JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 14, Number 2, Pages 279–296
S 0894-0347(00)00359-3
Article electronically published on November 30, 2000

THE AVERAGING LEMMA

RONALD DEVORE AND GUERGANA PETROVA

1. Introduction

Averaging lemmas arise in the study of regularity of solutions to nonlinear transport equations. The present paper shows how techniques from Harmonic Analysis, such as wavelet decompositions, maximal functions, and interpolation, can be used to prove averaging lemmas and to establish their sharpness.

Let $f(x,v)$ be a real-valued function defined on $\mathbb{R}^d \times \Omega$, where Ω is a bounded domain in $\mathbb{R}^d$. In applications Ω is a set of velocity vectors. Associated to f, we have the velocity average

$$\bar{f}(x) := \int_\Omega f(x,v)\,dv. \tag{1.1}$$

By an averaging lemma, we shall mean a result which deduces smoothness for $\bar{f}$ from assumptions about f and the function

$$g(x,v) := v \cdot \nabla_x f. \tag{1.2}$$

Note that $g(\cdot,v)$ only gives information about the smoothness of $f(\cdot,v)$ in the direction v. We shall restrict our attention to the nontrivial case $d \geq 2$. The simplest version of an averaging lemma is the following.

Theorem 1.1. *If $f, g \in L_2(\mathbb{R}^d \times \Omega)$, $d \geq 2$, then $\bar{f}$ is in the Sobolev space $W^{1/2}(L_2(\mathbb{R}^d))$ and*

$$\|\bar{f}\|_{W^{1/2}(L_2(\mathbb{R}^d))} \leq C\left[\|f\|_{L_2(\mathbb{R}^d\times\Omega)} + \|g\|_{L_2(\mathbb{R}^d\times\Omega)}\right], \tag{1.3}$$

where C depends only on d and Ω.

This theorem is easily proved using Fourier transforms [8].

We are interested in generalizations of the averaging lemma in which the role of L_2 is replaced by L_p, $p \neq 2$. We shall use the abbreviated notation $B^s_p := B^s_p(L_p(\mathbb{R}^d))$ for this Besov space. Note that $W^{1/2}(L_2(\mathbb{R}^d)) = B^{1/2}_2(L_2(\mathbb{R}^d))$. We shall prove the following averaging lemma in §3.

Received by the editors November 18, 1999 and, in revised form, July 7, 2000.

1991 *Mathematics Subject Classification.* Primary 35L60, 35L65, 35B65, 46B70; Secondary 46B45, 42B25.

Key words and phrases. Averaging lemma, regularity, transport equations, Besov spaces.

Both authors were supported in part by the Office of Naval Research Contract N0014-91-J1343.

The second author was also supported by the Rackham Grant and Fellowship Program.

©2000 American Mathematical Society

Theorem 1.2. *If $f, g \in L_p(\mathbb{R}^d \times \Omega)$, $d \geq 2$, then the average $\bar{f}$ is in the Besov space $B_p^{\min(1/p,1/p')}$ and*

$$\|\bar{f}\|_{B_p^{\min(1/p,1/p')}} \leq C \left[\|f\|_{L_p(\mathbb{R}^d \times \Omega)} + \|g\|_{L_p(\mathbb{R}^d \times \Omega)} \right], \tag{1.4}$$

where C depends only on d and Ω.

As noted above, the case $p = 2$ of this lemma was first proved by Golse, Lions, Perthame, and Sentis [8] and is a quite elementary application of Fourier transforms. This latter paper also proves that when $f, g \in L_p$, then $\bar{f}$ is in each of the Besov spaces B_p^s, $0 < s < \min(1/p, 1/p')$. DiPerna, Lions, and Meyer [7] remark that Theorem 1.2 holds for $p > 2$ and also show for $p < 2$ that $f, g \in L_p(\mathbb{R}^d \times \Omega)$ implies $\bar{f} \in B_2^{1/p'}(L_p(\mathbb{R}^d))$. Later, Bezard [2] showed that in the case $1 < p < 2$, the space $B_2^{1/p'}(L_p(\mathbb{R}^d))$ can be replaced by the (smaller) potential space $\mathcal{L}^{1/p'}(L_p(\mathbb{R}^d))$. The conclusion in Theorem 1.2 is slightly stronger than Bezard's result because $B_p^{1/p'}(L_p(\mathbb{R}^d))$ is properly embedded in $\mathcal{L}^{1/p'}(L_p(\mathbb{R}^d))$. The proofs of the previous authors for $1 < p < 2$ are based on Littlewood-Paley theory and in the case of Bezard on an analysis of the averaging lemma for $p = 1$ involving Hardy spaces.

One of the main points of the present paper is to show how Theorem 1.2 follows immediately from the following facts about the real method of interpolation:

$$(L_1(\mathbb{R}^d), B_2^{1/2}(L_2(\mathbb{R}^d)))_{2/p',p} = B_p^{1/p'}(L_p(\mathbb{R}^d)), \quad 1 < p < 2, \tag{1.5}$$

and

$$(L_\infty(\mathbb{R}^d), B_2^{1/2}(L_2(\mathbb{R}^d)))_{2/p,p} \subset B_p^{1/p}(L_p(\mathbb{R}^d)), \quad 2 < p < \infty. \tag{1.6}$$

The second of these interpolation theorems is a simple consequence of existing interpolation results and covers the case $2 < p < \infty$ in Theorem 1.2. On the other hand, (1.5) uses (simple) new ideas based on maximal functions in its proof. We feel that the new technique will prove useful in other settings.

We should note that to derive Theorem 1.2 from the above interpolation theorems is elementary and utilizes only the obvious result that $f \in L_p(\mathbb{R}^d \times \Omega)$ implies $\bar{f} \in L_p(\mathbb{R}^d)$, $p = 1, \infty$, together with the simple case $p = 2$ already noted above.

We shall also show that Theorem 1.2 is sharp in the following sense.

Theorem 1.3. *For each bounded domain $\Omega \subset \mathbb{R}^d$ and each $1 < p < \infty$ and $q < p$, there is a function f with $f, g \in L_p(\mathbb{R}^d \times \Omega)$, $d \geq 2$, such that $\bar{f}$ is not in $B_q^{\min(1/p,1/p')}(L_p(\mathbb{R}^d))$.*

In the case $1 < p \leq 2$, Lions [9] has shown the weaker result that for each $s > 1/p'$ there exists $f, g \in L_p$ for which $\bar{f}$ is not in B_p^s. He has also formulated some negative results in the case $2 < p < \infty$ but not in the above form.

We shall also go into a deeper analysis of the averaging lemma near the endpoint $p = 1$. We show that if Ω is bounded away from 0, then, whenever $f, g \in L_1(\mathbb{R}^d \times \Omega)$, the wavelet coefficients of $\bar{f}$ are in weak-ℓ_1. Note that for a function in L_1, it is generally not true that its wavelet coefficients are in weak-ℓ_1. We want to stress however that the proof of Theorem 1.2 does not require this weak-ℓ_1 theorem.

Our main vehicle for analyzing the averaging operation is wavelet decompositions. Wavelets give simpler characterizations of Besov spaces (in terms of wavelet coefficients) than those from the Littlewood-Paley decompositions.

We have restricted our attention in this paper to a specific setting for averaging lemmas. Many variants are possible such as replacing v by a more general function

$a(v)$ in $v \cdot \nabla_x f$ or taking a weighted average over $\mathbb{R}^d$ in place of the average over Ω. The techniques put forward in this paper can be applied to these variants as well. However, we feel that the main point of this paper is to understand how certain elements from Harmonic Analysis can be used in the analysis of averaging lemmas. Therefore, we do not strive to give the most general results but rather to illustrate these simple techniques and to show how they give sharp results.

An outline of this paper is as follows. In §2, we introduce the known results on wavelet decompositions and their characterization of Besov spaces that we shall need in this paper. In §3, we prove the interpolation results (1.5)–(1.6). In §4, we prove Theorem 1.2. In §5, we prove Theorem 1.3. The final section is devoted to the weak-ℓ_1 result.

