Nonlinear piecewise polynomial approximation and multivariate BV spaces of a Wiener–L. Young type. I.

Yu. Brudnyi

PII: S0021-9045(17)30026-6
DOI: http://dx.doi.org/10.1016/j.jat.2017.03.002
Reference: YJATH 5139

To appear in: *Journal of Approximation Theory*

Received date : 4 November 2015
Revised date : 3 March 2017
Accepted date : 7 March 2017

Nonlinear Piecewise Polynomial Approximation and Multivariate BV spaces of a Wiener–L. Young Type. I.

Yu. Brudnyi (Technion, Haifa)

Abstract

The named space denoted by V^k_{pq} consists of L_q functions on $[0,1)^d$ of bounded p-variation of order $k \in \mathbb{N}$. It generalizes the classical spaces $V_p(0,1)$ $(= V^1_{p\infty})$ and $BV([0,1)^d)$ (V^1_{1q} where $q := \frac{d}{d-1}$) and is closely related to several important smoothness spaces, e.g., to Sobolev spaces over L_p, BV and BMO and to Besov spaces.

The main approximation result concerns the space V^k_{pq} of *smoothness* $s := d\left(\frac{1}{p} - \frac{1}{q}\right) \in (0,k]$. It asserts the following:

Let $f \in V^k_{pq}$ be of smoothness $s \in (0,k]$, $1 \le p < q < \infty$ and $N \in \mathbb{N}$. There exist a family Δ_N of N dyadic subcubes of $[0,1)^d$ and a piecewise polynomial g_N over Δ_N of degree $k-1$ such that

$$\|f - g_N\|_q \leqslant CN^{-s/d}|f|_{V^k_{pq}}.$$

This implies similar results for the above mentioned smoothness spaces, in particular, solves the going back to the 1967 Birman–Solomyak paper [BS] problem of approximation of functions from $W^k_p([0,1)^d)$ in $L_q([0,1)^d)$ whenever $\frac{k}{d} = \frac{1}{p} - \frac{1}{q}$ and $q < \infty$.

Key words: N-term approximation. Piecewise polynomials. Dyadic cubes. Spaces of q-integrable functions of bounded variation. Sobolev spaces. Besov spaces.

§ 1. Introduction

1.1. The present paper is inspired by two results removed from each other by more than thirty years. The first, the 1967 pioneering paper [BS], asserts the following:

Theorem A. *Given $f \in W^k_p([0,1)^d)$, $N \in \mathbb{N}$ and $1 \le p < q < \infty$ satisfying*

$$\frac{k}{d} > \frac{1}{p} - \frac{1}{q}$$

there exist a partition Δ_N of $[0,1)^d$ into at most N dyadic subcubes and a piecewise polynomial g_N on Δ_N of degree $k-1$ such that

$$\|f - g_N\|_q \leqslant CN^{-k/d} \sup_{|\alpha|=k} \|D^\alpha f\|_p; \tag{1.1}$$

the constant $C > 0$ is independent of f and N and $C \to \infty$ as q tends to the Sobolev limiting exponent $q^* := \left(\frac{1}{p} - \frac{k}{d}\right)^{-1}$.

Using the compactness argument from [BS, § 5] one can prove that validity of (1.1) for $q = q^*$ implies (incorrect) compactness of embedding $W^k_p \subset L_{q^*}$.

This leads to the following:

Problem. Find the extension of Theorem A to $q = q^*$.

For the special case $k = p = 1$, $d = 2$ the answer was given in the 1999 paper [CDPX] by A. Cohen, DeVore, Petrushev and Hong Xu; the case $d > 2$ was than proved by Wojtaszczyk [W]. The result states:

Theorem B. *Given $f \in W_1^1([0,1)^d)$, $d \geqslant 2$, and $N \in \mathbb{N}$ there exist a partition Δ_N of $[0,1)^d$ into at most N d-rings (differences of two dyadic subcubes) and a piecewise constant function g_N on Δ_N such that*

$$\|f - g_N\|_{q^*} \leqslant c(d) N^{-1/d} \sup_{\|\alpha\|=1} \|D^\alpha f\|_1. \tag{1.2}$$

Hereafter $c(x, y, \dots)$ denotes a positive constant *depending only* on the parameters in the parentheses.

