Projections and nonlinear approximation in the space $BV(\mathbb{R}^d)$

P. Wojtaszczyk*
Institut for Applied Mathematics
Warsaw University
02-097 Warszawa ul. Banacha 2

June 30, 2002

1 Introduction

The aim of this paper is to provide an analysis of non-linear approximation in the L_p norm $p = d/(d-1)$ of functions of bounded variation on $\mathbb{R}^d$ with $d > 1$ by polynomials in the Haar system. The exponent p is the natural exponent as it is the correct exponent in the Sobolev inequality. The space of functions of bounded variation turned out to be quite useful in geometric measure theory, calculus of variation and differential equations, see [5]. More recently (but really only in the case $d = 2$) this space was used in the image processing, see e.g. [10].

The approximation schemes we are discussing in this paper are mostly related to Haar tresholding and m–term approximation. Those problems for $d = 2$ were studied in detail in [3]. The main aim of this paper is to extend results from [3] to the case $d \geq 2$. This task is justified by simple mathematical curiosity but also by the fact that minimization of various functionals related to variation

*This research was partially supported by KBN grant 5P03A 03620 located at the Institute of Mathematics of the Polish Academy of Sciences.

(also for dimension > 2) plays an important role in problems of differential equations (see [1]). This extension required two ingredients

1. To prove boundedness in $\mathrm{BV}(\mathbb{R}^d)$ of certain class of averaging operators.

2. To replace the approximation arguments in Hilbert space setting by approximation in L_p. For this we use (variants of) recent results on greedy approximation with respect to the Haar system (cf. [12], [15]).

The general structure of our arguments follows the pattern established in [3].

The main result of this paper is Theorem 7 which gives natural sufficient conditions on the sequence of sets $(K_j)_{j=0}$ to ensure that

$$Pf(x) = \begin{cases} \frac{1}{|K_j|}\int_{K_j} f(t)\,dt & \text{if } x \in K_j \\ f(x) & \text{if } x \notin \bigcup_{j=1} K_j \end{cases}$$

is a bounded projection in $\mathrm{BV}(\mathbb{R}^d)$. We use this result and Bernstein and Jackson type inequalities to show the stability of Haar tresholding in BV norm (see Theorem 10). We also show that Haar tresholding provides a near minimizer of certain non-linear functionals (see Section 7)

Acknowledgments. The author would like to express his gratitude to prof. Ron DeVore for calling his attention to the problem considered here and to prof. prof. P. Hajłasz, G. Kerkyacherian, A. Pełczyński and P. Strzelecki for useful (e-mail) conversations about the subject of this paper.

2 Basic properties of BV functions

Let $\Omega \subset \mathbb{R}^d$ be an open set. We say that a function $f \in L_1(\Omega)$ belongs to $\mathrm{BV}(\Omega)$ if $\mathsf{grad}\, f$, the distributional gradient of f, is a vector of measures of bounded variation on Ω. We define a semi-norm $|f|_{\mathrm{BV}(\Omega)} = \int_\Omega |\mathsf{grad}\, f|$. Here the symbol $\int_\Omega |\mathsf{grad}\ f|$ is a convenient shorthand but it has to be made precise. If $\mathsf{grad}\ f = (\mu_1, \ldots, \mu_d)$ where μ_i's are measures of finite variation, then we take a positive measure of finite variation μ such that $\mu_i = \varphi_i \mu$. Then by $|\mathsf{grad}\ f|$ we will mean the measure $\big(\sum_{i=1}^d |\varphi_i(x)|^2\big)^{1/2}\mu$ and we denote

$$\int_\Omega |\mathsf{grad}\ f| = \int_\Omega \big(\sum_{i=1}^d |\varphi_i(x)|^2\big)^{1/2}\, d\mu.$$

Clearly this does not depend on the particular choice of μ. An equivalent definition of $|f|_{\mathrm{BV}(\Omega)}$ is

$$|f|_{\mathrm{BV}(\Omega)} = \sup\big\{\int_\Omega f(x)\mathrm{div}\, g(x)\, dx\ :\ g \in C_0^1(\Omega, \mathbb{R}^d) \text{ and } |g(x)| \le 1\big\}.$$

It is important to remember that the quantity $|f|_{\mathrm{BV}(\Omega)}$ is a semi-norm, it vanishes on constant functions.

The following easy lemma will be used several times and will be referred to as *semicontinuity property* of BV norm. Its proof can be found in any book about BV functions, in particular [16, Th. 5.2.1] or [5, Th.1.9].

Lemma 1. *Let $\Omega \subset \mathbb{R}^d$ be an open set and let (f_j) be a sequence of functions in* $\mathrm{BV}(\Omega)$ *which converges in $L_1^{\mathrm{loc}}(\Omega)$ to f. Then*

$$\int_\Omega |\mathsf{grad}\ f| \le \liminf_j \int_\Omega |\mathsf{grad}\ f_j|.$$

Let us recall some geometric concepts which are of fundamental importance in the study of functions of bounded variation. Details can be found in [5, Chap. 2 and 3] and in [16].

We always work in $\mathbb{R}^d$ with dimension $d \ge 2$ and always $p = d/(d-1)$. Symbol ∂ denotes the boundary of the set i.e. $\partial K = \overline{K} \setminus \mathrm{Int} K$. It acts on the set immediately following it i.e. $\partial G \cap K$ means the intersection of boundary of G with the set K while $\partial(G \cap K)$ is used to denote the boundary of the intersection.

We will consider two natural measures on $\mathbb{R}^d$. The d-dimensional Lebesgue measure of a set A will be denoted by $|A|$ and the integral with respect to the Lebesgue measure will be denoted is $\int f\, dx$. Another measure which will be used extensively is the $(d-1)$–dimensional Hausdorff measure denoted $\mathcal{H}_{d-1}$. Definition and basic properties of Hausdorff measure can be found e.g. in [16, 1.4].

An open set $K \subset \mathbb{R}^d$ is said to have *Lipschitz boundary* if each point $x \in \partial K$ has a neighborhood U such that $U \cap K$ can be represented in some cartesian coordinate system by the inequality $x_d < f(x_1, \ldots, x_{d-1})$ where f is a Lipschitz function. For $f \in BV(K)$ we can define a trace. In order to explain this notion we will simply quote Theorem 2.10 from [5]. By $B(x, \rho)$ we will denote the open euclidian ball with center x and radius ρ.

Theorem 1. *Let K be a bounded open set in $\mathbb{R}^d$ with Lipschitz boundary ∂K and let $f \in \mathrm{BV}(K)$. Then there exists a function $\phi \in L_1(\partial K, \mathcal{H}_{d-1})$ such that for $\mathcal{H}_{d-1}$–almost all $x \in \partial K$*

$$\lim_{\rho \to 0} \rho^{-d} \int_{B(x,\rho)\cap K} |f(z) - \phi(x)|\, dz = 0. \tag{1}$$

Moreover, for every $g \in C_0^1(\mathbb{R}^d, \mathbb{R}^d)$,

$$\int_K f \mathrm{div} g\, dx = -\int_K \langle g, \mathsf{grad}\ f\rangle + \int_{\partial K} \phi\big(\langle g, \overrightarrow{n}\rangle\big) d\mathcal{H}_{d-1}(x) \tag{2}$$

where $\overrightarrow{n}$ is the unit outer normal to ∂K.

The function ϕ will be called a *trace* of f on ∂K. If we have a function $f \in \mathrm{BV}(\mathbb{R}^d)$ and a bounded open set K with Lipschitz boundary, then by *inner trace* of f on ∂K we will mean the trace of $f|K \in \mathrm{BV}(K)$. It is known (c.f. Theorem 5) that in that case $f|K$ really belongs to $\mathrm{BV}(K)$. We can also define an *outer trace* of f on ∂K as follows: Let B be a ball such that $\overline{K} \subset B$. Then $B \setminus \overline{K}$ is an open, bounded set with Lipschitz boundary and clearly $\partial(B \setminus \overline{K}) = \partial B \cup \partial K$ and this is a disjoint union. By the outer trace of f on ∂K we will mean the trace of $f|(B \setminus \overline{K})$ restricted to ∂K. The set E such that $\mathbf{1}_E \in \mathrm{BV}(\mathbb{R}^d)$ is called a set of *finite perimeter*. This means that $\int |\mathsf{grad}\ \mathbf{1}_E| < \infty$. Intuitively $\mathsf{grad}\ \mathbf{1}_E$ should live on the "boundary" of E. To make it precise, for a set E of finite perimeter we define its *reduced boundary* $\partial^* E$ (cf. [5, 3.3] or [16, 5.5.1]) as the set of points $x \in \mathbb{R}^d$ such that

$$\int_{B(x,\rho)} |\mathsf{grad}\ \mathbf{1}_E| > 0 \text{ for all } \rho > 0$$

and

$$\lim_{\rho \to 0} \left(\int_{B(x,\rho)} |\mathsf{grad}\ \mathbf{1}_E| \right)^{-1} \int_{B(x,\rho)} \mathsf{grad}\ \mathbf{1}_E = \nu_E(x)$$

exists and is a vector of length one. Clearly $\partial^* E \subset \partial E$. It is known (cf. Corollary 5.6.8 and Theorem 5.8.1 of [16]) that $|\mathsf{grad}\ \mathbf{1}_E| = \mathcal{H}_{d-1}|\partial^* E$ and $\mathsf{grad}\ \mathbf{1}_E = \nu_E(x) \cdot \mathcal{H}_{d-1}|\partial^* E$. If E has Lipschitz boundary then $\mathcal{H}_{d-1}(\partial E \setminus \partial^* E) = 0$.

The vector $\nu_E(x)$ plays a role of interior normal to E. To make it precise let us put

$$T^-(x) = \{z \in \mathbb{R}^d \ : \ \langle \nu_E(x), z - x\rangle < 0\}.$$

Then for $x \in \partial^* E$ we have (cf. [5, Th. 3.8] or [16, Th. 5.6.5])

$$\lim_{\rho \to 0} \rho^{-d} |E \cap B(x,\rho) \cap T^-(x)| = 0 \tag{3}$$

and

$$\lim_{\rho \to 0} \rho^{-d} |(B(x,\rho) \setminus E) \cap T^-(x)| = \tfrac{\alpha(d)}{2} \tag{4}$$

where $\alpha(d)$ denotes the volume of unit ball in $\mathbb{R}^d$. It is known (see [16, Cor. 5.6.8]) that the set of $x \in \mathbb{R}^d$ such that (3) and (4) hold with the set $T^-(x)$ replaced by some half-space, differs from $\partial^* E$ by a set of $\mathcal{H}_{d-1}$ measure 0.

Let us state the following easy properties of ∂^* which will be used without any comment in Section 5.

Proposition 2. *Let K be an open set with Lipschitz boundary and let A, B be open sets with finite perimeter such that $\mathbf{1}_K = \mathbf{1}_A + \mathbf{1}_B$ a.e. Then*

$$\begin{aligned} \partial^* A \cap K &= \partial^* B \cap K && (5)\\ \partial^* K \setminus \partial^* A &= \partial^* K \cap \partial^* B && (6)\\ \partial^* A &\subset K \cup \partial^* K && (7)\\ \big(\partial^* A \cap K\big) \cup \big(\partial^* K \setminus \partial^* A\big) &= \partial^* B && (8) \end{aligned}$$

where the equalities and inclusion are understood $\mathcal{H}_{d-1}$ almost everywhere.

Proof. All those facts follow easily from the equality

$$\nu_K(x) \cdot \mathcal{H}_{d-1} | \partial^* K = \nu_A(x) \cdot \mathcal{H}_{d-1} | \partial^* A + \nu_B(x) \cdot \mathcal{H}_{d-1} | \partial^* B.$$

We have to remember that $\nu_A(x)$ and $\nu_B(x)$ are vectors in $\mathbb{R}^d$ of length 1, and that $\mathcal{H}_{d-1}(\partial K \setminus \partial^* K) = 0$.