2. Besov spaces and wavelets

In this section, we introduce wavelet decompositions and explain how they characterize the classical smoothness spaces. General references for the material in this section are Meyer [10] and Daubechies [4]. Let E' denote the set of vertices of the cube $[0,1]^d$ and let E denote the set of nonzero vertices. Let ψ be a univariate continuously differentiable wavelet function with compact support which is obtained from the scaling function ϕ. Examples of such wavelets and scaling functions were given by Daubechies [3]. We could also use biorthogonal wavelets or even noncompactly supported wavelets but at the expense of technical complications. We shall use the notation $\psi^0 := \phi$ and $\psi^1 := \psi$. For each $e \in E'$, we define the multivariate wavelets

$$\psi^e(x_1, \dots, x_d) := \psi^{e_1}(x_1) \cdots \psi^{e_d}(x_d). \tag{2.1}$$

Let $\mathcal{D}$ denote the set of dyadic cubes in $\mathbb{R}^d$ and let $\mathcal{D}_j$ denote those dyadic cubes which have side length 2^{-j} and $\mathcal{D}_+ := \bigcup_{j\geq 0} \mathcal{D}_j$. For any dyadic cube $I = 2^{-j}(k + [0,1]^d)$ in $\mathcal{D}$ and any $e \in E'$, we define the wavelet

$$\psi_I^e(x) := \psi^e(2^j x - k), \tag{2.2}$$

which is a wavelet scaled to I. Notice that these functions are normalized in L_∞.

To simplify the notation that follows, we introduce the indexing set Δ which consists of all pairs (I, e) with $I \in \mathcal{D}_+$ and $e \in E$ ($e \in E'$ if $I \in \mathcal{D}_0$). We also let $\Delta_j := \{(I, e) \in \Delta : I \in \mathcal{D}_j\}$, $j \geq 0$. For any locally integrable function h on $\mathbb{R}^d$, we define its wavelet coefficients

$$a_I^e(h) := \langle h, \psi_I^e \rangle, \quad (I, e) \in \Delta. \tag{2.3}$$

Moreover, we let

$$a_I(h) := \begin{cases} \sum_{e\in E'} |a_I^e(h)|, & I \in \mathcal{D}_0, \\ \sum_{e\in E} |a_I^e(h)|, & I \in \mathcal{D}_j,\ j \geq 1. \end{cases} \tag{2.4}$$

These wavelet coefficients are normalized for $L_1(\mathbb{R}^d)$.

There are times when we shall need normalizations of the wavelet functions and coefficients for $L_p(\mathbb{R}^d)$. We define

$$\psi_{I,p}^e := |I|^{-1/p} \psi_I^e, \quad (I, e) \in \Delta, \tag{2.5}$$

and

$$a_{I,p}^e(h) := |I|^{-1/p'} a_I^e(h), \quad a_{I,p}(h) := |I|^{-1/p'} a_I(h). \tag{2.6}$$

Each function h which is locally integrable on $\mathbb{R}^d$ has the wavelet expansion

$$h = \sum_{(I,e)\in\Delta} a^e_{I,2}(h)\psi^e_{I,2}. \tag{2.7}$$

The functions appearing in (2.7) form a complete orthonormal system for $L_2(\mathbb{R}^d)$.

Let $s \geq 0$, $0 < q, p \leq \infty$. The Besov space $B^s_q(L_p(\mathbb{R}^d))$ is usually defined by means of moduli of smoothness or Fourier transforms. It is a smoothness space with s giving the order of smoothness (analogous to the number of derivatives), p giving the space in which smoothness is to be measured (namely $L_p(\mathbb{R}^d)$), and q giving a finer distinction of these spaces which is important in many applications. We refer the reader to any of the standard treatments of Besov spaces ([6], [10], [11], [1]).

There are equivalent characterizations of Besov spaces in terms of wavelet coefficients. We shall use these as our definition of the Besov spaces. Given $s > 0$, let $r > s$, let ϕ and ψ be in C^r, and let ψ have at least r vanishing moments. For $1 \leq p \leq \infty$, $0 < s < r$, $0 < q \leq \infty$, we define the Besov space $B^s_q(L_p(\mathbb{R}^d))$ by means of the (quasi-)norm

(2.8)
$$\|h\|_{B^s_q(L_p(\mathbb{R}^d))} := \begin{cases} \left(\sum_{j=0}^\infty 2^{jsq}\left(\sum_{I\in\mathcal{D}_j} a_{I,p}(h)^p\right)^{q/p}\right)^{1/q}, & 0 < q < \infty, \\ \sup_{j\geq 0} 2^{js}\left(\sum_{I\in\mathcal{D}_j} a_{I,p}(h)^p\right)^{1/p}, & q = \infty. \end{cases}$$

We note, but shall not use, the fact that the smoothness condition on ψ assumed for (2.8) can actually be weakened to only requiring that $\psi \in B^{s+\epsilon}_\infty(L_p)$ for some $\epsilon > 0$. Each choice of ϕ and ψ with the above-mentioned properties gives a norm in (2.8) which is equivalent to the Besov norm.

In going further, we shall use the abbreviated notation

$$B^s_p := B^s_p(L_p(\mathbb{R}^d)), \quad s > 0,\ 1 \leq p \leq \infty. \tag{2.9}$$

In the special case that $\phi = \chi_{[0,1]}$, the wavelet ψ is the Haar function and ψ^e_I is supported on I for each $I \in \mathcal{D}$, $e \in E'$. The Haar function is not completely sufficient for our purposes (it barely misses characterizing the Besov spaces of interest to us), because of its lack of smoothness. However, it does provide one direction of the characterization in (2.8).

Remark 2.1. If a function h is in the Besov space $B^s_q(L_p(\mathbb{R}^d))$, $0 < s < 1$, $0 < q, p \leq \infty$, then its Haar coefficients $a_{I,p}(h)$ satisfy

$$\left(\sum_{j=0}^\infty 2^{jsq}\left(\sum_{I\in\mathcal{D}_j} a_{I,p}(h)^p\right)^{q/p}\right)^{1/q} \leq C\|h\|_{B^s_q(L_p(\mathbb{R}^d))}, \tag{2.10}$$

with C depending only on s and d.

This follows from well-known results on approximation by piecewise constant functions. For example, let P_j denote the orthogonal projector which maps f into the piecewise constant function $P_j f$ which takes the value $f_I := \frac{1}{|I|}\int_I f$ on each dyadic cube I in $\mathcal{D}_j$. Then, it is well known and easy to prove that

$$\|f - P_j f\|_{L_p(\mathbb{R}^d)} \leq C|f|_{B^\alpha_\infty(L_p(\mathbb{R}^d))}2^{-j\alpha} \tag{2.11}$$

holds for all $f \in B^\alpha_\infty(L_p(\mathbb{R}^d))$, $0 < \alpha \le 1$. Obviously, the left side in (2.11) can be replaced by $\|P_j f - P_{j-1} f\|_{L_p(\mathbb{R}^d)}$. This latter expression is simply the ℓ_p norm of the L_p normalized Haar coefficients from the dyadic level j. This gives Remark 2.1 in the case $q = \infty$. Choosing $\alpha_1 < s < \alpha_2$ and interpolating between $B^{\alpha_1}_\infty(L_p(\mathbb{R}^d))$ and $B^{\alpha_1}_\infty(L_p(\mathbb{R}^d))$ gives Remark 2.1 for all $0 < s < 1$, $0 < q \le \infty$.

We can always assume that the scaling function ϕ and the wavelet ψ are supported in $[0, \ell]$ with ℓ an odd natural number. Then, when $I \in \mathcal{D}_j$ and $I = 2^{-j}(k + [0,1]^d)$, the functions ψ^e_I, $e \in E'$, are all supported in

$$\bar{I} := 2^{-j}(k + [0, \ell]^d). \tag{2.12}$$

The overlapping support causes technical difficulties which can be overcome with the following lemma (see Lemma 4.3 in [5]).

Lemma 2.2. *There exist disjoint sets of dyadic cubes $\Gamma_1, \dots, \Gamma_\gamma$, with $\gamma = \gamma(d)$, such that*

(i) $\mathcal{D} = \bigcup_{i=1}^{\gamma} \Gamma_i$,
(ii) $\Gamma_m \cap \Gamma_n = \emptyset, \quad n \ne m$,
(iii) *if I, $J \in \Gamma_m$, either $\bar{I} \subset \bar{J}$ or $\bar{J} \subset \bar{I}$ or $\operatorname{int}(\bar{I}) \cap \operatorname{int}(\bar{J}) = \emptyset$.*

3. An interpolation theorem

In this section, we shall prove the interpolation results (1.5) and (1.6). Given a pair of quasi-normed linear spaces X, Y which are continuously embedded in a Hausdorff space $\mathcal{X}$, the K-functional for this pair is defined by

$$K(f,t) := K(f,t;X,Y) := \inf_{f=f_0+f_1} \|f_0\|_X + t\|f_1\|_Y, \quad t > 0. \tag{3.1}$$

The real interpolation space $(X,Y)_{\theta,q}$, $0 < \theta < 1$, $0 < q \le \infty$, consists of all $f \in \mathcal{X}$ such that

$$\|f\|_{(X,Y)_{\theta,q}} := \begin{cases} \left(\int_0^\infty [t^{-\theta} K(f,t)]^q \, \frac{dt}{t}\right)^{1/q}, & 0 < q < \infty, \\ \sup_{t \ge 0} t^{-\theta} K(f,t), & q = \infty, \end{cases} \tag{3.2}$$

is finite. The expression in (3.2) defines the (quasi-)norm for this space.

