PDMI PREPRINT — 07/2005

STABILITY OF APPROXIMATION UNDER SINGULAR INTEGRALS, AND CALDERÓN–ZYGMUND TYPE DECOMPOSITIONS

S. V. Kislyakov and N. Ya. Kruglyak

POMI, Fontanka 27, St. Peterburg 191011, Russia
Luleå University of Technology, Luleå, Sweden

March 20, 2005

ABSTRACT

Let X be a space of "smooth" functions on $\mathbb{R}^n$ (like Lip_α, or $W_p^{(1)}$, or BMO). It is fairly well known that a specific Calderón–Zygmund type algorithm can be used to exhibit an element u of almost optimal L^1-approximation of a given function $f \in L^1$ by the ball of a fixed radius in X. We show that for most of the singular integral operators T the function Tu also approximates Tf in the same L^1-optimal sense provided Tf is integrable.

1. Introduction

0.1. Let X be a function space on $\mathbb{R}^n$ (we shall take $L^s(\mathbb{R}^n)$ with $s > 1$, BMO, homogeneous Lipschitz and Sobolev classes, etc. for the role of X). Suppose $f \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$, and consider an element of almost optimal L^p-approximation of f by functions in the ball $B_X(t) = B_X(0; t)$ of a given radius t in X. For technical reasons, often we do not insist on precisely "t" here allowing ct instead, that is, we are interested in functions u satisfying

$$u \in B_X(Ct), \quad \|f - u\|_1 \le C \operatorname{dist}_{L^p}(f, B_X(t/C)), \tag{1}$$

where t is a parameter and C is some constant independent of f and t.

Next, suppose T is a Calderón–Zygmund singular integral operator. Can u be chosen in such a way that Tu also approximates Tf in L^p in some optimal sense?

If $p > 1$, the inequality

$$\|Tf - Tu\|_{L^p} \le C\|f - u\|_{L^p} \le C' \operatorname{dist}_{L^p}(f, B_X(ct))$$

provides an immediate and trivial answer. For $p = 1$, the situation is different: L^1 is not taken by T to itself unless in degenerate cases, and the matter is not so easy. Nevertheless, very often the following can be ensured.

Suppose that $f \in L^1$ and Tf happens to be integrable. Then for every $t > 0$ there is a function u such that

$$u \in B_X(ct), \quad \|f - u\|_{L^1} \le C \operatorname{dist}_{L^1}(f, B_X(ct)) \tag{2}$$

and

$$\|Tf - Tu\|_{L^1} \le C_1(\operatorname{dist}_{L^1}(f, B_X(ct)) + \operatorname{dist}_{L^1}(Tf, B_X(t))), \tag{3}$$

where C_1 depends on T. (Surely, C, c must not depend on f and t; we shall assume this tacitly in similar situations to be treated below.)

The condition that Tf is integrable evokes thoughts about the real-variable Hardy space H^1 treated in [FS] by Fefferman and Stein:

$$H^1 = \{f \in L^1 : R_j f \in L^1, \quad 1 \le j \le n\},$$

where the R_j are the Riesz transformations, which are basic examples of singular integral operators. It is well known that, if $f \in H^1$, then in fact $Tf \in L^1$ for *all* singular integral operators T. However, in the above question we do not insist that f be an H^1-function. Still, we say a few words about the boundary class for the *analytic Hardy space* $H^1(\mathbb{R}^2_+)$. This space is related to the singular integral operator P, $Pf = (\hat{f}\chi_{\mathbb{R}^1_+})^\vee$ (the Riesz projection). Formulas (2), (3) with $T = P$ say that if $f \in L^1(\mathbb{R}^1)$ is *analytic* (in the sense that $Pf = f$), then $\operatorname{dist}_{L^1}(f, B_X(ct)) \asymp \operatorname{dist}_{L^1}(f, PB_X(c_1 t))$, i.e., an almost optimal approximation of an analytic function can always be realized by an analytic function. In the case where $X = L^r(\mathbb{R}^1)$, $r > 1$, some activity around this situation existed in relatively recent time, see, e.g., [Bo, Ki, KiX, P].

In this paper we want to explain that, should X be L^r or not, (2) and (3) can be ensured suprisingly often and in a surprisingly uniform way. The recipe is as follows: apply an appropriate Calderón–Zygmund type decomposition algorithm to f; it will give a function

u satisfying (2). Then for *arbitrary* singular integral operator T with Tf integrable we have (3).

In the most part of examples to be treated below T will be bounded on X. Therefore, along with u, Tu also belongs to a ball of radius roughly t in X.

In its classical form, Calderón–Zygmund decomposition was first presented in the famous paper [CZ] "On existence of certain singular integrals operators", published by Calderón and Zygmund in Acta Mathematica more than 50 years ago. This classical procedure is suitable for treatment of the above question in the case of $X = L^p$, $1 < p < \infty$. When X is a space of "smooth" functions (for uniformity, we refer this term also to BMO rather than merely to Lipschitz or Sobolev spaces), a different and more complicated algorithm is required. Refined Calderón–Zygmund type decomposition of this sort were constructed by N. Kruglyak [Kr1].

0.2. The layout of the paper. We include the treatment of the situation where $X = L^p$ with $1 < p < \infty$ (§1), though, basically, the result is known in this case: see Bourgain's paper [Bo], and also additions and refinements in [KiX, Ki]. One reason for which we do so is quite important for us, at least phychologically: strictly speaking, Bourgain did not follow the above recipe, namely, he applied the Calderón–Zygmund procedure *to some function related to* f rather than to f itself. Another reason is that in [Bo], only the case where T is a projection was treated. As was mentioned in [KiX, Ki], the general case can be reduced to this by considering the projection $(f_1, f_2) \mapsto (f_1, Tf_1)$ on pairs of functions. However, this formal way leads to u satisfying *only* conditions (4), (5) below instead of the *stronger* conditions (2), (3):

$$\|f - u\|_{L^1} \le Cd, \quad \|Tf - Tu\|_{L^1} \le Cd, \tag{4}$$

where

$$d = \max\{\mathrm{dist}_{L^1}(f, B_X(ct), \quad \mathrm{dist}(Tf, B_X(ct)). \tag{5}$$

Yet another reason is that we need a consequence of the results pertaining to the case of $X = L^p$ in order to analyze the case where X is a space of smooth functions. Namely, we shall prove the following duality lemma to be used in §2.

Lemma. *Suppose T is a Calderón–Zygmund singular integral operator and E is a measurable subset of $\mathbb{R}^n$. Let $f \in L^1$ and $\psi \in L^\infty \cap L^2$ be two functions such that $T^*\psi \in L^\infty$ and Tf is integrable over E. Then*

$$\int_E \psi T f = \int_{\mathbb{R}^n} (T^*\psi) f.$$

This statement will be deduced from the fact that the main results of §1 can be "localized": something similar to (2)–(3) can be said if we know only that Tf is integrable over some proper subset E of $\mathbb{R}^n$. Not quite explicitly, the possibility of such localization was mentioned in [KiX].

In §2 we pass to the main results of the paper and consider spaces X of smooth functions. Here the calculations are different from and more complicated than those in §1. However, there is a general pattern suitable for all spaces of smooth functions in question. We start §2 with description of this pattern in an abstract setting. Next, we show how this setting specializes in the cases of $X =$BMO, Lip_α, or $W_p^{(1)}$. This is prefaced by a description of the corresponding Calderón–Zygmund type decompositions. The section is finished with the case of $X = L^\infty$. Though a nontrivial singular integral operator does not take L^∞ to itself, our methods still give some information in this situation.

In §3 we briefly describe some interpolation consequences of the results of §2.

In many particular cases, even if we know a suitable Calderón–Zygmund type algorithm, the proof of the fact that it does give a function u satisfying (2) is rather bulky. Such a situation occurs already in §1 (see Theorem 1). Not to hide the essence of the proofs of our main results, we postpone the technicalities related to such statements till Appendix 1, especially as they can be presented in the framework of abstract interpolation theory.

All nonclassical Calderón–Zygmund type decompositions we need were described in the Thesis [Kr1], which exists in quite a restricted number of copies. The journal publication [Kr2] contains a discussion of Sobolev spaces but, for instance, the case of Lipschitz classes remains hardly accessible. So, in Appendix 2 we show how to construct the required decompositions for the couple $(L^1, \mathrm{Lip}_\alpha)$.

0.3. Preliminary information.

0.3.1. Spaces. When talking of BMO, Lip_α, and $W_p^{(1)}$, we mean *homogeneous* spaces, i.e., the spaces determined by the (semi)norms

$$\|f\|_{\mathrm{BMO}} = \sup_Q \frac{1}{|Q|}\int_Q |f - f_Q|,$$

$$\|f\|_{\mathrm{Lip}_\alpha} = \sup_{x\neq y}\frac{|f(x)-f(y)|}{|x-y|^\alpha},$$

$$\|f\|_{W_p^{(1)}} = \Big(\int\limits_{\mathbb{R}^n} |\nabla f|^p\Big)^{1/p}.$$

In particular, all constant functions have zero norms in these spaces. In the first line, Q stands for an arbitrary cube with edges parallel to coordinate axes, and f_Q is the average of f over Q: $f_Q = |Q|^{-1}\int\limits_Q f$. These agreements extend to the entire paper.

0.3.2. Singular integrals. We do not aim at writing out the least restrictive definition of a singular integral operator. However, for some reasons we want to avoid symmetry in the roles of the variables. So, we adopt the following definition.

An operator T is called a singular integral operator of Calderón–Zygmund type (abbriviated as SIO) if

(i) T is bounded on $L^r(\mathbb{R}^n)$ for all $r \in (1,\infty)$;

(ii) there exists a measurable function $K(x,y)$ defined for x, $y \in \mathbb{R}^n$, $x \neq y$, such that

$$Tf(x) = \int_{\mathbb{R}^n} K(x,y)f(y)dy$$

for a.e. x off the support of f;

(iii) $|K(x,y)| \leq C|x-y|^{-n}$ and $|\nabla_y K(x,y)| \leq C|x-y|^{-(n+1)}$ for $x \neq y$.

In particular, it is assumed that ∇_y applies to K, i.e., K is smooth in y for $y \neq x$. Usually, in (iii) also the symmetric estimate $|\nabla_x K(x,y)| \leq C|x-y|^{-(n+1)}$ is required. If we want this, we say that T^* is a SIO.

The above claim that (3) is satisfied for arbitrary SIO T is slightly deceiving: in the proofs we often need that T act in X, which is an additional restriction if X is a space of smooth functions. In principle, the reader may content himself with the statement that a *convolution type* SIO T (then $K(x,y) = k(x-y)$ for some finction k) always takes BMO, Lip_α, and $W_p^{(1)}$ to themselves. For the general (nonconvolution) operators, see [M] and [CM, Chapter 10] for appropriate criteria. For example, if T and T^* are SIO's, then T takes BMO to itself if and only if $T\mathbb{1} = 0$ (modulo constants).

In fact, we shall act in an "axiomatic" manner, i.e., we shall isolate conditions on T really needed in a particular proof. See [Ki, §3] for an attempt of a similar approach. The portion of properties of singular integrals which will suffice for our purposes in §1 can easily be formulated in the following terms.

Let $p \in (1,\infty)$. We denote by $\mathcal{S}_p$ the class of operators T such that

(A) T is continuous from $L^p(\mathbb{R}^n)$ to itself;

(B) T maps L^1 continuously to the space of measurable functions on $\mathbb{R}^n$ with convergence in measure;

(C) if Q is a cube and $g \in L^1$ is of mean value zero and is supported on Q, then

$$\int\limits_{\mathbb{R}^n\setminus(2Q)} |Tg| \leq C \int\limits_Q |g|$$

with C independent of g and Q.

(Throughout, αQ means the cube with the same center as Q but of diameter α times as much as that of Q.)