Remark 1.1. (a) Theorem B is proved for L_1 functions whose distributional derivatives of the first order are bounded Radon measures, see the cited papers. However, this more general result follows directly from (1.2), see Section 4.5 below.

(b) The partition Δ_N in Theorem B can be replaced by a cover of $[0,1)^d$ by at most N dyadic cubes. This result is equivalent to the previous, see Remark 4.9 below.

1.2. In this paper, we prove a general result on N-term piecewise polynomial approximation implying as consequences similar results for a number of spaces of smooth functions.

To illustrate the main result we formulate its consequence giving the solution of the Birman–Solomyak problem and the generalization of Theorem B to functions of higher smoothness.

Theorem 1.2. *(a) Given $f \in W_p^k([0,1)^d)$, $k \in \mathbb{N}$ and $1 \leq p < q^* < \infty$ such that*

$$\frac{k}{d} = \frac{1}{p} - \frac{1}{q^*}, \quad d \geqslant 2,$$

there exist a cover Δ_N of $[0,1)^d$ by at most N dyadic subcubes and a family of polynomials $\{P_Q\}_{Q \in \Delta_N} \subset \mathcal{P}_{k-1}$ (of degree $k-1$) such that[1]

$$\left\| f - \sum_{Q \in \Delta_N} P_Q \cdot 1_Q \right\|_{q^*} \leqslant c(k,d) N^{-k/d} \sup_{|\alpha|=k} \|D^\alpha f\|_p. \tag{1.3}$$

(b) For $p := 1$, hence, $q^* = \frac{d}{d-k}$, the previous holds for $f \in L_1$, whose derivatives of order k are bounded Radon measures.

The associated seminorm of the latter function space denoted by $BV^k([0,1)^d)$ is given by

$$|f|_{BV^k} := \sup_{|\alpha|=k} \operatorname*{var}_{[0,1)^d} (D^\alpha f). \tag{1.4}$$

Remark 1.3. The result can be reformulated equivalently with the cover Δ_N substituted for a *partition* of $[0,1)^d$ into at most N d-rings.

[1] hereafter 1_S stands for the *indicator* (characteristic function) of a set S

1.3. The formulated aim will be achieved by using the BV spaces of integrable on $[0,1)^d$ functions of arbitrary smoothness introduced in [B-71, sec. 4.5]. To motivate the definition of the corresponding space denoted by V_{pq}^k we begin with a model case, the Wiener–L. Young space V_p, whose associated seminorm is presented in the following equivalent form:

$$\operatorname{var}_p f := \sup_{\Delta} \left(\sum_{I \in \Delta} osc(f; I)^p \right)^{1/p} \tag{1.5}$$

where Δ runs over disjoint families of intervals $I = [a, b) \subset [0, 1)$ and

$$osc(f; I) := \sup_{x,y \in I} |f(x) - f(y)|. \tag{1.6}$$

To obtain the required seminorm of V_{pq}^k denoted by $\operatorname{var}_p^k(\cdot; L_q)$ we replace in (1.5) intervals by cubes $Q \subset [0,1)^d$, in (1.6) the first difference by the k-th one, and the uniform norm by L_q norm. This gives the following:

Definition 1.4. *The seminorm $f \mapsto \operatorname{var}_p^k(f; L_q)$ is a function on $L_q([0,1)^d)$ given by*

$$\operatorname{var}_p^k(f; L_q) := \sup_{\Delta} \left\{ \sum_{Q \in \Delta} osc_q^k(f; Q)^p \right\}^{1/p} \tag{1.7}$$

where Δ runs over disjoint families of cubes $Q \subset [0,1)^d$ and

$$osc_q^k(f; Q) := \sup_{h \in \mathbb{R}^d} \|\Delta_h^k f\|_{L_q(Q_{kh})}; \tag{1.8}$$

here $\Delta_h^k := \sum_{j=0}^{k} (-1)^{n-j} \binom{k}{j} \delta_{jh}$ *and* $Q_{kh} := \{x \in \mathbb{R}^d; x + jh \in Q, j = 0, 1, \dots, k\}$.