The fundamental interpolation theorem for the θ, q spaces is the following. If U is a linear operator which boundedly maps X_0 to X_1 and Y_0 to Y_1, then, for each $0 < \theta < 1$ and $0 < q \le \infty$, U also boundedly maps $(X_0, Y_0)_{\theta,q}$ into $(X_1, Y_1)_{\theta,q}$ and we have

$$\|Uf\|_{(X_1,Y_1)_{\theta,q}} \le C \|f\|_{(X_0,Y_0)_{\theta,q}} \tag{3.3}$$

for every $f \in (X_0, Y_0)_{\theta,q}$.

The real interpolation spaces have been characterized for many pairs (X,Y). We shall need such characterizations only for certain sequence spaces and we restrict the following discussion to the cases we need. Let w be a weight function defined on $(I,e) \in \Delta$ and consider the spaces $\ell_p(w)$ consisting of all sequences $(c^e_I)_{(I,e)\in\Delta}$ for which

$$\|(c^e_I)_{(I,e)\in\Delta}\|_{\ell_p(w)} := \begin{cases} \left(\sum_{(I,e)\in\Delta} w((I,e)) |c^e_I|^p\right)^{1/p}, & 0 < p < \infty, \\ \sup_{(I,e)\in\Delta} w((I,e)) |c^e_I|, & p = \infty, \end{cases} \tag{3.4}$$

is finite. We shall only be concerned with weights w of the form $w((I,e)) := |I|^s$, $s \in \mathbb{R}$.

If $w_0((I,e)) = |I|^{s_0}$ and $w_1((I,e)) = |I|^{s_1}$ are two weights of this form and $0 < \theta < 1$ and $0 < p_0, p_1 \le \infty$, then (see [1], p. 119)

$$(\ell_{p_0}(w_0), \ell_{p_1}(w_1))_{\theta,p} = \ell_p(w_\theta), \tag{3.5}$$

where

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad w_\theta(I) = |I|^{s_\theta}, \quad s_\theta = (1-\theta)p\frac{s_0}{p_0} + \theta p \frac{s_1}{p_1}. \tag{3.6}$$

It is also possible in the above interpolation results to replace the ℓ_p spaces by Lorentz spaces. We shall only need this in the case of the weak-$\ell_1(w)$ spaces which correspond to the Lorentz space $\ell_{1,\infty}(w)$. The quasi-norm on weak-$\ell_1(w)$ is given by

$$\begin{aligned} \|(c_I^e)_{(I,e)\in\Delta}\|_{\ell_{1,\infty}(w)} &:= \sup_{\epsilon>0} \ \epsilon w\{(I,e) \in \Delta : |c_I^e| > \epsilon\} \\ &= \sup_{\epsilon>0} \ \epsilon \sum_{(I,e):\, |c_I^e|>\epsilon} |I|^s. \end{aligned} \tag{3.7}$$

In the case $s = 0$ and $w((I,e)) = 1$, for all (I,e), we refer to the space weak-$\ell_1(w)$ as simply weak-ℓ_1. If $w((I,e)) = |I|^s$ for some $s \in \mathbb{R}$, then (see [1], p. 113)

$$(\ell_{1,\infty}(w), \ell_2(w))_{\theta,p} = \ell_p(w), \tag{3.8}$$

where

$$\frac{1}{p} = \frac{1-\theta}{1} + \frac{\theta}{2}. \tag{3.9}$$

Note that the weight w in (3.8) is not varying.

We shall prove (1.5) and (1.6) by using wavelet sequences. If a function $h \in L_1(\mathbb{R}^d)$, we can estimate its L_1 normalized wavelet coefficients by

$$|a_I^e(h)| \le C_0 \int_{\bar{I}} |h|, \tag{3.10}$$

where $\bar{I}$ is as defined in (2.12). It follows that the sequence of wavelet coefficients $(a_I^e(h))_{(I,e)\in\Delta_j}$ are in ℓ_1 for each $j = 0, 1, \dots$. However the combined sequence $(a_I^e(h))_{(I,e)\in\Delta}$ does not have this property. In fact, simple examples show that this latter sequence need not even be in weak-ℓ_1. The following lemma shows that a substitute for the weak-ℓ_1 property holds.

Lemma 3.1. *For any $s > 1$ and $w((I,e)) := |I|^s$, there exists a constant $C_s > 0$ such that for each $h \in L_1(\mathbb{R}^d)$, we have*

$$\epsilon \sum_{(I,e):|a_I^e(h)|>\epsilon|I|^s} w((I,e)) = \epsilon \sum_{(I,e):|a_I^e(h)|>\epsilon|I|^s} |I|^s \le C_s \|h\|_{L_1(\mathbb{R}^d)}, \quad \epsilon > 0. \tag{3.11}$$

Proof. From (3.10), for any $(I,e) \in \Delta$,

$$|a_I^e(h)| \le C_0 \int_{\bar{I}} |h|\, dx =: b_I^e(h) =: b_I(h), \tag{3.12}$$

where the equalities serve to define $b_I^e(h)$ and $b_I(h)$. We fix $\epsilon > 0$ and let $\Lambda := \{(I,e) \in \Delta : b_I^e(h) > \epsilon|I|^s\}$. For each m, let $\Lambda_m := \{(I,e) \in \Lambda : I \in \Gamma_m\}$, where the Γ_m are the sets of Lemma 2.2 which partition $\mathcal{D}$. It is enough to show that

$$\epsilon \sum_{(I,e)\in\Lambda_m} |I|^s \le C\|h\|_{L_1(\mathbb{R}^d)}. \tag{3.13}$$

We fix m and prove (3.13). We say a cube $I \in \Lambda_m$ is maximal if $\bar{I}$ is not contained in any other cube $\bar{J}$ with $(J,e) \in \Lambda_m$ for some e. We next show that for any maximal cube I, we have

$$\epsilon \sum_{(J,e)\in\Lambda_m, \bar{J}\subset\bar{I}} |J|^s \le C \int_{\bar{I}} |h|. \tag{3.14}$$

Indeed, for $k \ge 0$, there are at most $\ell^d 2^{kd} 2^d$ indices (J,e) such that $\bar{J} \subset \bar{I}$ and $|I| = 2^{kd}|J|$ and therefore

$$\sum_{(J,e)\in\Lambda_m, \bar{J}\subset\bar{I}} |J|^s \le (2\ell)^d \sum_{k=0}^{\infty} 2^{kd(1-s)} |I|^s \le C|I|^s \le C\epsilon^{-1} b_I = C\epsilon^{-1} \int_{\bar{I}} |h|, \tag{3.15}$$

where the next to last equality is the criterion for membership in Λ_m and the last equality is (3.12). This completes the proof of (3.14). We now add the inequalities (3.14) and use the fact that the maximal cubes are disjoint to arrive at (3.13). This completes the proof of the lemma. □

We can now prove our main interpolation theorem.