□

3 Discrete co-area decomposition

Let us recall the co-area formula (cf. [5] 1.23 or [16] 5.4.4)

Theorem 2. *Let $\Omega \subset \mathbb{R}^d$ be an open set and $u \in \mathrm{BV}(\Omega)$, then for $E_t = \{x : u(x) > t\}$ we have*

$$|u|_{\mathrm{BV}(\Omega)} = \int_{-\infty}^{\infty} |\mathbf{1}_{E_t}|_{\mathrm{BV}(\Omega)} \, dt. \tag{9}$$

Conversely if $u \in L_1(\Omega)$ *and* E_t *has finite perimeter in* Ω *for almost all* t *and* $\int_{-\infty}^{\infty} |\mathbf{1}_{E_t}|_{\mathrm{BV}(\Omega)}\, dt < \infty$ *then* $u \in \mathrm{BV}(\Omega)$.

Let us remark, that except for countably many t's, $\mathbf{1}_{E_t} = \mathbf{1}_{\{x\,:\,u(x)\geq t\}}$ up to a set of measure zero, so it represents the same element of $\mathrm{BV}(\Omega)$. Since for a function $u \in \mathrm{BV}(\Omega)$ we have

$$|\mathbf{1}_{\{x\,:\,u(x)\geq t\}}|_{\mathrm{BV}(\Omega)} = |\mathbf{1}_{\{x\,:\,u(x)<t\}}|_{\mathrm{BV}(\Omega)}$$

we see that in formula (9) instead of E_t we can use sets $\{x : u(x) \geq t\}$ or $\{x : u(x) < t\}$ as well.

Theorem 3. *For every* $f \in \mathrm{BV}(\Omega)$ *there exists a sequence of functions* $(g_n)_{n=1}^{\infty}$ *such that*

1. $\lim_{n\to\infty} |g_n|_{\mathrm{BV}(\Omega)} = |f|_{\mathrm{BV}(\Omega)}$
2. $\lim_{n\to\infty} g_n = f$ *in* $L_1(\Omega)$
3. $g_n = \sum_{j=1}^{\infty} a_j^n \mathbf{1}_{E_j^n}$ *where* E_j^n *are open subsets of* Ω *with finite perimeter and* C^∞ *boundary in* Ω *and* $\sum_{j=1}^{\infty} |a_j^n| |\mathbf{1}_{E_j^n}|_{\mathrm{BV}(\Omega)} \leq |f|_{\mathrm{BV}(\Omega)}$ *for* $n = 1, 2, \ldots$.

This Theorem is a very natural discretization of the co-area formula. Its proof not only uses the co-area formula itself but also uses many arguments from its proof.

Proof. From standard regularization results (see e.g. [5, Th. 1.17] or [16, 5.3.3]) we know that there exists a sequence f_n of bounded C^∞ functions such that $\lim f_n = f$ in $L_1(\Omega)$ and $\lim |f_n|_{\mathrm{BV}(\Omega)} = |f|_{\mathrm{BV}(\Omega)}$. Thus it clearly suffices to prove our Theorem under additional assumption that f is a bounded C^∞ function. Now let us consider the function

$$\varphi(t) = \begin{cases} |\mathbf{1}_{\{x\,:\,f(x)>t\}}|_{\mathrm{BV}(\Omega)} & \text{if } t > 0 \\ |\mathbf{1}_{\{x\,:\,f(x)<t\}}|_{\mathrm{BV}(\Omega)} & \text{if } t < 0. \end{cases}$$

From our assumption follows that φ has compact support, so the sums in what follows are really finite. Let us consider $t > 0$. If $t_n \nearrow t$ then $\{x : f(x) > t_n\} \supset \{x : f(x) > t\}$ and $\bigcap_{n=1}^{\infty} \{x : f(x) > t_n\} = \{x : f(x) \geq t\}$. Thus for all except a countable number of t's we have that $\mathbf{1}_{\{x\,:\,f(x)>t_n\}}$ converges in $L_1(\Omega)$ to $\mathbf{1}_{\{x\,:\,f(x)>t\}}$. By the semicontinuity property (Lemma 1) we get $\liminf \varphi(t_n) \geq \varphi(t)$. Analogously if $t_n \searrow t$ we get $\liminf \varphi(t_n) \geq \varphi(t)$ for all except countably many t's. In those exceptional points we modify φ and define $\varphi(t_0) = \liminf_{t \to t_0} \varphi(t)$. We argue analogously for $t < 0$ and we see that $\varphi(t)$ agrees with a lower semicontinuous function except on a countable set of points.

Note that by the implicit function theorem if $\mathsf{grad}\, f$ does not assume the value 0 on the set $\{x : f(x) = t\}$ then this set is a $(d-1)$-dimensional C^∞ manifold in Ω and

$$\partial\{x : f(x) > t\} = \partial\{x : f(x) < t\} = \{x : f(x) = t\}. \tag{10}$$

By Sard's Theorem the set $\{t : \exists x \in \Omega : \mathsf{grad}\, f(x) = 0 \text{ and } f(x) = t\}$ has Lebesgue measure zero. We also know by the classical Baire theorem that each lower semicontinous function is a pointwise limit of an increasing sequence of continuous functions.

Now let us fix a positive $\epsilon > 0$. Using previous remarks we find a continuous function $\psi(t)$ such that $0 \leq \psi(t) \leq \varphi(t)$ for $t \in \mathbb{R}$ and

$$\int_{-\infty}^{\infty} \varphi(t)\, dt \geq \int_{-\infty}^{\infty} \psi(t)\, dt \geq \int_{-\infty}^{\infty} \varphi(t)\, dt - \tfrac{\epsilon}{10}. \tag{11}$$

Thus we can find a sequence of points (partition) $\cdots < t_{-2} < t_{-1} < t_0 =$

$0 < t_1 < t_2 < \dots$ such that $(t_{j+1} - t_j)$ does not depend on j and

$$\sum_{-\infty}^{\infty} (t_{j+1} - t_j) \inf_{t \in [t_j, t_{j+1}]} \psi(t) \geq \int_{-\infty}^{\infty} \psi(t)\, dt - \tfrac{\epsilon}{10}. \tag{12}$$

From (11) and (12) we see that

$$\int_{-\infty}^{\infty} \varphi(t)\, dt \geq \sum_{-\infty}^{\infty} (t_{j+1} - t_j) \inf_{t \in [t_j, t_{j+1}]} \varphi(t) \geq \int_{-\infty}^{\infty} \varphi(t)\, dt - \tfrac{\epsilon}{5}. \tag{13}$$

Using (13) and the observations above we can find $\bar{t}_j \in (t_j, t_{j+1})$ such that

$$\int_{-\infty}^{\infty} \varphi(t)\, dt \geq \sum_{-\infty}^{\infty} (t_{j+1} - t_j) \varphi(\bar{t}_j) \geq \int_{-\infty}^{\infty} \varphi(t)\, dt - \tfrac{\epsilon}{4} \tag{14}$$

and

$$(t_0 - t_{-1}) \varphi(\bar{t}_{-1}) + (t_1 - t_0) \varphi(\bar{t}_0) \leq \int_{t_{-1}}^{t_1} \varphi(t)\, dt \tag{15}$$

and such that $\bar{t}_j$ satisfy (10) for $j = 0, \pm 1, \pm 2, \dots$.

Now we define the function

$$g = \sum_{j=1}^{\infty} (t_j - t_{j-1}) \mathbf{1}_{\{x \,:\, f(x) > \bar{t}_j\}} + \sum_{-\infty}^{-1} (t_j - t_{j+1}) \mathbf{1}_{\{x \,:\, f(x) < \bar{t}_{j-1}\}}.$$

Since f is continous and integrable it is an easy exercise to show that if the partition $(t_j)_{-\infty}^{\infty}$ is fine enough then $\|f - g\|_{L_1(\Omega)} < \epsilon$ and

$$\int_{t_{-1}}^{t_1} \varphi(t)\, dt < \tfrac{\epsilon}{10}. \tag{16}$$

We see directly from the definition of g that

$$\begin{aligned} \{x \,:\, g(x) > t\} &= \{x \,:\, f(x) > \bar{t}_j\} \text{ if } t \in (t_{j-1}, t_j) \text{ and } j = 1, 2, \dots \\ \{x \,:\, g(x) < t\} &= \{x \,:\, f(x) < \bar{t}_{j-1}\} \text{ if } t \in (t_j, t_{j+1}) \text{ and } j = -1, -2, \dots. \end{aligned}$$

So we get from Theorem 2

$$
\begin{aligned}
|g|_{\mathrm{BV}(\Omega)} &= \int_0^\infty |\mathbf{1}_{\{x\,:\,g(x)>t\}}|_{\mathrm{BV}(\Omega)}\,dt + \int_{-\infty}^0 |\mathbf{1}_{\{x\,:\,g(x)<t\}}|_{\mathrm{BV}(\Omega)}\,dt \\
&= \sum_{j=1}^{\infty}(t_j - t_{j-1})|\mathbf{1}_{\{x\,:\,f(x)>\bar{t}_j\}}|_{\mathrm{BV}(\Omega)} + \sum_{-\infty}^{-1}(t_{j+1} - t_j)|\mathbf{1}_{\{x\,:\,f(x)<\bar{t}_{j-1}\}}|_{\mathrm{BV}(\Omega)} \\
&= \sum_{j=1}^{\infty}(t_j - t_{j-1})\varphi(\bar{t}_j) + \sum_{-\infty}^{-1}(t_{j+1} - t_j)\varphi(\bar{t}_{j-1}) \\
&= \sum_{j=1}^{\infty}(t_{j+1} - t_j)\varphi(\bar{t}_j) + \sum_{-\infty}^{-2}(t_{j+1} - t_j)\varphi(\bar{t}_j) \\
&= \sum_{-\infty}^{\infty}(t_{j+1} - t_j)\varphi(\bar{t}_j) - (t_0 - t_{-1})\varphi(\bar{t}_{-1}) - (t_1 - t_0)\varphi(\bar{t}_0)
\end{aligned}
$$

From the above, and (14), (15) and (16) we get

$$
|g|_{\mathrm{BV}(\Omega)} \geq \int_{-\infty}^{\infty} \varphi(t)\,dt - \tfrac{\epsilon}{4} - \tfrac{\epsilon}{10} \geq |f|_{\mathrm{BV}(\Omega)} - \epsilon
$$

and

$$
|g|_{\mathrm{BV}(\Omega)} \leq |f|_{\mathrm{BV}(\Omega)}.
$$

This completes the proof of the Theorem.

□

Corollary 3. *Suppose that $T : L_1(\Omega) \longrightarrow L_1(\Omega)$ is an operator such that for some constant C and for every set E with bounded perimeter and C^∞ boundary we have*

$$
|T\mathbf{1}_E|_{\mathrm{BV}(\Omega)} \leq C|\mathbf{1}_E|_{\mathrm{BV}(\Omega)}.
$$

Then T is a continuous operator on $\mathrm{BV}(\Omega)$ with norm at most C.

Proof. Given $f \in \mathrm{BV}(\Omega)$ let g_n be the sequence given by Theorem 3. Then using the triangle inequality and 3. of Theorem 3 we get $|Tg_n|_{\mathrm{BV}(\Omega)} \leq C|f|_{\mathrm{BV}(\Omega)}$.

Clearly Tg_n converges to Tf in $L_1(\Omega)$. Using the semicontinuity property of $|.|_{BV}$ (Lemma 1) we get

$$|Tf|_{\mathrm{BV}(\Omega)} \leq \liminf_{n\to\infty} |Tg_n|_{\mathrm{BV}(\Omega)} \leq C|f|_{\mathrm{BV}(\Omega)}.$$

$\square$

4 Geometric conditions

Let us first recall the classical isoperimetric inequality

$$|G|^{1/p} \leq C(d)\mathcal{H}_{d-1}(\partial^* G) \tag{17}$$

for all $G \subset \mathbb{R}^d$ of finite perimeter. The best constant $C(d)$ equals $d^{-1}\alpha(d)$ where $\alpha(d)$ is the volume of d–dimensional unit ball. One easily checks that $C(d)$ is bounded in d.

Let us consider the following geometric property of set $K \subset \mathbb{R}^d$:

K is a bounded open connected set with Lipschitz boundary and there exists a constant C such that for every open set $G \subset K$ with finite perimeter and $|G| \leq \frac{1}{2}|K|$ we have (18)

$$\mathcal{H}_{d-1}(\partial^* G) \leq C\mathcal{H}_{d-1}(\partial^* G \cap K).$$

Remark: It follows directly from Theorem 2.11 and Theorem 1.17 from [5] that it suffices to assume (18) for sets G with C^∞ boundary in K.

