the distributional sense.*

(ii) *If* $T \in \mathrm{OpE}_\gamma$ *then there exists a constant* C *such that* $(1+|k-l|)^{n+\gamma}|a_{j,k,l}^{(\varepsilon_1,\varepsilon_2)}| \leq C$.
   The cancellation provided by (5) *implies* $\sum_m a_{j,m,k}^{(0,\varepsilon)} = 0$, $\sum_m a_{j,k,m}^{(\varepsilon,0)} = 0$, *for every* $\varepsilon \neq 0$, *every* j, *and every* k.

The proof can be found in [11]. One should observe that the estimates on the entries $a_{j,k,l}^{(\varepsilon_1,\varepsilon_2)}$ are uniform in j. In other words these estimates are scale-invariant. This scale invariance will be our guide line in the following section.

3. Pseudo-annular decompositions of operators

3.1. Pseudo-annular decompositions

Our goal is to prove L^2 or BMO estimates for singular integral operators with rough kernels. Using the David-Journé strategy it suffices to focus on the case where $T(1) = T^*(1) = 0$. When the distributional kernel $K(x,y)$ is not smooth enough Cotlar almost orthogonality lemma does not apply to the series $T = \sum_j T_j$ given by the BCR algorithm. Then an improved analysis is needed and new building blocks which will be used. Here is a sketchy presentation of these new building blocks. Keeping the notations of the preceding section we have $T = \sum_j T_j$. Let K_j be the distributional kernel of T_j and let $\tilde{K}_j(x,y) = 2^{-nj} K_j(2^{-j}x, 2^{-j}y)$. As was said before, the estimates on $\tilde{K}_j(x,y)$ are uniform in j. These estimates show that the kernels $\tilde{K}_j(x,y)$ have a poor localization in the space domain. That is why each $\tilde{K}_j(x,y)$ will be further split into a series of terms $\tilde{K}_j(x,y)\chi_R(x,y)$ where χ_R is the indicator function of $2^{R-1} \leq |x-y| \leq 2^R$, $R \in \mathbb{N}$. By an obvious rescaling $K_j(x,y)$ is then written as an expansion $\sum_{R \geq 1} K_j(x,y)\chi_R(2^j x, 2^j y)$. This brutal truncation destroys the cancellations given by $T(1) = T^*(1) = 0$ which reads $\int K_j(x,y)\,dy = \int K_j(x,y)\,dx = 0$. Some corrections are needed in our construction. We now end this line of thought and define the "band operators" by another approach where the construction is carried on the coefficients of the wavelet expansions of the kernels. We begin with the definition of some partial sums T_j^R inside the expansion of T_j. Then the "band operators" T_R are simply $T_R = \sum_j T_j^R$. One should instead write $T_R^{(\varepsilon_1, \varepsilon_2)}$ instead of T_R but this is only technical. The coefficients $a_{j,k,l}^{(\varepsilon_1, \varepsilon_2, R)}$ arising in the expansion of $T_j^R = T_j^{R,(\varepsilon_1,\varepsilon_2)}$ are given by three rules. Before stating these rules let us explain their *raison d'être*. We aim at defining $a_{j,k,l}^{(\varepsilon_1,\varepsilon_2,R)}$ by a plain truncation where $a_{j,k,l}^{(\varepsilon_1,\varepsilon_2,R)} = a_{j,k,l}^{(\varepsilon_1,\varepsilon_2)}$ if $2^{R-1} \leq |k-l| < 2^R$ and 0 if not. But this recipe destroys the cancellations given by $T(1) = T^*(1) = 0$. That is why an extra diagonal term is needed to restore the broken cancellations. Let us begin with the example of the Hilbert transform H. Then the non-standard entries can be ranged in three sequences. The first sequence is $a(j,k,l) = \langle H(\phi_{j,k}), \psi_{j,l} \rangle = \beta(k-l)$. The second sequence is $b(j,k,l) = \langle H(\psi_{j,k}), \phi_{j,l} \rangle = \gamma(k-l)$ and the third one is $c(j,k,l) = \langle H(\psi_{j,k}), \psi_{j,l} \rangle = \omega(k-l)$. The three sequences β, γ, and ω are rapidly decreasing at infinity. Then the Hilbert transform H can be written as a sum of three series $H_R^{(i)}$, $i = 1, 2$ or 3, and for the sake of simplicity we will focus on the first series. The distributional kernel of $H_R^{(1)}$ is $\sum_{j,k,l} \gamma_R(k-l)\psi_{j,k}(x)\phi_{j,l}(y)$ where $\gamma_R(k-l)$ is the product between $\gamma(k-l)$ and the indicator function of $2^{R-1} \leq |k-l| < 2^R$. This brutal truncation produces unbounded operators T_R. Indeed an obvious computation yields $T_R(1) = c_R \sum \sin(2^j x)$ where $c_R = \sum_l \gamma_R(l) \neq 0$. This function does not belong to BMO. Therefore H_R is not bounded on L^2. It is ridiculous to write a bounded operator as a series of unbounded pieces. This crucial remark paves the road to the

construction which follows.

Definition 3.1. The distributional kernel of $T_R^{(\varepsilon_1,\varepsilon_2)}$ is

$$K_R^{(\varepsilon_1,\varepsilon_2)}(x,y) = \sum_{j,k,l} a_{j,k,l}^{(\varepsilon_1,\varepsilon_2,R)}\Phi_{j,k}^{(\varepsilon_1)}(x)\Phi_{j,l}^{(\varepsilon_2)}(y)$$

and the coefficients $a_{j,k,l}^{(\varepsilon_1,\varepsilon_2,R)}$ are defined by the following rules where we first assume $R \geq 1$. (The case $R = 0$ will be considered afterwards.)

The non-diagonal entries are defined by $a_{j,k,l}^{(\varepsilon_1,\varepsilon_2,R)} = 0$ for $1 \leq |k-l| < 2^{R-1}$ or $|k-l| \geq 2^R$ and $a_{j,k,l}^{(\varepsilon_1,\varepsilon_2,R)} = a_{j,k,l}^{(\varepsilon_1,\varepsilon_2)}$ for $2^{R-1} \leq |k-l| < 2^R$. Then the diagonal entries are defined as follows:

(i) If $\varepsilon_1 \neq 0$ and $\varepsilon_2 \neq 0$, then $a_{j,k,k}^{(\varepsilon_1,\varepsilon_2,R)} = 0$.

(ii) If $\varepsilon_1 = 0$, then $a_{j,k,k}^{(\varepsilon_1,\varepsilon_2,R)} = -\sum_{m \neq k} a_{j,m,k}^{(\varepsilon_1,\varepsilon_2,R)}$.

(iii) If $\varepsilon_2 = 0$, then $a_{j,k,k}^{(\varepsilon_1,\varepsilon_2,R)} = -\sum_{m \neq k} a_{j,k,m}^{(\varepsilon_1,\varepsilon_2,R)}$.

One should observe that no diagonal corrections are needed in (i). Indeed the corresponding piece of the kernel is $\sum_{j,k,l} a_{j,k,l}^{(\varepsilon_1,\varepsilon_2,R)}\Phi_{j,k}^{(\varepsilon_1)}(x)\Phi_{j,l}^{(\varepsilon_2)}(y)$ and the cancellations are provided by the wavelets $\Phi_{j,k}^{(\varepsilon_1)}$ and $\Phi_{j,k}^{(\varepsilon_2)}$. Let us insist once more on the fact that the diagonal terms in (ii) and (iii) are needed to ensure:

$$\sum_{k} a_{j,k,l}^{(0,\varepsilon_2,R)} = \sum_{l} a_{j,k,l}^{(\varepsilon_1,0,R)} = 0 \tag{10}$$

and

$$T_R^{(\varepsilon_1,\varepsilon_2)}(1) = (T_R^{(\varepsilon_1,\varepsilon_2)})^*(1) = 0.$$

A third observation is playing an important role. We have

$$\sum_{R \geq 1} a_{j,k,k}^{(0,\varepsilon_2,R)} = -\sum_{m \neq k} a_{j,m,k}^{(0,\varepsilon_2,R)} = a_{j,k,k}^{(0,\varepsilon_2,R)}.$$

It implies that the diagonal terms which have been excluded in the truncation show up in the sum over $R \geq 1$.

These rules are completed by a trivial remark concerning the case $R = 0$. If $R = 0$, if $\varepsilon_1 \neq 0$ and $\varepsilon_2 \neq 0$ then we set:

(i) If $k \neq l$, $a_{j,k,l}^{(\varepsilon_1,\varepsilon_2,0)} = 0$. If $k = l$, $a_{j,k,k}^{(\varepsilon_1,\varepsilon_2,0)} = a_{j,k,k}^{(\varepsilon_1,\varepsilon_2)}$.

(ii) If $\varepsilon_1 = 0$ or $\varepsilon_2 = 0$, then $a_{j,k,l}^{(\varepsilon_1,\varepsilon_2,0)} = 0$. We then obtain

$$T = \sum_{\{R \geq 0,\ (\varepsilon_1,\varepsilon_2) \in \{0,1\}^{2n}\setminus\{0\}\}} T_R^{(\varepsilon_1,\varepsilon_2)}.$$

For any fixed R the continuity on L^2 of each T_R is trivial since T_R is a Calderón-Zygmund operators and the standard $T(1)$ theorem can be used. The weak boundedness property (3) suffices to ensure the kernel estimates if do not care about the dependence in R.

3.2. Modulus of continuity

We now introduce a new estimate $A(R)$ which will be defined by (12) and will be related to the modulus of continuity $\varepsilon(R)$ defined by Definition 1.1. The operator norms of T_R will be estimated by $A(R)$. Finally the spaces $\mathrm{Op\tilde{B}}_\mu$ and OpB_μ will be compared to the space OpM_μ which is studied by Meyer in [10]. We first introduce new norms on sequence spaces. For $1 \leq p \leq \infty, R \geq 0, (\varepsilon_1, \varepsilon_2) \in \{0,1\}^{2n} \setminus \{0\}$, one defines $\|A_R^{(\varepsilon_1,\varepsilon_2)}\|_p$ by the following:

When $R = 1, 2, \ldots$, one sets

$$A_{j,k}^{(\varepsilon_1,\varepsilon_2,R)} = \sum_{m:2^{R-1} \leq |k-m| < 2^R} \left\{ |a_{j,k,m}^{(\varepsilon_1,\varepsilon_2)}| + |a_{j,m,k}^{(\varepsilon_1,\varepsilon_2)}| \right\},$$

$$A_j^{(\varepsilon_1,\varepsilon_2,R)} = \sup_k A_{j,k}^{(\varepsilon_1,\varepsilon_2,R)},$$

$$A(R) = \sup_{(\varepsilon_1,\varepsilon_2,j)} A_j^{(\varepsilon_1,\varepsilon_2,R)}, \tag{11}$$

$$\|A_R^{(\varepsilon_1,\varepsilon_2)}\|_p = \sup_j \left\{ \sum_{s=0}^{R-1} (A_{j+s}^{(\varepsilon_1,\varepsilon_2,R)})^p \right\}^{\frac{1}{p}}. \tag{12}$$

The norm $\|A_R^{(\varepsilon_1,\varepsilon_2)}\|_p$ is decreasing as a function of the exponent p and we have $A(R) = \sup_{\varepsilon_1,\varepsilon_2} \|A_R^{(\varepsilon_1,\varepsilon_2)}\|_\infty$.