Theorem 3.2. *We have the following relations between interpolation spaces and Besov spaces on* $\mathbb{R}^d$*:*

$$(L_1, B_2^{1/2})_{2/p',p} = B_p^{1/p'}, \quad 1 < p < 2, \tag{3.16}$$

and

$$(L_\infty, B_2^{1/2})_{2/p,p} \subset B_p^{1/p}, \quad 2 < p < \infty. \tag{3.17}$$

Moreover, the norm for the interpolation space on the left side of (3.16) *is equivalent to the norm of the Besov space on the right side. Likewise, the norm of the Besov space on the right side of* (3.17) *is less than a fixed multiple of the norm of the interpolation space on the left side.*

Proof. We first prove (3.16). Let $w((I,e)) := |I|^s$, $s := \frac{d+1}{d}$, and let $c_I^e(h) := |I|^{-s} a_I^e(h)$ with $a_I^e(h)$ the L_1 normalized wavelet coefficients of h given by (2.3). We consider the linear operator U which maps h into $(c_I^e(h))_{(I,e)\in\Delta}$. Lemma 3.1 gives that U boundedly maps L_1 into weak-$\ell_1(w)$. The definition of the Besov space $B_2^{1/2}$ gives that U boundedly maps this Besov space into $\ell_2(w)$. By (3.3), we have that U boundedly maps $(L_1, B_2^{1/2})_{\theta,p}$ into $\ell_p(w)$ when θ satisfies $\frac{1}{p} = \frac{1-\theta}{1} + \frac{\theta}{2}$, i.e. $\theta = 2/p'$. From the definition (2.8) of the Besov spaces, we have

$$C\|(c_I^e(h))_{(I,e)\in\Delta}\|_{\ell_p(w)} \ge \|h\|_{B_p^{1/p'}}. \tag{3.18}$$

This gives that the interpolation space on the left of (3.16) is embedded in the Besov space on the right and

$$\|h\|_{B_p^{1/p'}} \le C\|h\|_{(L_1,B_2^{1/2})_{2/p',p}}. \tag{3.19}$$

We can reverse this embedding as follows. We consider the operator V which maps a given sequence $\mathbf{d} = (d_I^e)_{(I,e)\in\Delta}$ to the function $V(\mathbf{d}) = \sum_{(I,e)\in\Delta} d_I^e \psi_{I,1}^e$. Then V boundedly maps ℓ_1 into L_1 and $\ell_2(w_2)$ into $B_2^{1/2}$ when $w_2((I,e)) := |I|^{-\frac{d+1}{d}}$.

Thus, defining $w_p((I,e)) := |I|^{-\frac{d+1}{d}\frac{p}{p'}}$ and using (3.5), we obtain that V maps $\ell_p(w_p)$ into $(L_1, B_2^{1/2})_{2/p',p}$ and

$$\|V\mathbf{d}\|_{(L_1,B_2^{1/2})_{2/p'}} \le C\|\mathbf{d}\|_{\ell_p(w_p)} \le C\|V\mathbf{d}\|_{B_p^{1/p'}}. \tag{3.20}$$

The set of functions $V\mathbf{d}$, $\mathbf{d} \in \ell_p(w_p)$, equals $B_p^{1/p'}$ and therefore we have reversed the inequality (3.19). This completes the proof of (3.16).

The proof of the embedding (3.17) is similar to the proof of (3.19) except that we use (3.5) in place of (3.8). □

4. Proof of Theorem 1.2

In this section, we show how Theorem 1.2 can be proved by using the interpolation results of the previous section. Throughout this section, $\Omega \subset \mathbb{R}^d$ is a bounded set. We introduce below a linear operator T and examine its mapping properties. From these mapping properties we easily deduce the averaging lemma.

For any function $F \in L_1(\mathbb{R}^d \times \Omega) + L_\infty(\mathbb{R}^d \times \Omega)$, the differential equation

$$f + v \cdot \nabla_x f = F \tag{4.1}$$

has a solution

$$f(x,v) = \int_0^\infty e^{-s} F(x - sv, v)\, ds, \tag{4.2}$$

which is also in $L_1(\mathbb{R}^d \times \Omega) + L_\infty(\mathbb{R}^d \times \Omega)$. Note that if $F \in L_p(\mathbb{R}^d \times \Omega)$, $1 \le p \le \infty$, then so is f and

$$\|f\|_{L_p(\mathbb{R}^d \times \Omega)} \le \|F\|_{L_p(\mathbb{R}^d \times \Omega)}. \tag{4.3}$$

Also, we have

$$\|v \cdot \nabla_x f\|_{L_p(\mathbb{R}^d \times \Omega)} = \|F - f\|_{L_p(\mathbb{R}^d \times \Omega)} \le 2\|F\|_{L_p(\mathbb{R}^d \times \Omega)}. \tag{4.4}$$

We define the linear operator T by

$$(TF)(x) := \bar{f}(x) = \int_\Omega f(x,v)\, dv, \tag{4.5}$$

where f is given by (4.2). For each $1 \le p \le \infty$, we have

$$\|TF\|_{L_p(\mathbb{R}^d)} \le (\mathrm{meas}(\Omega))^{1/p'} \|f\|_{L_p(\mathbb{R}^d \times \Omega)} \le (\mathrm{meas}(\Omega))^{1/p'} \|F\|_{L_p(\mathbb{R}^d \times \Omega)}, \tag{4.6}$$

where the last inequality is (4.3).

Let us consider more closely the action of T on $L_2(\mathbb{R}^d \times \Omega)$. If $F \in L_2(\mathbb{R}^d \times \Omega)$, then from Theorem 1.1 (i.e. the case $p = 2$ of the averaging lemma), we have

$$\|TF\|_{B_2^{1/2}} \le C\left(\|f\|_{L_2(\mathbb{R}^d \times \Omega)} + \|v \cdot \nabla_x f\|_{L_2(\mathbb{R}^d \times \Omega)}\right) \le C\|F\|_{L_2(\mathbb{R}^d \times \Omega)}, \tag{4.7}$$

where the last inequality uses (4.3) and (4.4). In other words, T is a bounded mapping from $L_2(\mathbb{R}^d \times \Omega)$ into the Besov space $B_2^{1/2}$.

In summary, we have shown that T boundedly maps $L_p(\mathbb{R}^d \times \Omega)$ to $L_p(\mathbb{R}^d)$, $1 \le p \le \infty$, and boundedly maps $L_2(\mathbb{R}^d \times \Omega)$ into $B_2^{1/2}$. It follows that for $1 < p < \infty$, we have

$$\|TF\|_{B_p^{\min(1/p,1/p')}} \le C\|F\|_{L_p(\mathbb{R}^d \times \Omega)}, \quad F \in L_p(\mathbb{R}^d \times \Omega). \tag{4.8}$$

For example, the case $1 < p < 2$ follows from Theorem 3.2 and the fact that $(L_1(\mathbb{R}^d \times \Omega), L_2(\mathbb{R}^d \times \Omega))_{2/p',p} = L_p(\mathbb{R}^d \times \Omega)$.

Proof. Proof of Theorem 1.2 If $f \in L_p(\mathbb{R}^d \times \Omega)$ and $g := v \cdot \nabla_x f \in L_p(\mathbb{R}^d \times \Omega)$, then the function $F := f + g$ is also in $L_p(\mathbb{R}^d \times \Omega)$. The differential equation (4.1) with this choice of right side F has f as its unique solution. Thus $TF = \bar{f}$ and the theorem follows from (4.8). □

5. The sharpness of Theorem 1.2

In this section, we prove Theorem 1.3 in the case $2 \le p < \infty$. Similar techniques (which we do not give) will also prove the case $1 < p \le 2$ (see [12]).

Let $\Omega \subset \mathbb{R}^d$ be an arbitrary but fixed bounded domain. It follows that there is a constant $M > 0$ such that

$$|v| \le M, \quad v \in \Omega. \tag{5.1}$$

By rotating Ω if necessary, we can, without loss of generality, assume that $\Omega_\delta := \Omega \cap \{v = (v_1, \dots, v_d) : \ v_i \ge \delta, \ i = 1, \dots, d\}$ satisfies $\text{meas}(\Omega_\delta) \ne 0$ for some $\delta > 0$.

We shall use the piecewise linear function of one variable which satisfies

$$F(u) := \begin{cases} u + d + 1, & -d - 1 \le u \le -d, \\ 1, & -d \le u \le d, \\ -u + d + 1, & d \le u \le d + 1, \end{cases} \tag{5.2}$$

and is zero otherwise.

We shall build the function f called for in Theorem 1.3 as a sum of functions f_k which we now describe. Let

$$I_k := [9(k-1)d,\, 2^{-k} + 9(k-1)d]^d, \quad k = 1, 2, \dots,$$

be the family of dyadic cubes centered at $x^k = (x_1^k, \dots, x_d^k)$, where

$$x_i^k = 2^{-k-1} + 9(k-1)d, \quad i = 1, \dots, d.$$

From F we construct the family of multivariate functions

$$F_k(x) := F(x_1)\chi_{[-\delta_k, \delta_k]}(x_2) \dots \chi_{[-\delta_k, \delta_k]}(x_d), \quad \delta_k := 2^{-k}\sqrt{d}.$$

Note that F_k is one on the ball of radius $2^{-k}\sqrt{d}$ centered at the origin. For $v \in \Omega_\delta$ let $F_k(\cdot, v)$ denote the function whose graph is obtained from that of $F_k(\cdot)$, by a shift of the origin to x^k and then a rotation of the x_1-axis to the v-axis (there are many such rotations and the analysis that follows holds for any of them). Then, $F_k(\cdot, v)$ is one on the ball of radius $2^{-k}\sqrt{d}$ centered at x^k.