Theorem 4. *Let K be a set which satisfies (18) with constant C.*

1. *There exists constant $C_1 = C_1(C,d)$ such that*

$$|G|^{1/p} \leq C_1\mathcal{H}_{d-1}(\partial^* G \cap K) \tag{19}$$

for every open set with finite perimeter $G \subset K$ with $|G| \leq \frac{1}{2}|K|$.

2. There exists constant $C_2 = C_2(C,d)$ such that

$$\mathcal{H}_{d-1}(\partial^* K) = |\mathbf{1}_K|_{\mathrm{BV}} \le C_2|K|^{1/p} \tag{20}$$

Proof. Condition 1. follows directly from (18) and isoperimetric inequality (17) applied to the set G. The proof of 2. is more involved. For each axis direction e_i, $i = 1,\dots,d$ in $\mathbb{R}^d$ we define a function

$$\varphi^i(t) = \mathcal{H}_{d-1}(\{x \in K \,:\, \langle x, e_i\rangle = t\}).$$

From Fubini theorem we obtain

$$|K| = \int_{-\infty}^{\infty} \varphi^i(t)\,dt.$$

For each i we now define two numbers (the value of $\beta > 0$ is not important right now, later we will fix $\beta = 2^{1/d}$)

$$\begin{aligned} m_i &= \sup\{t \,:\, \varphi^i(t) \le \beta|K|^{1/p} \text{ and } \int_{-\infty}^{t} \varphi^i(s)\,ds \le \tfrac{1}{2}|K|\} \\ d_i &= \inf\{t \,:\, \varphi^i(t) \le \beta|K|^{1/p} \text{ and } \int_{t}^{\infty} \varphi^i(s)\,ds \le \tfrac{1}{2}|K|\}. \end{aligned}$$

If for some $i = 1,\dots d$ we have

$$\int_{m_i}^{d_i} \varphi^i(t)\,dt \le \tfrac{1}{2}|K| \tag{21}$$

then we apply (18) for the following three sets

$$\begin{aligned} G^1 &= \{x \in K \,:\, \langle x, e_i\rangle < m_i\} \\ G^2 &= \{x \in K \,:\, \langle x, e_i\rangle > d_i\} \\ G^3 &= \{x \in K \,:\, m_i < \langle x, e_i\rangle < d_i\}. \end{aligned}$$

Summing the resulting inequalities we get

$$\begin{aligned} \mathcal{H}_{d-1}(\partial^* K) &\le \sum_{j=1}^{3} \mathcal{H}_{d-1}(\partial^* G^j) \le C \sum_{j=1}^{3} \mathcal{H}_{d-1}(\partial^* G^j \cap K) \\ &\le 2C\big(\varphi^i(m_i) + \varphi^i(d_i)\big) \le 4C\beta|K|^{1/p} \end{aligned} \tag{22}$$

Now assume that (21) does not hold for any i. Then surely $\varphi^i(t) \geq \beta |K|^{1/p}$ for $m_i < t < d_i$ and $i = 1, \dots d$. Thus

$$\beta |K|^{1/p}(d_i - m_i) \leq \int_{m_i}^{d_i} \varphi^i(t)\, dt \leq |K|$$

so

$$(d_i - m_i) \leq \tfrac{1}{\beta} |K|^{1/d}. \tag{23}$$

Now consider the following subsets of K

$$\begin{aligned} G_i^1 &= \{x \in K \,:\, \langle x, e_i \rangle > d_i\} \\ G_i^2 &= \{x \in K \,:\, \langle x, e_i \rangle < m_i\} \end{aligned}$$

for $i = 1, \dots, d$ and also

$$R = \{x \in K \,:\, m_i < \langle x, e_i \rangle < d_i \text{ for } i = 1, \dots, d\}.$$

we have $|R| \leq \prod_{i=1}^d (d_i - m_i) \leq \beta^{-d}|K|$ so for $\beta = 2^{1/d}$ we get $|R| \leq \frac{1}{2}|K|$. Also sets G_i^1 and G_i^2 have measure at most $\frac{1}{2}|K|$. So using (18) we obtain

$$\begin{aligned} \mathcal{H}_{d-1}(\partial^* K) &\leq \mathcal{H}_{d-1}(\partial^* R) + \sum_{i=1}^d \mathcal{H}_{d-1}(\partial^* G_i^1) + \sum_{i=1}^d \mathcal{H}_{d-1}(\partial^* G_i^2) \\ &\leq C\big(\mathcal{H}_{d-1}(\partial^* R \cap K) + \sum_{i=1}^d \mathcal{H}_{d-1}(\partial^* G_i^1 \cap K) + \sum_{i=1}^d \mathcal{H}_{d-1}(\partial^* G_i^2 \cap K)\big). \end{aligned}$$

Clearly $\mathcal{H}_{d-1}(\partial^* G_i^1 \cap K) = \varphi^i(d_i)$ and $\mathcal{H}_{d-1}(\partial^* G_i^2 \cap K) = \varphi^i(m_i)$. Also using (23) we get

$$\mathcal{H}_{d-1}(\partial^* R \cap K) \leq \mathcal{H}_{d-1}(\partial^* \bigotimes_{i=1}^d (m_i, d_i)) \leq 2d\beta^{-(d-1)}|K|^{1/p}.$$

Thus

$$\mathcal{H}_{d-1}(\partial^* K) \leq C\big(2d\beta|K|^{1/p} + 2d\beta^{-(d-1)}|K|^{1/p}\big) \leq \tfrac{6d}{\sqrt{2}} C |K|^{1/p}$$

so we get 2. with $C_2 = \frac{6d}{\sqrt{2}}C$. $\square$

As an application of Theorem 3 let us show that on sets satisfying (18) Poincare inequalities hold with constants depending on C. Let f be a function defined (at least) on the set K of finite measure. Below we are using the standard notation $f_K = \frac{1}{|K|}\int_K f$.

Proposition 4. *Suppose that K satisfies condition (18) with constant C. Then there exists a constant $C' = C'(C,d)$ such that for every $f \in \mathrm{BV}(K)$ we have*

$$\|f - f_K\|_p \leq C' \int_K |\mathsf{grad}\ f|. \tag{24}$$

Proof. Let $G \subset K$ be an open set with C^∞ boundary in K. Then for $f = \mathbf{1}_G$ the left hand side of (24) becomes

$$\begin{aligned} L &= \Big(\big(1 - \tfrac{|G|}{|K|}\big)^p |G| + \Big(\tfrac{|G|}{|K|}\Big)^p (|K| - |G|)\Big)^{1/p} \\ &= |K|^{1/p}\Big(\big(\tfrac{|K|-|G|}{|K|}\big)^p \tfrac{|G|}{|K|} + \big(\tfrac{|G|}{|K|}\big)^p \tfrac{|K|-|G|}{|K|}\Big)^{1/p}. \end{aligned}$$

Since the right hand side of (24) equals $\mathcal{H}_{d-1}(\partial^* G \cap K)$ we see that if (24) holds for $\mathbf{1}_G$ it also holds (with the same constant) for $\mathbf{1}_{K\setminus G}$, thus we will assume that $|G| \leq \frac{1}{2}|K|$. Since we have the obvious inequality

$$(1-t)^p t + t^p(1-t) \leq 2t \ \text{ for } \ 0 \leq t \leq \tfrac{1}{2}$$

we infer that $L \leq 2^{1/p}|G|^{1/p}$. From (19) we get

$$L \leq 2^{1/p} C(d)\mathcal{H}_{d-1}(\partial^* G) \leq 2^{1/p} C(d) C \mathcal{H}_{d-1}(\partial^* G \cap K) \tag{25}$$

so (24) holds for $\mathbf{1}_G$ with uniform constant.

Now let f be arbitrary and let (g_n) be the sequence given by Theorem 3. Clearly

$$\begin{aligned} \|g_n - (g_n)_K\|_p &\leq \sum_j |a_j^n| \|\mathbf{1}_{E_j^n} - (\mathbf{1}_{E_j^n})_K\|_p \\ &\leq C \sum_j |a_j^n| |\mathbf{1}_{E_j^n}|_{\mathrm{BV}} \leq C|f|_{\mathrm{BV}}. \end{aligned} \tag{26}$$

Since g_n converges to f in $L_1(K)$ the sequence $g_n - (g_n)_K$ converges to $f - f_K$ in $L_1(K)$. From (26) we see that $g_n - (g_n)_K$ is a bounded sequence in $L_p(K)$ so it has a weakly convergent subsequence. The limit of this subsequence must equal $f - f_K$ so we get $f - f_K \in L_p(K)$ and

$$\|f - f_K\|_p \leq \liminf \|g_n - (g_n)_K\|_p \leq C|f|_{\mathrm{BV}}$$

which proves (24). □

Both Proposition 4 and its proof are basically known. It is known (cf. [8, § 3.2] or [6]) that (19) implies (and actually is equivalent to) the Poincare inequality with control of the constants. The proofs use co-area formula. Conditions on K ensuring that various versions of (24) hold are investigated in detail in the literature (see e.g. [6]). I was unable to decide if there exists a function K such that if (19) holds with the constant C then (18) holds with the constant $K(C)$.

5 Projections

In this Section we prove boundedness in the space $BV(\mathbb{R}^d)$ of some projections.

Proposition 5. *If $K \subset \mathbb{R}^d$ is a convex body, then the map $f \mapsto f\mathbf{1}_K$ is a norm one projection in the space $BV(\mathbb{R}^d)$.*

Proof. From Theorem 3 we see that it suffices to assume that $f = \mathbf{1}_E$ where E is a bounded set with C^∞ boundary. Since K is an intersection of halfspaces it suffices to check that for each such set E and each halfspace L we have

$$\mathcal{H}_{d-1}(\partial^*(L \cap E)) \leq \mathcal{H}_{d-1}(\partial^* E),$$

but this is geometrically obvious. □

This Proposition is surely well known. We give the simple proof of this Proposition since it is of some interest to us that such projection has norm 1. The following Theorem describes *projection sets* i.e. sets F such that the map $f \mapsto f\mathbf{1}_F$ is a bounded projection in $\mathrm{BV}(\mathbb{R}^d)$.

Theorem 5 ([9] Section 1.8). *The set F is a projection set if and only if F has locally finite perimeter and for some constant C' and each ball $B(x,\rho)$ we have*

$$\mathcal{H}_{d-1}(B(x,\rho)\cap\partial^*F) \le C'\rho^{d-1} \tag{27}$$

It follows immediately from Theorem 5 that every set with Lipschitz boundary is a projection set and the norm of the projection depends on the Lipschitz constant of the boundary. Also if the set F satisfies (18) with a constant C then the projection given by $\mathbf{1}_F$ has norm depending only on C. To see it let A equals $B(x,\rho)\cap F$ or $F\setminus B(x,\rho)$ in such a way that $|A|\le\frac{1}{2}|F|$. From (18) we get

$$\mathcal{H}_{d-1}(\partial^*A) \le C\mathcal{H}_{d-1}(\partial^*A\cap F). \tag{28}$$

Since

$$\partial^*F\cap B(x,\rho)\subset\partial^*A\subset\partial^*B(x,\rho)\cup\partial^*F$$

from (28) we get

$$\mathcal{H}_{d-1}(\partial^*F\cap B(x,\rho)) \le C\mathcal{H}_{d-1}(\partial^*B(x,\rho)\cap F) \le C'\rho^{d-1}.$$

Theorem 6. *Suppose that $(K_j)_{j=0}$ is a (finite or infinite) sequence of disjoint regular sets satisfying condition (18) with uniform constant C. Assume also that for all integers N we have*

$$\sum_{j\ge N}\mathcal{H}_{d-1}(\partial^*K_j) \le C\mathcal{H}_{d-1}(\partial^*K_N). \tag{29}$$

Then there exists a constant K=K(C,d) such that

$$\sum_{j=0}\left(\frac{1}{|K_j|}\left|\int_{K_j} f\right|\right)|\mathbf{1}_{K_j}|_{\mathrm{BV}} \leq K|f|_{\mathrm{BV}} \tag{30}$$

Proof. By Corollary 3 it suffices to consider $f = \mathbf{1}_E$ where E has C^∞ boundary and finite perimeter. Then

$$\tfrac{1}{|K_j|}\left|\int_{K_j} f\right||\mathbf{1}_{K_j}|_{\mathrm{BV}} = \tfrac{|E\cap K_j|}{|K_j|}\mathcal{H}_{d-1}(\partial^* K_j).$$