Using these norms on sequences, we define some new operator norms

$$\|T\|_{\tilde{B}_\mu} = \overset{(1)}{\sum_R} (1+R)^{\mu-\frac{1}{2}} \|A_R^{(\varepsilon_1,\varepsilon_2)}\|_2 + \overset{(2)}{\sum_R} \|A_R^{(\varepsilon_1,\varepsilon_2)}\|_\infty, \tag{13}$$

$$\|T\|_{B_\mu} = \overset{(3)}{\sum_R} (1+R)^{\mu-\frac{1}{2}} \|A_R^{(\varepsilon_1,0)}\|_2$$

$$+ \overset{(4)}{\sum_R} (1+R)^{\mu-\frac{3}{2}} \|A_R^{(0,\varepsilon_2)}\|_1 + \overset{(5)}{\sum_R} \|A_R^{(\varepsilon_1,\varepsilon_2)}\|_2, \tag{14}$$

where the first sum runs over $R = 0, 1, 2, \ldots$ and $(\varepsilon_1, \varepsilon_2) \in \{0,1\}^{2n} \setminus \{0\}, |\varepsilon_1||\varepsilon_2| = 0$, the second one runs over $R = 0, 1, 2, \ldots$ and $(\varepsilon_1, \varepsilon_2) \in \{0,1\}^{2n}, \varepsilon_1 \neq 0, \varepsilon_2 \neq 0$, the third one runs over $R = 1, 2, \ldots$ and $\varepsilon_1 \in \{0,1\}^n \setminus \{0\}$, the fourth one runs over $R = 1, 2, \ldots$ and $\varepsilon_2 \in \{0,1\}^n \setminus \{0\}$, and the last one runs over $R = 0, 1, 2, \ldots$, $(\varepsilon_1, \varepsilon_2) \in \{0,1\}^{2n}, \varepsilon_1 \neq 0, \varepsilon_2 \neq 0$.

Finally some new operator spaces are defined by

Definition 3.2. One writes $T \in \mathrm{Op}\tilde{\mathrm{B}}_\mu$ if $\|T\|_{\tilde{B}_\mu}$ is finite when it is defined by (13). One writes $T \in \mathrm{OpB}_\mu$ if $\|T\|_{B_\mu}$ is finite when it is defined by (14) and if $A_0^{(\varepsilon_1,\varepsilon_2)} = 0$ when $\varepsilon_1 = 0$ or $\varepsilon_2 = 0$.

We then have

Theorem 3.3. *Let $A(R)$ be defined by* (11) *and $\varepsilon(R)$ be defined by Definition 1.1. Then we have:*

(i) $A(R) \leq C[\varepsilon(R-1) + \varepsilon(R) + \varepsilon(R+1)]$, $R \geq M+5$.

(ii) *Moreover for $\mu \geq 1/2$ one has* $\mathrm{OpM}_\mu \subset \mathrm{Op}\tilde{\mathrm{B}}_\mu$ *and* $\mathrm{OpM}_\mu \subset \mathrm{OpB}_\mu$.

(iii) *Every operator $T \in \mathrm{OpB}_{3/2}$ is bounded on* BMO.

The values $R < M+5$ can be ignored since we already know that the corresponding pieces T_R are bounded on L^2.

The third assertion will be proved in section 5. This is a slight improvement on Theorem 1.6.

The proof of (i) in Theorem 3.3 begins with a few simple remarks.

(i) If $\varepsilon_1 \neq 0$, $\Phi^{(\varepsilon_1)}$ is a wavelet, then (ii) in Theorem 2.1 implies

$$
\begin{aligned}
I &= \sum_{k:2^{R-1}\leq|k-l|<2^R} |a_{j,k,l}^{(\varepsilon_1,\varepsilon_2)}| \\
&= \sum_{k:2^{R-1}\leq|k-l|<2^R} 2^{jn} \left| \iint \Phi^{(\varepsilon_1)}(2^jx-k) K(x,y) \Phi^{(\varepsilon_2)}(2^jy-l)\, dx\, dy \right| \\
&\leq \sum_{k:2^{R-1}\leq|k-l|<2^R} 2^{jn} \sum_u |\tilde{h}_u| \left| \iint \left\{ \tilde{\Phi}_u^{(\varepsilon_1)}(2^jx-k) - \tilde{\Phi}_u^{(\varepsilon_1)}\left(2^jx-k-\frac{1}{2}e^{\varepsilon_1}\right)\right\} K(x,y) \right. \\
&\qquad \left. \times \Phi^{(\varepsilon_2)}(2^jy-l)\, dx\, dy \right| \\
&= \sum_{k:2^{R-1}\leq|k-l|<2^R} 2^{jn} \sum_u |\tilde{h}_u| \left| \iint \tilde{\Phi}_u^{(\varepsilon_1)}(2^jx-k)\{K(x,y) - K(x+2^{-j-1}e^{\varepsilon_1},y)\} \right. \\
&\qquad \left. \times \Phi^{(\varepsilon_2)}(2^jy-l)\, dx\, dy \right|.
\end{aligned}
$$

Since $R \geq M+5$ and $|k-l| \geq 2^{R-1}$, the supports of $\tilde{\Phi}^{(\varepsilon_1)}(2^jx-k)$ and $\Phi^{(\varepsilon_2)}(2^jx-l)$ are disjoint. Therefore the double integral is performed away form the diagonal and

the smoothness of the kernel is playing a role. We have

$$I \leq C 2^{jn} \sum_u |\tilde{h}_u| \iint_{\substack{|y-2^{-j}l|<2^{-j+M} \\ 2^{-j}(2^{R-1}-2^{M+2})<|x-y|<2^{-j}(2^R+2^{M+2})}} |K(x,y) - K(x + 2^{-j-1}e^{\varepsilon_1}, y)| \, dx \, dy.$$

Using once more $R \geq M + 5$, we have

$$I \leq C \sup_{y \in \mathbb{R}^n} \int_{2^{R-2-j}<|x-y|<2^{R+1-j}} |K(x,y) - K(x + 2^{-j-1}e^{\varepsilon_1}, y)| \, dx.$$

For every frozen y, this integral will be split into three parts where the integration is performed on $2^{R-2-j} < |x - y| \leq 2^{R-1-j}$, $2^{R-1-j} < |x - y| \leq 2^{R-j}$, and $2^{R-j} < |x - y| < 2^{R+1-j}$. We then obtain $I \leq C(\varepsilon(R-1) + \varepsilon(R) + \varepsilon(R+1))$. In fact we proved a better estimate where $\epsilon(R)$ is replaced by $\eta(R)$. If $\varepsilon_1 = 0$, then $\varepsilon_2 \neq 0$ and the same arguments are used to obtain $A(R) \leq C[\varepsilon(R-1) + \varepsilon(R) + \varepsilon(R+1)]$, $R \geq M + 5$.

(ii) Let us check the inclusion $\mathrm{OpM}_\mu \subset \mathrm{OpB}_\mu$, $\mu \geq 0$.

If $R \geq 1$, we then have $\|A_R^{(\varepsilon_1,\varepsilon_2)}\|_2 \leq CR^{\frac{1}{2}} A(R)$, $\|A_R^{(\varepsilon_1,\varepsilon_2)}\|_1 \leq CRA(R)$. It implies the following:

$$\sum_{\substack{R \geq M+5 \\ \varepsilon_1 \in \{0,1\}^n \backslash \{0\}}} (1+R)^{\mu-\frac{1}{2}} \|A_R^{(\varepsilon_1,0)}\|_2 + \sum_{\substack{R \geq M+5 \\ \varepsilon_2 \in \{0,1\}^n \backslash \{0\}}} (1+R)^{\mu-\frac{3}{2}} \|A_R^{(0,\varepsilon_2)}\|_1$$

$$+ \sum_{\substack{R \geq M+5 \\ (\varepsilon_1,\varepsilon_2) \in \{0,1\}^{2n} \\ |\varepsilon_1||\varepsilon_2| \neq 0}} \|A_R^{(\varepsilon_1,\varepsilon_2)}\|_2 \leq C \sum_{R \geq M+5} R^\mu A(R) \leq C \sum_{R \geq M+4} R^\mu \varepsilon(R). \quad \text{(a)}$$

We then use (3). If $R = 0, |\varepsilon_1||\varepsilon_2| \neq 0$, then $\|A_0^{(\varepsilon_1,\varepsilon_2)}\|_p \leq C$. Next if $1 \leq R \leq M+4$, then we have $A_R^{(\varepsilon_1,\varepsilon_2)} \leq C$, a property which will be named (b).

Properties (a), (b) together with $T \in \mathrm{OpM}_\mu$ imply $T \in \mathrm{OpB}_\mu$.

(iii) Similar arguments yield $T \in \mathrm{Op\tilde{B}}_\mu$.

4. L^2 estimates for $T_R^{(\varepsilon_1,\varepsilon_2)}$ and $\|A_R^{(\varepsilon_1,\varepsilon_2)}\|_p$

As announced in the title, our goal is to estimate the operator norm of $T_R^{(\varepsilon_1,\varepsilon_2)}$ acting on L^2.