In this section, we use the notation $\psi_k := H_{I_k, p'}$ for the following $L_{p'}$ normalized multivariate Haar function with support I_k:

$$\psi_k(x) := 2^{kd/p'} H(2^k x_1 - 9d(k-1)2^k)\chi(2^k x_2 - 9d(k-1)2^k) \\ \dots \chi(2^k x_d - 9d(k-1)2^k)$$

where $\chi := \chi_{[0,1]}$. Note that for all $v \in \Omega_\delta$, $F_k(\cdot, v) = 1$ on the support of ψ_k.

We denote by $\mathcal{H}_v$ the hyperplane which is orthogonal to v and passes through the origin, and we denote by P_v the projector onto $\mathcal{H}_v$. For a multivariate function h, we let $\int_{L(P_v x)} h$ be the line integral of h along the line $L(P_v x)$, which consists of all points in $\mathbb{R}^d$ which project onto $P_v x$.

For each $k \ge 1$, let

$$f_k(x, v) := \begin{cases} F_k(x, v)\, \text{sgn}(\int_{L(P_v x)} \psi_k), & v \in \Omega_\delta, \\ 0, & v \in \Omega \setminus \Omega_\delta. \end{cases} \tag{5.3}$$

Given any sequence $\beta := (\beta_k)$ of positive real numbers from ℓ_p, we define

$$f(\cdot, v) := f(\cdot, v; \beta) := \sum_{k=1}^{\infty} 2^{k(d-1)/p} \beta_k f_k(\cdot, v). \tag{5.4}$$

As usual, g is defined via

$$g := v \cdot \nabla_x f. \tag{5.5}$$

In this construction every sequence $\beta \in \ell_p$ generates a corresponding function $f(\cdot, \beta)$. The smoothness of $\bar{f}$ is governed by the behavior of β. We investigate this relation closer and show how different requirements, imposed on β, lead to different regularity of $\bar{f}$. We will need the following lemma.

Lemma 5.1. *Let Ω_δ, I_k, and ψ_k be defined as above. Then, for every $v \in \Omega_\delta$ and $k \geq 1$, we can find a set $W(v,k) \subset P_v I_k \subset \mathcal{H}_v$ with the properties:*

(a) $\operatorname{meas}_{d-1}(W(v,k)) \geq c 2^{-k(d-1)}$,

(b) *for every $v \in \Omega_\delta$ and $w \in W(v,k)$ we have*

$$|\int_{L(w)} \psi_k| \geq c 2^{-k(1-d/p')},$$

with the constant $c > 0$ depending only on δ and Ω.

Proof. We shall prove a result similar to (a) and (b) for the unit cube. Then, the lemma follows by dilating and translating. Let $Q := [0,1]^d$ and let H_0 be the following Haar function supported on Q:

$$H_0(x) = \begin{cases} 1, & 0 \leq x_1 < 1/2, \quad 0 \leq x_i < 1, \quad i = 2, \dots, d, \\ -1, & 1/2 \leq x_1 < 1, \quad 0 \leq x_i < 1, \quad i = 2, \dots, d. \end{cases} \tag{5.6}$$

Let $W := \{x \in Q : \ 1/2 \leq x_1 \leq 3/4, \ 0 < x_i \leq 3/4, \ i = 2, \dots, d-1, \ x_d = 0\}$. For each $v \in \Omega_\delta$ and $x \in W$, $x + tv \in Q$ if and only if $t \geq 0$ and

$$x_i + t v_i \leq 1, \quad i = 1, \dots, d. \tag{5.7}$$

This means that if $v \in \Omega_\delta$, $x \in W$, and $x + tv \in Q$, then $H_0(x + tv) = -1$. Note that (5.7) holds for

$$0 \leq t \leq \min_{i=1,\dots,d} \frac{1 - x_i}{v_i}.$$

Since for $x \in W$ and $v \in \Omega_\delta$

$$\frac{1 - x_i}{v_i} \geq \frac{1}{4 v_i} \geq \frac{1}{4M},$$

we have (5.7) for $0 \leq t \leq \frac{1}{4M}$. Hence for each $x \in W$ the line segment $L(P_v x) \cap Q$ has length $\geq \frac{1}{4M}\delta$, because $|v| \geq \delta$. Now, let $W(v)$ be the projection (by P_v) of W onto $\mathcal{H}_v$. Then, $\operatorname{meas}_{d-1}(W(v)) \geq c \operatorname{meas}_{d-1}(W)$ with c depending only on δ and M, because $\delta \leq v_i \leq M$, $i = 1, \dots, d$. This shows property (a) for Q. Also, for $w \in W(v)$,

$$|\int_{L(w)} H_0| = \int_{L(w)} 1 \geq \frac{\delta}{4M},$$

which verifies (b). ∎

The following theorem implies Theorem 1.3 for $2 \le p < \infty$.

Theorem 5.2. *Let $\Omega \subset \mathbb{R}^d$ be a fixed bounded domain. For every sequence $\beta \in \ell_p$, $2 \le p < \infty$, the functions f and g, defined in (5.4) and (5.5), are in $L_p(\mathbb{R}^d \times \Omega)$ and the velocity average $\bar{f}$ is in $B_p^{1/p}(L_p(\mathbb{R}^d))$. If $\beta \in \ell_p \setminus \ell_q$, $q < p$, then $\bar{f} \notin B_q^{1/p}(L_p(\mathbb{R}^d))$.*

Proof. We start with a direct calculation of the L_p norm of f. We use the fact that for each fixed $v \in \Omega_\delta$ the supports of $f_k(\cdot, v)$ are disjoint since the points x^k are sufficiently separated from each other. This gives

$$||f||^p_{L_p(\mathbb{R}^d \times \Omega)} = \int_{\Omega_\delta} \int_{\mathbb{R}^d} |f(x,v)|^p \, dxdv = \sum_{k=1}^{\infty} 2^{k(d-1)} \beta_k^p \int_{\Omega_\delta} \int_{\mathbb{R}^d} |f_k(x,v)|^p \, dxdv.$$

Since $\|f_k(\cdot, v)\|_{L_\infty(\mathbb{R}^d)} \le 1$, we have that for every $v \in \Omega_\delta$,

$$\int_{\mathbb{R}^d} |f_k(x,v)|^p \, dx \le \text{meas}(\text{supp } \ f_k(\cdot, v)) \le C 2^{-k(d-1)}. \tag{5.8}$$

Therefore, we derive

$$||f||^p_{L_p(\mathbb{R}^d \times \Omega)} \le C \, \text{meas}(\Omega_\delta) \sum_{k=1}^{\infty} \beta_k^p < \infty, \tag{5.9}$$

and hence $f \in L_p(\mathbb{R}^d \times \Omega)$.

We next prove that g is also in $L_p(\mathbb{R}^d \times \Omega)$. For each fixed $v \in \Omega_\delta$, $g(\cdot, v) = |v| D_v(f(\cdot, v))$, where $D_v(f(\cdot, v)) := \frac{v}{|v|} \cdot \nabla_x f$ is the derivative of f in direction $\frac{v}{|v|}$. Thus,

$$\begin{aligned} ||g||^p_{L_p(\mathbb{R}^d \times \Omega)} &= \int_{\Omega_\delta} |v|^p \|D_v(f(\cdot, v))\|^p_{L_p(\mathbb{R}^d)} \, dv \\ &\le C \sum_{k=1}^{\infty} 2^{k(d-1)} \beta_k^p \int_{\Omega_\delta} \|D_v(f_k(\cdot, v))\|^p_{L_p(\mathbb{R}^d)} \, dv. \end{aligned}$$

Now, for each $v \in \Omega_\delta$, we have $\|D_v(f_k(\cdot, v))\|_{L_\infty(\mathbb{R}^d)} \le 1$, and $D_v(f_k(\cdot, v)) = 0$ outside of the support of $f_k(\cdot, v)$. Therefore, as in (5.8), we get the inequality $\|D_v(f_k(\cdot, v))\|^p_{L_p(\mathbb{R}^d)} \le C 2^{-k(d-1)}$. This result and the previous estimate for $\|g\|^p_{L_p(\mathbb{R}^d \times \Omega)}$ lead to

$$||g||^p_{L_p(\mathbb{R}^d \times \Omega)} \le C \, \text{meas}(\Omega_\delta) \sum_{k=1}^{\infty} \beta_k^p < \infty. \tag{5.10}$$

Therefore, by Theorem 1.2, $\bar{f} \in B_p^{1/p}(L_p(\mathbb{R}^d))$ and the first part of the theorem is proved.