It is standard that any SIO belongs to $\mathcal{S}_p$ for every p (see, e.g., [M]). Observe, however, that if $T \in \mathcal{S}_p$, $h \in L^\infty(\mathbb{R}^n)$, and the operator T_1 is defined by $T_1 f = hTf$, then, trivially, $T_1 \in \mathcal{S}_p$. In particular, the characteristic function of any measurable set E can be taken for h.

In §2 more restrictions on T will be required. We shall return to this later.

§1. The L^p-spaces

1.1. Calderón–Zygmund algorithm for the couple (L^1, L^∞). Recall that by a "cube" we always mean a cube with sides parallel to coordinate axes. In the classical paper [CZ], for every nonzero function $f \in L^1$ and every $t > 0$ a family of cubes Q_i with disjoint interiors and with the following properties was constructed:

(i) $t < |Q_i|^{-1} \int_{Q_i} |f| \leq 2^n t$;

(ii) $|f| \leq t$ a.e. off $\cup Q_i$.

Since then, this construction was reproduced in many handbooks, see, e.g., [S, GR].

With the family $\{Q_i\}$, we associate an averaging operator $P_t = P_t^f$ as follows:

$$P_t^f h = P_t h = \sum_i \left(\frac{1}{|Q_i|} \int\limits_{Q_i} h \right) \chi_{Q_i} + h\chi_{\mathbb{R}^n \setminus \cup Q_i}. \tag{6}$$

Quite often, we omit the superscript f in the notation.

Next, put $f_t = P_t^f f$. Formula (6) shows that

$$f_t = \sum c_i \chi_{Q_i} + f\chi_{\mathbb{R}^n \setminus (\cup Q_i)},$$

where c_i is a unique solution of the equation $\int\limits_{Q_i} (f - c_i) = 0$. Then $f - f_t = \sum\limits_i (f - c_i)\chi_{Q_i}$. The ith summand on the right in the last formula is supported on Q_i and has zero average. The following simple observation is very important.

Lemma 1. *We have*

$$\|f - f_t\|_{L^1} \leq 2^{n+2} \operatorname{dist}_{L^1} \left(f, B_{L^\infty}(\frac{t}{2})\right).$$

Since f_t itself belongs to $B_{L^\infty}(2^n t)$, we see that the Calderón–Zygmund procedure really gives an element of almost optimal approximation in the sense of formula (1) *with $p = 1$ and $X = L^\infty$.*

Proof of Lemma 1. First,

$$\|f - f_t\|_{L^1} \leq \sum \|(f - c_i)\chi_{Q_i}\| \leq 2 \sum \int\limits_{Q_i} |f| \leq 2^{n+1} \sum |Q_i|$$

by (i). Next, if $g \in B_{L^\infty}(t/2)$, then

$$\|f - g\|_{L^1} \geq \int\limits_{\cup Q_i} |f - g| \geq \int\limits_{\cup Q_i} |f| - \int\limits_{\cup Q_i} |g| \geq t \sum |Q_i| - \frac{t}{2} \sum |Q_i| = \frac{t}{2} \sum |Q_i|.$$

Taking the infimum over g, by the preceding estimate we obtain

$$\|f - f_t\|_{L^1} \leq 2^{n+2} \left(\frac{t}{2} \sum |Q_i|\right) \leq 2^{n+2} \operatorname{dist}_{L^1}(f, B_{L^\infty}(\frac{t}{2})). \qquad \square$$

1.2. Almost optimal approximation for the couple (L^1, L^p), $1 < p < \infty$. It turns out that, after a parameter change, the *same* classical Calderón–Zygmund procedure makes it possible to explicitly exhibit an element of almost optimal L^1-approximation of a function $f \in L^1$ by the ball of a given radius in L^p.

Theorem 1. *If $f \in L^1$ and $f \notin B_{L^p}(s)$, then there exists $t = t(s)$ such that*

$$\|f_{t(s)}\|_{L^p} \leq Cs \quad \text{and} \quad \|f - f_{t(s)}\|_{L^1} \leq C \operatorname{dist}_{L^1}(f, B_{L^p}(s)),$$

where $C > 0$ is independent of f and s. Moreover, $t(s)$ can be calculated as follows: put $\lambda(s) = (s^{-1} \operatorname{dist}_{L^1}(f, B_{L^p}(s)))^q$ $(q^{-1} + p^{-1} = 1)$, then $t(s) = 2\lambda(s)^{-1} \int_0^{\lambda(s)} f^*(\tau)d\tau$, where f^* is the nonincreasing rearrangement of $|f|$. With this choice of $\lambda(s)$ and $t(s)$, we have

$$\sum |Q_i| \leq C\lambda(s) \tag{7}$$

for the Calderón–Zygmund cubes invoved in the definition of $f_{t(s)}$.

Remark. The construction is flexible in the sense that, in place of the precise identity, we may take $\lambda(s) \asymp (s^{-1}\operatorname{dist}_{L^1}(f, B_{L^p}(s))^q$.

Though Theorem 1 involves a minimum of "raw material", it seems to have passed unnotices up to now. The proof of it is surprisingly bulky, so it will be postponed till Subsection 3.1 (in fact, even till Appendix 1). Note that Lemma 1 is a crucial element of that proof.

Now we take Theorem 1 for granted and pass to the first approximation results that involve singular integrals.

1.3. Up to the end of this section, $p \in (1, \infty)$ and T is an operator of class $\mathcal{S}_p$ (see Subsection 0.3.2).

Theorem 2. *Let $f \in L^1$, let $s > 0$, and suppose that $Tf \in L^1 + L^p$. Then there exists a function $g \in B_{L^p}(Cs)$ such that*

$$\|f - g\|_{L^1} \leq C \operatorname{dist}_{L^1}(f, B_{L^p}(s))$$

and

$$\|Tf - Tg\|_{L^1} \leq C(\operatorname{dist}_{L^1}(f, B_{L^p}(s)) + \operatorname{dist}_{L^1}(Tf, B_{L^p}(s))).$$

Specifically, we may take $g = f$ if $f \in B_{L^p}(s)$ and $g = f_{t(s)}$ otherwise, where $t(s)$ is defined in Theorem 1.

Proof. The calculations below are motivated by an argument in Bourgain's paper [Bo]. After the proof, we shall comment on a point of difference. We may assume that $f \notin B_{L^p}(s)$. Then denote $g = f_{t(s)}$ and pick a function $u \in B_{L^p}(s)$ satisfying

$$\|Tf - u\|_{L^1} \leq 2 \operatorname{dist}_{L^1}(Tf, B_{L^p}(s)). \tag{8}$$

We have

$$\|Tf - Tg\|_{L^1} \leq \int\limits_{\mathbb{R}^n \setminus (\cup 2Q_i)} |Tf - Tg| + \int\limits_{\cup(2Q_i)} |Tf - u| + \int\limits_{\cup(2Q_i)} |Tg - u|,$$

where $\{Q_i\}$ is the colection of Calderón–Zygmund cubes arising when we decompose f with the parameter $t = t(s)$. The middle term on the right is controlled with the help of (8). Since $f - g = f - f_{t(s)} = \sum (f - c_i)\chi_{Q_i}$ and the summands are of mean value zero, by property (C) in Subsection 0.3.2 the first term is estimated as follows:

$$\int\limits_{\mathbb{R}^n \setminus (\cup 2Q_i)} |Tf - Tg| \leq \sum_i \int\limits_{\mathbb{R}^n \setminus (2Q_i)} |T((f - c_i)\chi_{Q_i})| \leq C \sum_i \int\limits_{Q_i} |f - c_i|$$
$$\leq C\|f - g\|_{L^1} \leq C' \operatorname{dist}_{L^1}(f, B_{L^p}(s))$$

because g is equal to c_i on Q_i and by Theorem 1. The third summand is estimated with the help of the Hölder inequality:

$$\int\limits_{\cup(2Q_i)} |Tg - u| \leq |\cup (2Q_i)|^{1/q}(\|Tg\|_{L^p} + \|u\|_{L^p}) \leq C(\sum |Q_i|)^{1/q} s$$

by the choice of u and the estimate $\|f_{t(s)}\|_{L^p} \leq Cs$ in Theorem 1: $\|Tg\|_{L^p} = \|Tf_{t(s)}\|_{L^p} \leq C\|f_{t(s)}\|_{L^p}$. By (7), we see that

$$\int\limits_{\cup(2Q_i)} |Tg - u| \leq C'\lambda(s)^{1/q}s = C' \operatorname{dist}_{L^1}(f, B_{L^p}(s)).$$

It remains to collect the estimates. $\square$

Remark. When adapted to our situation, Bourgain's argument from [Bo] differs in the choice of g: it dictates to put $g = h + P_t^{(f-h)}(f-h)$, where h is an arbitrary function satisfying

$$h \in B_{L^q}(s), \quad \|f-h\|_{L^1} \leq 2 \operatorname{dist}_{L^1}(f, B_{L^p}(s))$$

(as before, the latter distance is assumed to be positive). The parameter t is determined by the equation $t^{1/q}\|f-h\|_{L^1}^{1/p} = s$. Slight changes in the above calculations are required to show that this choice of g also does the job (see [Bo] for the pattern).

This way allows one to bypass the bulky Theorem 1, but, psychologically, it has a disadvantage, namely, an "unknown" fucntion $f-h$ rather than the "known" function f undergoes Calderón–Zygmund decomposition.

1.4. Localization. We formulate an immediate consequence of Theorem 1.

Corollary 1. *Let $1 < p < \infty$, and let $T \in \mathcal{S}_p$. If $f \in L^1$, $s > 0$, and a measurable set $E \subset \mathbb{R}^n$ is such that $(Tf)\chi_E \in L^1 + L^p$, then there exists a function $g \in B_{L^p(\mathbb{R}^n)}(cs)$ with*

$$\|f-g\|_{L^1} \leq C \operatorname{dist}_{L^1}(f, B_{L^p}(s))$$

and

$$\int\limits_E |Tf - Tg| \leq C(\operatorname{dist}_{L^1}(f, B_{L^p}(s)) + \operatorname{dist}_{L^1(E)}(\chi_E Tf, B_{L^p(E)}(s))).$$

Specifically, if $f \notin B_{L^p}(s)$, we may take $g = f_{t(s)}$ with the same choice of $t(s)$ as in Theorems 1 and 2; otherwise, we may take $g = f$.

Proof. Apply Theorem 1 to the operator $T_1 : h \mapsto \chi_E Th$ in place of T. We have already mentioned that T_1 is in $\mathcal{S}_p$ along with T. $\square$

Corollary 2. *If $T \in \mathcal{S}_p$, $f \in L^1$, and Tf is integrable over a set $E \subset \mathbb{R}^n$, then there exists a sequence $f_n \in L^1 \cap L^p$ such that $f_n \to f$ in L^1, all functions Tf_n are integrable over E, and $\int\limits_E |Tf_n - Tf| \to 0$.*

Proof. Combine Corollary 1 with the fact that f belongs to the closure of $L^1 \cap L^p$ and $\chi_E Tf$ to the closure of $L^1(E) \cap L^p(E)$. $\square$

Apparently, Corollary 2 is nontrivial even under the assumption that $E = \mathbb{R}^n$ (and $Tf \in L^p$), unless T commutes with translations, in which case convolution of f with some approximate identity gives the result.

The following statement will be useful in the next section.

Lemma 2. *Suppose $T \in \mathcal{S}_p$ and E is a measurable subset of $\mathbb{R}^n$. Let $\psi \in L^\infty \cap L^q$ $(q^{-1} + p^{-1} = 1)$ be a function supported on E and such that $T^*\psi \in L^\infty$. Then*

$$\int_E \psi T f = \int_{\mathbb{R}^n} (T^*\psi) f.$$

Proof. Taking a sequence f_n as in Corollary 2, we can write

$$\int_E \psi T f = \lim_n \int_{\mathbb{R}^n} \psi T f_n = \lim_n \int_{\mathbb{R}^n} (T^*\psi) f_n = \int_{\mathbb{R}^n} (T^*\psi) f.$$

Passage to the adjoint in the second identity is justified because T is bounded on L^p, $\psi \in L^q$, and $f_n \in L^p$. $\square$

§2. Spaces of smooth functions

2.1. We pass to the main subject of the paper, i.e., to couples (L^1, X), where X is a space of functions "smooth" in a sense. Calculations will not be as easy as in the proof of Theorem 2. However, in all cases they fit into an abstract pattern, which we are going to describe.