The important characteristic of the space V_{pq}^k is its *smoothness* introduced by the following:

Definition 1.5. *Smoothness of the space V_{pq}^k denoted by $s(V_{pq}^k)$ is a real number given by*

$$s(V_{pq}^k) := d\left(\frac{1}{p} - \frac{1}{q}\right). \tag{1.9}$$

This concept is closely related to differential and approximation properties of V_{pq}^k functions. In fact, a function with $s(V_{pq}^k) = s$ belongs to the Taylor class $T_q^s(x)$ a.e. if $0 < s < k$ and $t_q^k(x)$ a.e. if $s = k$, see [B-94, § 2].

Moreover, as we will see, its order of N-term approximation in $L_q([0,1)^d)$ by piecewise polynomial is $N^{-s/d}$ for $0 < s \leqslant k$.

In particular, the proof of Theorem 1.2 is based on the equality

$$V_{pq^*}^k = W_p^k$$

and the fact that

$$s(V^k_{pq^*}) := d\left(\frac{1}{p} - \frac{1}{q^*}\right) = k$$

that allow to derive it directly from the corresponding result for $V^k_{pq^*}$.

1.4. The outline of the paper is the following:

The main result, Theorem 2.1, is formulated in § 2 along with its consequences describing similar approximation results for classical smoothness spaces (one of them is Theorem 1.2 formulated above).

In § 3, we prove properties of V^k_{pq} spaces essential for the proof of Theorem 2.1. The first result asserts that a function $f \in V^k_{pq}$ can be weakly approximated in L_q by C^∞ functions whose (k,p)-variations are bounded by $\mathrm{var}^k_p(f; L_q)$. For the special case of the space $BV([0,1)^d)$ $(= V^1_{1\,d/d-1})$, see, e.g., [Z, Thm. 5.3.3].

The second result estimates polynomial approximation of order $k-1$ for $f \in V^k_{pq}(Q' \setminus Q'')$ via its (k,p)-variation; here $Q'' \subsetneqq Q'$ are dyadic cubes.

The latter result essentially uses a cover lemma proved in collaboration with V. Dolnikov; its proof is presented in Appendix I.

The main result, Theorem 2.1, is proved in § 4 and its consequences for the classical smoothness spaces in § 5. The approximation algorithm used in the construction of the family of dyadic cubes for Theorem 2.1 is presented in Appendix II.

Its primary version was developed to prove the similar to Theorem 1.2 result for the N-term approximation of functions from Besov spaces by B–splines; the result is announced in [BI-87] and proved in [BI-92].

The special case of Theorem 2.1 for functions with absolutely continuous (k,p)-variation was proved a long time ago and announced in [B-2002]. This result allows to prove all consequences of Theorem 2.1 presented in § 2 but only much later the author understood how to derive Theorem 2.1 from this special case, see Section 4.5.

Acknowledgement. I am very grateful to my friends Irina Asekritova, Vladimir Dolnikov and Natan Kruglyak for their lively interest to and useful discussions of this paper.

§ 2. Formulations of the Main Results

Throughout the paper we use the following notions and notations.

A *cube* denoted by Q, Q', K etc. is a set of $\mathbb{R}^d$ homothetic to the (half–open) *unit cube*

$$Q^d := [0,1)^d. \tag{2.1}$$

$\mathcal{D}(Q)$ denotes the family of *dyadic cubes* of Q, i.e., cubes of the form

$$K := 2^{-j}(Q + \alpha) \tag{2.2}$$

where $j \in \mathbb{Z}_+ := \{0, 1, 2, \dots, \}$ and $\alpha \in \mathbb{Z}^d$.

Further, $\mathcal{P}_l = \mathcal{P}_l(\mathbb{R}^d)$ is *the space of polynomials in* $x := (x_1, x_2, \dots, x_d)$ *of degree* l while $\mathcal{P}_l(\Delta)$ denotes the *space of piecewise polynomials on a set* $\Delta \subset \mathcal{D}(Q)$ *of degree* l.

In other words,

$$\mathcal{P}_l(\Delta) := \{f \in L_\infty(Q); f = \sum_{K\in\Delta} P_K \cdot 1_K\} \tag{2.3}$$

where $\{P_K\}_{K\in\Delta} \subset \mathcal{P}_l$.