Going further, we consider only sequences $\beta \in \ell_p$. We shall show that $\bar{f} \notin B_q^{1/p}(L_p(\mathbb{R}^d))$, $q < p$, whenever β, used in the definition of f, is not an element of ℓ_q. For this purpose, we calculate $a_{I_k,p}(\bar{f})$ (see (2.6)) which is the sum of the L_p normalized Haar coefficients of $\bar{f}$ for I_k. One of the Haar coefficients that appears in the sum defining $a_{I_k,p}$ is obtained from the inner product of $\bar{f}$ with ψ_k which is the same as the inner product of $2^{k(d-1)/p} \beta_k \bar{f}_k$ with ψ_k. Therefore, using the

definition (5.3) of f_k, we obtain

$$|a_{I_k,p}(\bar f)| \ge |\int_{v\in\Omega_\delta}\int_{x\in I_k} 2^{k(d-1)/p}\beta_k f_k(x,v)\psi_k(x)\,dxdv|$$
$$= |\int_{v\in\Omega_\delta} 2^{k(d-1)/p}\beta_k \int_{w\in P_v I_k}\left[\int_{L(P_v x)} f_k\psi_k\right] dwdv|.$$

By construction, $f_k \equiv 1$ or $f_k \equiv -1$ along the line segment $L(P_v x)\cap I_k$, where the sign of f_k is determined by the sign of $\int_{L(P_v x)}\psi_k$. Therefore we have

$$|a_{I_k,p}(\bar f)| \ge \int_{v\in\Omega_\delta} 2^{k(d-1)/p}\beta_k \int_{w\in P_v I_k} |\int_{L(P_v x)}\psi_k|\,dwdv.$$

Now, we use Lemma 5.1 and derive that

$$|a_{I_k,p}(\bar f)| \ge C\int_{v\in\Omega_\delta} 2^{k(d-1)/p}\beta_k 2^{-k}2^{kd/p'}\,\mathrm{meas}(W(v,k))dv$$
$$\ge C\,\mathrm{meas}(\Omega_\delta)\beta_k 2^{k(d-1)/p}2^{-k}2^{kd/p'}2^{-k(d-1)} = C\beta_k 2^{-k/p}.$$

Hence, we get

$$(\sum_{I\in\mathcal{D}_k}|a_{I,p}(\bar f)|^p)^{1/p} \ge |a_{I_k,p}(\bar f)| \ge C\beta_k 2^{-k/p},$$

and then

$$\sum_{k=0}^{\infty}(2^{k/p}(\sum_{I\in\mathcal{D}_k}|a_{I,p}(\bar f)|^p)^{1/p})^q \ge C\sum_{k=1}^{\infty}\beta_k^q.$$

For every $\beta\in\ell_p$, but not in ℓ_q, the right side is infinite and Remark 2.1 shows that $\bar f \notin B_q^{1/p}(L_p(\mathbb{R}^d))$. This proves the theorem. □

6. Weak-ℓ_1 estimates

In the proof of Theorem 1.2, we have not used any special properties of averaging near L_1. We have only used the fact that $f\in L_1(\mathbb{R}^d\times\Omega)$ implies that $\bar f$ is in $L_1(\mathbb{R}^d)$. In this section, we want to show that the condition $f,g\in L_1(\mathbb{R}^d\times\Omega)$ gives extra information about $\bar f$. We first consider domains Ω which satisfy

$$\Omega\subset\{v\in\mathbb{R}^d:\ 0<M\le|v|<M'\}, \tag{6.1}$$

for some positive numbers $0<M<M'$.

To a function $h\in L_1(\mathbb{R}^d)$ we associate the sequence

$$\lambda := \lambda(h) := (a_I(h))_{I\in\mathcal{D}^+},$$

where a_I is defined by (2.4). As we remarked earlier, for a general function $h\in L_1(\mathbb{R}^d)$ and for any $j\ge 0$, the sequence $(a_I(h))_{I\in\mathcal{D}_j}$ is in ℓ_1. However the combined sequence $\lambda(h)$ is not necessarily in ℓ_1 and in fact need not even be in weak-ℓ_1. We shall show however that when $f,g\in L_1(\mathbb{R}^d\times\Omega)$, then $\lambda(\bar f)$ is in weak-ℓ_1.

The main new ingredient in the weak-ℓ_1 estimate for wavelet coefficients is given in the following lemma. For its formulation, we need to introduce some notation. For any cube Q, we let $\ell(Q)$ denote its side length. Given a vector v and a cube Q, we denote by $Q(v)$ the smallest (finite) cylinder which contains Q and has axis passing through the center of Q and in the direction v . We recall our notation $\bar I$ for the support cube associated to I (see (2.12)) and P_v for the projector along v (see the paragraph preceding (5.3)). Given a dyadic cube I, the cube $\bar I$ together

with some of its translates form a tiling of $\mathbb{R}^d$. We let I' denote the union of $\bar{I}$ with all of the neighboring cubes of $\bar{I}$ in this tiling. Each neighboring cube can be written as $\bar{J}$ for some dyadic cube J with $|J| = |I|$. Therefore,

$$I' = \bigcup_{j=1}^{\mu} \bar{I}_j \tag{6.2}$$

with the number μ of dyadic cubes in (6.2) depending only on d. We note that $\bar{I}(v) \subset I'$ for all $v \in \Omega$.

Lemma 6.1. *Let Ω be any bounded domain in $\mathbb{R}^d$ and $f, g \in L_1(\mathbb{R}^d \times \Omega)$. Then for every I and $J \in \mathcal{D}$, for which $\bar{J} \subset \bar{I}$, we have*

$$\int_\Omega \int_{\bar{J}} |f(x,v)|\,dxdv \le C\frac{\ell(\bar{J})}{\ell(\bar{I})} \int_\Omega \int_{I'} \left(|f(x,v)| + \ell(\bar{I})\frac{|g(x,v)|}{|v|} \right) dxdv, \tag{6.3}$$

where C depends only on d and M.

Proof. We fix dyadic cubes I and $J \in \mathcal{D}$ for which $\bar{J} \subset \bar{I}$. Note that $\bar{J}(v) \subset \bar{I}(v)$ for all $v \in \Omega$. We fix v and consider any translate $K := \alpha v + \bar{J}(v)$ of $\bar{J}(v)$ which is contained in $\bar{I}(v)$. For every $v \in \Omega \setminus \{(0, \dots, 0)\}$, we can write g as $g(\cdot, v) = |v| D_v f$, with $D_v f$ the directional derivative of $f(\cdot, v)$ in the direction $\frac{v}{|v|}$. Then, for each $x \in \bar{J}(v)$, we have

$$|f(x,v)| \le |f(x + \alpha v, v)| + \int_{L(P_v x)} \frac{|g(\cdot, v)|}{|v|}, \tag{6.4}$$

where $P_v x$ is the projection of x onto $\mathcal{H}_v$ and $L(P_v x)$ is the line segment consisting of all points in $\bar{I}(v)$ which project onto $P_v x$. When we integrate (6.4) over $\bar{J}(v)$ we get

$$\int_{\bar{J}(v)} |f(x,v)|\,dx \le \int_K |f(x,v)|\,dx + C\ell(\bar{J}) \int_{\bar{I}(v)} \frac{|g(x,v)|}{|v|}\,dx. \tag{6.5}$$

There are at least $C\frac{\ell(\bar{I})}{\ell(\bar{J})}$ choices of α such that the sets K are disjoint. Therefore, by summing the corresponding inequalities (6.5), we obtain

$$C\frac{\ell(\bar{I})}{\ell(\bar{J})} \int_{\bar{J}(v)} |f(x,v)|\,dx \le \int_{\bar{I}(v)} |f(x,v)|\,dx + C\ell(\bar{I}) \int_{\bar{I}(v)} \frac{|g(x,v)|}{|v|}\,dx.$$