In all specific examples, there will be two natural spaces that contain $L^1 + X$: the space of measurable functions (with convergence in measure) and the space of locally integrable functions. We need both. In the abstract setting this will be formalized as follows.

Let (X_0, X_1) be a couple of spaces included in some linear topological space $\mathcal{E}$ (in applications, L^1 will play the role of X_0, and $\mathcal{E}$ will be the space of measurable functions). Moreover, we suppose that $X_0 + X_1$ is included in some linear subspace $\mathcal{F}$ of $\mathcal{E}$, and for every $x \in X_0$ and every $t > 0$ some linear operator $P_t = P_t^x : \mathcal{F} \to \mathcal{F}$ is defined (in specific examples, P_t will be similar to but more complicated than the operator given by (6); it will always be possible to take the space of locally integrable functions for $\mathcal{F}$). Next, the space X_0 is assumed to be Banach, but X_1 may only be endowed with a complete seminorm (as the spaces mentioned in Subsection 0.3.1). We subject the operators P_t^x to the following conditions:

(a) $\|P_t^x x\|_{X_1} \le Ct$, $\|(I - P_t^x)x\|_{X_0} \le C \operatorname{dist}_{X_0}(x, B_{X_1}(t))$;
(b) $\|P_t^x\|_{X_0 \to X_0} \le C$, $\|I - P_t^x\|_{X_1 \to X_0} \le C \frac{\operatorname{dist}_{X_0}(x, B_{X_1}(t))}{t}$.

Note that (b) implies that $I - P_t^x$ vanishes on the kernel of the seminorm on X_1.

Moreover, suppose T is a linear operator that maps X_0 to $\mathcal{E}$ continuously and takes $X_0 \cap X_1$ to $\mathcal{F}$ (a singular integral operator in applications). We impose yet another condition, involving both T and P_t:

(c) if $Tx \in X_0 + X_1$, then
 (c1) $\|P_t^x T(I - P_t^x)x\|_{X_0} \le C \operatorname{dist}_{X_0}(x, B_{X_1}(t))$,
 (c2) $\|(I - P_t^x) T P_t^x x\|_{X_0} \le C \operatorname{dist}_{X_0}(x, B_{X_1}(t))$.

Note that $P_t^x x \in X_0 \cap X_1$ by (a) and (b); since $X_0 \cap X_1$ is taken by T to $\mathcal{F}$, we see that the expressions under the norm sign in (c1) and (c2) make sense. Next, if T is a bounded operator from X_1 to itself (which will often be the case), then (c2) is a consequence of the first inequality in (a) and the second inequality in (b). Finaly, if $x \in B_{X_1}(t)$, we can always put $P_t^x = I$, so that in specific examples we should only construct the operators P_t^x for $x \notin B_{X_1}(t)$.

Theorem 3. *Under the above assumptions, if $x \in X_0$ and $Tx \in X_0 + X_1$, then, putting $x_t = P_t^x x$, we have*

$$\begin{aligned}\|Tx - Tx_t\|_{X_0} &\le C(\|x - x_t\|_{X_0} + \operatorname{dist}_{X_0}(Tx, B_{X_1}(t)) \\ &\le C'(\operatorname{dist}_{X_0}(x, B_{X_1}(ct)) + \operatorname{dist}_{X_0}(Tx, B_{X_1}(t))).\end{aligned}$$

Remark. The above inequality should be considered in combination with the first inequality in (a).

Proof. Take $u \in B_{X_1}(t)$ with $\|Tx - u\|_{X_0} \le 2 \operatorname{dist}_{X_0}(Tx, B_{X_1}(t))$ and write

$$\begin{aligned}&\|Tx - Tx_t\|_{X_0} \le \|P_t^x T(x - x_t)\|_{X_0} + \|(I - P_t^x)T(x - x_t)\|_{X_0} \\ &\le \|P_t^x T(I - P_t^x)x\|_{X_0} + \|(I - P_t^x)(Tx - u)\|_{X_0} \\ &+ \|(I - P_t^x)u\|_{X_0} + \|(I - P_t^x)TP_t^x x\|_{X_0}.\end{aligned}$$

The first and fourth summands in this expression are controlled with the help of (c). Next, $\|(I - P_t^x)(Tx - u)\|_{X_0} \le C'\|Tx - u\|_{X_0}$ by the first inequality in (b), and $\|(I - P_t^x)u\|_{X_0} \le \|u\|_{X_1}\|I - P_t^x\|_{X_1 \to X_0} \le C \operatorname{dist}_{X_0}(x, B_{X_1}(t))$ by the second inequality in (b). $\square$

2.2. The couples $(L^1$, BMO$)$ and $(L^1, \mathrm{Lip}_\alpha)$. Our immediate goal is to construct "averaging" (or, rather, "smoothing") operators $P_t = P_t^x$ satisfying (a)–(c) for these couples. To do this, we need Calderón–Zygmund type decompositions described in the Thesis [Kr1]. As has already been mentioned, we reproduce the constructions for Lip_α in Appendix 2 in order to make them more accessible. Now we only state the results.

Let X be either BMO $(\mathbb{R}^n)$ or $\mathrm{Lip}_\alpha(\mathbb{R}^n)$, $0 < \alpha < 1$. The formulas will be quite similar in both cases, and we present them simultaneously. The difference is in a certain parameter σ, which is equal to 0 if $X =$ BMO and to α/n if $X = \mathrm{Lip}_\alpha$. By a *packing* we mean any collection of cubes with disjoint interiors.

Suppose $f \in L^1$, $t > 0$, and $f \notin B_X(t)$. The Calderón–Zygmund type procedure we are talking about yields a collection $\{Q_i\}$ of cubes and a collection $\{\psi_i\}$ of nonnegative C^∞-functions with $\operatorname{supp} \psi_i \subset \frac{1}{2}Q_i$ that satisfy the conditions listed below.

Cubes. The cubes Q_i cover $\mathbb{R}^n$ in the case of Lip_α and do not cover $\mathbb{R}^n$ in the case of BMO; however, $\cup(\frac{1}{2}Q_i) = \cup Q_i$ in any case. Next,

(I) the collection $\{Q_i\}$ splits into at most B_n packings, where B_n depends only on the dimension n;

(II) the cubes Q_i are *linked*: if $Q_i \cap Q_j \ne \varnothing$, then $|Q_i \cap Q_j| \ge \varepsilon(n) \max(|Q_i|, |Q_j|)$, where $\varepsilon(n)$ depends only on n. Combining (I) and (II) we obtain

(III) $\operatorname{card}\{i : Q_i \cap Q_j \ne \varnothing\} \le \gamma(n)$ for all j, where $\gamma(n)$ depends on the dimension n only. Indeed, by (II), in every packing formed by Q_i's there may be at most $\varepsilon(n)^{-1}$ cubes that differ from Q_j and intersect Q_j.

Associated functions. The collection $\{\psi_i\}$ is a partition of unity in the sense that

(CZ1) $\sum \psi_j = 1$ on $\cup Q_i$.

The functions satisfy the size conditions

(CZ2) $\operatorname{supp} \psi_i \subset \frac{1}{2}Q_i$, $|\nabla \psi_i| \le C|Q_i|^{-1/n}$, $\int \psi_i \ge c|Q_i|$.

We put $\psi_\infty = \chi_{\mathbb{R}^n \setminus (\cup Q_i)}$ and define a "smoothing" operator $P_t = P_t^f$ in the following way (compare with formula (6)):

$$P_t^f h = \sum \left(\frac{1}{\int \psi_i} \int h\psi_i \right) \psi_i + h\psi_\infty.$$

This definition makes sense for all locally integrable functions h. It turns out (see Appendix 2 for the case of Lipschitz classes) that

(CZ3) $\|P_t^f f\|_X \le ct$;

(CZ4) $\sum |Q_i|^{1+\sigma} \le \frac{c}{t} \operatorname{dist}_{L^1}(f, B_X(t))$.

Next, put $c_i = \frac{1}{\psi_i} \int f\psi_i$ (so that $P_t^f f = \sum c_i\psi_i + f\psi_\infty$). Then (see Appendix 2 for the Lipschitz classes)

(CZ5) $\frac{1}{|Q_i|^{1+\sigma}} \int_{Q_i} |f - c_i| \le Ct$.

From (CZ4) and (CZ5) we easily deduce the following property:

(CZ6) $\|f - P_t^f f\|_{L^1} \le C \operatorname{dist}_{L^1}(f, B_X(t))$.

Indeed,

$$\|f - P_t^f f\| \le \sum \int_{Q_i} |f - c_i| \le Ct \sum |Q_i|^{1+\sigma} \le C' \operatorname{dist}_{L^1}(f, B_X(t)). \tag{9}$$

We remind the reader that the condition $f \notin B_X(t)$ was imposed to avoid degeneration.

We are going to check properties (a)–(c) in Subsection 2.1 for the operators P_t^f and a singular integral operator T satisfying some additional assumptions. Again, we can forget in fact about the origin of T and act in an "axiomatic" manner. We suppose that $T \in \mathcal{S}_p$ for some $p \in (1, \infty)$ (see Subsection 0.3.2), that T maps boundedly X to itself, and that

$$|T^*\psi| \le C = C(A, T, n) \tag{10}$$

whenever ψ is a C^∞-function supported on some cube Q and satisfying $|\psi| \le A$, $|\nabla\psi| \le A|Q|^{-1/n}$.

To be specific, we note at once that (10) *is fulfilled if* T^* *is a* SIO *with* $T^*1 = 0$. (The reader is referred to [M, CM] for the meaning of the condition $T^*1 = 0$ and for its consequences.) This can be proved by a direct calculation involving the kernel of T, but we prefer to deduce this formally from the following lemma.