We split the sum in (30) into three sums as follows

(α) $|E\cap K_j| > \frac{1}{2}|K_j|$

(β) $|E\cap K_j| \leq \frac{1}{2}|K_j|$ and $\mathcal{H}_{d-1}(\partial^* E\cap K_j) \geq \mathcal{H}_{d-1}(\partial^* K_j)$

(γ) $|E\cap K_j| \leq \frac{1}{2}|K_j|$ and $\mathcal{H}_{d-1}(\partial^* E\cap K_j) < \mathcal{H}_{d-1}(\partial^* K_j)$

In the first case we have using first (30) and (20) and next the isoperimetric inequality (17) and the Remark after Theorem 5 that

$$\begin{aligned}
\sum_{j\in(\alpha)}\frac{|E\cap K_j|}{|K_j|}\mathcal{H}_{d-1}(\partial^* K_j) &\leq \sum_{j\in(\alpha)}\mathcal{H}_{d-1}(\partial^* K_j)\\
&\leq C\mathcal{H}_{d-1}(\partial^* K_{j_o}) \leq C^2|K_{j_o}|^{1/p}\\
&\leq 2^{1/p}C^2|E\cap K_{j_o}|^{1/p}\\
&\leq 2C^2C(d)\mathcal{H}_{d-1}(\partial^*(E\cap K_{j_o}))\\
&= 2C^2C(d)|\mathbf{1}_{E\cap K_{j_o}}|_{\mathrm{BV}} \leq 2C^3C(d)|\mathbf{1}_E|_{\mathrm{BV}}\\
&= 2C^3C(d)|f|_{\mathrm{BV}}.
\end{aligned}$$

In the second case we use disjointness of K_j's and we get

$$\begin{aligned}
\sum_{j\in(\beta)}\frac{|E\cap K_j|}{|K_j|}\mathcal{H}_{d-1}(\partial^* K_j) &\leq \sum_{j\in(\beta)}\frac{|E\cap K_j|}{|K_j|}\mathcal{H}_{d-1}(\partial^* E\cap K_j)\\
&\leq \tfrac{1}{2}\sum_{j\in(\beta)}\mathcal{H}_{d-1}(\partial^* E\cap K_j)\\
&\leq \tfrac{1}{2}\mathcal{H}_{d-1}(\partial^* E) = \tfrac{1}{2}|f|_{\mathrm{BV}}
\end{aligned}$$

In the third case we use (17), (18), (20) and the fact that K_j's are disjoint and get

$$\begin{aligned}
&\sum_{j\in(\gamma)} \frac{|E\cap K_j|}{|K_j|}\mathcal{H}_{d-1}(\partial^* K_j) \le \sum_{j\in(\gamma)} \mathcal{H}_{d-1}(\partial^*(E\cap K_j))^p \frac{\mathcal{H}_{d-1}(\partial^* K_j)}{|K_j|} \\
= \;& C^p \sum_{j\in(\gamma)} \mathcal{H}_{d-1}(\partial^* E\cap K_j)\mathcal{H}_{d-1}(\partial^* E\cap K_j)^{1/(d-1)} \frac{\mathcal{H}_{d-1}(\partial^* K_j)}{|K_j|} \\
\le \;& C^p \sum_{j\in(\gamma)} \mathcal{H}_{d-1}(\partial^* E\cap K_j) \frac{\mathcal{H}_{d-1}(\partial^* K_j)^p}{|K_j|} \\
\le \;& C' \sum_{j\in(\gamma)} \mathcal{H}_{d-1}(\partial^* E\cap K_j) \\
\le \;& C'\mathcal{H}_{d-1}(\partial^* E) = C'|f|_{\mathrm{B}V}
\end{aligned}$$

Putting the above three cases together we obtain the claim.

□

The following is immediate by Theorem 6.

Corollary 6. *Suppose that sets $(K_j)_{j=0}$ satisfy assumptions of Theorem 6. The averaging projection*

$$P(f) = \sum_{j=0} \Big(\tfrac{1}{|K_j|}\int_{K_j} f\Big)\mathbf{1}_{K_j} \tag{31}$$

is a bounded projection in the space $BV(\mathbb{R}^d)$ and its range is isomorphic to ℓ_1. More precisely the sequence $\big(|\mathbf{1}_{K_j}|^{-1}_{\mathrm{B}V}\mathbf{1}_{K_j}\big)_{j=0}$ is equivalent to the unit vector basis in ℓ_1.

Theorem 7. *Let $(K_j)_{j=0}$ be a (finite or infinite) sequence of disjoint sets satisfying (18) with uniform constants. Then the operator P defined by*

$$Pf(x) = \begin{cases} \frac{1}{|K_j|}\int_{K_j} f(x)\,dx & \text{if } x\in K_j \\ f(x) & \text{if } x\notin \bigcup_{j=0} K_j \end{cases} \tag{32}$$

is a continuous projection in the space $BV(\mathbb{R}^d)$.

For the proof we will need some lemmas.

Lemma 7. *Let K be a set satisfying (18). Then for any set with finite perimeter $G \subset K$ we have*

$$\int_{\partial^* K} \left|\mathbf{1}_{\partial^* G}(x) - \tfrac{|G|}{|K|}\right| d\mathcal{H}_{d-1}(x) \leq C\mathcal{H}_{d-1}(\partial^* G \cap K). \tag{33}$$

Proof. We rewrite (33) as

$$\left|1 - \tfrac{|G|}{|K|}\right| \mathcal{H}_{d-1}(\partial^* G \cap \partial^* K) + \tfrac{|G|}{|K|}\mathcal{H}_{d-1}(\partial^* K \setminus \partial^* G) \leq C\mathcal{H}_{d-1}(\partial^* G \cap K). \tag{34}$$

This inequality is clearly invariant under dilations so we can assume $|K| = 1$ and then (34) becomes

$$\big(1 - |G|\big)\mathcal{H}_{d-1}(\partial^* G \cap \partial^* K) + |G|\mathcal{H}_{d-1}(\partial^* K \setminus \partial^* G) \leq C\mathcal{H}_{d-1}(\partial^* G \cap K) \tag{35}$$

Let us remark that if (35) holds for G then it also holds for $K \setminus G$ so we can assume $|G| \geq \frac{1}{2}$. From (19) applied to $(K \setminus G)$ we get

$$|K \setminus G|^{1/p} \leq C\mathcal{H}_{d-1}(\partial^* G \cap K) \tag{36}$$

so

$$\big(1 - |G|\big) \leq C^p \mathcal{H}_{d-1}(\partial^* G \cap K)^p. \tag{37}$$

Applying (18) to the set $K \setminus G$ we obtain

$$\mathcal{H}_{d-1}(\partial^*(K \setminus G)) \leq C\mathcal{H}_{d-1}(\partial^*(K \setminus G) \cap K) = C\mathcal{H}_{d-1}(\partial^* G \cap K) \tag{38}$$

Thus if we denote the left hand side of (35) by L, using (37) we obtain

$$L \leq C^p \mathcal{H}_{d-1}(\partial^* G \cap K)^p \cdot \mathcal{H}_{d-1}(\partial^* G \cap \partial^* K) + |G|\mathcal{H}_{d-1}(\partial^* K \setminus \partial^* G) \tag{39}$$

from (20) we get $\mathcal{H}_{d-1}(\partial^* G \cap \partial^* K) \leq \mathcal{H}_{d-1}(\partial^* K) \leq C|K|^{1/p} = C$ so

$$L \leq C' \mathcal{H}_{d-1}(\partial^* G \cap K)^p + \mathcal{H}_{d-1}(\partial^* K \setminus \partial^* G). \tag{40}$$

If $\mathcal{H}_{d-1}(\partial^* G \cap K) \leq 1$ we continue

$$L \leq C' \mathcal{H}_{d-1}(\partial^* G \cap K) + \mathcal{H}_{d-1}(\partial^* K \setminus \partial^* G) \tag{41}$$

$$\leq C' \mathcal{H}_{d-1}(\partial^* (K \setminus G)) \tag{42}$$

and from (38) we get

$$\leq C'' \mathcal{H}_{d-1}(\partial^* G \cap K). \tag{43}$$

Thus (35) holds under the additional assumption $\mathcal{H}_{d-1}(\partial^* G \cap K) \leq 1$. On the other hand if $\mathcal{H}_{d-1}(\partial^* G \cap K) \geq 1$ then (35) also holds since always $L \leq \mathcal{H}_{d-1}(\partial^* K)$ and this is bounded by C from (20). $\square$

Proposition 8. *Let us fix a set K satisfying assumptions of Theorem 7 and let $f \in \mathrm{BV}(\mathbb{R}^d)$ assumes finite number of values and $f|K = \mathbf{1}_G$ for a set $G \subset K$ with C^∞ boundary in K. Let us define a function f_o by the formula*

$$f_o(x) = \begin{cases} \frac{1}{|K|} \int_K f(t)\, dt & \text{if } x \in K \\ f(x) & \text{if } x \in \mathbb{R}^d \setminus K. \end{cases} \tag{44}$$

Then

$$|f_o|_{\mathrm{BV}} \leq |f|_{\mathrm{BV}} + C \int_K |\mathsf{grad}\ f| = |f|_{\mathrm{BV}} + C \mathcal{H}_{d-1}(\partial^* G \cap K). \tag{45}$$

Proof. Obviously $\mathsf{grad}\ f = \mathsf{grad}\ f_o$ on $\mathbb{R}^d \setminus \overline{K}$, so it suffices to work on $\overline{K}$. Let φ denotes the outer trace of f on $\partial^* K$. Also let $\xi(x)$ denotes the inner trace of $\mathbf{1}_G$ on $\partial^* G$ i.e. $\xi(x) = \mathbf{1}_{\partial^* G}$. Since $\mathsf{grad}\ f_o$ is zero on K it suffices to show

$$\int_{\partial^* K} \left|\varphi(x) - \tfrac{|G|}{|K|}\right| d\mathcal{H}_{d-1}(x) \leq \int_{\partial^* K} \left|\varphi(x) - \xi(x)\right| d\mathcal{H}_{d-1}(x) + C\mathcal{H}_{d-1}(\partial^* G \cap K). \tag{46}$$

But

$$\begin{aligned} \int_{\partial^* K} \left|\varphi(x) - \tfrac{|G|}{|K|}\right| d\mathcal{H}_{d-1}(x) \quad &\leq \quad \int_{\partial^* K} \left|\varphi(x) - \xi(x)\right| d\mathcal{H}_{d-1}(x) \\ &\quad + \int_{\partial^* K} \left|\xi(x) - \tfrac{|G|}{|K|}\right| d\mathcal{H}_{d-1}(x) \end{aligned}$$

so it suffices to show

$$\int_{\partial^* K} \left|\xi(x) - \tfrac{|G|}{|K|}\right| d\mathcal{H}_{d-1}(x) \leq C\mathcal{H}_{d-1}(\partial^* G \cap K). \tag{47}$$

This immediately follows from Lemma 7.

$\square$

From Proposition 8 Theorem 7 follows by an easy induction. We define operators $P_j f$ by the formula

$$P_j f(x) = \begin{cases} f(x) & \text{if } x \notin K_j \\ \frac{1}{|K_j|}\int_{K_j} f & \text{if } x \in K_j \end{cases} \tag{48}$$

and we put $f_o = P_o f$ and then inductively $f_{j+1} = P_{j+1} f_j$ for $f = \mathbf{1}_E$ where E is a bounded set with C^∞ boundary. From Proposition 8 we get

$$|f_{j+1}|_{\mathrm{BV}} \leq |f_j|_{\mathrm{BV}} + C\int_{K_{j+1}} |\mathsf{grad}\ f_j|$$

since K_j's are open and disjoint

$$= |f_j|_{\mathrm{BV}} + C\int_{K_{j+1}} |\mathsf{grad}\ f| = |P_j f_{j-1}|_{\mathrm{BV}} + C\int_{K_{j+1}} |\mathsf{grad}\ f|$$

once more from Proposition 8

$$\leq |f_{j-1}|_{\mathrm{BV}} + C\int_{K_j} |\mathsf{grad}\ f| + C\int_{K_{j+1}} |\mathsf{grad}\ f|$$

and continuing in this manner

$$\leq |f|_{\mathrm{BV}} + C\sum_{s=0}^{j+1}\int_{K_s} |\mathsf{grad}\ f|$$

because K_j's are disjoint we get

$$\leq (C+1)|f|_{\mathrm{BV}}.$$

Since $f_j \to Pf$ in L_1 we get $|Pf|_{\mathrm{BV}} \leq (C+1)|f|_{\mathrm{BV}}$. An appeal to Corollary 3 completes the proof of Theorem 7.