Theorem 4.1. *For every $R \in \mathbb{N}$, we have:*

(i) *If neither ε_1 nor ε_2 is 0, then $\|T_R^{(\varepsilon_1, \varepsilon_2)}\|_{L^2 \to L^2} \le C\|A_R^{(\varepsilon_1, \varepsilon_2)}\|_\infty$.*

(ii) *If either ε_1 or ε_2 is 0, then $\|T_R^{(\varepsilon_1, \varepsilon_2)}\|_{L^2 \to L^2} \le C\|A_R^{(\varepsilon_1, \varepsilon_2)}\|_2$.*

(iii) *Every $T \in \mathrm{OpM}_{1/2}$ is bounded on L^2.*

One observes that the first estimate is obvious since in (i) the operator is fully decoupled in frequency channels. In other words, $T_R^{(\varepsilon_1, \varepsilon_2)}$ maps W_j into itself. Moreover (iii) immediately follows from (i) and (ii). Indeed we have $\|A_R^{(\varepsilon_1, \varepsilon_2)}\|_\infty \le A(R)$ and $\|A_R^{(\varepsilon_1, \varepsilon_2)}\|_2 \le R^{1/2}A(R)$. Therefore we will focus on (ii). The author was using Haar wavelets to prove similar estimates in [4]. One should observe that the definition of $A(R)$ depends on the wavelet basis. Therefore the proof given in the Haar case does not yield the result for general wavelets. Here smooth wavelets are being used to define $\|A_R^{(\varepsilon_1, \varepsilon_2)}\|_p$. These smooth wavelets are needed in the characterization of BMO. That is why the proof which is given here differs from the one in [4].

Studying the adjoint operator when needed, one can limit our attention to $\varepsilon_1 = 0$, $\varepsilon_2 \ne 0$. We will prove the following estimate: $\|T_R^{(0, \varepsilon_2)}f(x)\|_2 \le C\|A_R^{(0, \varepsilon)}\|_2\|f(x)\|_2$.

Let $f(x) = \sum_{\lambda \in \Lambda_n} \alpha_{j,k}^{(\varepsilon)}\Phi_{j,k}^{(\varepsilon)}(x) \in L^2$. Let us consider the partial sum

$$f_j^{(\varepsilon)}(x) = \sum_{k \in \mathbb{Z}^n} \alpha_{j,k}^{(\varepsilon)}\Phi_{j,k}^{(\varepsilon)}(x).$$

We then have $T_R^{(0, \varepsilon_2)}f_j^\varepsilon(x) = 0$ if $\varepsilon_2 \ne \varepsilon$. We now assume $\varepsilon = \varepsilon_2$ and write $T_R^{(0, \varepsilon)}f_j^\varepsilon(x) = g_j^\varepsilon(x)$. Then $T_R^{(0, \varepsilon)}f(x) = \sum_{j, \varepsilon} g_j^\varepsilon(x)$ and an almost-orthogonality lemma will be used to estimate the L^2 norm of the sum of this series.

Since $\varepsilon \ne 0$, we are in the first case of Lemma 2.5 and the non-standard representation yields $g_j^\varepsilon(x) = 2^{nj/2}\tilde{g}_j^{\varepsilon}(2^j x)$ where

$$\tilde{g}_j^\varepsilon(x) = \sum_{k,l \in \mathbb{Z}^n} a_{j,l,k}^{(0,\varepsilon,R)}\alpha_{jk}^{(\varepsilon)}\Phi^{(0)}(x - l). \tag{16}$$

It is easily seen that

$$\|g_j^\varepsilon(x)\|_2 \le CA_j^{(0,\varepsilon,R)}\|f_j^\varepsilon\|_2. \tag{17}$$

Indeed, for each frozen j, the expansion given by the right-hand side of (16) is orthogonal.

Next the inner products $\gamma = \langle g_j^\varepsilon, g_{j'}^\varepsilon \rangle$ need to be estimated. Without losing generality one can assume $j \ge j'$. We have

$$\gamma = \langle g_j^\varepsilon, g_{j'}^\varepsilon \rangle = 2^{\frac{j'-j}{2}n}\langle \tilde{g}_j^\varepsilon(x), \tilde{g}_{j'}^\varepsilon(2^{j'-j}x) \rangle.$$

It implies

$$\gamma = 2^{\frac{j'-j}{2}n}\left\langle \sum_{k,l\in\mathbb{Z}^n} a_{j,l,k}^{(0,\varepsilon,R)} a_{j,k}^{(\varepsilon)}\Phi^{(0)}(x-l),\ \tilde{g}_{j'}^{\varepsilon}(2^{j'-j}x)\right\rangle.$$

The property $\sum_l a_{j,l,k}^{0,\varepsilon,R} = 0$ will now play a key role. Using this cancellation, one writes

$$\tilde{g}_j^{\varepsilon}(x) = \sum_{2^{R-1}\le|m|<2^R}\sum_k a_{j,k+m,k}^{(0,\varepsilon,R)}\alpha_{j,k}^{(\varepsilon)}(\Phi^{(0)}(x-k-m)-\Phi^{(0)}(x-k)),$$

which implies

$$2^{\frac{j-j'}{2}n}\gamma = \left\langle \sum_{2^{R-1}\le|m|<2^R}\sum_k a_{j,k+m,k}^{(0,\varepsilon,R)}\alpha_{j,k}^{(\varepsilon)}(\Phi^{(0)}(x-k-m)-\Phi^{(0)}(x-k)),\right.$$

$$\left. \tilde{g}_{j'}^{\varepsilon}(2^{j'-j}x)\right\rangle.$$

An obvious change of variables gives

$$2^{\frac{j-j'}{2}n}\gamma = \sum_{2^{R-1}\le|m|<2^R}\left\langle \sum_k a_{j,k+m,k}^{(0,\varepsilon,R)}\alpha_{j,k}^{(\varepsilon)}\Phi^{(0)}(x-k),\right.$$

$$\left. \tilde{g}_{j'}^{\varepsilon}(2^{j'-j}x+2^{j'-j}m)-\tilde{g}_{j'}^{\varepsilon}(2^{j'-j}x)\right\rangle.$$

We then obtain

$$2^{\frac{j-j'}{2}n}\gamma = \sum_{2^{R-1}\le|m|<2^R}\left\langle \sum_k a_{j,k+m,k}^{(0,\varepsilon,R)}\alpha_{j,k}^{(\varepsilon)}\Phi^{(0)}(x-k),\right.$$

$$\sum_{k'}\sum_{l'} a_{j',l',k'}^{(0,\varepsilon,R)}\alpha_{j',k'}^{(\varepsilon)}\Phi^{(0)}(2^{j'-j}x+2^{j'-j}m-l')$$

$$\left. -\sum_{k'}\sum_{l'} a_{j',l',k'}^{(0,\varepsilon,R)}\alpha_{j',k'}^{(\varepsilon)}\Phi^{(0)}(2^{j'-j}x-l')\right\rangle.$$

Finally

$$2^{\frac{j-j'}{2}n}\gamma = \sum_{2^{R-1}\le|m|<2^R}\left\langle \sum_k a_{j,k+m,k}^{(0,\varepsilon,R)}\alpha_{j,k}^{(\varepsilon)}\Phi^{(0)}(x-k),\right.$$

$$\left. \sum_{k'}\sum_{l'} a_{j',l',k'}^{(0,\varepsilon,R)}\alpha_{j',k'}^{(\varepsilon)}(\Phi^{(0)}(2^{j'-j}x+2^{j'-j}m-l')-\Phi^{(0)}(2^{j'-j}x-l'))\right\rangle.$$

The fundamental theorem of calculus yields

$$2^{\frac{j-j'}{2}n}|\langle g_j^\varepsilon, g_{j'}^\varepsilon\rangle|$$

$$\leq \sum_{2^{R-1}\leq|m|<2^R}\sum_k |a_{j,k+m,k}^{(0,\varepsilon,R)}||\alpha_{j,k}^{(\varepsilon)}| \int |\Phi^{(0)}(x-k)| \sum_{k'}\sum_{l'} |a_{j',l',k'}^{(0,\varepsilon,R)}||\alpha_{j',k'}^{(\varepsilon)}|$$

$$\times |2^{j'-j}m||\nabla\Phi^{(0)}(2^{j'-j}x-l'+2^{j'-j}z)|\,dx,$$

where $z = z(x,j,j',l,m)$, $0 \leq |z| \leq |m|$, $\nabla\Phi^{(0)}(x)$ is the gradient of $\Phi^{(0)}(x)$. On the other hand we obviously have $\sum_{l'}|\nabla\Phi^{(0)}(x-l'+2^{j'-j}z)| \leq C$ uniformly in x.

By a repeated use of Cauchy-Schwarz inequality we obtain

$$\sum_{l'}\sum_{k'}|a_{j',l',k'}^{(0,\varepsilon,R)}||a_{j',k'}^{(\varepsilon)}||\nabla\Phi^{(0)}(2^{j'-j}x-l'+2^{j'-j}z)|$$

$$\leq \sum_{l'}\Big(\sum_{k'}|a_{j',l',k'}^{(0,\varepsilon,R)}|\Big)^{\frac{1}{2}}\Big(\sum_{k'}|a_{j',l',k'}^{(0,\varepsilon,R)}||a_{j',k'}^{(\varepsilon)}|^2\Big)^{\frac{1}{2}}|\nabla\Phi^{(0)}(2^{j'-j}x-l'+2^{j'-j}\triangle m)|$$

$$\leq C A_{j'}^{(0,\varepsilon,R)}\Big(\sum_{k'}|a_{j',k'}^{(\varepsilon)}|^2\Big)^{\frac{1}{2}}.$$

We finally denote by c_0 the integral $\int|\Phi^{(0)}(x-k)|dx$. Since $|m| \leq 2^R$, we have

$$|\langle g_j^\varepsilon, g_{j'}^\varepsilon\rangle| \leq Cc_0 2^{\frac{j'-j}{2}n}\sum_{2^{R-1}\leq|m|<2^R}\sum_k |a_{j,k+m,k}^{(0,\varepsilon,R)}||a_{j,k}^{(\varepsilon)}||2^{j'-j}m|A_{j'}^{(0,\varepsilon,R)}\Big(\sum_{k'}|a_{j',k'}^{(\varepsilon)}|^2\Big)^{\frac{1}{2}}$$

$$\leq C2^{j'-j+R}A_{j'}^{(0,\varepsilon,R)}\Big(\sum_{k'}|a_{j',k'}^{(\varepsilon)}|^2\Big)^{\frac{1}{2}}2^{\frac{j'-j}{2}n}\sum_{2^{R-1}\leq|m|<2^R}\sum_k |a_{j,k+m,k}^{(0,\varepsilon,R)}||a_{j,k}^{(\varepsilon)}|.$$

This discussion can be rewritten as an almost-orthogonality estimate given by the following lemma:

Lemma 4.2. *Keeping the same notations as above and assuming* $j \geq j'+R$, *we have*

$$|\langle g_j^\varepsilon, g_{j'}^\varepsilon\rangle| \leq C2^{(j'-j+R)(\frac{n}{2}+1)}A_{j'}^{(0,\varepsilon,R)}A_j^{(0,\varepsilon,R)}\|f_{j'}^\varepsilon\|_2\|f_j^\varepsilon\|_2. \tag{18}$$

Finally (17) and (18) imply Theorem 4.1. Indeed,

$$\|T_R^{(0,\varepsilon)}f\|_2^2 = \|T_R^{(0,\varepsilon)}f^\varepsilon\|_2^2 = \Big\langle \sum_j g_j^\varepsilon(x), \sum_{j'} g_{j'}^\varepsilon(x)\Big\rangle$$

$$\leq 2\sum_{0\leq j-j'\leq R}\sum \langle |g_j^\varepsilon(x)|, |g_{j'}^\varepsilon(x)|\rangle + 2\sum_{j-j'\geq R+1}\sum |\langle g_j^\varepsilon(x), g_{j'}^\varepsilon(x)\rangle|$$

$$\leq 2 \sum_{0 \leq j-j' \leq R} \sum A_j^{(0,\varepsilon,R)} A_{j'}^{(0,\varepsilon,R)} \|f_j^\varepsilon\|_2 \|f_{j'}^\varepsilon\|_2$$

$$+ C \sum_{j-j' \geq R+1} \sum 2^{(j'-j+R)(\frac{n}{2}+1)} A_j^{(0,\varepsilon,R)} A_{j'}^{(0,\varepsilon,R)} \|f_j^\varepsilon\|_2 \|f_{j'}^\varepsilon\|_2 = S_1 + S_2.$$

To estimate the first sum S_1, one writes $j' - j = q$ and freezes q while summing over j. Cauchy-Schwarz inequality is used and we obtain an estimate given by $\sum_0^\infty (A_j^{(0,\varepsilon,R)})^2 \|f_j\|_2^2 \leq A(R)^2 \|f\|_2^2$. These estimates are summed over $q \in [0, R]$ which amounts to multiply everything by $O(R)$. This yields a bound of the operator norm by $R^{1/2} A(R)$ but a closer look gives a better estimate. For each j, Cauchy-Schwarz inequality yields $\sum_{0 \leq q \leq R} A_{j+q}^{(0,\varepsilon,R)} \|f_{j+q}\|_2 \leq \|A_R\|_2 (\sum_{0 \leq q \leq R} \|f_{j+q}\|_2^2)^{1/2}$. We then need to estimate $\|A_R\|_2 \sum_j A_j^{(0,\varepsilon,R)} \|f_j\|_2 (\sum_{0 \leq q \leq R} \|f_{j+q}\|_2^2)^{1/2}$. We apply Cauchy-Schwarz inequality once more and obtain $\|A_R\|_2 \|f\|_2 (\sum_j \sum_{0 \leq q \leq R} (A_j^{(0,\varepsilon,R)})^2 \|f_{j+q}\|_2^2)^{1/2}$. We make the change of variable $j' = j + q$ we freeze j' and sum over q. We conclude as above. We now turn to S_2, which is estimated by Schur's lemma in its most primitive version. It says that a convolution product between a sequence belonging to l^1 and a sequence belonging to l^2 still belongs to l^2. We then estimate S_2 by $\sum_0^\infty (A_j^{(0,\varepsilon,R)})^2 \|f_j\|_2^2$ which ends the proof.

5. BMO estimates for $T_R^{(\varepsilon_1,\varepsilon_2)}$

In this section we relate the operator norm of $T_R^{(\varepsilon_1,\varepsilon_2)}$ acting on BMO to $\|A_R^{(\varepsilon_1,\varepsilon_2)}\|_p$. Let us denote by H^1 the Hardy space whose dual is BMO. The following lemma is an obvious corollary of Lemma 2.3:

Lemma 5.1. *If* $a_{j,k,l}^{(\varepsilon_1,0,R)}$ *fulfills* (10) *we then have*

$$\left\| \sum_l a_{j,k,l}^{(\varepsilon_1,0,R)} \Phi_{j,l}^{(0)} \right\|_{H^1} \leq \left\| \sum_l a_{j,k,l}^{(\varepsilon_1,0,R)} \Phi_{j,l}^{(0)} \right\|_{\dot{B}_1^{0,1}} \leq C 2^{-\frac{n}{2}j} R A_j^{(\varepsilon_1,0,R)}.$$

For keeping notations as simple as possible all wavelets will be written Ψ^ϵ (they were denoted by Φ^ϵ in the preceding calculations) and Φ will denote the scaling function.

Next new norms $B_R^{(\varepsilon_1,\varepsilon_2)}$ will be defined in terms of $\|A_R^{(\varepsilon_1,\varepsilon_2)}\|_p$. To be specific we write

$$B_R^{(\varepsilon_1,\varepsilon_2)} = \begin{cases} R\|A_R^{(\varepsilon_1,0)}\|_2 & \text{if } R = 1, 2, \ldots \text{ and } \varepsilon_1 \in \{0,1\}^n \setminus \{0\}, \ \varepsilon_2 = 0, \\ \|A_R^{(0,\varepsilon_2)}\|_1 & \text{if } R = 1, 2, \ldots \text{ and } \varepsilon_1 = 0, \ \varepsilon_2 \in \{0,1\}^n \setminus \{0\}, \\ \|A_R^{(\varepsilon_1,\varepsilon_2)}\|_2 & \text{if } R = 0, 1, 2, \ldots \text{ and } (\varepsilon_1,\varepsilon_2) \in \{0,1\}^{2n}, \ \varepsilon_1 \neq 0, \ \varepsilon_2 \neq 0. \end{cases}$$

We then have

Theorem 5.2.

(i) $\|T_R^{(\varepsilon_1,\varepsilon_2)}\|_{\mathrm{BMO}\to\mathrm{BMO}} \le CB_R^{(\varepsilon_1,\varepsilon_2)}$.

(ii) *This estimate is optimal.*

Let us first sketch the proof. Let $f(x) = \sum a_{j,k}^{(\eta)}\Psi_{j,k}^{(\eta)}(x) \in \mathrm{BMO}$. We then have

$$g_R^{(\varepsilon_1,\varepsilon_2)}(x) = T_R^{(\varepsilon_1,\varepsilon_2)}f(x) = \sum b_{j,k}^{(\varepsilon_1,\varepsilon_2,\varepsilon)}\Psi_{j,k}^{(\varepsilon)}(x). \tag{19}$$

Notations used in (19) will be kept in the proof. The indices ε_1 and ε_2 are labelling the operator. When both indices are not 0, the operator $T_R^{(\varepsilon_1,\varepsilon_2)}$ maps W_j into itself and this case is easier than the two other cases where $\varepsilon_1 = 0$ or $\varepsilon_2 = 0$. In the former case, the operator $T_R^{(0,\varepsilon_2)}$ maps W_j into V_j, and in the latter case the operator $T_R^{(\varepsilon_1,0)}$ maps V_j into W_j. These two cases need to be treated independently since BMO is not a Hilbert space and the duality argument used in the L^2 setting does not apply any more. The index ε will be used for labelling the frequency channels which are occurring in the wavelet expansion of $T_R^{(\varepsilon_1,\varepsilon_2)}(f)$. We compute the wavelet coefficients $b_{j,k}^{(\varepsilon_1,\varepsilon_2,\varepsilon)}$ as in Lemma 2.5. When ε_2 is not 0, we have $\varepsilon_2 = \eta$ (see Lemma 2.5). Similarly we have $\varepsilon_1 = \epsilon$ when ε_1 is not 0.

The characterization of BMO given by Theorem 2.2 is now used. To prove (i) in Theorem 5.2, it suffices to show that for each $\varepsilon \in \{0,1\}^n \setminus \{0\}$ and each dyadic cube $Q_{s,p}$, $s \in \mathbb{Z}$, $p \in \mathbb{Z}^n$, one has

$$\Lambda_{s,p}^{(\varepsilon_1,\varepsilon_2)} = \sum_{Q_{j,k}\subset Q_{s,p}} \left|b_{j,k}^{(\varepsilon_1,\varepsilon_2,\varepsilon)}\right|^2 \le C(B_R^{(\varepsilon_1,\varepsilon_2)})^2\|f\|_{\mathrm{BMO}}^2|Q_{s,p}|.$$

To estimate $|b_{j,k}^{(\varepsilon_1,\varepsilon_2,\varepsilon)}|$, one decomposes $f(x)$ into a sum of three functions $f_i^{s,p}(x)$, $i = 1,2,3$, where $f_i^{s,p}(x) = \sum a^i(\lambda)\Psi_\lambda(x)$ and where the coefficients $a^i(\lambda)$ are defined as indicated below. These three functions are partial sums in the wavelet expansion of f. Roughly speaking, the first function is the partial sum of the series corresponding to the wavelets which are supported by the enlarged cube $Q_{At,s,p}$. The second function is given by the wavelets which are located far away from $Q_{t,s,p}$. The third one is given by wavelets with a large support which intersects $Q_{t,s,p}$. For each given size length the cardinality of these large cubes does not exceed a constant C'. The notations which are needed for defining there three functions are given now. First $\tilde{Q}$ denotes the enlarged cube which is centered as Q is and whose size is C_0 times the size of Q. The constant C_0 only depends on the size of the support of the mother wavelet ψ. A new constant A will be defined in the proof of Lemma 5.4. Let us state that A only depends on C_0. All wavelets are denoted by Ψ. Let us set $t = 2^{M+2}$. The wavelet coefficients $a^i(\lambda) = a_{j,i}^{(\varepsilon,i)}$ of $f_i^{s,p}$ are now unveiled:

Definition 5.3. We set:

(i) $a_{j,k}^{(\varepsilon,1)} = a_{j,k}^{(\varepsilon)}$ if $Q_{j,k} \subset Q_{At,s,p}$ and $a_{j,k}^{(\varepsilon,1)} = 0$ if not.

(ii) $a_{j,k}^{(\varepsilon,2)} = a_{j,k}^{(\varepsilon)}$ if $Q_{j,k}$ is not contained in $Q_{At,s,p}$ and if $\tilde{Q}_{j,k} \cap Q_{t,s,p} = \emptyset$ while $a_{j,k}^{(\varepsilon,2)} = 0$ if $Q_{j,k}$ is not contained in $Q_{At,s,p}$ and if $\tilde{Q}_{j,k} \cap Q_{t,s,p} \neq \emptyset$.