Theorem 2.1. *(a) Let $f \in V^k_{pq}(Q^d)$ where the smoothness $s := s(V^k_{pq})$, see (1.9), and d, p, q be such that*

$$d \geqslant 2,\ 0 < s \leqslant k \text{ and } 1 \leqslant p < q < \infty. \tag{2.4}$$

Given $N \in \mathbb{N}$ there exist a cover $\Delta_N \subset \mathcal{D}(Q^d)$ of Q^d with card $\Delta_N \leqslant N$ and $g_N \in \mathcal{P}_{k-1}(\Delta_N)$ such that

$$\|f - g_N\|_q \leqslant c(d)N^{-s/d}|f|_{V^k_{pq}}. \tag{2.5}$$

The same is true for $q = \infty$, i.e., for $s/d = 1/p$, if f is uniformly continuous on Q^d.

(b) The cover Δ_N can be replaced by a partition of Q^d into at most N dyadic d-rings.

Stipulation 2.2. We drop the symbol Q^d from the next notations writing, e.g., $\mathcal{D}$, V^k_{pq}, L_q instead of $\mathcal{D}(Q^d)$, $V^k_{pq}(Q^d)$, $L_q(Q^d)$, if it does not lead to misunderstanding.

The first consequence of the main result, Theorem 1.2, immediately follows from Theorem 2.1(a) and the inequality

$$\operatorname{var}^k_p(f; L_{q^*}) \leqslant c \begin{cases} |f|_{W^k_p} & \text{if} \quad p > 1 \\ |f|_{BV^k} & \text{if} \quad p = 1, \end{cases} \tag{2.6}$$

here $c = c(k, d, q^*)$ and $q^* := \left(\frac{k}{d} - \frac{1}{p}\right)^{-1}$.

This and analogous embedding results for Besov spaces are presented in § 5.

Let now $\dot{B}^{\lambda\theta}_p := \dot{B}^{\lambda\theta}_p(Q^d)$ be the homogeneous Besov space defined by the semi-norm

$$|f|_{B^{\lambda\theta}_p} := \left\{ \int_0^1 \left(\frac{\omega_k(t; f; L_p)}{t^\lambda} \right)^\theta \frac{dt}{t} \right\}^{1/\theta} \tag{2.7}$$

where $k = k(\lambda) := \min\{n \in \mathbb{N}; n > \lambda\}$ and $\omega_k(\cdot; f; L_p)$ is the k-th modulus of continuity of $f \in L_p$, see e.g., [N] or [DL] for its definition.

The first result concerns the "diagonal" Besov space $\dot{B}^\lambda_p := \dot{B}^{\lambda p}_p$, $1 \leqslant p < \infty$.

Theorem 2.3. *Let $f \in \dot{B}^\lambda_p$ and d, p, q, λ be such that*

$$d \geqslant 2, 1 \leqslant p < q < \infty \quad \text{and} \quad \frac{\lambda}{d} = \frac{1}{p} - \frac{1}{q}.$$

Given $N \in \mathbb{N}$ there exist a cover $\Delta_N \subset \mathcal{D}$ of Q^d by at most N cubes and $g_N \in \mathcal{P}_{k-1}(\Delta_N)$ such that

$$\|f - g_N\|_q \leqslant c(k, d)N^{-\lambda/d}|f|_{B^\lambda_p}.$$

The second result concerns approximation in the uniform norm ($q = \infty$).

Theorem 2.4. *Let $f \in \dot{B}_p^{\lambda 1}$ and*

$$d \geqslant 2, \quad 1 \leqslant p < \infty, \quad \frac{\lambda}{d} = \frac{1}{p}.$$

Given $N \in \mathbb{N}$ there exist $\Delta_N \subset \mathcal{D}$ satisfying the condition of Theorem 2.3 and $g_N \in \mathcal{P}_{k-1}(\Delta_N)$ such that

$$\|f - g_N\|_\infty \leqslant c(\lambda, d) N^{-\lambda/d} |f|_{B_p^{\lambda 1}}.$$

Remark 2.5. (a) The definitions of seminorms $|f|_{V_{pq}^k}$ and $|f|_{B_p^{\lambda\theta}}$ are naturally extended to $p < 1$, see (1.7) and (2.7). Moreover, unlike (2.6), the corresponding embedding results are true for this case. For instance,

$$|f|_{B_p^\lambda} \leqslant c(\lambda, d, p) |f|_{V_{p\infty}^k}$$

for $0 < p \leqslant 1$ and $\frac{\lambda}{d} = \frac{1}{p}$.