Now we would like to investigate what subsets K of $\mathbb{R}^d$ have property (18). We will need the following

Lemma 9. *Let $K \subset \mathbb{R}^d$ be the open unit cube. There exists a constant C such that for each set $E \subset K$ with C^∞ boundary in K and $|E| \leq \frac{1}{2}$ we have*

$$|E| \leq C\mathcal{H}_{d-1}(\partial^* E \cap K)^p. \tag{49}$$

For direct proof of this lemma see [2, Theorem 2.3]. In general it is known that such inequality for a set K follows (with the same constant) from Poincare inequality for the set K. That the Poincare inequality is valid for the cube is classical (cf. [16, Th. 5.12.7], see also [6, Corollary 5]) and is known in much greater generality.

Proposition 10. *There exists a constant C depending only on dimension d such that if K is a cube in $\mathbb{R}^d$ and $G \subset K$ has a C^∞ boundary in K and $|G| \leq \frac{1}{2}|K|$ then*

$$\mathcal{H}_{d-1}(\partial^* G) \leq C\mathcal{H}_{d-1}(\partial^* G \cap K) \tag{50}$$

so the cube has property (18) with constant C.

Proof. Clearly equation (50) is invariant under dilations, so we can assume $|K| = 1$. We may also note that for the cube the extended boundary equals the boundary up to a set of $\mathcal{H}_{d-1}$ measure 0. We will consider two cases:

If $\mathcal{H}_{d-1}(\partial^* G \cap K) \geq 1$ then

$$\begin{aligned}
\mathcal{H}_{d-1}(\partial^* G) &= \mathcal{H}_{d-1}(\partial^* G \cap \partial^* K) + \mathcal{H}_{d-1}(\partial^* G \cap K) \\
&\leq \mathcal{H}_{d-1}(\partial^* K) + \mathcal{H}_{d-1}(\partial^* G \cap K) \leq 2^d + \mathcal{H}_{d-1}(\partial^* G \cap K) \\
&\leq (2^d + 1)\mathcal{H}_{d-1}(\partial^* G \cap K).
\end{aligned}$$

If $\mathcal{H}_{d-1}(\partial^* G \cap K) \leq 1$ then for an open face S of the cube K let P_S denotes the orthogonal projection from the cube onto this face. If a point $x \in P_S(G)$ then either there is a point $y \in \partial^* G \cap K = \partial G \cap K$ such that $P_S(y) = x$ or the whole interval of length 1 which is perpendicular to S and joins x with the face opposite S is contained in G. From this we infer that $\mathcal{H}_{d-1}(P_S(G)) \leq \mathcal{H}_{d-1}(\partial^* G \cap K) + |G|$. Since clearly $\mathcal{H}_{d-1}(\partial^* G \cap S) \leq \mathcal{H}_{d-1}(P_S(G))$, we get

$$\mathcal{H}_{d-1}(\partial^* G \cap S) \leq \mathcal{H}_{d-1}(\partial^* G \cap K) + |G|. \tag{51}$$

Now we estimate $|G|$ using Lemma 9 and we get

$$\begin{aligned}
\mathcal{H}_{d-1}(\partial^* G \cap S) &\leq \mathcal{H}_{d-1}(\partial^* G \cap K) + C(d)\mathcal{H}_{d-1}(\partial^* G \cap K)^p \\
&\leq (1 + C(d))\mathcal{H}_{d-1}(\partial^* G \cap K).
\end{aligned}$$

If we sum this estimate over all faces we get

$$\mathcal{H}_{d-1}(\partial^* G \cap \partial^* K) \leq 2^d (C(d) + 1)\mathcal{H}_{d-1}(\partial^* G \cap K)$$

which clearly yields

$$\mathcal{H}_{d-1}(\partial^* G) \leq \left[2^d (C(d) + 1) + 1\right]\mathcal{H}_{d-1}(\partial^* G \cap K).$$

$\square$

A *dyadic ring* is a set of the form $R = \mathrm{Int}\ (K_1 \setminus K_2)$ where K_1 and K_2 are dyadic cubes and $K_2 \not\subseteq K_1$. We allow $K_2 = \emptyset$ so a dyadic cube is also a (degenerate) dyadic ring.

Proposition 11. *There exists a constant $K = K(d)$ such that each dyadic ring R satisfies (18) with constant K.*

Proof. Let us take G as in (18). We will prove the inequality

$$\mathcal{H}_{d-1}(\partial^* G \cap R) \geq \alpha \mathcal{H}_{d-1}(\partial^* G \cap \partial^* K_2). \tag{52}$$

First consider the case $|K_2| \geq 2|G|$. Let $Q_1, \ldots, Q_N$ denote all dyadic subcubes of K_1 different from K_2 and such that $|Q_j| = |K_2|$ and K_2 and Q_j have common face F_j. Clearly $\bigcup_{j=1}^N F_j = \partial^* K_2 \cap K_1$. Depending on the position of K_2 in K_1 the number N can vary but always $2^{d-1} \leq N \leq 2^d$. Clearly for each j

$$|G \cap Q_j| \leq |G| \leq \tfrac{1}{2}|K_2| = \tfrac{1}{2}|Q_j|$$

so from Proposition 10 we get

$$\mathcal{H}_{d-1}(\partial^* G \cap Q_j) \geq \tfrac{1}{C}\mathcal{H}_{d-1}(\partial^*(G \cap Q_j)) \geq \tfrac{1}{C}\mathcal{H}_{d-1}(\partial^* G \cap F_j).$$

Summing this over j we get

$$\begin{aligned}
\mathcal{H}_{d-1}(\partial^* G \cap R) \geq & \sum_{j=1}^N \mathcal{H}_{d-1}(\partial^* G \cap Q_j) \geq \tfrac{1}{C}\sum_{j=1}^N \mathcal{H}_{d-1}(\partial^* G \cap F_j) \\
= & \quad 2^{d-1}C^{-1}\mathcal{H}_{d-1}(\partial^* G \cap \partial^* K_2).
\end{aligned}$$

Now consider the case $|K_2| < 2|G|$ and let $Q_1, \ldots Q_N$ denote all dyadic subcubes of K_1 different from K_2 and such that $|Q_j| = |K_2|$ for all j.

If there exists a cube Q_{j_o} such that $|Q_{j_o} \cap G| > \frac{1}{2}|Q_{j_o}|$ then using the fact that $|G| \leq \frac{1}{2}|R|$ by continuity we can find a cube Q (not necessarily dyadic) such that $Q \subset R$, $|Q| = |K_2|$ and $|G \cap Q| = \frac{1}{2}|Q|$. Using Lemma 9 we obtain

$$\begin{aligned}
\mathcal{H}_{d-1}(\partial^* G \cap Q) &\geq C^{-1}|Q \cap G|^{1/p} = C^{-1}2^{-1/p}|Q|^{1/p} \\
&= C^{-1}2^{-1/p}2^{-d}\mathcal{H}_{d-1}(\partial^* Q) = C^{-1}2^{-1/p}2^{-d}\mathcal{H}_{d-1}(\partial^* K_2) \\
&\geq C^{-1}2^{-1/p}2^{-d}\mathcal{H}_{d-1}(\partial^* G \cap \partial^* K_2)
\end{aligned}$$

so (52) holds in this situation.

If on the other hand for all cubes Q_j we have $|Q_j \cap G| \le \frac{1}{2}|Q_j|$ then Lemma 9 gives

$$\mathcal{H}_{d-1}(\partial^* G \cap Q_j) \ge C^{-1}|G \cap Q_j|^{1/p}.$$

Summing this over j we get

$$\begin{aligned}
\mathcal{H}_{d-1}(\partial^* G \cap R)^p &\ge \Big(\sum_{j=1}^{N} \mathcal{H}_{d-1}(\partial^* G \cap Q_j)\Big)^p \\
&\ge C^{-p}\Big(\sum_{j=1}^{N} |G \cap Q_j|^{1/p}\Big)^p \\
&\ge C^{-p}\sum_{j=1}^{N} |G \cap Q_j| = C^{-p}|G| \\
&\ge \tfrac{1}{2}C^{-p}|K_2| = \tfrac{1}{2}C^{-p}2^{-dp}\mathcal{H}_{d-1}(\partial^* K_2)^p \\
&\ge \tfrac{1}{2}C^{-p}2^{-dp}\mathcal{H}_{d-1}(\partial^* K_2 \cap \partial^* G)^p
\end{aligned}$$

so

$$\mathcal{H}_{d-1}(\partial^* G \cap R) \ge 2^{-1/p}C^{-1}2^{-d}\mathcal{H}_{d-1}(\partial^* K_2 \cap \partial^* G)$$

and the proof of (52) is finished with the constant $\alpha = 2^{-1/p}C^{-1}2^{-d}$.

In order to show that (52) implies (18) we consider G as a subset of K_1 and from Proposition 10 we get

$$\begin{aligned}
\mathcal{H}_{d-1}(\partial^* G) &\le C\mathcal{H}_{d-1}(\partial^* G \cap K_1) \\
&= C\Big(\mathcal{H}_{d-1}(\partial^* G \cap R) + \mathcal{H}_{d-1}(\partial^* G \cap \partial^* K_2)\Big) \\
&\le C\Big(\mathcal{H}_{d-1}(\partial^* G \cap R) + \alpha^{-1}\mathcal{H}_{d-1}(\partial^* G \cap \partial^* R)\Big) \\
&= C(1+\alpha^{-1})\mathcal{H}_{d-1}(\partial^* G \cap R).
\end{aligned}$$

□

REMARK: It is easy to see that we need some assumptions on sets $(K_j)_{j=0}$ for Theorem 7 to hold. Let us consider the following examples:

1. If we consider one of the sets $K_n = \{(x,y) : 0 < x < 1 \text{ and } \frac{1}{2}\sin nx < y < 2\} \subset \mathbb{R}^2$ we easily see that for big n the norm the projection P corresponding to this set K_n can be arbitrarily big, simply because $|\mathbf{1}_{K_n}|_{BV} \geq n$. If K is a set such that $\mathbf{1}_K \notin \mathrm{BV}$ then the corresponding projection is clearly unbounded in BV.
2. Consider the rectangle $K_\alpha = \{(x,y) : 0 < x < \alpha \text{ and } 0 < y < 1/\alpha\} \subset \mathbb{R}^2$ with $\alpha \geq 1$. The set K_α satisfies (18) but with big constant when α is big. However the corresponding projections P_α have uniformly (in α) bounded norms. To see it let us take a set G with C^∞ boundary and finite perimeter. It is geometrically clear and follows from Proposition 5 that $\mathcal{H}_{d-1}(\partial^*(G \cap K_\alpha)) \leq \mathcal{H}_{d-1}(\partial^* G)$. Also from Green's formula $|K_\alpha \cap G| \leq C\mathcal{H}_{d-1}(\partial^*(G \cap K_\alpha))\frac{1}{\alpha}$. Thus
$$\begin{aligned} |P_\alpha(\mathbf{1}_G)|_{BV} &= |G \cap K_\alpha| 2(\alpha + \tfrac{1}{\alpha}) \\ &\leq C\mathcal{H}_{d-1}(\partial^*(G \cap K_\alpha))\tfrac{2}{\alpha}(\alpha + \tfrac{1}{\alpha}) \\ &\leq C'|\mathbf{1}_{G \cap K_\alpha}|_{BV} \leq C|\mathbf{1}_G|_{BV}. \end{aligned}$$
3. For a fixed N and $n = 1, 2, \ldots, N$ we consider the following sequence of sets:
$$K_n = \{(x,y) : \tfrac{n-1}{N} < y < \tfrac{n}{N} \text{ and } -\tfrac{1}{2} < (-1)^n x < N - \tfrac{1}{2}\}.$$
From the above we know that projections corresponding to individual sets K_n have uniformly bounded norm. Also each set satisfies (18) but with the constant increasing with N. Let us take $G = (-\frac{1}{2}, \frac{1}{2}) \times (0,1)$. Clearly $|\mathbf{1}_G|_{BV} = 4$. On the other hand if we apply projection P corresponding to the sequence $(K_n)_{n=1}^N$ to $\mathbf{1}_G$ we get $|P\mathbf{1}_G|_{BV} \geq \frac{1}{N}N(2N-2)$ so we get $\|P\| \geq \frac{N-1}{2}$.