(iii) $a_{j,k}^{(\varepsilon,3)} = a_{j,k}^{(\varepsilon)} - a_{j,k}^{(\varepsilon,1)} - a_{j,k}^{(\varepsilon,2)}$.

In (iii) $Q_{j,k}$ is not contained in $Q_{At,s,p}$ and $\tilde{Q}_{j,k} \cap Q_{t,s,p} \neq \emptyset$.

Lemma 5.4. *If $j \geq s$ and if $\tilde{Q}_{j,k} \cap Q_{t,s,p} \neq \emptyset$, then $Q_{j,k} \subset Q_{At,s,p}$. This implies $j \leq s - 1$ in* (iii).

In other words we cannot have (iii) when $j \geq s$. Indeed, if the "small cube" $Q_{j,k}$ is not contained in $Q_{At,s,p}$, this small cube cannot intersect $Q_{t,s,p}$. Once C_0 (which defines the enlarged cube) is defined, A is fixed by this simple remark which ends the proof of Lemma 5.4. Theorem 2.2 implies $\|f_i^{s,p}\|_{\mathrm{BMO}} \leq C\|f\|_{\mathrm{BMO}}$ and this estimate will be seminal in the proof.

The coefficients $b_{j,k}^{(\varepsilon_1,\varepsilon_2,\varepsilon,i)}$, which need to be estimated, show up in the wavelet expansion of

$$g_i^{(\varepsilon_1,\varepsilon_2)} = T_R^{(\varepsilon_1,\varepsilon_2)} f_i^{s,p} = \sum_{\lambda \in \Lambda_n} b_{j,k}^{(\varepsilon_1,\varepsilon_2,\varepsilon,i)} \Psi_{j,k}^{(\varepsilon)}.$$

We then write

$$\Lambda_{s,p}^{(\varepsilon_1,\varepsilon_2,i)} = \sum_{Q_{jk} \subset Q_{s,p}} |b_{j,k}^{(\varepsilon_1,\varepsilon_2,\varepsilon,i)}|^2$$

and we have $\Lambda_{s,p}^{(\varepsilon_1,\varepsilon_2)} \leq \sum_{i=1}^3 \Lambda_{s,p}^{(\varepsilon_1,\varepsilon_2,i)}$.

To prove the first assertion of Theorem 5.2, it suffices to show that, for $i = 1, 2, 3$, one has

$$\Lambda_{s,p}^{(\varepsilon_1,\varepsilon_2,i)} \leq C(B_R^{(\varepsilon_1,\varepsilon_2)})^2 \|f\|_{\mathrm{BMO}}^2 |Q_{s,p}|. \tag{20}$$

These fundamental estimates (20) will be proved in subsections (I), (II), and (III), corresponding to $i = 1, 2, 3$ respectively. Three cases (denoted by A, B, and C) are occurring in our treatment of the operator $T_R^{(\varepsilon_1,\varepsilon_2)}$ and altogether we are unfortunately faced to nine cases. For instance II.A means that neither ε_1 nor ε_1 is 0 and we are applying our operator to the piece $f_2^{s,p}(x)$ of f. Some cases are easier than others and will be only sketched. Three counterexamples will show that the estimates given by Theorem 5.2 are optimal.

(I) To estimate $\Lambda_{s,p}^{(\varepsilon_1,\varepsilon_2,1)}$, it suffices to use Theorem 4.1. The three cases A, B, and C are identical. Theorem 2.2 is used and yields $\|f_1^{s,p}\|_2 \leq C\|f\|_{\mathrm{BMO}}|Q_{s,p}|^{\frac{1}{2}}$. Finally one obtains

$$\Lambda_{s,p}^{(\varepsilon_1,\varepsilon_2,1)} \leq C(B_R^{(\varepsilon_1,\varepsilon_2)})^2 \|f\|_{\mathrm{BMO}}^2 |Q_{s,p}|.$$

(II) One then turns to $\Lambda_{s,p}^{(\varepsilon_1,\varepsilon_2,2)}$. Depending on the values of $(\varepsilon_1,\varepsilon_2)$, three cases are occurring which are denoted by II.A, II.B and II.C.

The case II.A occurs when neither ε_1 nor ε_2 is 0. This case is easy since the non-standard decomposition yields an obvious calculation of the action of the given operator. One then has $b_{j,k}^{(\varepsilon_1,\varepsilon_2,\varepsilon,2)} = 0$ if $\varepsilon_1 \neq \varepsilon$. If $\varepsilon_1 = \varepsilon$, one has $b_{j,k}^{(\varepsilon_1,\varepsilon_2,\varepsilon,2)} = \sum_l a_{j,k,l}^{(\varepsilon_1,\varepsilon_2,R)} a_{j,l}^{(\varepsilon_2,2)}$. This says that the action of the operator is fully decoupled in the frequency channels W_j. Let us remind the reader that $f_2^{s,p}(x) = \sum a_{j,l}^{(\varepsilon,2)} \Psi_{j,l}^{(\varepsilon)}(x)$ We now prove a lemma:

Lemma 5.5. *If $\varepsilon_1 \neq 0$ and $\varepsilon_2 \neq 0$ one has $b_{j,k}^{(\varepsilon_1,\varepsilon_2,2)} = \sum_l a_{j,k,l}^{(\varepsilon_1,\varepsilon_2,R)} a_{j,l}^{(\varepsilon_2,2)}$. Moreover if $j - s \geq R$ and if $Q_{j,k} \subset Q_{s,p}$, then $b_{j,k}^{(\varepsilon_1,\varepsilon_2,\varepsilon,2)} = 0$.*

Indeed, if $Q_{j,k} \subset Q_{s,p}$, then, for $i = 1, \ldots, n$, we have $2^{j-s} p_i \leq k_i \leq 2^{j-s}(p_i+1)-1$. If $\tilde{Q}_{j,l} \cap Q_{t,s,p} = \emptyset$, then this is a fortiori true when $\tilde{Q}_{j,l}$ is replaced by $Q_{j,l}$. Therefore there exists an index $i \in \{1, \ldots, n\}$ such that $l_i \geq 2^{j-s}(p_i+t+1)$ or $l_i \leq 2^{j-s}(p_i-t)$. It implies $|k-l| \geq 2^{j-s}$. But if $|k-l| \geq 2^R$, then $a_{j,k,l}^{(\varepsilon_1,\varepsilon_2,R)} = 0$. Lemma 5.5 is proved.

We now estimate $\Lambda_{(s,p)}^{(\varepsilon_1,\varepsilon_2,2)}$. This estimate will remain valid if $f \in$ BMO is replaced by the weaker hypothesis $f \in \dot{B}_\infty^{0,\infty}$. Using Lemma 5.5 and the definition of the coefficients $b_{j,k}^{(\varepsilon_1,\varepsilon_2,\varepsilon,2)}$ we have

$$\Lambda_{(s,p)}^{(\varepsilon_1,\varepsilon_2,2)} = \sum_{Q_{jk} \subset Q_{s,p}} \left| b_{j,k}^{(\varepsilon_1,\varepsilon_2,\varepsilon,2)} \right|^2$$

$$= \sum_{\substack{Q_{j,k} \subset Q_{s,p} \\ 0 \leq j-s \leq R-1}} \left| \sum_l a_{j,k,l}^{(\varepsilon_1,\varepsilon_2,R)} a_{j,l}^{(\varepsilon_2,2)} \right|^2$$

$$\leq \sum_{0 \leq j-s \leq R-1} \sum_{\substack{2^{j-s} p_i \leq k_i \leq 2^{j-s}(p_i+1)-1 \\ i=1,\ldots,n}} \left(\sum_l \left| a_{j,k,l}^{(\varepsilon_1,\varepsilon_2,R)} \right| \left| a_{j,l}^{\varepsilon_2,2} \right| \right)^2$$

$$\leq \sum_{0 \leq j-s \leq R-1} \sum_{\substack{2^{j-s} p_i \leq k_i \leq 2^{j-s}(p_i+1)-1 \\ i=1,\ldots,n}} \left(\sup_k \sum_l \left| a_{j,k,l}^{(\varepsilon_1,\varepsilon_2,R)} \right| \sup_l \left| a_{j,l}^{(\varepsilon_2,2)} \right| \right)^2$$

$$\leq C \sum_{0 \leq j-s \leq R-1} 2^{(j-s)n} \left| A_j^{(\varepsilon_1,\varepsilon_2,R)} \right|^2 \sup_l \left| a_{j,l}^{(\varepsilon_2)} \right|^2$$

$$\leq C \sum_{0 \leq j-s \leq R-1} \left| A_j^{(\varepsilon_1,\varepsilon_2,R)} \right|^2 2^{-sn} \sup_{\varepsilon_2,j,l} 2^{jn} \left| a_{j,l}^{(\varepsilon_2,2)} \right|^2$$

$$\leq C \left\| A_R^{(\varepsilon_1,\varepsilon_2)} \right\|_2^2 \|f\|_{\text{BMO}}^2 |Q_{s,p}|.$$

Here $f \in \dot{B}_\infty^{0,\infty}$ would suffice to conclude.

The case II.B is defined by $\varepsilon_1 \neq 0, \varepsilon_2 = 0$. Here also, the estimate we are going to prove will remain valid if $f \in$ BMO is replaced by the much weaker hypothesis $f \in \dot{B}_\infty^{0,\infty}$. If $\varepsilon_1 \neq \varepsilon$, then we have $b_{j,k}^{(\varepsilon_1,\varepsilon_2,\varepsilon,2)} = 0$. Since $\varepsilon_2 = 0$ the non-standard algebra does not coincide with standard matrix algebra and we have instead

$$b_{j,k}^{(\varepsilon,0,\varepsilon,2)} = \left\langle \sum_l a_{j,k,l}^{(\varepsilon,0,R)} \Phi_{j,l}^{(0)}, \ f_2^{s,p} \right\rangle = \left\langle \sum_l a_{j,k,l}^{(\varepsilon,0,R)} \Phi_{j,l}^{(0)}, \ \sum_{j' \leq j-1} a_{j',k'}^{(\varepsilon',2)} \Psi_{j',k'}^{(\varepsilon')} \right\rangle.$$

The limitation $j' \leq j - 1$ comes from the orthogonality between wavelets and scaling functions. But the support of the function $\Psi_{j',k'}^{(\varepsilon')}$ is contained in the cube $2^{-j'}k' + 2^{-j'}[-2^M, 2^M]^n \subset \tilde{Q}_{j',k'}$. By assumption this cube does not intersect $Q_{t,s,p}$. On the other hand the support of $\sum_l a_{j,k,l}^{(\varepsilon,0,R)} \Phi_{j,l}^{(0)}$ is contained in $2^{-j}k + 2^{-j}[-2^R, 2^R]^n + 2^{-j}[-2^M, 2^M]^n$. This comes from $|k - l| \leq 2^R$ in the preceding series.