Therefore Theorems 2.3 and 2.4 are true for $p < 1$ as well.

(b) For B–splines and, more generally, for refinable functions with the Strang–Fix condition the analog of Theorem 2.3 was proved in [BI] and [DJP], respectively. The latter result was extended to $p < 1$ in [DPY].

Formally, Theorem 2.3 does not follow from these results but it can be surely proved by the approximation algorithms used in the cited papers.

§ 3. Spaces V_{pq}^k

3.1. (k,p)*-variation.* The named object is a set-function defined by (1.7) with Q^d substituted for a measurable set $S \subset \mathbb{R}^d$ of nonempty interior.

However, we prefer to use an equivalent definition more suitable for our purpose. In fact, we replace osc_p^k by *local best approximation*, a set-function given for $f \in L_q^{loc}(\mathbb{R}^d)$ and $S \subset \mathbb{R}^d$ by

$$E_k(f; S; L_q) := \inf_{m \in \mathcal{P}_{k-1}} \|f - m\|_{L_q(S)}.$$

This replacement gives the following:

Definition 3.1. (k,p)*-variation of a function $f \in L_q^{loc}(\mathbb{R}^d)$ is a set-function on subsets $S \subset \mathbb{R}^d$ with nonempty interior given by*

$$\operatorname{var}_p^k(f; S; L_q) := \sup_\Delta \left\{ \sum_{Q \in \Delta} E_k(f; Q; L_q)^p \right\}^{1/p} \tag{3.1}$$

where Δ runs over all disjoint families of cubes $Q \subset S$.

Equivalence of this definition to the previous one follows from the main result of [B-70] implying, e.g., the next two-sided inequality with constants depending only on k:

$$osc_p^k(f;Q;L_q) \approx E_k(f;Q;L_q).$$

It should be pointed out that in what follows all definitions and results involving the space V_{pq}^k use Definition 3.1. In particular, the associated seminorm of this space is

$$\|f\|_{V_{pq}^k} := \sup_{\Delta} \left\{ \sum_{Q\in\Delta} E_k(f;Q;L_q)^p \right\}^{1/p}$$

where Δ runs over all disjoint families of $Q \subset Q^d$.

Now we present some basic properties of (k,p)-variation starting with those following directly from Definition 3.1.

Proposition 3.1$^{\mathbf{a}}$ (*Subadditivity*)**.** *Let $\{S_i\}$ be a disjoint families of measurable sets with nonempty interiors. Then*

$$\left\{ \sum_i \operatorname{var}_p^k(f;S_i:L_q)^p \right\}^{1/p} \leqslant \operatorname{var}_p^k(f;\bigcup_i S_i;L_q). \tag{3.2}$$

(Lower semicontinuity) If $\{f_j\}$ converges in L_q to a function f, then

$$\operatorname{var}_p^k(f;S;L_q) \leqslant \varliminf_{j\to\infty} \operatorname{var}_p^k(f_j;S;L_q). \tag{3.3}$$

Proof. Let $\Delta := \{Q\}$ be a disjoint family of cubes and

$$\operatorname{var}_p^k(f;\Delta;L_q) := \left\{ \sum_{Q\in\Delta} E_k(f;Q;L_q)^p \right\}^{1/p}. \tag{3.4}$$

If $\{\Delta_i\}$ is disjoint and $\bigcup\limits_{Q\in\Delta_i} Q \subset S_i$, then

$$\sum_i \operatorname{var}_p^k(f;\Delta_i;L_q)^p = \operatorname{var}_p^k(f;\bigcup \Delta_i;L_q) \leqslant \operatorname{var}_p^k(f;\bigcup S_i;L_q)$$

and it remains to take supremum over each Δ_i to prove (3.2).

The property (3.3) is proved similarly. □

A more substantive property of (k,p)-variation gives the next result.

Proposition 3.1$^{\mathbf{b}}$. *Let a C^∞ function f belong to the space V_{pq}^k of smoothness $s \leqslant k$. Then uniformly in S*

$$\lim_{|S|\to 0} \operatorname{var}_p^k(f;S;L_q) = 0. \tag{3.5}$$

Hereafter $|S|$ denotes the Lebesgue d-measure of S.