The following is an immediate Corollary to Theorem 7 and Proposition 11.

Corollary 12. *Suppose that $(K_j)_{j=1}^\infty$ is a partition of $\mathbb{R}^d$ into (essentially disjoint) dyadic rings (in particular dyadic cubes) and let P denotes the averaging projection with respect to $(K_j)_{j=1}^\infty$. Then $|Pf|_{BV} \leq C|f|_{BV}$.*

In particular for $f \in \mathrm{BV}(\mathbb{R}^d)$ let f_n denotes the average of f with respect to partition of $\mathbb{R}^d$ into dyadic cubes of side length 2^{-n}. Then we have the following version of Theorem 3: *for every $f \in \mathrm{BV}(\mathbb{R}^d)$ the sequence f_n satisfies*

1. $\lim_{n\to\infty} f_n = f$ *in* L_1
2. $\lim_{n\to\infty} |f_n|_{BV} \leq C|f|_{BV}$
3. $f_n = \sum_j a_j^n \mathbf{1}_{E_j^n}$ *where E_j^n are finite unions of disjoint dyadic cubes (each of side length 2^{-n}) and* $\sum_j |a_j^n||\mathbf{1}_{E_j^n}|_{BV} = |f_n|_{BV} \leq C|f|_{BV}$

6 Approximation by Haar expansion

Our first aim is to recall the definition of the Haar system on $\mathbb{R}^d$. We start with two functions on $\mathbb{R}$ defined as $h^0 = \mathbf{1}_{[0,1]}$ and

$$h^1(t) = \begin{cases} 1 & \text{if } t \in [0, 1/2) \\ -1 & \text{if } t \in [1/2, 1) \\ 0 & \text{if } t \notin [0, 1) \end{cases} \tag{53}$$

For $\epsilon = (\epsilon_1, \ldots, \epsilon_d)$ with $\epsilon_j = 0, 1$ and $\sum_{j=1}^d \epsilon_j > 0$ (the set of all such ϵ's will be denoted by $\mathcal{E}$) we define

$$h^\epsilon(t_1, \ldots, t_d) = h^{\epsilon_1}(t_1) \cdots h^{\epsilon_d}(t_d). \tag{54}$$

For $j \in \mathbb{Z}$ and $k \in \mathbb{Z}^d$ we define

$$h^\epsilon_{j,k}(t) = 2^{j(d-1)} h^\epsilon(2^j t - k) \tag{55}$$

and

$$\widetilde{h^\epsilon_{j,k}}(t) = 2^j h^\epsilon(2^j t - k) \tag{56}$$

where $t \in \mathbb{R}^d$. It is easy to check that $h^\epsilon_{j,k}$ with $\epsilon \in \mathcal{E}$, $j \in \mathbb{Z}$ and $k \in \mathbb{Z}^d$ is a complete orthogonal system in $L_2(\mathbb{R}^d)$ and that $|h^\epsilon_{j,k}|_{\mathrm{BV}} \sim$ const. Let us denote $\mathfrak{S} = \mathcal{E} \times \mathbb{Z} \times \mathbb{Z}^d$ and we will denote a generic element of $\mathfrak{S}$ by $\mathfrak{s}$. To shorten the notation we will denote $h^\epsilon_{j,k}$ by $h_\mathfrak{s}$ and $\langle f, \widetilde{h^\epsilon_{j,k}} \rangle$ by $c_\mathfrak{s}(f)$.

For each $f \in L_1(\mathbb{R}^d)$ (in particular for each $f \in \mathrm{BV}(\mathbb{R}^d)$) we have the (formal) expansion

$$f \sim \sum_{(\epsilon,j,k) \in \mathfrak{S}} \langle f, \widetilde{h^\epsilon_{j,k}} \rangle h^\epsilon_{j,k} = \sum_{\mathfrak{s} \in \mathfrak{S}} c_\mathfrak{s}(f) h_\mathfrak{s}. \tag{57}$$

It is well known (cf. e.g. [13, 8.3]) that the Haar system is an *unconditional basis* in L_p for $1 < p < \infty$. This means that there exists a constant $C = C(p)$ such that for each $f \in L_p$ the series $\sum_{\mathfrak{s} \in \mathfrak{S}} c_\mathfrak{s}(f) h_\mathfrak{s}$ converges to f in the L_p norm and for each bounded sequence of numbers $(\lambda_{\mathfrak{s}})_{\mathfrak{s}\in\mathfrak{S}}$ the series $\sum_{\mathfrak{s}\in\mathfrak{S}}\lambda_{\mathfrak{s}}c_{\mathfrak{s}}(f)h_{\mathfrak{s}}$ is an expansion of another L_p function and we have

$$\Big\|\sum_{\mathfrak{s}\in\mathfrak{S}}\lambda_{\mathfrak{s}}c_{\mathfrak{s}}(f)h_{\mathfrak{s}}\Big\|_p \le C\|f\|_p \sup_{\mathfrak{s}\in\mathfrak{S}}|\lambda_{\mathfrak{s}}|. \tag{58}$$

For general information about unconditional bases in Banach spaces see e.g. [14].

Now we define three families of non-linear spaces. The family

$$\Sigma_N^h = \Big\{\sum_{\mathfrak{s}\in E} a_{\mathfrak{s}}h_{\mathfrak{s}} \,:\, E\subset\mathfrak{S} \text{ and } |E|\le N\Big\} \tag{59}$$

is associated with the N–term approximation by Haar system. Let $\mathcal{D}$ denotes the family of all dyadic cubes in $\mathbb{R}^d$. We define a family of piecewise constant functions

$$\Sigma_N^c = \Big\{\sum_{I\in E} a_I\mathbf{1}_I \,:\, E\subset\mathcal{D} \text{ and } |E|\le N\Big\}. \tag{60}$$

For a given $Q\in\mathcal{D}$ we form the family of spaces

$$\Sigma_N^r(Q) = \Big\{\sum_{R\in\mathcal{P}} c_R\mathbf{1}_R\Big\} \tag{61}$$

where $\mathcal{P}$ is a set of at most N dyadic rings that form a partition of Q.

The nonlinearity of those spaces is controlled, namely we have $\Sigma_N^h+\Sigma_M^h\subset\Sigma_{N+M}^h$ and $\Sigma_N^c+\Sigma_M^c\subset\Sigma_{N+M}^c$ and also $\Sigma_N^r(Q)+\Sigma_M^r(Q)\subset\Sigma_{a(N+M)}^r(Q)$ for some a. Those spaces are related namely $\Sigma_N^h\subset\Sigma_{2^dN}^c$ and $\Sigma_N^r\subset\Sigma_{2N}^c$.

We will also define two non–linear operators related to Haar expansion. For $\epsilon>0$ we put

$$\mathcal{H}_\epsilon(f) = \sum_{|c_{\mathfrak{s}}(f)|>\epsilon} c_{\mathfrak{s}}(f)h_{\mathfrak{s}} \tag{62}$$

and for $N=1,2,\dots$ we put

$$\mathcal{G}_N(f) = \sum_{\mathfrak{s}\in E} c_{\mathfrak{s}}(f)h_{\mathfrak{s}} \tag{63}$$

where $E \subset \mathfrak{S}$ is any subset of cardinality N such that $|c_{\mathfrak{s}}(f)| \geq |c_{\mathfrak{l}}(f)|$ whenever $\mathfrak{s} \in E$ and $\mathfrak{l} \in \mathfrak{S} \setminus E$.

Note also that our normalization ensures that $\|h_{\mathfrak{s}}\|_p \sim$ const. The following Lemma is basically known.

Lemma 13. *There exist constants $0 < c < C$ (depending on p) such that for each finite set $E \subset \mathfrak{S}$ we have*

$$c|E|^{1/p} \leq \Big\| \sum_{\mathfrak{s} \in E} h_{\mathfrak{s}} \Big\|_p \leq C|E|^{1/p}. \tag{64}$$

The upper estimate is just a type of the space L_p (see [14, III.A.23]) and the lower estimate for $d = 1$ is proved in [15, Lemma 9] and [12]. The proof in our case is the same. From Lemma 13 and the unconditionality of the Haar system it follows (see [7] or [15, Theorem 4]) that the Haar system is a *greedy basis* in L_p, $1 < p < \infty$. This means that there exists a constant $C = C(p)$ such that for each f and N we have

$$\|f - \mathcal{G}_N f\|_p \leq C \inf \big\{ \|f - h\|_p \,:\, h \in \Sigma_N^h \big\}. \tag{65}$$

Now let us recall the fundamental theorem proved in [3] (see also [2]).

Theorem 8 ([3]). *There exists a constant C such that for all $f \in \mathrm{BV}(\mathbb{R}^d)$ and for every $\lambda > 0$*

$$\big| \big\{ \mathfrak{s} \in \mathfrak{S} \,:\, |c_{\mathfrak{s}}(f)| > \lambda \big\} \big| \leq C\lambda^{-1} |f|_{\mathrm{BV}} \tag{66}$$

i.e. the sequence of Haar coefficients $\{c_{\mathfrak{s}}(f)\}$ belongs to the space weak-$\ell_1(\mathfrak{S})$.

Proposition 14. *If $f \in \mathrm{BV}(Q)$ (Q is a dyadic cube) and have mean zero or $f \in \mathrm{BV}(\mathbb{R}^d)$ then*

$$\|f - \mathcal{H}_\epsilon f\|_p \leq C\epsilon^{1/d} |f|_{\mathrm{BV}}^{1/p} \tag{67}$$

and

$$\|f - \mathcal{G}_N f\|_p \leq C N^{-1/d} |f|_{\mathrm{BV}}. \tag{68}$$

Proof. Let us recall (cf. e.g. [13, 8.3]) that the Haar system is an unconditional basis in L_p. Thus from Theorem 8 we infer that

$$\|f - \mathcal{G}_N f\|_p \leq C \Big\| \sum_{n \geq N} \tfrac{1}{n} h_{\pi(n)} \Big\|_p |f|_{\mathrm{BV}} \tag{69}$$

where π is an appropriate 1–1 map from $\mathbb{N}$ into $\mathfrak{S}$. Without loss of generality (but with slight loss of the constant) we can assume $N = 2^\mu$. Using unconditionality of the Haar system we get

$$\begin{aligned}
\Big\| \sum_{n \geq N} \tfrac{1}{n} h_{\pi(n)} \Big\|_p &= \Big\| \sum_{j=\mu}^{\infty} \sum_{n=2^j}^{2^{j+1}-1} \tfrac{1}{n} h_{\pi(n)} \Big\|_p \\
&\leq \sum_{j=\mu}^{\infty} \Big\| \sum_{n=2^j}^{2^{j+1}-1} \tfrac{1}{n} h_{\pi(n)} \Big\|_p \\
&\leq C \sum_{j=\mu}^{\infty} 2^{-j} \Big\| \sum_{n=2^j}^{2^{j+1}-1} h_{\pi(n)} \Big\|_p .
\end{aligned}$$

Using Lemma 13 we can majorise the above by

$$C \sum_{j=\mu}^{\infty} 2^{-j} 2^{j/p} = C N^{-1/d}.$$

Thus we get (68). In order to prove (67) let us note that $\mathcal{H}_\epsilon f = \mathcal{G}_{N(\epsilon)} f$ for appropriate integer $N(\epsilon)$. From Theorem 8 we infer that $N(\epsilon) \leq C \epsilon^{-1} |f|_{BV}$, thus using (68) we obtain

$$\|f - \mathcal{H}_\epsilon f\|_p = \|f - \mathcal{G}_{N(\epsilon)} f\|_p \leq C N(\epsilon)^{-1/d} |f|_{\mathrm{BV}} \leq C \epsilon^{1/d} |f|_{\mathrm{BV}}^{1/p}.$$

□

REMARK: One may note that the order in the estimate (68) is optimal. It follows from the observation that for functions $h_{\mathfrak{s}}$ whose supports are far apart we have $|\sum a_{\mathfrak{s}} h_{\mathfrak{s}}|_{\mathrm{BV}} \sim \sum |a_{\mathfrak{s}}|$ and $\|\sum a_{\mathfrak{s}} h_{\mathfrak{s}}\|_p \sim \big(\sum |a_{\mathfrak{s}}|^p\big)^{1/p}$.