Lemma 3. *If* $\alpha \in (0, 1)$, *then*

$$\|g\|_{L^\infty} \le C_{n,\alpha} \|g\|_{L^{1,\infty}}^{\frac{\alpha}{\alpha+n}} \|g\|_{\mathrm{Lip}\,\alpha}^{\frac{n}{\alpha+n}}$$

for every function g *such that the right-hand side makes sense.*

Proof. Denote $a = \|g\|_{L^{1,\infty}}$, $b = \|g\|_{\mathrm{Lip}\,\alpha}$. Let $\tau > 0$, and suppose that $|g(x)| > \tau$ for some x. Since $|g(y)| \ge |g(x)| - b|x - y|^\alpha$ for every y, we see that the set $e = \{y : |g(y)| > \tau/2\}$ contains the ball

$$B = \{y : b|x - y|^\alpha \le \tau/2\} = \{y : |x - y| \le 2^{-1/\alpha}(\tau/b)^{1/\alpha}\}.$$

Consequently, $|e| \geq |B| \geq C(\tau/b)^{n\alpha}$. Since $|e| \leq (2/\tau)a$, it follows that $a/\tau \geq C(\tau/b)^{n/\alpha}$, whence $\tau \leq (ab^{n/\alpha})^{\frac{1}{1+n/\alpha}} = a^{\frac{\alpha}{\alpha+n}} b^{\frac{n}{\alpha+n}}$. $\square$

Now, suppose S is an operator of weak type $(1,1)$ taking Lip_α to itself boundedly for some $\alpha \in (0,1)$, and let ψ be a function satisfying $|\psi| \leq A$, $|\nabla\psi| \leq A|Q|^{-1/n}$ (as in (10)). Clearly, $|\psi(x) - \psi(y)| \leq A|Q|^{1/n}|x-y| \leq A'|Q|^{-1/n}|Q|^{\frac{1-\alpha}{n}}|x-y|^\alpha$, whence $\|\psi\|_{\mathrm{Lip}_\alpha} \leq A'|Q|^{-\frac{\alpha}{n}}$. Consequently, for $g = S\psi$ we have $\|g\|_{\mathrm{Lip}\,\alpha} \leq C\|\psi\|_{\mathrm{Lip}\,\alpha} \leq C_1|Q|^{-\alpha/n}$ and $\|g\|_{L^{1,\infty}} \leq C\|\psi\|_{L^1} \leq C_2|Q|$. Application of Lemma 3 shows that $\|\psi\|_{L^\infty} \leq C(A,S,n)$. It remains to recall that if T^* is a SIO as defined in Subsection 0.3.2 and $T^*1 = 0$, then T^* is bounded on Lip_α and (surely) is of weak type $(1,1)$. $\square$

Theorem 4. *Let X be either* BMO *or* Lip_α $(0 < \alpha < 1)$*, and let $T \in \mathcal{S}_p$ for some $p \in (1,\infty)$. Suppose T maps X to itself boundedly and satisfies* (10)*. Then for every $t > 0$ there exists a function $g \in B_X(ct)$ such that $\|f-g\|_{L^1} \leq C\,\mathrm{dist}_{L^1}(f, B_X(t))$ and*

$$\|Tf - Tg\|_{L^1} \leq C(\mathrm{dist}_{L^1}(f, B_X(t)) + \mathrm{dist}_{L^1}(Tf, B_X(t))).$$

Moreover, we can take $g = f$ if $f \in B_X(t)$ and $g = P_t^f f$ otherwise. (Note also that $Tg \in B_X(ct)$ in our case.)

Proof. The case where $f \in B_X(t)$ is trivial. In the opposite case, by Theorem 3, it suffices to verify conditions (a)–(c) in Subsection 2.1 for the operators T and P_t^f.

We start with verification of (c1), which is the most difficult part of the proof. For further needs, we formulate a lemma.

Lemma 4. *We have*

$$\|P_t^f T(I - P_t^f)f\|_{L^1} \leq c\sum_i \int_{Q_i} |f - c_i|,$$

where $c_i = \frac{1}{\psi_i}\int f\psi_i$. Moreover, only (I), (III) *and* (CZ1) *are used in the proof of this inequality.*

In the lemma it is also assumed implicitly that the function $T(I - P_t^f)f$ is locally integrable. This is true indeed under the conditions of Theorem 4. Note also that $T(f - P_t^f f)$ can be treated as a genuine function rather than a class modulo constants because $f - P_t^f f \in L^1$. If the lemma is proved, (c1) follows from an inequality in the chain (9).

Proof of Lemma 4. For short, we write P_t instead of P_t^f. We have

$$P_t(T(f - P_t f)) = \sum_j \left(\frac{1}{\int \psi_j}\int \psi_j T(f - P_t f)\right)\psi_j + \psi_\infty T(f - P_t f).$$

The second term on the right is treated easily:

$$\int_{\mathbb{R}^n \setminus \cup Q_k} |T(f - P_t f)| \leq \sum_k \int_{\mathbb{R}^n \setminus Q_k} |T((f - c_k)\psi_k)| \leq C\sum \int_{Q_k} |f - c_k|$$

by the first relation in (CZ2) and property (C) in Subsection 0.3.2. To estimate the remaining sum, we first consider the jth summand separately. In this summand, we split $f - f_t = f - P_t f$ as follows:

$$f - f_t = \sum_{i: Q_i \cap Q_j \neq \varnothing} (f - c_i)\psi_i + \sum_{i: Q_i \cap Q_j = \varnothing} (f - c_i)\psi_i \stackrel{def}{=} u_j + v_j.$$

The sum defining u_j is finite by (III), and the series defining v_i converges to v_j in L^1; hence, $Tv_j = \sum\limits_{i: Q_i \cap Q_j = \varnothing} T((f - c_i)\psi_i)$, where the series converges in measure by property (B) in Subsection 0.3.2.

Consider the measurable function

$$V_j = \sum_{i: Q_i \cap Q_j} |T((f - c_i)\psi_i)|$$

(probably, infinite somewhere). We have

$$\sum_j \int_{Q_j} V_j = \sum_i \sum_{j: Q_i \cap Q_j = \varnothing} \int_{Q_j} |T((f - c_i)\psi_i)|$$

$$\leq C \sum_i \int_{\mathbb{R}^n \setminus Q_i} |T((f - c_i)\psi_i)|$$

by the finite multiplicity of the covering (see (I)). Applying (C) in Subsection 0.3.2 once again, we arrive at

$$\sum_j \int_{Q_j} V_j \leq C \sum_i \int_{Q_j} |f - c_i|.$$

Since $|Tv_j| \leq V_j$, it follows that Tv_j is integrable over Q_j; moreover, (12) implies the inequality

$$\sum_j \int_{Q_j} |Tv_j| \leq C \sum_i \int_{Q_i} |f - c_i|. \tag{12}$$

Next, $Tu_j = T(f - P_t f) - Tv_j$ is also integrable over Q_j (we have already mentioned that $T(f - f_t)$ is locally integrable). By (10), we can apply Lemma 2 to obtain

$$\int_{Q_j} \psi_j T u_j = \int (T^* \psi_j) u_j = \sum_{i: Q_i \cap Q_j \neq 0} \int (T^* \psi_j)(f - c_i)\psi_i,$$

whence, again by (10),

$$\sum_j |\int_{Q_j} \psi_j T u_j| \leq C \sum_i \sum_{j: Q_i \cap Q_j \neq \varnothing} \int_{Q_i} |f - c_i| \leq c\gamma(n) \sum_i \int_{Q_i} |f - c_i| \tag{13}$$

by (III).

Now, we write

$$|P_tT(f-P_tf)| = |\sum_j \left[\frac{1}{\int \psi_j}\int (T(u_j+v_j))\psi_j\right]\psi_j|$$

$$\leq \sum_j \frac{\psi_j}{\int \psi_j}(|\int_{Q_j} \psi_j T u_j| + \int_{Q_j} |Tv_j|)$$

Integrating this over $\mathbb{R}^n$ and taking (12) and (13) into account, we prove Lemma 4. $\square$

Verification of the other conditions among (a)–(c) in Subsection 2.1 is more straightforward. Condition (a) follows from (CZ3) and (CZ6). The first inequality in (b) is easy:

$$\|P^f h\|_{L^1} \leq \|\sum \left(\frac{1}{\int \psi_i}\int h\psi_i\right)\psi_i\|_{L^1} + \|h\psi_\infty\|_{L^1}$$

$$\leq \sum \int_{Q_i} |h| + \int_{\mathbb{R}^n\setminus(\cup Q_i)} |h| \leq B_n\|h\|_{L^1}$$

by (I).

To prove the second inequality in (b), we observe that for every function $h\in X$ and every cube Q we have

$$\int_Q |h-h_Q| \leq C\|h\|_X|Q|^{1+\sigma},$$

where h_Q is the average of h over Q. This is a direct consequence of the definitions of the norms in BMO and Lip_α. Also, $I-P_t^f$ is zero on the constant functions, so, applied to a function in X, it gives a genuine function (rather than an equivalence class modolo constants). Now for $h\in X$ we can write

$$\|(I-P_t^f)h\|_{L^1} = \int_{\cup Q_i} |h(x) - \sum_i \frac{1}{\int\psi_i}\int h(y)\psi_i(y)dy\psi_i(x)|dx$$

$$\leq \sum_i \frac{1}{\int\psi_i}\iint |h(x)-h(y)|\psi_i(x)\psi_i(y)dxdy$$

$$\leq \sum_i \frac{1}{\int\psi_i}\iint |h(x)-h_{Q_i}|\psi_i(x)\psi_i(y)dxdy+$$

$$+\sum_i \frac{1}{\int\psi_i}\iint |h(y)-h_{Q_i}|\psi_i(x)\psi_i(y)dxdy$$

$$\leq C\|h\|_X\sum |Q_i|^{1+\sigma} \leq \frac{C'}{t}\,\mathrm{dist}_{L^1}(f, B_X(t))$$

by (CZ4).

Finally, (c2) is a consequence of (b) because T is a bounded operator from X to itself in our case. With this, Theorem 4 is proved. $\square$

2.3. Sobolev spaces. We pass to the couples $(L^1, W_p^{(1)})$, $1 < p < \infty$. Not too much additional work should be done in this case.

It is known (see [Br]) that an equivalent norm in $W_p^{(1)}$ is given by the formula

$$\|f\|_{1,p} = \sup_{\pi=\{Q_i\}} \left(\sum_i \frac{\left(\int_{Q_i} |f - f_{Q_i}| \right)^p}{|Q_i|^\lambda} \right)^{1/p}, \quad 1 < p < \infty,$$

where $\lambda = 1 - \frac{1}{p} + \frac{1}{n}$ and the supremum is taken over arbitrary packings π of cubes. This formula can be rewritten as follows:

$$\|f\|_{1,p} = \sup_{\pi=\{Q_i\}} \left(\sum_i \left(\frac{\int_{Q_i} |f - f_{Q_i}|}{|Q_i|^{1+\sigma}} \right)^p |Q_i|^{1+\sigma} \right)^{1/p} \tag{14}$$

with $\sigma = q/n$, $q^{-1} + p^{-1} = 1$. (The parameter σ will play the same role it played in the preceding subsection.)

We use the Calderón–Zygmund type algorithm for the couple $(L^1, W_p^{(1)})$ that was elaborated in [Kr1, Kr2]. It will result in the following.[1] Let $f \in L^1$ with $F \notin B_{W_p^{(1)}}(t)$. Then there exists a collection $\{Q_i\}$ of cubes satisfying $\cup Q_i = \mathbb{R}^n$ and a collection ψ_j of C^∞-functions such that properties (I)–(III), (CZ1)–(CZ3), and (CZ6) are fulfilled with $X = W_p^{(1)}$ and with P_t^f defined as before: $P_t^f h = \sum \left(\frac{1}{\psi_j} \int h\psi_j \right) \psi_j$. Instead of (CZ4) and (CZ5) we have

(CZ4a) $\quad \left(\sum_i |Q_i|^{1+\sigma} \right)^{1/q} \le \frac{c}{t} \operatorname{dist}_{L^1}(f, B_{W_p^{(1)}}(t)),$

(CZ5a) $\quad \frac{1}{|Q_i|^{1+\sigma}} \int_{Q_i} |f - c_i| \le C \frac{t^q}{\operatorname{dist}_{L^1}(f, B_{W_p^{(1)}}(t))^{q-1}}.$

Theorem 5. *Suppose T belongs to $\mathcal{S}_r$ for some $r \in (1, \infty)$, maps W_p^1 boundedly to itself, and satisfies (10). If $f \in L^1$, $Tf \in L^1 + W_p^{(1)}$, and $f \notin B_{W_p^{(1)}}(t)$, then the function $g = P_t^f f$ satisfies*

$$\|Tf - Tg\|_{L^1} \le C(\|f - g\|_{L^1} + \operatorname{dist}_{L^1}(Tf, B_{W_p^{(1)}}(t))).$$

Proof. Again, we check conditions (a), (b), and (c) in Subsection 2.1 for the operators P_t^f. Only (c1) and the second inequality in (b) need a proof. We start with the latter. The same calculation as in the case of BMO or Lip_α yields

$$\|(I - P_t^f)h\|_{L^1} \le 2 \sum_i \frac{1}{\int \psi_i} \iint |h(x) - h_{Q_i}| \psi_i(x)\psi_i(y) dx dy \le$$

$$\le C \sum_i \int_{Q_i} |h - h_{Q_i}|.$$

[1] In fact, in [Kr1–Kr2] similar estimates in terms of the K-functional were presented. However, the K-functional is in an intimate relationship with the distance to the ball of radius t (see, e.g., the proof of Theorem A in Appendix 1), whence the required inequalities follow.