If $Q_{j,k} \subset Q_{s,p}$ and $j \geq R + s$, the support of $\sum_l a_{j,k,l}^{(\varepsilon,0,R)} \Phi_{j,l}^{(0)}$ is contained in $Q_{s,p} + 2^{-s-R}[-(2^R + 2^M), 2^R + 2^M]^n \subset Q_{t,s,p}$. Then (ii) implies that the support of $\sum_{j' \leq j-1} a_{j',k'}^{(\varepsilon',2)} \Psi_{j',k'}^{(\varepsilon')}$ does not intersect $Q_{t,s,p}$. We just proved the following:

Lemma 5.6. *If* $Q_{j,k} \subset Q_{s,p}$ *and* $j \geq R + s$ *then* $b_{j,k}^{(\varepsilon,0,\varepsilon,2)} = 0$ *in the case II.B.*

Therefore we have

$$\Lambda_{s,p}^{(\varepsilon,0,2)} = \sum_{s \leq j \leq R+s-1} \ \sum_{\substack{2^{j-s}p_i < k_i < 2^{j-s}(p_i+1) \\ i=1,\ldots,n}} \left| \left\langle \sum_l a_{j,k,l}^{(\varepsilon,0,R)} \Phi_{j,l}^{(0)}, \ f_2^{s,p} \right\rangle \right|^2$$

$$\leq \sum_{s \leq j \leq R+s-1} \ \sum_{\substack{2^{j-s}p_i < k_i < 2^{j-s}(p_i+1) \\ i=1,\ldots,n}} \left\| \sum_l a_{j,k,l}^{(\varepsilon,0,R)} \Phi_{j,l}^{(0)} \right\|_{H^1}^2 \|f_2^{s,p}\|_{\text{BMO}}^2.$$

Lemma 5.1 is now applied and yields

$$\Lambda_{s,p}^{(\varepsilon,0,2)} \leq C \sum_{s \leq j \leq R+s-1} \ \sum_{\substack{2^{j-s}p_i < k_i < 2^{j-s}(p_i+1) \\ i=1,\ldots,n}} 2^{-nj} \cdot R^2 \cdot (A_j^{(\varepsilon,0,R)})^2 \|f\|_{\text{BMO}}^2$$

$$\leq CR^2 \sum_{s \leq j \leq R+s-1} \left(A_j^{(\varepsilon,0,R)}\right)^2 2^{-ns} \|f\|_{\text{BMO}}^2$$

$$\leq CR^2 \|A_R^{(\varepsilon,0)}\|_2^2 \|f\|_{\text{BMO}}^2 |Q_{s,p}|.$$

The case II.C is defined by $\varepsilon_1 = 0, \ \varepsilon_2 \neq 0$. We have

$$b_{j,k}^{(0,\varepsilon_2,\varepsilon,2)} = \left\langle \sum_{\substack{j' \in \mathbb{Z} \\ k',l' \in \mathbb{Z}^n}} a_{j',k',l'}^{(0,\varepsilon_2,R)} a_{j',l'}^{(\varepsilon_2,2)} \Phi_{j',k'}^{(0)}, \ \Psi_{jk}^{(\varepsilon)} \right\rangle.$$

If $Q_{j,k} \subset Q_{s,p}$, the support of $\Phi_{j,k}^{(\varepsilon)}$ is contained in $Q_{\frac{t}{4},s,p}$. Moreover as in II.B the support of $\sum_{j' \geq R+s} a_{j',k',l'}^{(0,\varepsilon_2,R)} \Phi_{j',k'}^{(0)} a_{j',l'}^{(\varepsilon_2,2)}$ does not intersect $Q_{\frac{t}{2},s,p}$. It implies:

$$b_{j,k}^{(0,\varepsilon_2,\varepsilon,2)} = \Big\langle \sum_{\substack{j+1 \leq j' < R+s \\ k',l' \in \mathbb{Z}^n}} a_{j',k',l'}^{(0,\varepsilon_2,R)} a_{j',l'}^{(\varepsilon_2,2)} \Phi_{j',k'}^{(0)}, \ \Psi_{j,k}^{(\varepsilon)} \Big\rangle.$$

The limitation $j' \geq j+1$ is coming from the orthogonality between wavelets and scaling functions. The constraint $j \geq s$ is coming from $Q_{j,k} \subset Q_{s,p}$.

We now use a local Plancherel identity given by the following lemma:

Lemma 5.7. *Let E be a Borel set and let W be the collection of all $\lambda \in \Lambda_n$ such that E contains the support of the wavelet $\Psi_\lambda = \Psi_{jk}^{(\varepsilon)}$ labelled by λ. Let us denote by $a(\lambda) = a_{j,k}^\epsilon$ the wavelet coefficients of a function f. Then we have*

$$\sum_{\lambda \in W} |a(\lambda)|^2 \leq \int_E |f(x) - m_E(f)|^2 \, dx \leq \int_E |f(x)|^2 \, dx$$

where $m_E(f)$ is the mean value of f over E.

The proof is easy. The indicator function of E is denoted by χ_E. Writing $\tilde{f} = (f - m_E(f))\chi_E$, we observe that if $\lambda \in W$, we have $a_{j,k}^\epsilon = \int f \Psi_{j,k}^\epsilon \, dx = \int \tilde{f} \Psi_{j,k}^\epsilon \, dx = \tilde{a}_{j,k}^\epsilon$. This implies Lemma 5.7.

Lemma 5.7 is applied to $E = Q_{\frac{t}{2},s,p}$ and to

$$g(x) = \sum_{s+1 \leq j' < R+s} \sum_{k' \in \mathbb{Z}^n} \sum_{l' \in \mathbb{Z}^n} a_{j',k',l'}^{(0,\varepsilon_2,R)} a_{j',l'}^{(\varepsilon_2,2)} \Phi_{j',k'}^{(0)}(x).$$

Lemma 5.7 implies

$$\Lambda_{s,p}^{(0,\varepsilon_2,2)} = \sum_{Q_{j,k} \subset Q_{s,p}} \Big| \Big\langle \sum_{s+1 \leq j' < R+s} \sum_{k' \in \mathbb{Z}^n} \sum_{l' \in \mathbb{Z}^n} a_{j',k',l'}^{(0,\varepsilon_2,R)} a_{j',l'}^{(\varepsilon_2,2)} \Phi_{j',k'}^{(0)}, \Psi_{j,k}^{(\varepsilon)} \Big\rangle \Big|^2$$

$$\leq C \int_E \Big| \sum_{s+1 \leq j' < R+s-1} \sum_{k' \in \mathbb{Z}^n} \sum_{l' \in \mathbb{Z}^n} a_{j',k',l'}^{(0,\varepsilon_2,R)} a_{j',l'}^{(\varepsilon_2,2)} \Phi_{j',k'}^{(0)}(x) \Big|^2 \, dx.$$

The following bounds are obvious:

$$
\begin{aligned}
\Lambda_{s,p}^{(0,\varepsilon_2,2)} &\leq C \int_E \Big[\sum_{s+1\leq j'\leq R+s-1} \sum_{k'\in\mathbb{Z}^n} \sum_{l'\in\mathbb{Z}^n} \big|a_{j',k',l'}^{(0,\varepsilon_2,R)}\big|\big|a_{j',l'}^{(\varepsilon_2,2)}\big|\big|\Phi_{j',k'}^{(0)}(x)\big| \Big]^2 dx \\
&\leq C \int_E \Big[\sum_{s+1\leq j'\leq R+s-1} \sum_{k'\in\mathbb{Z}^n} \sum_{l'\in\mathbb{Z}^n} \big|a_{j',k',l'}^{(0,\varepsilon_2,R)}\big|\big|a_{j',l'}^{(\varepsilon_2)}\big|\big|\Phi_{j',k'}^{(0)}(x)\big| \Big]^2 dx \\
&\leq C \int_E \Big[\sum_{s+1\leq j'\leq R+s-1} \sup_{k'} \sum_{l'\in\mathbb{Z}^n} \big|a_{j',k',l}^{(0,\varepsilon_2,R)}\big| \\
&\qquad\quad \times \sup_{j',l',\varepsilon_2} 2^{\frac{j'n}{2}} \big|a_{j',l'}^{(\varepsilon_2)}\big| \sum_{k'} \big|\Phi^{(0)}(2^{j'}x-k')\big| \Big]^2 dx \\
&\leq C \Big(\sum_{s+1\leq j'\leq R+s-1} A_{j'}^{(0,\varepsilon_2,R)} \Big)^2 \|f\|_{\mathrm{BMO}}^2 |Q_{\frac{t}{2},s,p}| \\
&\leq C \big\| A_R^{(0,\varepsilon_2)} \big\|_1^2 \|f\|_{\mathrm{BMO}}^2 |Q_{\frac{t}{2},s,p}| \\
&\leq C \big\| A_R^{(0,\varepsilon_2)} \big\|_1^2 \|f\|_{\mathrm{BMO}}^2 |Q_{s,p}|.
\end{aligned}
$$

(III) We need to compute $\Lambda_{s,p}^{(\varepsilon_1,\varepsilon_2,3)}$. The same arguments which were used in (II) will now apply again. Three cases are considered. The case III.A is defined by $\varepsilon_1 \neq 0$, $\varepsilon_2 \neq 0$. Then we have

$$
\Lambda_{s,p}^{(\varepsilon_1,\varepsilon_2,3)} = \sum_{Q_{j,k}\subset Q_{s,p}} \Big| \sum_l a_{j,k,l}^{(\varepsilon_1,\varepsilon_2,R)} a_{j,l}^{(\varepsilon_2,3)} \Big|^2.
$$

If $Q_{j,k}\subset Q_{s,p}$, we have $j\geq s$. Moreover if $a_{j,l}^{(\varepsilon_2,3)}\neq 0$, we have $j\leq s-1$. It yields $\Lambda_{s,p}^{(\varepsilon_1,\varepsilon_2,3)}=0$.