Proof. Let Δ be a disjoint family of cubes $Q \subset S$. By the Taylor formula

$$E_k(f;Q;L_q) \leqslant |Q|^{1/q} E_k(f;Q;C) \leqslant c(k,d)|Q|^{1/q+k/d} \max_{|\alpha|=k} \max_Q |D^\alpha f|;$$

this implies

$$\operatorname{var}_p^k(f;\Delta;L_q) \leqslant c(k,d) \left\{ \sum_{Q\in\Delta} |Q|^{p\left(\frac{1}{q}+\frac{k}{d}\right)} \right\}^{1/p} |f|_{C^k(Q^d)}.$$

Since $p\left(\frac{1}{q}+\frac{k}{d}\right) \geqslant p\left(\frac{1}{q}+\frac{s}{d}\right) = 1$, the sum here is bounded by $\left\{ \sum_{Q\in\Delta} |Q| \right\}^{1/p} \leqslant |S|^{1/p}$, and therefore

$$\operatorname{var}_p^k(f;S;L_q) := \sup_\Delta \operatorname{var}_p^k(f;\Delta;L_q) \to 0 \text{ as } |S| \to 0.$$

□

3.2. *C^∞ Approximation of V_{pq}^k Functions.* Since the space V_{pq}^k is, in general, nonseparable, C^∞ approximated functions form a proper subspace of V_{pq}^k. However, a weaker form of C^∞ approximation is true.

Theorem 3.2. *Let a function f belong to V_{pq}^k if $q < \infty$ and to $V_{p\infty}^k \bigcap C(\mathbb{R}^d)$, otherwise. Assume that Q is a subcube of Q^d such that*

(3.6) $$dist(Q;\mathbb{R} \setminus Q^d) > 0.$$

Then there exists a sequence $\{f_j\} \subset C^\infty(\mathbb{R}^d)$ such that

(3.7) $$\lim_{j\to\infty} f_j = f \text{ (convergence in } L_q(Q))$$

and, moreover,

(3.8) $$\sup_j \operatorname{var}_p^k(f_j;Q;L_q) \leqslant \operatorname{var}_p^k(f;L_q).$$

Proof. Let f_ε be a *regularizer* of f given by

(3.9) $$f_\varepsilon(x) := \int_{Q_\varepsilon} f(x-\varepsilon y)\varphi(y)\,dy, \quad x \in Q,$$

where $\varphi \in C^\infty(\mathbb{R}^d)$ is a test function, i.e.,

(3.10) $$\varphi \geqslant 0, \int \varphi\,dx = 1 \text{ and } supp\,\varphi \subseteq [-1,1]^d;$$

here $\varepsilon > 0$ is such that

(3.11) $$Q_\varepsilon := Q + [-\varepsilon,\varepsilon]^d \subseteq Q^d,$$

see (3.6).

Now (3.10) and the Minkowski inequality yield

$$\|f - f_\varepsilon\|_{L_q(Q)} \leqslant \sup_{|y|\leqslant 1} \|f(\cdot - \varepsilon y) - f\|_{L_q(Q)}.$$

Since the right-hand side tends to 0 as $\varepsilon \to 0$ for $q < \infty$ and for $q = \infty$ if $f \in C(\mathbb{R}^d)$, (3.7) follows.

To proceed we need the following:

Lemma 3.3. *It is true that*

$$E_k(f_\varepsilon; Q; L_q) \leqslant \int\limits_{|y|\leqslant 1} E_k(f; Q - \varepsilon y; L_q)\varphi(y)\, dy. \tag{3.12}$$

Proof. It suffices to prove (3.12) for $q < \infty$ and then pass q to $+\infty$.

Let $q < \infty$ and q' denote the conjugate exponents. By $\mathcal{P}^\perp_{k-1}(Q)$ we denote the set of functions $g \in L_{q'}(\mathbb{R}^d)$ such that

$$\|g\|_{L_{q'}} = 1,\ \mathit{supp}\, g \subset Q,\ \int x^\alpha g(x)\, dx = 0,\ |\alpha| \leqslant k - 1. \tag{3.13}$$

By the duality of L_q and $L_{q'}$

$$E_k(f_\varepsilon; Q; L_q) = \sup \left\{ \int\limits_Q f_\varepsilon g\, dx; g \in \mathcal{P}^\perp_{k-1}(Q) \right\}. \tag{3.14}$$