Proposition 15. *If $f \in \Sigma_N^c$ then we have*

$$|f|_{\mathrm{BV}} \leq CN^{1/d}\|f\|_p. \tag{70}$$

Proof. Let us start with some notation. If $F \subset \mathcal{D}$ is a finite set of dyadic cubes then for $I \in F$ by $B(I)$ we denote the set of maximal elements of the set $\{J \in F \,:\, J \subset I \text{ and } J \neq I\}$. We write $f \in \Sigma_N^c$ as $f = \sum_{I \in F} c_I \mathbf{1}_I$ with $E \subset \mathcal{D}$ and $|E| \leq N$. From Lemma 4.1 of [4] we get that there exists a set $F \subset \mathcal{D}$ such that $E \subset F$, $|F| \leq 2^d N$ and $|B(I)| \leq 2^d$ for all $I \in F$. For $I \in F$ we write $I' = I \setminus \bigcup_{J \in B(I)} J$. Functions $\mathbf{1}_{I'}$ with $I \in F$ have disjoint supports and we can write

$$f = \sum_{I \in F} d_I \mathbf{1}_{I'}. \tag{71}$$

Let us prove the following estimate

$$|\mathbf{1}_{I'}|_{\mathrm{BV}} \leq C\|\mathbf{1}_{I'}\|_p \tag{72}$$

where as C we can take $3 \cdot 2^{3d-2}(2^d-1)^{1/p}$. Clearly

$$\begin{aligned} |\mathbf{1}_{I'}|_{\mathrm{BV}} &\leq |\mathbf{1}_I|_{\mathrm{BV}} + \sum_{J \in B(I)} |\mathbf{1}_J|_{\mathrm{BV}} \\ &\leq \mathcal{H}_{d-1}(\partial^* I) + 2^d \mathcal{H}_{d-1}(\partial^*(\tfrac{1}{2}I)) = 3\mathcal{H}_{d-1}(\partial^* I). \end{aligned}$$

On the other hand

$$|I'| \geq |I| - (2^d-1)|\tfrac{1}{2}I| - |\tfrac{1}{4}I| = 4^{-d}(2^d-1)|I|.$$

Since $\mathcal{H}_{d-1}(\partial^* I) = 2^d|I|^{1/p}$ we get

$$\begin{aligned} \|\mathbf{1}_{I'}\|_p &= |I'|^{1/p} \geq \big(4^{-d}(2^d-1)|I|\big)^{1/p} \\ &= 2^{-d}\big(4^{-d}(2^d-1)\big)^{1/p}\mathcal{H}_{d-1}(\partial^* I) \geq \tfrac{1}{3}2^{-d}\big(4^{-d}(2^d-1)\big)^{1/p}|\mathbf{1}_{I'}|_{\mathrm{BV}} \end{aligned}$$

which is (72). Using (72) and Hölder's inequality we obtain

$$
\begin{aligned}
|f|_{\mathrm{BV}} &\leq \sum_{I\in F} |d_I||\mathbf{1}_{I'}|_{\mathrm{BV}} \leq C\sum_{I\in F} |d_I|\|\mathbf{1}_{I'}\|_p \\
&\leq C|F|^{1/d}\Big(\sum_{I\in F}|d_I|^p\|\mathbf{1}_{I'}\|_p^p\Big)^{1/p} \\
&\leq C2|E|^{1/d}\|f\|_p.
\end{aligned}
$$

□

Corollary 16. *If $f\in\Sigma_N^h$ then we have*

$$|f|_{\mathrm{BV}} \leq CN^{1/d}\|f\|_p. \tag{73}$$

Since $\Sigma_N^h\subset\Sigma_{2^dN}^c$ this follows immediately from Proposition 15.

Corollary 17. *If $f=\sum_{\mathfrak{s}\in E} c_{\mathfrak{s}}h_{\mathfrak{s}}$ and $|c_{\mathfrak{s}}|\geq t>0$ for all $\mathfrak{s}\in E$ then*

$$|f|_{\mathrm{BV}} \leq Ct^{-\frac{1}{d-1}}\|f\|_p^p. \tag{74}$$

Proof. Using the unconditionality of the Haar system and Lemma 13 we get

$$\|f\|_p \geq C\|\sum_{\mathfrak{s}\in E} th_{\mathfrak{s}}\|_p \geq Ct|E|^{1/p}$$

so

$$|E| \leq Ct^{-p}\|f\|_p^p.$$

Thus from (73) we get

$$|f|_{\mathrm{BV}} \leq Ct^{-p/d}\|f\|_p^{p/d}\|f\|_p = Ct^{\frac{-1}{d-1}}\|f\|_p^p.$$

□

Now we want to provide a direct estimate of approximation by elements from Σ_N^r. The main element in the proof is the algorithm choosing an appropriate partition. We have the following result whose proof can be obtained by repeating the proof of [3, Theorem 6.1]. Let us fix a dyadic cube Q (we can as well assume that it is a unit cube). Let $\mathcal{A}$ denotes the algebra of sets (not a σ-algebra) generated by dyadic subcubes of Q. Clearly dyadic rings are in $\mathcal{A}$. Let $\Phi : \mathcal{A} \to \mathbb{R}^+$ be a function satisfying

a) If $K_1, K_2 \in \mathcal{A}$ are disjoint, then

$$\Phi(K_1) + \Phi(K_2) \leq \Phi(K_1 \cup K_2) \tag{75}$$

b) Let $\mathcal{D}_k$ denotes the set of all dyadic subcubes of Q whose side has length 2^{-k}. We assume that

$$\lim_{k\to\infty} \sup_{K\in\mathcal{D}_k} \Phi(K) = 0 \tag{76}$$

Theorem 9. *Let Φ be as above and let $\epsilon > 0$ be such that*

$$\mathcal{T}_\epsilon = \big\{ I \in \mathcal{D},\ I \subset Q \ :\ \Phi(I) > \epsilon \big\} \neq \emptyset.$$

There exist a partition $\mathcal{P}_\epsilon$ of Q into dyadic rings such that $\Phi(K) \leq \epsilon$ for $K \in \mathcal{P}_\epsilon$, and a set $\widetilde{\mathcal{P}_\epsilon}$ of pairwise disjoint dyadic rings such that $\Phi(K) > \epsilon$ for $K \in \widetilde{\mathcal{P}_\epsilon}$ and $|\mathcal{P}_\epsilon| \leq 2^{d+1}|\widetilde{\mathcal{P}_\epsilon}|$.

Proof. As mentioned above the proof is the same as the proof of Theorem 6.1 from [3] but one has to remember that each dyadic cube has 2^d children. $\square$

Proposition 18. *If $f \in \mathrm{BV}(Q)$ where Q is a cube and $\epsilon > 0$ then there exists a partition $\mathcal{P}_\epsilon$ of Q into dyadic rings such that*

$$\big|\mathcal{P}_\epsilon\big| \leq C\epsilon^{-1/p}|f|_{\mathrm{BV}} \tag{77}$$

and for averaging projection P associated with $\mathcal{P}_\epsilon$ we have

$$\big\|f - Pf\big\|_p^p \leq C\epsilon^{1/d}|f|_{\mathrm{BV}}. \tag{78}$$

Proof. For a given f we define a set function $\Phi(K) = \inf_{\lambda\in\mathbb{R}} \int_K |f-\lambda|^p$. One easily checks that Φ satisfies assumptions of Theorem 9. Let us take the set $\widetilde{\mathcal{P}_\epsilon}$ given by this Theorem.

$$\begin{aligned}
\epsilon^{1/p}\left|\mathcal{P}_\epsilon\right| &\leq 2^{d+1}\epsilon^{1/p}\left|\widetilde{\mathcal{P}_\epsilon}\right| \leq 2^{d+1}\sum_{K\in\widetilde{\mathcal{P}_\epsilon}}\Phi(K)^{1/p} \\
&= 2^{d+1}\sum_{K\in\widetilde{\mathcal{P}_\epsilon}}\inf_{\lambda\in\mathbb{R}}\big(\int_K|f-\lambda|^p\big)^{1/p} \\
&\leq 2^{d+1}\sum_{K\in\widetilde{\mathcal{P}_\epsilon}}\big(\int_K|f-f_K|^p\big)^{1/p} \\
&\qquad \text{from Poincare inequality (24)} \\
&\leq 2^{d+1}C\sum_{K\in\widetilde{\mathcal{P}_\epsilon}}\int_K|\mathsf{grad}\ f| \leq 2^{d+1}C\int_Q|\mathsf{grad}\ f|
\end{aligned}$$

so we get (77).

For partition $\mathcal{P}_\epsilon$ we define $A_\epsilon f = \sum_{K\in\mathcal{P}_\epsilon}\alpha_K\mathbf{1}_K$ where α_K is chosen so that $\int_K|f-\alpha_K|^p = \inf_\lambda\int_K|f-\lambda|^p$. Then using (77) we get

$$\left\|f-A_\epsilon f\right\|_p^p = \sum_{K\in\mathcal{P}_\epsilon}\Phi(K) \leq \epsilon|\mathcal{P}_\epsilon| \leq C\epsilon^{1/d}|f|_{\mathrm{B}V}. \tag{79}$$

Since the projection P has norm 1 in $L_p(Q)$ and $P(A_\epsilon f) = A_\epsilon f$ we get $\|f - Pf\|_p \leq 2\|f-A_\epsilon\|_p$ so (78) holds. ∎

Remark: We also have

$$\inf\left\{\|f-g\|_p\ :\ g\in\Sigma_N^r(Q)\right\} \leq CN^{-1/d}|f|_{\mathrm{BV}}. \tag{80}$$

This follows immediately from (77) and (79) when we take $\epsilon = \left(CN^{-1}|f|_{\mathrm{BV}}\right)^p$.

Now we are ready for one of the main results of this paper.

Theorem 10. *There exists a constant C such that for each N and each $f\in$ $\mathrm{BV}(Q)$ where Q is a dyadic cube we have*

$$|\mathcal{G}_N f|_{\mathrm{B}V} \leq C|f|_{\mathrm{B}V} \tag{81}$$

Proof. Let g be (almost) the best approximation to f from Σ_N^r in the $L_p(Q)$ norm and let P be the averaging projection onto partition of Q into rings which is given by the choice of g. Then

$$|g|_{\mathrm{B}V} \leq |g - Pf|_{\mathrm{B}V} + |Pf|_{\mathrm{B}V}. \tag{82}$$

Since $g - Pf \in \Sigma_N^r \subset \Sigma_{2N}^c$ from Proposition 15 and (80) we get

$$|g - Pf|_{\mathrm{B}V} \leq CN^{1/d}\|g - Pf\|_p \leq CN^{1/d}\|g - f\|_p \leq C|f|_{\mathrm{B}V}. \tag{83}$$

Since $\mathcal{G}_N f \in \Sigma_N^h \subset \Sigma_{2^d N}^c$ and $g \in \Sigma_N^r \subset \Sigma_{2N}^c$ we get $\mathcal{G}_N f - g \in \Sigma_{(2^d+2)N}^c$, so from Proposition 15 we get

$$|\mathcal{G}_N f - g|_{\mathrm{B}V} \leq CN^{1/d}\|\mathcal{G}_N f - g\|_p \tag{84}$$

Using (68) and (80) we get

$$CN^{1/d}\big(\|\mathcal{G}_N f - f\|_p + \|f - g\|_p\big) \leq C|f|_{\mathrm{B}V} \tag{85}$$

Writing $|\mathcal{G}_N f|_{\mathrm{B}V} \leq |g|_{\mathrm{B}V} + |\mathcal{G}_N f - g|_{\mathrm{B}V}$ and using (83), (82) and Theorem 7 for the first summand and (84) and (85) for the second summand we get (81). □

REMARK: Although the above Theorem was proved and formulated for the space $\mathrm{BV}(Q)$ where Q is a dyadic cube it also remains valid for the space $\mathrm{BV}(\mathbb{R}^d)$. This follows from the fact that the constant in (81) does not depend on Q and that we can exhaust $\mathbb{R}^d$ by a sequence of 2^d dyadic cubes.