Suppose J is a subset of the index set such that the collection $\{Q_i\}_{i\in J}$ is a packing. Then

$$\sum_{i\in J}\int_{Q_i}|h-h_{Q_i}| = \sum_{i\in J}\frac{\int_{Q_i}|h-h_{Q_i}|}{|Q_i|^{1+\sigma}}|Q_i|^{1+\sigma} \le$$

$$\le \left(\sum_{i\in J}\left(\frac{\int_{Q_i}|h-h_{Q_i}|}{|Q_i|^{1+\sigma}}\right)^p|Q_i|^{1+\sigma}\right)^{1/p}\left(\sum|Q_i|^{1+\sigma}\right)^{1/q} \le$$

$$\le C\|h\|_{1,p}t^{-1}\operatorname{dist}_{L^1}(f,B_{W_p^{(1)}}(t))$$

by (14) and (CZ4a). Since the collection $\{Q_i\}$ splits in B_n packings, we are done.

To prove (c1), we observe that Lemma 4 is applicable and use (CZ5a) and (CZ4a) to continue the estimate in that lemma as follows:

$$\sum_i\int_{Q_i}|f-c_i| \le c\big(\sum_i|Q_i|^{1+\sigma}\big)t^q(\operatorname{dist}_{L^1}(f,B_{W_p^{(1)}}(t)))^{1-q} \le$$

$$\le C'\operatorname{dist}_{L_1}(f,B_{W_p^{(1)}}(t)). \quad \square$$

2.4. Some remarks on the couple (L^1,L^∞). In this subsection we assume that both T and T^* are SIO's, as defined in Subsection 0.3.2. Then our direct methods give something even for the couple (L^1,L^∞), though this is not quite satisfactory.

Proposition 1. *If $f\in L^1$, $Tf\in L^1+L^\infty$, and $t>0$, then there exists a function $g\in B_{L^\infty}(t)$ with $\|f-g\|_{L^1}\le C\operatorname{dist}_{L^1}(f,B_{L^\infty}(t))$ and*

$$\|Tf-Tg\|_{L^1}\le C(\operatorname{dist}_{L^1}(f,B_{L^\infty}(t))+\operatorname{dist}_{L^1}(Tf,B_{L^\infty}(t))).$$

Moreover, if $g\notin B_{L^\infty}(t)$, then g can be obtained from f with the help of a Calderón–Zygmund type procedure.

The disadvantage of this result is that, in general, Tg falls out of the space L^∞; we can only assert that it is in $B_{\mathrm{BMO}}(ct)$. It should be noted that an indirect argument involving the Hahn–Banach theorem yields some related results with Tg bounded, see [Bo, KiX, Ki].

To prove Proposition 1, we use a *nonclassical* Calderón–Zygmund type procedure. The collection $\{Q_i\}$ of cubes we need can be obtained either by application the same covering theorem that will be used in Appendix 2 to all *limitiny* cubes satisfying $|Q|^{-1}\int_Q|f|>t$, or simply by taking the Whitney cubes for the open set $\{Mf>t\}$ (Mf is the Hardy–Littlewood maximal function for $|f|$). In either case the cubes Q_i give rise to a partition $\{\psi_i\}$ of unity, and conditions (I)–(III), (CZ1)–(CZ3), and (CZ6) are satisfied with $X=L^\infty$. Conditions (CZ4) and (CZ5) will be replaced by

(CZ4b) $\sum|Q_i|\le\frac{c}{t}\operatorname{dist}_{L^1}(f,B_{L^\infty}(t))$,
(CZ5b) $\frac{1}{|Q_i|}\int_{Q_i}|f-c_i|\le Ct$.

With these objects at hand, we again verify conditions (a)–(c) in Subsection 2.1 for the corresponding operators P_t^f (note that $\psi_\infty \neq 0$ this time, i.e., $\cup Q_i \neq \mathbb{R}^n$). The proof develops much as in the case of $X = \mathrm{BMO}$ (Subsection 2.2) with one exception: since T does not take L^∞ to itself, (c2) is no longer fulfilled automatically. But the function $h = TP_t^f f$ satisfies $\|h\|_{\mathrm{BMO}} \leq Ct$ by the first inequality in (a), Subsection 2.1. Putting $d_i = \frac{1}{\psi_i}\int h\psi_i$, we see that

$$\int |(I - P_t^f)h| = \int\limits_{\cup Q_i} |(I - P_t^f)h| \leq \sum_i \int\limits_{Q_i} |h - d_i| \leq C\|h\|_{\mathrm{BMO}} \sum_i |Q_i|$$

(because $|d_i - h_{Q_i}| \leq \frac{1}{\psi_i}\int\limits_{Q_i} |h - h_{Q_i}|\psi_i \leq C\|h\|_{\mathrm{BMO}}|Q_i|$), and (c2) follows from (CZ4b). $\square$

§3. Interpolation

3.1. Approximation by elements of intermediate spaces. Let X be either L^r $(1 < r < \infty)$, or BMO, or Lip_α $(0 < \alpha < 1)$, or $W_p^{(1)}$ $(1 < p < \infty)$. Suppose T is an operator of class $\mathcal{S}_r$ if $X = L^r$ and that T maps X to itself, satisfies (10) and belongs to S_u for some $u \in (1, \infty)$ if $X = \mathrm{BMO}$, Lip_α, or W_p^1.

In §1,2 we showed that in all cases the following is true: if $f \in L^1$ is such that $Tf \in L^1 + X$, then for every $t > 0$ there exists an element $f_t \in X$ satisfying $\|f_t\|_X \leq ct$ and

$$\|Tf - Tf_t\|_{L^1} \leq C(\mathrm{dist}_{L^1}(f, B_X(t)) + \mathrm{dist}_{L^1}(Tf, B_X(t))); \tag{15}$$

moreover, the family $\{f_t\}_{t>0}$ is constructed with the help of certain averaging (or smoothing) operators $P_t = P_t^f$ for the couple (L^1, X).

It turns out that real interpolation makes it possible to prove the same for a much wider class of spaces X. Furthermore, after a change of variables $t = t(s)$, the operators $P_{t(s)}$ will serve as appropriate smoothing operators for the couple (L^1, Y), where $Y = (L^1, X)_{\theta,q}$. Since

$$(L^1, \mathrm{Lip}_\alpha)_{\theta,\tau} = B_\tau^{\alpha\theta} \quad \text{for} \quad \frac{1}{\tau} = \frac{1-\theta}{1} + \frac{\theta}{\infty} \tag{16}$$

(see [Kr1]), and

$$(L^1, W_p^{(1)})_{\theta,\tau} = B_\tau^\theta \quad \text{for} \quad \frac{1}{\tau} = \frac{1-\theta}{1} + \frac{\theta}{p} \tag{17}$$

(see [BL]), we shall be able to extend (15) to the case where X is one of the Besov spaces occurring in (16), (17). We shall need a general result, which extends Theorem 1 in §1 to an abstract setting.

Let (X_0, X_1) be a compatible couple of spaces, and suppose that for every element $x \in X_0 + X_1$ a family x_t, $t > 0$, of its "almost optimal approximations" is given, i.e., $x_t \in X_1$ $(t > 0)$ and

$$\|x_t\|_{X_1} \leq Ct, \quad \|x - x_t\|_{X_0} \leq C\,\mathrm{dist}_{X_0}(x, B_{X_1}(t)).$$

Next, let X be an intermediate space for the couple (X_0, X_1) with norm equivalent to the norm of $X_{\theta,q} = (X_0, X_1)_{\theta,q}$.

Theorem A. *If* $\operatorname{dist}_{X_0}(x, B_X(s)) > 0$, *then there exists* $t = t(s)$ *such that*

(*i*) $\|x_{t(s)}\|_X \le Cs, \quad \|x - x_{t(s)}\|_{X_0} \le C \operatorname{dist}_{X_0}(x, B_X(s))$;

(*ii*) $\frac{1}{t(s)} \operatorname{dist}_{X_0}(x, B_{X_1}(t(s)) \le C(\frac{1}{s} \operatorname{dist}_{X_0}(x, B_X(s))^{1/\theta}$;

(*iii*) $\operatorname{dist}_{X_0}(x, B_{X_1}(t(s))) \le C \operatorname{dist}_{X_0}(x, B_X(s))$.

Remark. There is an explicit formula for $t(s)$. Put

$$\lambda(s) = \big(\frac{1}{s} \operatorname{dist}_{X_0}(x, B_X(s))\big)^{1/\theta}.$$

Then we may take

$$t(s) = \frac{1}{\lambda(s)} K(\lambda(s), x; X_0, X_1).$$

Here $K(\sigma, x; X_0, X_1) = \inf\{\|x_0\|_{X_0} + \sigma \|x_1\|_{X_1} : x = x_0 + x_1, x_i \in X_i, i = 1, 2,\}$ is the K-functional for the couple (X_0, X_1).

We observe at once that Theorem 1 (see Subsection 1.2) is a particular case of the above statements. It suffices to use Lemma 1 (Subsection 1.1) and put $X_0 = L^1$, $X_1 = L^\infty$ in Theorem A. Then $L^p = (L^1, L^\infty)_{\theta,p}$ with $\theta = 1 - 1/p$. The formula for $t(s)$ in Theorem 1 says the same as that for $t(s)$ in the above remark because of the explicit formula for the K-functional for the couple (L^1, L^∞), see [BL]. Finally, inequality (7) in Theorem 1 is true because the calculations in the proof of Lemma 1 show in fact that $\sum |Q_i| \asymp t(s)^{-1} \operatorname{dist}_{L^1}(f, B_{L^\infty}(t(s)))$, which is dominated by $C\lambda(s)$ in view of (ii) in Theorem A.

Theorem A and the remark to it will be proved in Appendix 1. Here we apply them to show how the interpolation spaces (16), (17) can be included in the general framework. Moreover, we show that the abstract setting in Theorem 3 is stable under interpolation.

Corollary. *Suppose the assumptions of Theorem 3 are satisfied, i.e.,* $Tx \in X_0 + X_1$ *and there is a family* $P_t = P_t^x$ *of linear operators obeying conditions* (a), (b), *and* (c) *in Subsection* 2.1. *Let* X *be an intermediate space for the couple* (X_0, X_1) *endowed with a norm equivalent to that of the space* $X_{\theta,q} = (X_0, X_1)_{\theta,q}$. *Consider the family* $\tilde{P}_s = P_{t(s)}$ *with* $t(s)$ *given by Theorem A. Then* $\{\tilde{P}_s\}$ *satisfies conditions similar to* (a), (b), *and* (c), *but for the couple* (X_0, X):

($\tilde{a}$) $\|\tilde{P}_s x\|_X \le Cs$, $\|(I - \tilde{P}_s)x\|_{X_0} \le C \operatorname{dist}_{X_0}(x, B_X(s))$;

($\tilde{b}$) $\|\tilde{P}_s\|_{X_0 \to X_0} \le C$, $\|I - \tilde{P}_s\|_{X \to X_0} \le C \frac{\operatorname{dist}_{X_0}(x, B_X(s))}{s}$;

($\tilde{c}1$) $\|\tilde{P}_s T(I - \tilde{P}_s)x\|_{X_0} \le C \operatorname{dist}_{X_0}(x, B_X(s))$;

($\tilde{c}2$) $\|(I - \tilde{P}_s)T\tilde{P}_s x\|_{X_0} \le C \operatorname{dist}_{X_0}(x, B_X(s))$.