The case III.B is defined by $\varepsilon_1 = 0$, $\varepsilon_2 \neq 0$. If $a_{j',l'}^{(\varepsilon_2,3)}\neq 0$, we have $j'\leq s-1$. Moreover if $Q_{j,k}\subset Q_{s,p}$, we have $j\geq s$. It implies $j'\leq j-1$. Since wavelets are orthogonal to scaling functions we have $\langle \sum_{k'} a_{j',k',l'}^{(0,\varepsilon_2,R)} \Phi_{j',k'}^{(0)}, \Phi_{j,k}^{(\varepsilon)}\rangle = 0$ and $\Lambda_{s,p}^{(0,\varepsilon_2,3)}=0$.

The case III.C is defined by $\varepsilon_1 \neq 0$, $\varepsilon_2 = 0$. We have

$$
\begin{aligned}
\Lambda_{s,p}^{(\varepsilon_1,0,3)} &= \sum_{\substack{s\leq j\leq R+s-1 \\ Q_{j,k}\subset Q_{s,p}}} \Big| \Big\langle \sum_l a_{j,k,l}^{(\varepsilon_1,0,R)} \Phi_{j,l}^{(0)}, f_3^{s,p} \Big\rangle \Big|^2 + \sum_{\substack{j\geq R+s \\ Q_{j,k}\subset Q_{s,p}}} \Big| \Big\langle \sum_l a_{j,k,l}^{(\varepsilon_1,0,R)} \Phi_{j,l}^{(0)}, f_3^{s,p} \Big\rangle \Big|^2 \\
&= \Lambda_{s,p}^{(\varepsilon_1,0,3,1)} + \Lambda_{s,p}^{(\varepsilon_1,0,3,2)}.
\end{aligned}
$$

Let us consider the first term. Lemma 5.1 is used again and we have

$$
\begin{aligned}
\Lambda_{s,p}^{(\varepsilon_1,0,3,1)} &\leq \sum_{\substack{s \leq j \leq R+s-1 \\ Q_{j,k} \subset Q_{s,p}}} \Big(\Big\| \sum_l a_{j,k,l}^{(\varepsilon_1,0,R)} \Phi_{j,l}^{(0)}(x) \Big\|_{B_1^{0,1}} \Big)^2 (\|f_3^{s,p}\|_{B_\infty^{0,\infty}})^2 \\
&\leq C \sum_{\substack{s \leq j \leq R+s-1 \\ Q_{j,k} \subset Q_{s,p}}} 2^{-nj} R^2 (A_j^{(\varepsilon_1,0,R)})^2 \|f_3^{s,p}\|_{\mathrm{BMO}}^2 \\
&\leq C R^2 \|A_R^{(\varepsilon_1,0)}\|_2^2 \|f\|_{\mathrm{BMO}}^2 |Q_{s,p}|.
\end{aligned}
$$

Let us consider the second term now. We know that $j' \leq s-1$ in (iii). Therefore we have

$$
\begin{aligned}
\Lambda_{s,p}^{(\varepsilon_1,0,3,2)} &= \sum_{\substack{j \geq R+s \\ Q_{j,k} \subset Q_{s,p}}} \Big| \Big\langle \sum_l a_{j,k,l}^{(\varepsilon_1,0,R)} \Phi_{j,l}^{(0)}, \ \sum_{j' \leq s-1} \sum_{k'} \sum_\varepsilon a_{j',k'}^{(\varepsilon,3)} \Psi_{j',k'}^{(\varepsilon)} \Big\rangle \Big|^2 \\
&\leq \sum_{\substack{j \geq R+s \\ Q_{j,k} \subset Q_{s,p}}} \Big(\sum_{j' \leq s-1} \sum_{k'} \sum_\varepsilon |a_{j',k'}^{(\varepsilon,3)}| \Big| \Big\langle \sum_l a_{j,k,l}^{(\varepsilon_1,0,R)} \Phi_{j,l}^{(0)}, \ \Psi_{j',k'}^{(\varepsilon)} \Big\rangle \Big| \Big)^2.
\end{aligned}
$$

The cancellations built in the definition of the operators $T_R^{(\epsilon_1,\epsilon_2)}$ yield

$$
\begin{aligned}
&\Big| \Big\langle \sum_l a_{j,k,l}^{(\varepsilon_1,0,R)} \Phi_{j,l}^{(0)}, \ \Psi_{j',k'}^{(\varepsilon)} \Big\rangle \Big| \\
&= 2^{\frac{j'-j}{2}n} \Big| \Big\langle \sum_l a_{j,k,l}^{(\varepsilon_1,0,R)} \Phi^{(0)}(x-l), \ \Psi^{(\varepsilon)}(2^{j'-j}x-k') \Big\rangle \Big| \\
&= 2^{\frac{j'-j}{2}n} \Big| \Big\langle \sum_{2^{R-1} \leq |m| < 2^R} a_{j,k,k+m}^{(\varepsilon_1,0,R)} \big(\Phi^{(0)}(x-k-m) - \Phi^{(0)}(x-k) \big), \\
&\qquad \Psi^{(\varepsilon)}(2^{j'-j}x-k') \Big\rangle \Big| \\
&= 2^{\frac{j'-j}{2}n} \Big| \sum_{2^{R-1} \leq |m| < 2^R} \Big\langle a_{j,k,k+m}^{(\varepsilon_1,0,R)} \Phi^{(0)}(x-k)), \\
&\qquad \Psi^{(\varepsilon)}(2^{j'-j}x-k'+2^{j'-j}m) - \Psi^{(\varepsilon)}(2^{j'-j}x-k') \Big\rangle \Big| \\
&\leq C 2^{\frac{j'-j}{2}n} \\
&\quad \times \sum_{2^{R-1} \leq |m| < 2^R} |a_{j,k,k+m}^{(\varepsilon_1,0,R)}| \|\Psi^{(\varepsilon)}(2^{j'-j}x-k'+2^{j'-j}m) - \Psi^{(\varepsilon)}(2^{j'-j}x-k')\|_\infty \\
&\leq C 2^{(j'-j)(R+\frac{n}{2}+1)} A_j^{(\varepsilon_1,0,R)}.
\end{aligned}
$$

Theorem 2.2 implies that for j' frozen we have $\sum_{k'} \sum_\varepsilon |a_{j',k'}^{(\varepsilon,3)}| 2^{nj'/2} \leq \|f\|_{\mathrm{BMO}}$. Indeed each individual term is estimated as indicated and their cardinality is uniformly bounded. Here $f \in \dot{B}_\infty^{0,\infty}$ would suffice. We then have

$$\Lambda_{s,p}^{(\varepsilon_1,0,3,2)} \leq \sum_{\substack{j \geq R+s \\ Q_{j,k} \subset Q_{s,p}}} \left(\sum_{j' \leq s-1} 2^{-(\frac{n}{2}+1)j+j'+R} A_j^{(\varepsilon_1,0,R)} \|f\|_{\mathrm{BMO}} \right)^2$$

$$\leq \sum_{\substack{j \geq R+s \\ Q_{j,k} \subset Q_{s,p}}} 2^{-nj-2(j-R-s)} \left(A_j^{(\varepsilon_1,0,R)} \right)^2 \|f\|_{\mathrm{BMO}}^2$$

$$\leq C \|A_R^{(\varepsilon,0)}\|_\infty^2 \|f\|_{\mathrm{BMO}}^2 |Q_{s,p}|.$$

Theorem 5.2 is now proved and our last task is to construct some counter-examples. This will prove that the given estimates are optimal ones. To simplify notations, we stick to the one dimensional case. We consider three cases

(a) $|\varepsilon_1||\varepsilon_2| \neq 0$,   (b) $\varepsilon_1 \neq 0$, and $\varepsilon_2 = 0$,   (c) $\varepsilon_1 = 0$ with $\varepsilon_2 \neq 0$.

In case (a), we set

$$a_{j,l} = \begin{cases} 2^{-\frac{j}{2}} & \text{if } 2^{R-1} + 2^{j+m} \leq l \leq 2^R - 2^{j+m} \text{ and } 0 \leq j \leq \frac{R}{3}, \\ 0 & \text{if not.} \end{cases}$$

Since $2^{R-1-j} + 2^M \leq 2^{-j}l \leq 2^{R-j} - 2^M$, one checks that the BMO norm of $f_R(x) = \sum_j \sum_l a_{j,l} \psi_{j,l}(x)$ is $O(1)$, as is the L^∞ norm. Then we set $a_{j,k,l}^{(1,1,R)} = 2^{-R}$ if $2^{R-1} \leq l-k \leq 2^R - 1$, $0 \leq j \leq \frac{R}{3}$, and $0 \leq k \leq 2^j - 1$, 0 if not. We then have $B_R^{(1,1)} \sim R^{\frac{1}{2}}$. Moreover we have $T_R^{(1,1)} f_R(x) = \sum_{j,k} b_{j,k} \psi_{j,k}$ and an easy computation proves that $b_{j,k}$ has the same order of magnitude as $2^{-j/2}$ if $0 \leq j \leq R/3$ and $0 \leq k < 2^j$. Then one compute the squared $L^2[-2^{M+2}, 2^{M+2}]$ norm of $T_R^{(1,1)} f_R(x)$. One obtains $C \sum_{0 \leq j \leq \frac{R}{3}} \sum_{0 \leq k \leq 2^j - 1} 2^{-j} \geq CR$. This implies $\|T_R^{(1,1)}\|_{\mathrm{BMO} \to \mathrm{BMO}} \geq C B_R^{(1,1)}$.