7 Minimizing K and U functionals

For $f \in L_p$ and $t \geq 0$ we define

$$K(f,t) = \inf_{g \in \mathrm{BV}} \big(\|f - g\|_p + t|g|_{\mathrm{B}V}\big)$$

and

$$U(f,t) = \inf_{g \in \mathrm{BV}} \big(\|f - g\|_p^p + t|g|_{\mathrm{B}V}\big).$$

Our aim in this section is to show that Haar tresholding provides approximate minimizers for those functionals.

The functional $K(f,t)$ defined above is well known in theory of interpolation of linear operators. It is the basis of the real method of interpolation. It is natural to expect that our result will provide some information about the interpolation between L_p and BV.

The functional U for $d = 2$ was introduced in [11] in the context of image processing. For $d > 2$ it seems to be considered here for the first time. However it is a special case of variational functionals considered in [1]. It should be noted that Σ_N^c is contained in the space SBV considered in [1], so Haar tresholding provide an approximation by elements of SBV i.e. is related to the main problem considered in [1]. The functional U may also have some statistical interpretation.

Theorem 11. *There exists a constant C such that for all $t > 0$ and all $f \in L_p$ we have*

$$\|f - \mathcal{H}_\epsilon f\|_p^p + t|\mathcal{H}_\epsilon f|_{\mathrm{BV}} \leq CU(f,t) \tag{86}$$

where $\epsilon = t^{d-1}$.

Proof. First let us observe that for any g and $\epsilon > 0$ we have $\|f - \mathcal{H}_\epsilon f\|_p \leq C\|f - \mathcal{H}_{2\epsilon} g\|_p$ because, as is easily checked, absolute values of all Haar coefficients of $f - \mathcal{H}_{2\epsilon} g$ are greater than absolute values of corresponding Haar coefficients of $f - \mathcal{H}_\epsilon f$. So using (67) we get

$$\begin{aligned} \|f - \mathcal{H}_\epsilon f\|_p &\leq C\|f - \mathcal{H}_{2\epsilon} g\|_p \\ &\leq C\|f - g\|_p + C\|g - \mathcal{H}_{2\epsilon} g\|_p \qquad (87) \\ &\leq C\|f - g\|_p + C\epsilon^{1/d}|g|_{\mathrm{BV}}^{1/p}. \end{aligned}$$

Now we will estimate $|\mathcal{H}_\epsilon f|_{\mathrm{BV}}$ as

$$|\mathcal{H}_\epsilon f|_{\mathrm{BV}} \leq |\mathcal{H}_\epsilon f - \mathcal{H}_\epsilon g|_{\mathrm{BV}} + |\mathcal{H}_\epsilon g|_{\mathrm{BV}}. \tag{88}$$

We write $\widetilde{\mathcal{H}_\epsilon} = \mathrm{Id} - \mathcal{H}_\epsilon$ and we have

$$\mathcal{H}_\epsilon f - \mathcal{H}_\epsilon g = \mathcal{H}_\epsilon\big(\mathcal{H}_\epsilon f - \mathcal{H}_\epsilon g\big) + \widetilde{\mathcal{H}_\epsilon}\big(\mathcal{H}_\epsilon f - \mathcal{H}_\epsilon g\big). \tag{89}$$

Using (74) and then (87) we have

$$\begin{aligned}
|\mathcal{H}_\epsilon(\mathcal{H}_\epsilon f - \mathcal{H}_\epsilon g)|_{\mathrm{BV}} &\leq C\epsilon^{-1/(d-1)}\|\mathcal{H}_\epsilon f - \mathcal{H}_\epsilon g\|_p^p \\
&\leq C\epsilon^{-1/(d-1)}\big(\|\mathcal{H}_\epsilon f - f\|_p^p + \|f - \mathcal{H}_\epsilon g\|_p^p\big) \\
&\leq C\epsilon^{-1/(d-1)}\big(\|f-g\|_p^p + \epsilon^{p/d}|g|_{\mathrm{BV}}\big) \\
&\quad + C\epsilon^{-1/(d-1)}\big(\|f-g\|_p^p + \|g - \mathcal{H}_\epsilon g\|_p^p\big) \\
&\leq C\epsilon^{-1/(d-1)}\|f-g\|_p^p + C|g|_{\mathrm{BV}}
\end{aligned} \tag{90}$$

where in the last inequality we used (67).

We still need to estimate $|\widetilde{\mathcal{H}_\epsilon}(\mathcal{H}_\epsilon f - \mathcal{H}_\epsilon g)|_{\mathrm{BV}}$. For this we remark that if $0 < |c_{\mathfrak{s}}(\mathcal{H}_\epsilon f - \mathcal{H}_\epsilon g)| \leq \epsilon$ then necessarily $|c_{\mathfrak{s}}(g)| > \epsilon$. Thus we infer that $\widetilde{\mathcal{H}_\epsilon}(\mathcal{H}_\epsilon f - \mathcal{H}_\epsilon g)$ has at most N non zero coefficients and from Theorem 8 we have an estimate $N \leq C\epsilon^{-1}|g|_{\mathrm{BV}}$. Using this and (73) we obtain

$$\begin{aligned}
|\widetilde{\mathcal{H}_\epsilon}(\mathcal{H}_\epsilon f - \mathcal{H}_\epsilon g)|_{\mathrm{BV}} &\leq CN^{1/d}\|\mathcal{H}_\epsilon f - \mathcal{H}_\epsilon g\|_p \\
&\leq C\epsilon^{-1/d}|g|_{\mathrm{BV}}^{1/d}\|\mathcal{H}_\epsilon f - \mathcal{H}_\epsilon g\|_p \\
&\qquad \text{using an estimate } a^{1/p}\cdot b^{1/d} \leq a+b \\
&\leq C\Big(|g|_{\mathrm{BV}} + \epsilon^{-1/(d-1)}\|\mathcal{H}_\epsilon f - \mathcal{H}_\epsilon g\|_p^p\Big).
\end{aligned} \tag{91}$$

Using (87) and (67) we get

$$\begin{aligned}
\|\mathcal{H}_\epsilon f - \mathcal{H}_\epsilon g\|_p &\leq \|f - \mathcal{H}_\epsilon f\|_p + \|g - \mathcal{H}_\epsilon g\|_p + \|f-g\|_p \\
&\leq C\|f-g\|_p + C\epsilon^{1/d}|g|_{\mathrm{BV}}^{1/p}.
\end{aligned} \tag{92}$$

Substituting (92) into (91) we get

$$|\widetilde{\mathcal{H}}_\epsilon\big(\mathcal{H}_\epsilon f - \mathcal{H}_\epsilon g\big)|_{\mathrm{BV}} \le C\big(\epsilon^{-1/(d-1)}\|f-g\|_p^p + |g|_{\mathrm{BV}}\big). \tag{93}$$

Combining (87)–(90), (93) and Theorem 10 we get

$$\begin{aligned}\|f-\mathcal{H}_\epsilon f\|_p^p + \epsilon^{1/(d-1)}|\mathcal{H}_\epsilon f|_{\mathrm{BV}} &\le C\big(\|f-g\|_p^p + \epsilon^{1/(d-1)}|g|_{\mathrm{BV}}\big)\\ &\quad + \epsilon^{1/(d-1)}\big(C\epsilon^{-1/(d-1)}\|f-g\|_p^p + C|g|_{\mathrm{BV}}\big)\\ &\le C\big(\|f-g\|_p^p + \epsilon^{1/(d-1)}|g|_{\mathrm{BV}}\big).\end{aligned}$$

Taking the inf over g we get (86). □

Theorem 12. *There exists a constant C such that for each $N = 1, 2, \ldots$ and each $f \in L_p$ we have*

$$\|f - \mathcal{G}_N f\|_p + N^{-1/d}|\mathcal{G}_N f|_{\mathrm{BV}} \le CK(f, N^{-1/d}). \tag{94}$$

Proof. Let $g \in \mathrm{BV}$ be arbitrary. Since the Haar system is greedy in L_p we have

$$\begin{aligned}\|f-\mathcal{G}_N f\|_p &\le C\|f-\mathcal{G}_N g\|_p \le C\big(\|f-g\|_p + \|g-\mathcal{G}_N g\|_p\big)\\ &\le C\big(\|f-g\|_p + N^{-1/d}|g|_{\mathrm{BV}}\big)\end{aligned} \tag{95}$$

where in the last inequality we used (68). From Corollary 16 and Theorem 10 we infer that

$$\begin{aligned}|\mathcal{G}_N f|_{\mathrm{BV}} &\le |\mathcal{G}_N f - \mathcal{G}_N g|_{\mathrm{BV}} + |\mathcal{G}_N g|_{\mathrm{BV}}\\ &\le C\,N^{1/d}\|\mathcal{G}_N f - \mathcal{G}_N g\|_p + C|g|_{\mathrm{BV}}.\end{aligned} \tag{96}$$

Using once more the fact that the Haar system is greedy we get

$$\|\mathcal{G}_N f - \mathcal{G}_N g\|_p \le \|f-\mathcal{G}_N f\|_p + \|f - \mathcal{G}_N g\|_p \le C\|f-\mathcal{G}_N g\|_p. \tag{97}$$

We estimate (97) using the estimate contained in (95) and we get

$$\|\mathcal{G}_N f - \mathcal{G}_N g\|_p \le C\big(\|f-g\|_p + N^{-1/d}|g|_{\mathrm{B}V}\big)$$

and substituting this into (96) we get

$$|\mathcal{G}_N f|_{\mathrm{B}V} \le C\big(N^{1/d}\|f-g\|_p + |g|_{\mathrm{B}V}\big). \tag{98}$$

Putting together (95) and (98) we get

$$\|f - \mathcal{G}_N f\|_p + N^{-1/d}|\mathcal{G}_N f|_{\mathrm{B}V} \le C\big(\|f-g\|_p + N^{-1/d}|g|_{\mathrm{B}V}\big)$$

and taking the infimum over g we get (94).

$\square$

References

[1] L. Ambrosio, *Existence Theory for a New Class of Variational Problems,* Arch.Rational Mech. Anal. 111(1990) N.4 pp.291-322

[2] A. Cohen, W. Dahmen, Ingrid Daubechies and R. DeVore, *Harmonic Analysis of the space BV*, (preprint)

[3] A. Cohen, R. DeVore, P. Petrushev and H. Xu, *Nonlinear approximation and the space $BV(\mathbb{R}^2)$*, Amer. J. Math. 121(1999) pp.587-628

[4] R. DeVore and V. Popov, *Free multivariate splines*, Constr. Approx. 3 (1987) pp.239-248

[5] E. Giusti, **Minimal Surfaces and Functions of Bounded Variation,** Monographs in Math. 80, Birkhäuse, Boston 1984

[6] P. Hajłasz and P. Koskela, *Isoperimetric inequalities and imbedding theorems in irregular domains*, J. London Math. Soc. (2)58 (1998) pp.425-50

[7] S.V. Konyagin and V.N. Temlyakov, *A remark on greedy approximation in Banach spaces*, East J. Approx. 5 (3) (1999) pp.365-79

[8] V.G. Mazia, **Sobolev spaces,** Izd. Leningrad. Univ. 1985 (in Russian)

[9] V.G. Mazia and T.O. Shaposhnikova, **Multipliers in spaces of differentiable functions,** Izd. Leningrad. Univ. 1986 (in Russian)

[10] Y. Meyer, *Wavelets and functions with bounded variation, from image processing to pure mathematics,* Rend. Mat. Acc. Lincei s.9, fasc. spec. 2000 pp.77-105

[11] L.I. Rudin and S.J. Osher, *Total variation bases restoration with free local constraints,* Proc. ICIP, IEEE International Conference on Image Processing, Austin, TX, 1994, pp.31-35

[12] V.N. Temlyakov, *The best m–term approximation and greedy algorithms* Adv. in Comp. Math. 8 (1998) pp.249-65

[13] P. Wojtaszczyk, **A mathematical introduction to wavelets**, London Math. Soc Student Texts 37, Cambridge University Press, Cambridge 1997

[14] P. Wojtaszczyk, **Banach spaces for analysts**, Cambridge Studies in Advanced Math. 25, Cambridge University Press, Cambridge 1991

[15] P. Wojtaszczyk, *Greedy Algorithm for general Biorthogonal Systems*, J. Approx. Theory 107 (2000) pp.293-314

[16] W.P. Ziemer, **Weakly Differentiable Functions,** Graduate Texts in Math. 120, Springer Verlag, New York 1989