Proof. Denoting $x_t = P_t x$, we see that all properties are immediate consequences of Theorem A, except for the second inequality in ($\tilde{b}$). But this inequality follows from (ii) in Theorem A by interpolation:

$$\|I - \tilde{P}_s\|_{X \to X_0} \le C\|I - P_{t(s)}\|_{X_{\theta,q} \to X_0} \le C\|I - P_{t(s)}\|_{X_0 \to X_0}^{1-\theta} \|I - P_{t(s)}\|_{X_1 - X_0}^{\theta}$$

$$\le C' \Big(\frac{\operatorname{dist}_{X_0}(x, B_{X_1}(t(s)))}{t(s)} \Big)^\theta \le C'' \frac{\operatorname{dist}_{X_0}(x, B_X(s))}{s}. \quad \square$$

3.2. K-closedness. Recall that a subcouple (Y_0, Y_1) of an interpolation couple (X_0, X_1) is said to be *K-closed* if

$$K(\sigma, y; Y_0, Y_1) \asymp K(\sigma, y, X_0, X_1)$$

uniformly in $y \in Y_0 + Y_1$ and $\sigma > 0$. In the presence of this property, we clearly have the convenient formula $(Y_1, Y_1)_{\theta,q} = (Y_0 + Y_1) \cap (X_0, X_1)_{\theta,q}$.

The statement below is a direct consequence of the preceding material. Suppose X is BMO, Lip_α, $W_p^{(1)}$, or else X is one of interpolation spaces occurring in (16) or (17), and let T be an operator bounded on X, belonging to $\mathcal{S}_t$ for some $t \in (1, \infty)$, and satisfying (10). Put $\mathcal{L}^1 = L^1 \oplus L^1$, $\mathfrak{X} = X \oplus X$, and consider the subspaces $\mathcal{H}^1(T)$ and $\mathcal{H}_X(T)$ in $\mathcal{L}^1$ and $\mathfrak{X}$, respectively, that consist of the pairs of functions of the form (f, Tf).

Proposition. *The couple $(\mathcal{H}^1(T), \mathcal{H}_X(T))$ is K-closed in $(\mathcal{L}^1, \mathfrak{X})$.* $\square$

Specific interpolation implications of this proposition are standard and are left to the reader. We have used the letter $\mathcal{H}$ because of a clear ideological relationship with the definition of Hardy spaces. Note that f may be an arbitrary function in X in the definition of $\mathcal{H}_X(T)$ (then Tf is in X automatically), but in that of $\mathcal{L}^1(T)$ we consider only those $f \in L^1$ for which Tf is also integrable.

Counterparts of the above proposition with L^p in place of X have been explored quite thoroughly, see, e.g., [P, Bo, KiX, Ki].

Appendix 1. Proof of Theorem A

Theorem A stated in Subsection 3.1 played an important part in the entire paper — starting with §1, where it appeared under disguise of a statement about the classical Calderón–Zygmund procedure, and finishing with §3. We pass to its proof, which will be split into several steps.

Step 1. *If a family $y_\tau (\tau > 0)$ satisfies*

$$\|y - y_\tau\|_{X_0} + \tau \|y_\tau\|_X \le CK(\tau, x; X_0, X), \tag{18}$$

then for

$$\tau(s) = \frac{1}{s} \operatorname{dist}_{X_0}(x, B_X(s)) \tag{19}$$

we have $\|y_{\tau(s)}\|_X \le Cs$ and $\|x - y_{\tau(s)}\|_{X_0} \le C \operatorname{dist}_{X_0}(x, B_X(s))$.

Indeed, since

$$K(\tau, x; X_0, X) = \inf_{\rho > 0} (\operatorname{dist}_{X_0}(x, B_X(\rho)) + \tau\rho),$$

by putting $\tau = \tau(s)$ and $\rho = s$ we obtain

$$K(\tau(s), x; X_0, X) \le 2 \operatorname{dist}_{X_0}(x, B_X(s)). \tag{20}$$

Step 2. *If a family y_τ $(\tau > 0)$ satisfies*

$$\|x - y_\tau\|_{X_0} + \tau \|y_\tau\|_{X_{\theta,q}} \le CK(\tau, x; X_0, X_{\theta,q}), \tag{21}$$

then

$$\|x - y_\tau\|_{X_0} + \tau \|y_\tau\|_X \le C' K(\tau, x; X_0, X).$$

Indeed, the norms of X and $X_{\theta,q}$ being equivalent, so are the K-functionals for the couples (X_0, X) and $(X_0, X_{\theta,q})$, whence the claim.

Steps 1–2 show that it suffices to construct some vectors y_τ that almost realize K-functionals for the couple $(X_0, X_{\theta,q})$, as in (21).

Step 3. *If a family y_τ satisfies*

$$\|x - y_\tau\|_{X_0} + \tau^{1/\theta}\|y_\tau\|_{X_1} \le CK(\tau^{1/\theta}, x; X_0, X_1),$$

then the same family satisfies (21).

The proof of this claim is contained in fact in the proof of the so-called Holmstedt formula and involves somewhat bulky calculations. We postpone it till the end of this Appendix and proceed with further steps.

Step 4. *If a family x_t $(t > 0)$ satisfies $\|x_t\|_{X_1} \le Ct$ and $\|x - x_t\|_{X_0} \le C\,\mathrm{dist}_{X_0}(x, B_X(t))$, then for*

$$t = t(\tau) = \tau^{-1/\theta}K(\tau^{1/\theta}, x; X_0, X_1)$$

we have

$$\|x - x_{t(\tau)}\|_{X_0} + \tau^{1/\theta}\|x_{t(\tau)}\|_{X_1} \le CK(\tau^{1/\theta}, x; X_0, X_1). \tag{23}$$

Indeed,

$$K(\tau^{1/\theta}, x; X_0, X_1) = \inf_{\rho>0}(\mathrm{dist}_{X_0}(x, B_{X_1}(\rho)) + \tau^{1/\theta}\rho).$$

For $\rho > t(\tau)$ we have $\tau^{1/\theta}\rho > K(\tau^{1/\theta}, x; X_0, X_1)$, therefore

$$K(\tau^{1/\theta}, x; X_0, X_1) = \inf_{0<\rho\le t(\tau)}(\mathrm{dist}_{X_0}(x, B_{X_1}(\rho)) + \tau^{1/\theta}\rho).$$

Since the function $\rho \mapsto \mathrm{dist}_{X_0}(x, B_{X_1}(\rho))$ is monotone nonincreasing, we have

$$\mathrm{dist}_{X_0}(x, B_{X_1}(t(\tau))) \le K(\tau^{1/\theta}, x; X_0, X_1). \tag{24}$$

Consequently,

$$\|x_{t(\tau)}\|_{X_1} \le Ct(\tau) = c\tau^{-1/\theta}K(\tau^{1/\theta}, x; X_0, X_1)$$

and

$$\|x - x_{t(\tau)}\|_{X_0} \le c\,\mathrm{dist}_{X_0}(x, B_{X_1}(t(\tau))) \le CK(\tau^{1/\theta}, x; X_0, X_1).$$

These two inequalities imply (23).

Step 5. *At Steps 1–4 we have reduced the problem of constructing almost optimal approximations for the couple (X_0, X) to the same problem for the couple (X_0, X_1). It remains to trace the parameter changes.*

Given s, at Step 1 we defined $\tau(s)$ by (19) and showed that it suffices to produce decompositions of x at which $K(\tau(s), x; X_0, X)$ is almost attained. At Step 2 we replaced X with $X_{\theta,q}$ leaving $\tau(s)$ in peace. At Step 3 we showed that it suffices to construct decompositions of x almost realizing $K(\tau(s)^{1/\theta}, x; X_0, X_1)$. Finaly, at Step 4 we showed that the latter can be exhibited if we know almost optimal approximations of x by the ball $B_t(X_1)$ with

$$t = t(s) = \frac{1}{\tau(s)^{1/\theta}}K(\tau(s)^{1/\theta}, x; X_0, X_1),$$

where $\tau(s)$ is still given by (19). This proves statement (i) in Theorem A and the formula for $t(s)$ in the remark to that theorem (note that for the function $\lambda(s)$ in that remark we have $\lambda(s) = \tau(s)^{1/\theta}$).

To prove (ii) in Theorem A, we refer to (24) and the formula for $t(s)$ to obtain

$$\frac{\operatorname{dist}_{X_0}(x, B_{X_1} B_{X_1}(t(s)))}{t(s)} \leq \tau(s) = \big(\frac{1}{s} \operatorname{dist}_{X_0}(x, B_X(s))\big)^{1/\theta}.$$

To prove (iii) in Theorem A, we observe that

$$K(\tau(s)^{1/\theta}, x; X_0, X_1) \leq C\tau\Bigg(\int\limits_{\tau(s)^{1/\theta}}^{\infty} [t^{-\theta} K(t, x; X_0, X_1)]^q \frac{dt}{t}\Bigg)^{1/q}$$

$$\leq CK(\tau(s), x; X_0, X_{\theta,q}) \leq C'K(\tau(s), x; X_0, X).$$

The first term in this chain of inequalities dominates $\operatorname{dist}_{X_0}(x, B_{X_1}(t(s))$ by (24), and the last term is dominated by $2\operatorname{dist}_{X_0}(x, B_X(s))$ (see (20)). $\square$

Now we present the skipped proof of the claim at Step 3. First, note that if $x = x_0 + x_1$, then

$$\tau\Bigg(\int\limits_{\tau^{1/\theta}}^{\infty} [t^{-\theta} K(t, x; X_0, X_1)]^q \frac{dt}{t}\Bigg)^{1/q} \leq$$

$$\leq c\tau\Bigg(\int\limits_{\tau^{1/\theta}}^{\infty} [t^{-\theta} K(t, x_0; X_0, X_1)]^q \frac{dt}{t}\Bigg)^{1/q} + c\tau\Bigg(\int\limits_{\tau^{1/\theta}}^{\infty} [t^{-\theta} K(t, x_1; X_0, X_1)]^q \frac{dt}{t}\Bigg)^{1/q}$$

$$\leq C(\|x_0\|_{X_0} + \tau\|x_1\|_{\theta,q}),$$

and the infimum of the latter quantity over all decompositions $x = x_0 + x_1$ is $CK(\tau, x; X_0, X_{\theta,q})$. So, it suffices to verify the inequality

$$\|x - y_\tau\|_{X_0} + \tau\|y_\tau\|_{X_{\theta,q}} \leq \tau\Bigg(\int\limits_{\tau^{1/\theta}}^{\infty} [t^{-\theta} K(t, x; X_0, X_1)]^q \frac{dt}{t}\Bigg)^{1/q}.$$

By the assumption of the claim at Step 3, we have

$$\|x - y_\tau\| \leq CK(\tau^{1/\theta}, x; X_0, X_1) \leq c\tau\Bigg(\int\limits_{\tau^{1/\theta}}^{\infty} [t^{-\theta} K(t, x; X_0, X_1)]^q \frac{dt}{t}\Bigg)^{1/q}.$$

Therefore, it remains to show that

$$\tau\|y_\tau\|_{X_{\theta,q}} \leq \tau\Bigg(\int\limits_{\tau^{1/q}}^{\infty} [t^{-\theta} K(t, x; X_0, X_1)]^q \frac{dt}{t}\Bigg)^{1/q}.$$