We have that $\bar{J} \subset \bar{J}(v)$ and $\bar{I}(v) \subset I'$ for all v, and therefore

$$C\frac{\ell(\bar{I})}{\ell(\bar{J})} \int_{\bar{J}} |f(x,v)|\,dx \le \int_{I'} \left(|f(x,v)| + C\ell(\bar{I})\frac{|g(x,v)|}{|v|} \right) dx. \tag{6.6}$$

We integrate over $v \in \Omega$ and arrive at (6.3). ∎

We shall use the two auxiliary sequences

$$\tilde{b}_I := \tilde{b}_I(f) := c_0 \int_\Omega \int_{\bar{I}} \left(|f(x,v)| + \ell(\bar{I})\frac{|g(x,v)|}{|v|} \right) dx\,dv \tag{6.7}$$

and

$$b_I := b_I(f) := c_0 \int_\Omega \int_{I'} \left(|f(x,v)| + \ell(\bar{I})\frac{|g(x,v)|}{|v|} \right) dx\,dv, \tag{6.8}$$

where the constant $c_0 := 2^d \max(\|\phi\|^d_{L_\infty(\mathbb{R})}, \|\psi\|^d_{L_\infty(\mathbb{R})})$. It follows that

$$a_I(\bar{f}) \le \tilde{b}_I(f) \le b_I(f). \tag{6.9}$$

To prove that $(a_I(\bar{f}))_{I\in\mathcal{D}_+}$ is in weak-ℓ_1, it will be enough to estimate the number of elements in the set

$$\Lambda(f,\epsilon) := \{I \in \mathcal{D}^+ : \ \tilde{b}_I > \epsilon\}.$$

For this purpose, we define a cube $I \in \Lambda(f,\epsilon)$ to be *special* if

$$b_I < 2\mu\tilde{b}_I, \tag{6.10}$$

where the constant μ is defined by (6.2). We denote by $\Lambda^s(f,\epsilon)$ the collection of special cubes in $\Lambda(f,\epsilon)$. The following lemma gives an estimate for the cardinality of $\Lambda(f,\epsilon)$ in terms of special cubes.

Lemma 6.2. *Let f and $g \in L_1(\mathbb{R}^d \times \Omega)$ where Ω satisfies* (6.1). *Then for every $\epsilon > 0$ we have*

$$\#\Lambda(f,\epsilon) \le C \sum_{I\in\Lambda^s(f,\epsilon)} (\log \frac{\tilde{b}_I}{\epsilon})^d, \tag{6.11}$$

where C depends only on d. Here and later log *denotes logarithm to the base* 2.

Proof. Let us observe that for every cube $J \in \mathcal{D}^+$ we have

$$\begin{aligned} \tilde{b}_J &\le \frac{C}{M} \int_\Omega \int_{\bar{J}} |f(x,v)| + |g(x,v)|\, dxdv \\ &\le \frac{C}{M}(\|f\|_{L_1(\mathbb{R}^d\times\Omega)} + \|g\|_{L_1(\mathbb{R}^d\times\Omega)}), \end{aligned} \tag{6.12}$$

because $\ell(\bar{J}) \le C$ and $|v| \ge M$ for $v \in \Omega$.

If $J \in \Lambda(f,\epsilon)$ is not a special cube, we have that

$$b_J \ge 2\mu\tilde{b}_J. \tag{6.13}$$

It follows from (6.13) that there is a cube $\bar{J}_1$ with $|J_1| = |J|$, which is one of the cubes, participating in J' (see (6.2)), such that

$$\tilde{b}_{J_1} \ge 2\tilde{b}_J. \tag{6.14}$$

If J_1 is a special cube, we stop. If not, we repeat the same procedure for J_1 and get a cube J_2 ($J_2 \notin \{J, J_1\}$) such that

$$\tilde{b}_{J_2} \ge 2\tilde{b}_{J_1} \ge 2^2\tilde{b}_J,$$

where in the last inequality we use (6.14). This process will terminate after a finite number of steps, because from (6.12) it follows that on the n-th step we have

$$\frac{C}{M}(\|f\|_{L_1(\mathbb{R}^d\times\mathbb{R}^d)} + \|g\|_{L_1(\mathbb{R}^d\times\mathbb{R}^d)}) \ge \tilde{b}_{J_n} \ge 2^n\tilde{b}_J.$$

Let I be the special cube which terminates this sequence, i.e. $I = J_n$ and

$$\tilde{b}_I \ge 2^n\tilde{b}_J > 2^n\epsilon.$$

Hence we have

$$n < \log \frac{\tilde{b}_I}{\epsilon}. \tag{6.15}$$

Therefore, every $J \in \Lambda(f,\epsilon)$ is contained in a ball with center, the center of a special cube $I = I(J)$, and radius $\leq Cn\ell(J) \leq C\ell(J)\log\frac{\tilde{b}_I}{\epsilon}$ with C depending only on d. This gives

$$\#\Lambda(f,\epsilon) \leq C \sum_{I\in\Lambda^s(f,\epsilon)} (\log\frac{\tilde{b}_I}{\epsilon})^d,$$

and the proof is complete. □

To each special cube $I \in \Lambda^s(f,\epsilon)$ we associate an index $m(I) := m(I,f,\epsilon)$ as follows. Let J be the smallest special cube in $\Lambda(f,\epsilon)$, such that

$$\bar{J} \subseteq \bar{I}. \tag{6.16}$$

Note that $\tilde{b}_J$ tends to 0 as $|J| \to 0$ because $f, g \in L_1(\mathbb{R}^d \times \Omega)$. Therefore, a smallest J does exist. We define

$$m(I) := \log_2\left(\frac{|I|^{1/d}}{|J|^{1/d}}\right).$$

Then, $m(I)$ is an integer which represents the difference between the dyadic level of I and J. Recall that since J is a special cube, we have $\tilde{b}_J > \epsilon$ and $2\mu\tilde{b}_J > b_J$.

Lemma 6.3. *Let $f, g \in L_1(\mathbb{R}^d \times \Omega)$. If $I \in \Lambda^s(f,\epsilon)$ is a special cube, then*

$$\tilde{b}_I > C2^{m(I)}\epsilon, \tag{6.17}$$

where C is a constant, depending only on c_0, ℓ and d.

Proof. Let $I \in \Lambda^s(f,\epsilon)$ and let J be the smallest special cube in $\Lambda^s(f,\epsilon)$, such that $\bar{J} \subset \bar{I}$. Then $\ell(\bar{J})2^{m(I)} = \ell(\bar{I})$ and $\tilde{b}_J > \epsilon$. We multiply this last inequality by $2^{m(I)}$ and obtain

$$\frac{\ell(\bar{I})}{\ell(\bar{J})}c_0\int_\Omega\int_{\bar{J}} |f(x,v)|\,dxdv + c_0\int_\Omega\int_{\bar{J}} \ell(\bar{I})\frac{|g(x,v)|}{|v|}\,dxdv > 2^{m(I)}\epsilon.$$

Now we use (6.3) of Lemma 6.1 and derive

$$C\int_\Omega\int_{I'}\left[|f(x,v)| + \ell(\bar{I})\frac{|g(x,v)|}{|v|}\right]dxdv + c_0\int_\Omega\int_{\bar{J}} \ell(\bar{I})\frac{|g(x,v)|}{|v|}\,dxdv > 2^{m(I)}\epsilon.$$

Since $\bar{J} \subset \bar{I} \subset I'$, the above inequality gives

$$C\int_\Omega\int_{I'}\left[|f(x,v)| + \ell(\bar{I})\frac{|g(x,v)|}{|v|}\right]dxdv > 2^{m(I)}\epsilon.$$

But the cube I is special (see (6.10)) and therefore

$$\tilde{b}_I \geq (2\mu)^{-1}b_I = (2\mu)^{-1}c_0\int_\Omega\int_{I'}\left[|f(x,v)| + \ell(\bar{I})\frac{|g(x,v)|}{|v|}\right]dxdv > C2^{m(I)}\epsilon.$$

This proves the lemma. □

The next theorem gives us the weak-ℓ_1 estimate.

Theorem 6.4. *If $f \in L_1(\mathbb{R}^d \times \Omega)$, $g := v\cdot\nabla_x f \in L_1(\mathbb{R}^d \times \Omega)$ and Ω satisfies assumption* (6.1)*, then $\lambda(\bar{f})$ is in weak-ℓ_1 and*

$$\|\lambda(\bar{f})\|_{\ell_{1,\infty}} \leq \frac{C}{M}\left[\|f\|_{L_1(\mathbb{R}^d\times\Omega)} + \|g\|_{L_1(\mathbb{R}^d\times\Omega)}\right], \tag{6.18}$$

where C is a constant, depending only on c_0, ℓ and d.