In the second case, one sets $g_j^R(x) = \sum_{s=4+M+j-R}^{j-M} \psi(2^s x - 11 \cdot 2^{R-4-j+s})$ and $f_R(x) = \sum_{j=M+3}^{[\frac{R}{3}]-M} g_j^R(x)$. Since the support of $g_j^R(x)$ is contained in $[2^{R-1-j}, 2^{R-j}]$, we have

$$\|f_R(x)\|_{\mathrm{BMO}} \leq \max_j \|g_j^R(x)\|_{\mathrm{BMO}} \leq C.$$

One considers the operator $T_R^{(1,0)}$ whose distributional kernel is

$$K_{(x,y)}^{(1,0,R)} = \sum_{j=M+3}^{[\frac{R}{3}]-M} \sum_{0 \leq k \leq 2^j - 1} \psi_{j,k}(x)(\varphi_{j,11.2^{R-4}}(y) - \varphi_{j,13.2^{R-4}}(y)).$$

We then have $B_R^{(1,0)} \sim R^{\frac{3}{2}}$. Moreover one has

$$
\begin{aligned}
a_j &= \int f_R(y)(\varphi_{j,11 \cdot 2^{R-4}}(y) - \varphi_{j,13 \cdot 2^{R-4}}(y))\, dy \\
&= \int g_j^R(y)(\varphi_{j,11 \cdot 2^{R-4}}(y) - \varphi_{j,13 \cdot 2^{R-4}}(y)\, dy \\
&= \int g_j(y)\varphi_{j,11 \cdot 2^{R-4}}(y)\, dy \geq CR \cdot 2^{-\frac{j}{2}}.
\end{aligned}
$$

Therefore we have

$$
T_R^{(1,0)} f_R(x) = \sum_{j=M+3}^{[\frac{R}{3}]-M} \sum_{0 \leq k \leq 2^{j-1}} a_j \psi_{j,k}(x).
$$

Finally Theorem 2.2 implies

$$
\left\| T_R^{(1,0)} f_R(x) \right\|_{\text{BMO}}^2 \geq C \sum_{\substack{Q_{j,k} \subset [0,1] \\ M+3 \leq j \leq [\frac{R}{3}]-M}} R^2 \cdot 2^{-j} \geq CR^3 \sim \left(B_R^{(1,0)} \right)^2.
$$

In the last case, one sets

$$
a_{j,k}^R = \begin{cases} 2^{-\frac{j}{2}}, & \text{if } -2^{R-1} + 2^{\frac{1}{4}j} \leq k \leq -2^R - 1 + 2^{\frac{3}{4}j} \text{ and } 4M \leq j \leq [\frac{R}{3}] + M, \\ 0, & \text{if not.} \end{cases}
$$

Since the support of $\sum_k a_{j,k}^R \psi_{j,k}(x)$ is contained in $[-2^{R-1-j}, -2^{R-2-j}]$, the BMO norm of $f_R(x) = \sum_j \sum_k a_{j,k}^R \psi_{jk}(x)$ is $O(1)$ as is the L^∞ norm. Then one sets

$$
K_{(x,y)}^{(0,1,R)} = \sum_{j=4M}^{[\frac{R}{3}]+M} \sum_k (\varphi_{j,k}(x) - \varphi_{j,k+2^{R-1}}(x))\psi_{j,k}(y).
$$

We obviously have $B_R^{(0.1)} \sim R$. But we have

$$
\begin{aligned}
T_R^{(0,1)} f_R(x) &= \sum_{j=4M}^{[\frac{R}{3}]+M} \sum_{-2^{R-1}+2^{\frac{1}{4}j} \leq k \leq -2^{R-1}+2^{\frac{3}{4}j}} 2^{-\frac{j}{2}}(\varphi_{j,k}(x) - \varphi_{j,k+2^{R-1}}(x)) \\
&= \sum_{j=4M}^{[\frac{R}{3}]+M} \sum_{-2^{R-1}+2^{\frac{1}{4}j} \leq k \leq -2^{R-1}+2^{\frac{3}{4}j}} 2^{-\frac{j}{2}}\varphi_{j,k}(x) - \sum_{j=4M}^{[\frac{R}{3}]+M} \sum_{2^{\frac{1}{4}j} \leq k \leq 2^{\frac{3}{4}j}} 2^{-\frac{j}{2}}\varphi_{j,k}(x) \\
&= \sum_{j=4M}^{[\frac{R}{3}]+M} g_j(x) - g(x).
\end{aligned}
$$

We have $\operatorname{supp} g_j(x) \subset [-2^{R-1-j}, -2^{R-2-j}]$ and $\operatorname{supp} g(x) \subset [\frac{3}{16}, 1]$. Then, for $0 < t < \frac{3}{16}$, we define $B_t = [t, 1]$, $\int_{[0,1]} g(x)\,dx = m_0$. We then have

$$m_0 = \sum_{j=4M}^{[\frac{R}{3}]+M} \sum_{2^{\frac{1}{4}j} \le k \le 2^{\frac{3}{4}j}} 2^{-j} \le C$$

and

$$m_{B_t} g = \frac{1}{|B_t|} \int_{B_t} g(x)\,dx = \frac{m_0}{1-t}.$$

Moreover we have $\int g^2(x)dx \ge CR^2$. Therefore we have

$$\|g\|_{\mathrm{BMO}}^2 \ge \sup_{0<t<\frac{3}{16}} \frac{1}{|B_t|} \int (g - m_{B_t} g)^2\,dx$$

$$= \sup_{0<t<\frac{3}{16}} \frac{1}{|B_t|} \int g^2(x)\,dx - \frac{1}{|B_t|}(m_{B_t} g)^2$$

$$= \sup_{0<t<\frac{3}{16}} \frac{1}{1-t} \int g^2(x)\,dx - \frac{m_0^2}{(1-t)^3}$$

$$\ge CR^2.$$

Finally we can conclude with

$$\|T_R^{(0,1)} f_R(x)\|_{\mathrm{BMO}} \ge \|g\|_{\mathrm{BMO}} \ge CB_R^{(0,1)}.$$

6. The proof of Theorem 1.6 and some open problems

We first prove Theorem 1.6.

(i) Theorem 5.2 implies that if $T \in \mathrm{OpB}_{\frac{3}{2}}$, then T and T^* are bounded on BMO. Using Theorem 3.3, we know that if $T \in \mathrm{OpM}_{\frac{3}{2}}$, then $T \in \mathrm{OpB}_{\frac{3}{2}}$. Therefore if $T \in \mathrm{OpM}_{\frac{3}{2}}$, then T and T^* are continuous on BMO.

(ii) One defines $b(x) = \sum_k a(k)\Phi^{(0)}(x-k)$ where (a), $\sum_k a(k) = 0$; (b), $a(k) = |k|^{-1}(\log(2+|k|))^{-2}$, if $|k| \ge 1$. Next we define $K(x,y) = \Phi^{(1,0,\cdots,0)}(x)b(y)$, and $K(x,y)$ is the distributional kernel of an operator $T \in \mathrm{OpM}_\mu$ ($\forall \mu < 1$). For $1 \le p < \infty$, $1 \le q \le \infty$, T is bounded on $\dot{F}_p^{0,q}$ and also on $\dot{B}_p^{0,q}$. But $T^*\Phi^{(1,0,\cdots,0)}(x) = b(x) \notin H^1$. Therefore T cannot be bounded on BMO.

Three open problems are now discussed.

(i) From Theorem 1.6, we know that every $T \in \mathrm{OpM}_{\frac{3}{2}}$ is bounded on BMO. On the other hand for $\mu < 1$, there exists an operator $T \in \mathrm{OpM}_\mu$ which is not bounded on BMO. These remarks are raising the following issue. What is the smallest $\mu \in [1, \frac{3}{2})$ such that every $T \in \mathrm{OpM}_\mu$ is bounded on BMO?

(ii) The operator class OpE_γ is characterized by a simple decay estimation on the non-standard coefficients. Is it true for OpH?

(iii) If T belongs to $\mathrm{OpM}_{\frac{1}{2}}$, then T is bounded on L^2 (Theorem 4.1). If T is given by a convolution and if $T \in \mathrm{OpH}$, then T is bounded on L^2. These two remarks are raising the following problem. Is it true that every operator $T \in \mathrm{OpM}_0$ or $T \in \mathrm{OpH}$ is bounded on L^2? The author proved in [14] that for any $T \in$ OpH the two following properties are equivalent ones: (a) T is bounded on L^2, (b) $\|T\chi_E(x)\|_{WL^1(F)} \le C|F|$ for all cubes F and all Borel sets $E \subset F$. Here $|F|$ denotes the volume of F and $WL^1(F)$ is denoting the "weak L^1 norm."

References

[1] G. Beylkin, R. Coifman, and V. Rokhlin, *Fast wavelet transforms and numerical algorithms*, I, Comm. Pure Appl. Math. **44** (1991), no. 2, 141–183.

[2] J.-M. Bony, *Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires*, Ann. Sci. École Norm. Sup. (4) **14** (1981), no. 2, 209–246.

[3] G. David and J.-L. Journé, *A boundedness criterion for generalized Calderón-Zygmund operators*, Ann. of Math. (2) **120** (1984), no. 2, 371–397.

[4] D. G. Deng, L. X. Yan, and Q. X. Yang, *Blocking analysis and $T(1)$ theorem*, Sci. China Ser. A **41** (1998), no. 8, 801–808.

[5] ———, *On Hörmander condition*, Chinese Sci. Bull. **42** (1997), no. 16, 1341–1345.

[6] M. Frazier, Y.-S. Han, B. Jawerth, and G. Weiss, *The $T1$ theorem for Triebel-Lizorkin spaces*, Harmonic Analysis and Partial Differential Equations (El Escorial, 1987), Lecture Notes in Math., vol. 1384, Springer, Berlin, 1989, pp. 168–181.

[7] F. John and L. Nirenberg, *On functions of bounded mean oscillation*, Comm. Pure Appl. Math. **14** (1961), 415–426.

[8] Y.-S. Han and S. Hofmann, *T1 theorems for Besov and Triebel-Lizorkin spaces*, Trans. Amer. Math. Soc. **337** (1993), no. 2, 839–853.

[9] P. G. Lemarié, *Continuité sur les espaces de Besov des opérateurs définis par des intégrales singulières*, Ann. Inst. Fourier (Grenoble) **35** (1985), no. 4, 175–187 (French, with English summary).

[10] Y. Meyer, *La mimimalité de l'espace de Besov $B_1^{0.1}$ et la continuité des opérateurs définis par des intégrales singulières*, Monografías de Matemáticas, vol. 4, Univ. Autónoma de Madrid, 1987.

[11] ———, *Wavelets and operators*, Cambridge Studies in Advanced Mathematics, vol. 37, Cambridge University Press, Cambridge, 1992.

[12] Y. Meyer and R. Coifman, *Wavelets: Calderón-Zygmund and multilinear operators*, Cambridge Studies in Advanced Mathematics, vol. 48, Cambridge University Press, Cambridge, 1997.

[13] Q. X. Yang, *Fast algorithms for Calderón-Zygmund singular integral operators*, Appl. Comput. Harmon. Anal. **3** (1996), no. 2, 120–126.

[14] _____ , *From weak continuity to strong continuity*, J. Jilin Univ. Sci. **40** (2002), no. 4, 331–338 (Chinese, with English and Chinese summaries).

[15] _____ , *Symbol operator and the speed of approximation by compact operator*, Acta Math. Sinica, to appear.