This is proved as follows:

$$\tau\|y_\tau\|_{X_{\theta,q}} = \tau\bigg(\int\limits_0^\infty [t^{-\theta}K(t,y_\tau;X_0,X_1)]^q\frac{dt}{t}\bigg)^{1/q} \le$$

$$\le c\tau\bigg(\int\limits_0^{\tau^{1/\theta}} [t^{-\theta}K(t,y_\tau;X_0,X_1)]^q\frac{dt}{t}\bigg)^{1/q} + c\tau\bigg(\int\limits_{\tau^{1/\theta}}^\infty [t^{-\theta}K(t,y_\tau;X_0,X_1)]^q\frac{dt}{t}\bigg)^{1/q}$$

$$\le c\tau\bigg(\int\limits_0^{\tau^{1/\theta}} [t^{-\theta}t\|y_\tau\|_{X_1}]^q\frac{dt}{t}\bigg)^{1/q} + c\tau\bigg(\int\limits_{\tau^{1/\theta}}^\infty [t^{-\theta}K(t,x;X_0,X_1)]^q\frac{dt}{t}\bigg)^{1/q} +$$

$$+c\tau\bigg(\int\limits_{\tau^{1/\theta}}^\infty [t^{-\theta}K(t,x-y_\tau;X_0,X_1)]^q\frac{dt}{t}\bigg)^{1/q} \le$$

$$\le C\tau\|y_\tau\|_{X_1}\tau^{\frac{1-\theta}{\theta}} + c\tau\bigg(\int\limits_{\tau^{1/\theta}}^\infty [t^{-\theta}K(t,x;X_0,X_1)]^q\frac{dt}{t}\bigg)^{1/q} +$$

$$+c\tau\bigg(\int\limits_{\tau^{1/\theta}}^\infty [t^{-\theta}\|x-y_\tau\|_{X_0}]^q\frac{dt}{t}\bigg)^{1/q} \le$$

$$\le c\tau^{1/\theta}\|y_\tau\|_{X_1} + c\tau\bigg(\int\limits_{\tau^{1/\theta}}^\infty [t^{-\theta}K(t,x;X_0,X_1)]^q\frac{dt}{t}\bigg)^{1/q} +$$

$$+c\|x-y_\tau\|_{X_0} \le K(\tau^{1/\theta},x;X_0,X_1) + c\tau\bigg(\int\limits_{\tau^{1/\theta}}^\infty [t^{-\theta}K(t,x;X_0,X_1)]^q\frac{dt}{t}\bigg)^{1/q} \le$$

$$\le c\tau\bigg(\int\limits_{\tau^{1/\theta}}^\infty [t^{-\theta}K(t,x;X_0,X_1)]^q\frac{dt}{t}\bigg)^{1/q}. \quad \square$$

Appendix 2. Calderón–Zygmund type decompositions for the couple $(L^1, \mathrm{Lip}_\alpha)$

This appendix consists of two parts. In the first part we formulate a covering theorem proved in [Kr1, Kr2] and show how it is used to construct C^∞-partitions of unity with "good" estimates of derivatives. In the second part such a partition of unity is applied to produce a Calderón–Zygmund type decomposition for the couple $(L^1, \mathrm{Lip}_\alpha)$. This result can be found in the Thesis [Kr1], but the Thesis itself is hardly accessible, to the extent that the result can be classified as unpublished up to now. Here we simplify somewhat the proof presented in [Kr1].

I. Whitney–Besicovich coverings and partitions of unity

Definition. A covering of $\mathbb{R}^n$ by cubes Q_i, $i \in I$, is called a Whitney–Besicovich covering (abbreviated as WB-covering) if

(a) the family $\{Q_i\}_{i\in I}$ splits into finitely many packings (the number of such packings will be denoted by B);

(b) for some $q < 1$, the family $\{qQ_i\}$ is also a covering of $\mathbb{R}^n$;

(c) the cubes Q_i are linked, i.e., for some positive ε (called the linking coefficient) the relation $Q_i \cap Q_j \neq \varnothing$ implies

$$|Q_i \cap Q_j| \geq \varepsilon \max(|Q_i|, |Q_j|).$$

In all specific cases to be considered below, only coverings with $q = 1/4$ and $B = B(n)$, $\varepsilon = \varepsilon(n)$ depending only on the dimension n will occur.

Remark. Properties (a) and (c) imply that each cube Q_{i_0} intersects at most B/ε cubes Q_j.

Indeed, let π be a packing formed by cubes of the covering. If $Q \in \pi$ and $Q \cap Q_{i_0} \neq \varnothing$, then $|Q \cap Q_{i_0}| \leq \varepsilon |Q_{i_0}|$. Consequently, π contains at most ε^{-1} cubes Q with these properties. Since by (a) the family $\{Q_i\}_{i\in I}$ splits in at most B packings, we are done.

In a natural way, a specific C^∞-partition of unity is related to every WB-covering. This partition of unity is constructed as follows. First, we take some C^∞-functions φ_i with $0 \leq \varphi_i \leq 1$, $\varphi_1 \equiv 1$ on qQ_i, and $\varphi_i \equiv 0$ off $\left(\frac{1+q}{2}\right)Q_i$. Next, put $\psi_i = \varphi_i / \left(\sum\limits_j \varphi_j\right)$. It is easily seen that the family $\{\psi_i\}_{i\in I}$ has the following properties:

(A) $\operatorname{supp}\psi_i \subset \left(\frac{1+q}{2}\right)Q_i$, $\quad \varphi_i(x) \geq B/\varepsilon \quad$ for $\quad x \in qQ_i$;

(B) $0 \leq \psi_i \leq 1, \sum \psi_i = 1$;

(C) for every multiindex $k = (k_1, \ldots, k_n)$ we have

$$\left| \frac{\partial^k}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}} \psi_i \right| \leq C |Q_i|^{-\frac{|k|}{n}},$$

where, as usual, $|k| = k_1 + \cdots + k_n$, and $C > 0$ depends only on k.

In the sequel, we refer to such a partition of unity as a *partition corresponding to the* BW-*covering* $\{Q_i\}_{i\in I}$.

We need yet another definition. Suppose we are given a family $\{Q_x\}_{x\in\mathbb{R}^n}$ such that each Q_x is centered at x (some cubes may be degenerate, i.e., it may happen that $|Q_x| = 0$).

Definition. The family $\{Q_x\}$ is said to have *bounded α-capacity* if for every packing π formed by some of the Q_x we have

$$\sum_{Q\in\pi} |Q|^\alpha \leq C < \infty,$$

where $C > 0$ does not depend on π. The smallest C is called the *α-capacity of the family* $\{Q_x\}$ and is denoted by $|\{Q_x\}|_\alpha$.

Theorem I. *Suppose a family $\{Q_x\}_{x\in\mathbb{R}^n}$ (Q_x is centered at x) has bounded β-capacity for some $\beta > 1$. Then for every $q < 1$ there is a WB-covering $\{Q_i\}$ such that the coefficient in property* (b) *for it is q, and*

(i) *if x_i is the center of Q_i, then $Q_{x_i} \subset Q_i$;*

(ii) *for every x there exists $i = i(x)$ such that $Q_x \subset Q_{i(x)}$;*

(iii) $\sum |Q_i|^\beta \le c(n)|\{Q_x\}|_\beta$.

Here $c(n)$, as well as the constants B and ε in the definition of a WB-covering depend only on the dimension n.

See [Kr2] for the proof.

II. A Calderón–Zygmund type decomposition for the couple $(L^1, \mathrm{Lip}_\alpha), 0 < \alpha < 1$

We remind the reader (see [C, Meyers]) that the norm

$$\sup_Q \frac{1}{|Q|^{1+\frac{\alpha}{n}}} \int_Q |f - f_Q|$$

is equivalent to the usual norm on the homogeneous space Lip_α.

Theorem II. *Let $f \in L^1$, and let $t > 0$. Then there exists a WB-covering $\{Q_i\}$ (and, consequently, a C^∞-partition $\{\psi_i\}$ of unity corresponding to this covering) such that*

1) $\sum |Q_i|^{1+\frac{\alpha}{n}} \le \frac{c}{t} \operatorname{dist}_{L^1}(f, B_{\mathrm{Lip}_\alpha}(t))$,

and for the function $f_t = \sum c_i\psi_i$, where the c_i are determined by the equations

$$\int (f - c_i)\psi_i = 0, \quad \text{i.e.,} \quad c_i = \frac{1}{\int \psi_i} \int f\psi_i, \tag{*}$$

we have

2) $\|f_t\|_{\mathrm{Lip}_\alpha} \le ct$;

3) $\frac{1}{|Q|^{1+\frac{\alpha}{n}}} \int\limits_{Q_i} |f - c_i| \le ct$.

All estimational constants depend only on the dimension n.

Remark. We saw in §2 (see the proof of (CZ8) there) that the family $\{f_t\}$ provides almost optimal approximations for f in the couple $(L^1, \mathrm{Lip}_\alpha)$: $\|f - f_t\|_{L^1} \le C \operatorname{dist}_{L^1}(f, B_{\mathrm{Lip}_\alpha}(t))$.

Proof of the theorem.

Step 1. *The family $\{Q_x\}_{x\in\mathbb{R}^n}$ of limiting cubes.* For every $x \in \mathbb{R}^n$, consider the function

$$\varphi_x(r) = \frac{1}{|Q(x,r)|^{1+\frac{\alpha}{n}}} \int\limits_{Q(x,r)} |f - f_{Q(x,r)}|, \quad r > 0.$$

Here and below, $Q(x, r)$ is the cube $\{y : \max\limits_{1\le i\le n} |x_i - y_i| \le r\}$. If $\sup\limits_{r>0} \varphi_x(r) \le t$, we put $Q_x = \{x\}$ (a degenerate cube, $|Q_x| = 0$). But if $\sup\limits_{r>0} \varphi_x(r) > t$, we put $r_x = \sup\{r : \varphi_x(r) > t\}$ and define $Q_x = Q(x, r_x)$. Since φ_x is continuous and $\lim_{r\to\infty} \varphi_x(r) = 0$, we see that r_x is finite and is a unique number satisfying the following properties:

a) $\varphi_x(r) \le t$ for all $r \ge r_x$,
b) there exist r smaller than r_x and arbitrarily close to r_x for which $\varphi_x(r) > t$,
c) $\varphi_x(r_x) = \frac{1}{|Q(x,r_x)|^{1+\frac{\alpha}{n}}} \int\limits_{Q(x,r_x)} |f - f_{Q(x,r_x)}| = t.$

The cubes Q_x constructed above are called *limiting cubes.*

Step 2. *Estimating the* $\left(1+\frac{\alpha}{n}\right)$*-capacity of the family of limiting cubes.* Let $\pi = \{Q_{x_i}\}$ be a packing formed by limiting cubes. By c), we have

$$t \sum_{Q_{x_i} \in \pi} |Q_{x_i}|^{1+\frac{\alpha}{n}} = \sum_i \int\limits_{Q_{x_i}} |f - f_{Q_{x_i}}| \le 2\|f\|_{L^1},$$

i.e., $|\{Q_x\}|_{1+\frac{\alpha}{n}} \le 2t^{-1}\|f\|_{L^1}$. However, a finer estimate is available.