Proof. Given any $\epsilon > 0$, we need to show that

$$\#\Lambda(f,\epsilon) \le \frac{C}{M\epsilon}\left[\|f\|_{L_1(\mathbb{R}^d\times\Omega)} + \|g\|_{L_1(\mathbb{R}^d\times\Omega)}\right]. \tag{6.19}$$

Lemma 6.2 gives that

$$\#\Lambda(f,\epsilon) \le C\sum_{I\in\Lambda^s(f,\epsilon)} (\log\frac{\tilde b_I}{\epsilon})^d = \frac{C}{\epsilon}\sum_{I\in\Lambda^s(f,\epsilon)} \frac{(\log\frac{\tilde b_I}{\epsilon})^d}{(\frac{\tilde b_I}{\epsilon})}\tilde b_I.$$

Since for $I\in\Lambda(f,\epsilon)$, we have $\tilde b_I > \epsilon$, there is an absolute constant C such that

$$\frac{(\log\frac{\tilde b_I}{\epsilon})^d}{(\frac{\tilde b_I}{\epsilon})} \le C\left(\frac{\tilde b_I}{\epsilon}\right)^{-1/2},$$

and therefore

$$\#\Lambda(f,\epsilon) \le \frac{C}{\epsilon}\sum_{I\in\Lambda^s(f,\epsilon)} \left(\frac{\tilde b_I}{\epsilon}\right)^{-1/2}\tilde b_I. \tag{6.20}$$

For each $I\in\Lambda^s(f,\epsilon)$, Lemma 6.3 gives

$$\left(\frac{\tilde b_I}{\epsilon}\right)^{-1/2} < C2^{-m(I)/2}. \tag{6.21}$$

Also, for $v\in\Omega$ and $I\in\mathcal{D}^+$, we have

$$\tilde b_I \le \frac{C}{M}\int_\Omega\int_{\mathbb{R}^d} (|f(x,v)| + |g(x,v)|)\chi_{\bar I}(x)\,dxdv. \tag{6.22}$$

When we combine (6.20), (6.21), and (6.22), we obtain

$$\#\Lambda(f,\epsilon) \le \frac{C}{\epsilon M}\int_\Omega\int_{\mathbb{R}^d} (|f(x,v)| + |g(x,v)|)K(x)\,dxdv,$$

where

$$K(x) := \sum_{I\in\Lambda^s(f,\epsilon)} 2^{-m(I)/2}\chi_{\overline{I}}(x).$$

We will use Lemma 2.2 to show that K is bounded, which will complete the proof. We let $\Gamma(i,f,\epsilon) := \Lambda^s(f,\epsilon)\cap\Gamma_i$, $i=1,\dots,\gamma$. Then, we can write

$$K = \sum_{i=1}^{\gamma} K_i,$$

where

$$K_i := \sum_{I\in\Gamma(i,f,\epsilon)} 2^{-m(I)/2}\chi_{\overline{I}},$$

and it is enough to show that each of the functions K_i is bounded. In other words, we need to show that for every $x\in\mathbb{R}^d$

$$\sum_{I\in\Gamma^*(i,f,\epsilon,x)} 2^{-m(I)/2} \le C, \quad i=1,2,\dots,\gamma, \tag{6.23}$$

where $\Gamma^*(i,f,\epsilon,x) := \{I\in\Gamma^*(i,f,\epsilon) :\ x\in\bar I\}$. To prove this inequality, we fix a

point x for which the sum in (6.23) is not zero. We make the following observations about the cubes appearing in this sum:

(i) There is a smallest cube $I^* \in \Gamma^*(i, f, \epsilon, x)$ (because $f,\ g \in L_1(\mathbb{R}^d \times \mathbb{R}^d)$).

(ii) There is at most one $I \in \mathcal{D}_j$, $j \geq 0$, in $\Gamma^*(i, f, \epsilon, x)$ because of the defining property of the sets Γ_i.

(iii) If $I \in \Gamma^*(i, f, \epsilon, x)$, then $m(I) \geq m(I^*) + q$, where q is the difference between the dyadic level of $\bar{I}$ and $\bar{I}^*$.

It follows from (i)–(iii) that

$$\sum_{I \in \Gamma^*(i,f,\epsilon,x)} 2^{-m(I)/2} \leq \sum_{n=0}^{\infty} 2^{-n/2} \leq C,$$

as desired. □

Notice that the smaller the M in (6.1), the bigger the constant $\frac{C}{M}$ in (6.18). However, one can allow M to go to zero if we restrict our attention only to certain wavelet coefficients as the following result shows.

We fix an arbitrary $n \geq 0$ and consider domains Ω such that

$$\Omega \subset \{v \in \mathbb{R}^d :\ 2^{-n} \leq |v| \leq M'\}, \quad n \geq 0. \tag{6.24}$$

For $f \in L_1(\mathbb{R}^d \times \Omega)$, let

$$\lambda_n(\bar{f}) := (a_I(\bar{f}))_{I \in \mathcal{D}_n^+},\ \text{with } \mathcal{D}_n^+ := \bigcup_{j \geq n} \mathcal{D}_j.$$

Corollary 6.5. *Let $f \in L_1(\mathbb{R}^d \times \Omega)$ and $g := v \cdot \nabla_x f \in L_1(\mathbb{R}^d \times \Omega)$ where Ω satisfies assumption* (6.24)*. Then, $\lambda_n(\bar{f})$ is in weak-ℓ_1 and*

$$\|\lambda_n(\bar{f})\|_{\ell_{1,\infty}} \leq C \left[\|f\|_{L_1(\mathbb{R}^d \times \Omega)} + \|g\|_{L_1(\mathbb{R}^d \times \Omega)} \right],$$

where C is an absolute constant.

Proof. This can be proved by dilation or alternatively by repeating the proof of Theorem 6.4 and using the fact that for $v \in \Omega$ and $I \in \mathcal{D}_n^+$, we have $\frac{\ell(\bar{I})}{|v|} \leq C$. □

Acknowledgement

The authors wish to thank Professor Eitan Tadmor for many useful discussions about averaging lemmas. They also thank Professors A. Cohen, W. Dahmen, and I. Daubechies for valuable conversations concerning the results of this paper which in particular are reflected in the proof of the interpolation theorem in §3.

References

1. J. Bergh and J. Löfström, *Interpolation Spaces*, Springer-Verlag, 1976. MR **58:**2349
2. M. Bezard, *Regularite L^p precisee des moyennes dans les equations de transport*, Bull. Soc. Math. France **22** (1994), 29–76. MR **95g:**82083
3. I. Daubechies, *Orthonormal bases of compactly supported wavelets*, Comm. Pure and Appl. Math. **41** (1988), 909–996. MR **90m:**42039
4. I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1992. MR **93e:**42045
5. R. DeVore, P. Petrushev, and X. Yu, *Wavelet approximation in the space $C(\mathbb{R}^d)$*, Progress in Approximation Theory, Springer Verlag, New York, 1992, 261–283. MR **94h:**41070
6. R. DeVore and R. Sharpley, *Besov spaces on domains in $\mathbb{R}^d$*, TAMS **335** (1993), 843–864. MR **93d:**46051

7. R. DiPerna, P-L. Lions, and Y. Meyer, *L_p regularity of velocity averages*, Ann. Inst. H. Poincaré Anal. Non Linéaire **8** (1991), 271–288. MR **92g:**35036
8. F. Golse, P-L. Lions, B. Perthame, and R. Sentis, *Regularity of the moments of the solution of a transport equation*, J. Funct. Anal. **76** (1988), 110–125. MR **89a:**35179
9. P-L. Lions, *Regularite optimale des moyennes en vitesses*, C. R. Acad. Sci. Paris **320** (1995), 911–915. MR **96c:**35184
10. Y. Meyer, *Ondelettes et Opérateurs*, Hermann, Paris, 1990. MR **93i:**42002
11. J. Peetre, *A Theory of Interpolation Spaces*, Notes, Universidade de Brasilia, 1963.
12. G. Petrova, *Transport Equations and Velocity Averages*, Ph.D. Thesis, University of South Carolina, 1999.

Department of Mathematics, University of South Carolina, Columbia, South Carolina 29208
E-mail address: devore@math.sc.edu

Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48109
E-mail address: petrova@math.lsa.umich.edu