To obtain this finer estimate, we note first that for an arbitrary cube Q and an arbitrary constant c we have

$$\int\limits_Q |f - f_Q| \le \int\limits_Q |f - c| + |Q||c - f_Q| = \int\limits_Q |f - c| + |\int\limits_Q (c - f)| \le 2\int\limits_Q |f - c|.$$

Hence,

$$\int\limits_Q |f - c| \ge \frac{1}{2} \int\limits_Q |f - f_Q| \tag{**}$$

for every constant c. Now, let $g \in \mathrm{Lip}_\alpha$ with $\|g\|_{\mathrm{Lip}_\alpha} \le \varepsilon t$, where ε will be fixed later. Then for every packing $\pi = \{Q_{x_i}\}$ formed by limiting cubes we have

$$\|f - g\|_{L^1} \ge \sum_i \int\limits_{Q_{x_i}} |f(x) - g(x)|dx \ge \sum_i \int\limits_{Q_{x_i}} |f(x) - g_{Q_{x_i}}|dx - \sum_i \int\limits_{Q_{x_i}} |g(x) - g_{Q_{x_i}}|dx$$

$$\ge \sum_i \frac{1}{2} \int\limits_{Q_{x_i}} |f - f_{Q_{x_i}}| - \sum_i \int\limits_{Q_{x_i}} |g - g_{Q_{x_i}}| \ge \frac{1}{2} t \sum_i |Q_{x_i}|^{1+\frac{\alpha}{n}} - \sum_i \int\limits_{Q_{x_i}} |g - g_{Q_{x_i}}|.$$

Since

$$\int |g - g_Q| \le C(n)|Q|^{1+\frac{\alpha}{n}} \|g\|_{\mathrm{Lip}_\alpha}$$

for every cube Q with $C(n)$ depending only on n, it follows that

$$\sum_i \int\limits_{Q_{x_i}} |g - g_{Q_{x_i}}| \le C(n)\varepsilon t \Big(\sum |Q_{x_i}|^{1+\frac{\alpha}{n}} \Big)$$

because $\|g\|_{\mathrm{Lip}_\alpha} \le \varepsilon t$. We fix ε by the requirement $C(n)\varepsilon = 1/4$, then the above calculations yield

$$\|f - g\|_{L^1} \ge \frac{1}{4} t \Big(\sum_{Q_{x_i} \in \pi} |Q_{x_i}|^{1+\frac{\alpha}{n}} \Big).$$

Since a packing π is arbitrary, we arrive at the estimate

$$|\{Q_x\}|_{1+\frac{\alpha}{n}} \leq \frac{c}{t} \operatorname{dist}_{L^1}(f, B_{\mathrm{Lip}_\alpha}(\frac{t}{c})).$$

Now, to the family $\{Q_x\}$ we apply the covering theorem stated in Part I of this Appendix. This will result in a WB-covering $\{Q_i\}$ satisfying (i), (ii), and (iii) in that theorem; more precisely, (iii) will take the form

$$\sum |Q_i|^{1+\frac{\alpha}{n}} \leq \frac{c}{t} \operatorname{dist}_{L^1}\left(f, B_{\mathrm{Lip}_\alpha}(\frac{t}{c})\right), \tag{***}$$

which is nearly property 1) in Theorem II. We may and do put $q = 1/4$ in property (b) of a WB-covering.

As was described in Part I, the family $\{Q_i\}$ gives rise to a C^∞-partition $\{\psi_i\}$ of unity satisfying (A)–(C) in Part I with $q = 1/4$.

Step 3. *Verification of the estimates.* First, we prove estimate 3) in Theorem II. By the definition of the coefficients c_i (see $(*)$ in the statement of Theorem II), we have

$$\int\limits_{Q_i} |f - c_i| = \int\limits_{Q_i} |f(x) - \frac{1}{\int \psi_i} \int\limits_{Q_i} f(y)\psi_i(y)dy|dx \leq$$

$$\leq \int\limits_{Q_i}\int\limits_{Q_i} \frac{1}{\int \psi_i}|f(x) - f(y)|dxdy \leq \frac{C}{|Q_i|} \int\limits_{Q_i}\int\limits_{Q_i} |f(x) - f(y)|dxdy.$$

Writing $|f(x) - f(y)| \leq |f(x) - f_{Q_i}| + |f_{Q_i} - f(y)|$, we see that

$$\int\limits_{Q_i} |f - c_i| \leq 2C \int\limits_{Q_i} |f(\xi) - f_{Q_i}|d\xi \leq C'|Q_i|^{1+\frac{\alpha}{n}}t.$$

In the last estimate we have used the fact that, by (i) in Theorem I, Q_i contains a limiting cube with the same center as Q_i.

It remains to prove inequality 2) in Theorem II, i.e., to show that $\|f_t\|_{\mathrm{Lip}_\alpha} \leq ct$. As has already been noted, for this it suffices to show that

$$\frac{1}{|Q|^{1+\frac{\alpha}{n}}} \int\limits_{Q} |f_t - (f_t)_Q| \leq Ct \tag{****}$$

for all cubes Q with C depending on the dimension only. We analyze the following two cases separately.

Case 1. Q is contained in some Q_i (then Q is said to be *small*).

Case 2. Q is not contained in any Q_i (then Q is said to be *big*).

First, suppose Q is small. Since $f_t \in C^\infty$, we have

$$\int\limits_{Q} |f_t - (f_t)_Q| \leq \frac{1}{|Q|} \int\limits_{Q}\int\limits_{Q} |f_t(x) - f_t(y)|dxdy \leq |Q|^{1+\frac{1}{n}} \max_{Q} |\nabla f_t|.$$

Thus, $(****)$ will follow if we prove the inequality

$$\max_Q |\nabla f_t| \le Ct \frac{1}{|Q|^{\frac{1-\alpha}{n}}}.$$

Since $Q \subset Q_{i_0}$ for some i_0, it suffices to prove that

$$\max_{Q_{i_0}} |\nabla f_t| \le Ct \frac{1}{|Q_{i_0}|^{\frac{1-\alpha}{n}}}.$$

Now, $f_t = \sum c_i \psi_i$ and $\sum \psi_i \equiv 1$, whence

$$\frac{\partial}{\partial x_k} f = \sum_i c_i \psi_i \quad \text{and} \quad \sum_i \frac{\partial}{\partial x_k} \psi_i = 0.$$

The second relation implies that $\frac{\partial}{\partial x_k}\psi_{i_0} = -\sum\limits_{i \ne i_0} \frac{\partial}{\partial x_k}\psi_i$; therefore

$$\frac{\partial}{\partial x_k} f_z = \sum_i (c_i - c_{i_0}) \frac{\partial}{\partial x_k} \psi_i.$$

On Q_{i_0}, only finitely many summands on the right-hand side do not vanish (see the remark after the definition of a WB-covering and recall that $\operatorname{supp}\psi_i \subset Q_i$). Since also $\left|\frac{\partial}{\partial x_k}\psi_i\right| \le C|Q|^{-1/n}$, we have

$$\max_{Q_{i_0}} \left|\frac{\partial}{\partial x_k} f_t\right| \le C \sum_{i: Q_i \cap Q_{i_0} \ne 0} |c_i - c_{i_0}||Q_i|^{-\frac{1}{n}}.$$

But the relation $Q_i \cap Q_{i_0} \ne \varnothing$ implies that the volumes of Q_i, Q_{i_0}, and $Q_i \cap Q_{i_0}$ are comparable; so,

$$\max_{Q_{i_0}} \left|\frac{\partial}{\partial x_k} f_t\right| \le C \frac{1}{|Q_{i_0} \cap Q_i|} \int\limits_{Q_{i_0} \cap Q_i} |c_i - c_{i_0}||Q_{i_0}|^{-1/n}$$

$$\le \frac{C}{|Q_{i_0}|^{1+\frac{1}{n}}} \left(\int\limits_{Q_i} |f - c_i| + \int\limits_{Q_{i_0}} |f - c_{i_0}| \right)$$

$$\le \frac{C}{|Q_{i_0}|^{1+\frac{1}{n}}} \left(t|Q_i|^{1+\frac{\alpha}{n}} + t|Q_{i_0}|^{1+\frac{\alpha}{n}} \right) \le Ct|Q_{i_0}|^{\frac{1-\alpha}{n}},$$

and we are done in the case of a small cube Q.

We pass to the case of a big cube Q. We claim that

1) $\frac{1}{|Q|^{1+\frac{\alpha}{n}}} \int\limits_Q |f - f_Q| \le t$,

2) $\sum\limits_{Q_i \cap Q \ne \varnothing} |Q_i| \le C(n)|Q|$.

To prove 1), we denote by x the center of Q and consider the limiting cube Q_x. By construction, $Q_x \subset Q_i$ for some $i = i(x)$, and we see that $Q \not\subset Q_x$ because otherwise $Q \subset Q_i$, i.e., Q is small. Therefore, $Q_x \subset Q$, and 1) follows from the definition of a limiting cube.

We pass to the proof of 2). Since $\{Q_i\}$ splits in $B(n)$ packings, it suffices to prove that $Q_i \subset c(n)Q$ whenever $Q_i \cap Q \neq \varnothing$, where $c(n)$ depends only on n. Since $\{\frac{1}{4}Q_i\}$ is still a covering of $\mathbb{R}^n$ by construction, every $y \in Q_i \cap Q$ is contained in $\frac{1}{4}Q_{i_0}$ for some index i_0 (depending on y). But $Q \not\subset Q_{i_0}$, therefore $\frac{3}{4}r(Q_{i_0}) \leq 2r(Q)$, whence $r(Q_{i_0}) \leq 3r(Q)$. Since the volumes of Q_{i_0} and Q_i are comparable, we see that $Q_i \subset c(n)Q$, as claimed.

The remaining part of the proof of $(*\!*\!*\!*)$ in the case of a big cube Q is fairly easy. By $(**)$ and 1) above, we have

$$\int_Q |f_t - (f_t)_Q| \leq 2\int_Q |f_t - f_Q| \leq 2\Big(\int_Q |f - f_Q| + \int_Q |f - f_t|\Big) \leq$$

$$2t|Q|^{1+\frac{\alpha}{n}} + 2\sum_{Q_i \cap Q \neq \varnothing} \int_{Q_i} |f - c_i| \leq 2t|Q|^{1+\frac{\alpha}{n}} + ct\sum_{Q_i \cap Q \neq \varnothing} |Q_i|^{1+\alpha/n},$$

and it suffices to refer to 2) above in order to estimate the sum in the last term.

Step 4. Strictly speaking, it remains to explain a triviality: inequality $(***)$ is not quite property 1) in Theorem 2; specifically, in $(***)$ the radius of the ball on the right-hand side is t/c rather than t. To correct this, it suffices to change the parameter by putting $\tau = t/c$ and renaming $f_{c\tau}$ by f_τ. $\square$

References

[BL] J. Bergh, J. Löfström, *Interpolation spaces. An introduction. Grundlehren der mathematischen Wissenschaften*, v. 223. Springer-Verlag, Berlin etc., 1976.

[Bou] J. Bourgain, *Some consequences to Pisier's approach to interpolation*, Isr. Math. J., **77** (1992), 165–185.

[Br] Yu. A. Brudnyĭ, *Spaces determined by local approximations*, Trudy Mosk. Matem. Ob-va, **24** (1971), 69–132.

[CZ] A. P. Calderón and A. Zygmund, *On existence of certain singular integrals*, Acta Math., **88** (1952), 85–139.

[C] S. Campanato, *Proprieta di holderianita di alcune classi di fimzioni*, Ann. Scuola Norm. Sup. Pisa, **17** (1963), 175–188.

[CM] R. R. Coifman and Y. Meyer, *Ondelettes et Opépateurs*, vol. III. Hermann, Paris, 1991.

[FS] C. Fefferman and E. M. Stein, *H^p spaces of several variables*, Acta Math., **129** (1972), 137–193.

[GR] J. Garcia-Cuerva and J. L. Rubio de Francia, *Weighted norm inequalities and related topics*, North Holland, Amsterdam, 1985.

[Ki] S. V. Kislyakov, *Interpolation of H^p-spaces: some recent developments*, Israel Math. Conf. Proceddings, **13** (1999), 102–140.

[KiX] S. V. Kislyakov and Quanhua Xu, *Real interpolation and singular integrals*, St. Petersburg Math. J., **8** (1997), no. 4.

[Kr1] Kruglyak N. Ya., *Investigations on the real interpolation method*, Doctor of Science Thesis, POMI, 1996.

[Kr2] Kruglyak N. Ya., *Smooth analogs of the Calderon–Zygmund decomposition quantitative covering theorem, and the K-functional*, Algebra i Analiz, **8**, no. 4 (1996), 110–160; English transl.: St. Petersburg Math. J., **8**, no. 4 (1997), 617–649.

[M] Y. Meyer, *Ondelettes et Opérateurs*, vol. I, II. Hermann, Paris, 1990.

[Meyers] G. N. Meyers, *Mean oscillation over cubes and Hölder continuity*, Proc. Amer. Math. Soc., **15** (1964), 717–721.

[P] G.Pisier, *Interpolation between H^p-spaces and noncommutative generalization.* I, Pacific J. Math., **155** (1992), 341–368.

[S] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton University Press, Princeton, New Jersey, 1970.