On the other hand,

$$\left| \int\limits_Q f_\varepsilon g\, dx \right| \leqslant \int\limits_{\mathbb{R}^d} \left| \int\limits_{Q-\varepsilon y} f(x) g(x + \varepsilon y)\, dx \right| \varphi(y)\, dy \tag{3.15}$$

and the function $x \mapsto g(x + \varepsilon y)$, $x \in Q$, clearly, belongs to the set $\mathcal{P}^\perp_{k-1}(Q - \varepsilon y)$. Therefore for every polynomial m of degree $k - 1$

$$\int\limits_{Q-\varepsilon y} f_\varepsilon(x) g(x + \varepsilon y)\, dx = \int\limits_{Q-\varepsilon y} f(x)(g - m)(x + \varepsilon y)\, dx.$$

Combining this with (3.15) and using the Hölder inequality we obtain

$$\left| \int\limits_Q f_\varepsilon g\, dx \right| \leqslant \int\limits_{\mathbb{R}^d} \varphi(y) \|f - m\|_{L_q(Q-\varepsilon y)} \|g\|_{L_{q'}(\mathbb{R}^d)}\, dy.$$

Taking here infimum over all polynomials m and then supremum over all $g \in \mathcal{P}^\perp_{k-1}(Q)$ we get by (3.14)

$$E_k(f_\varepsilon; Q; L_q) \leqslant \int\limits_{\mathbb{R}^d} E_k(f; Q - \varepsilon y; L_q)\varphi(y)\, dy.$$

The proof is complete. □

Now we prove (3.8). To this end, first we estimate $\operatorname{var}_p^k(f;\Delta;L_q)$, see (3.4), for the disjoint family of cubes of Q. Due to (3.12) the Minkowski inequality gives for such Δ

$$\operatorname{var}_p^k(f_\varepsilon;\Delta;L_q) \leqslant \int\limits_{\mathbb{R}^d} \operatorname{var}_p^k(f;\Delta-\varepsilon y;L_q)\varphi(y)\,dy.$$

Since $\Delta - \varepsilon y := \{\widehat{Q}-\varepsilon y; \widehat{Q}\in\Delta\}$ is the disjoint family of cubes containing for small ε in Q^d, the right-hand side is bounded by $\operatorname{var}_p^k(f;L_q)$. Taking then supremum over all such Δ, we obtain the inequality

$$\operatorname{var}_p^k(f_\varepsilon;Q;L_q) \leqslant \operatorname{var}_p^k(f;L_q)$$

that clearly implies (3.8). □

Remark 3.4. Let Q^d be the extension domain of V_{pq}^k, i.e., there exists a linear continuous extension operator from V_{pq}^k int $V_{pq}^k(\widetilde{Q})$ where $Q^d \subset \widetilde{Q}$ and $dist(Q^d;\mathbb{R}\setminus\widetilde{Q}) > 0$. In this case, f_ε can be defined on the unit cube Q^d and therefore Theorem 3.2 holds for Q^d substituted for Q.

Unfortunately, the corresponding extension theorem is unknown though it exists for some spaces V_{pq}^k, e.g., for $s(V_{pq}^k)=k$. The special case of the last assertion for the space $BV(Q^d)$ and even for more general class of domains is presented, e.g., in [Z, Th, 5.6].

This remark leads to the following:

Conjecture. For every $f\in V_{pq}^k$ there is a sequence $\{f_j\}\subset C^\infty(\mathbb{R}^d)$ such that

$$\|f-f_j\|_{L_q}\to 0 \text{ as } j\to\infty$$

and, moreover,

$$\lim_{i\to\infty} \operatorname{var}_p^k(f_j;L_q) \leqslant \operatorname{var}_p^k(f;L_q) \tag{3.16}$$

3.3. *Polynomial Approximation on d-Rings.*

Theorem 3.5. *Let $Q\subset Q^*$ be dyadic subcubes of Q^d. Then it is true that*

$$E_k(f;Q^*\setminus Q;L_q) \leqslant c(k,d)\operatorname{var}_p^k(f;Q^*\setminus Q;L_q). \tag{3.17}$$

Proof. We need several auxiliary results.

Lemma 3.6. *Let $S_1,S_2\subset\mathbb{R}^d$ be subsets of finite measure such that for $\varepsilon>0$*

$$|S_1\bigcap S_2| \geqslant \varepsilon\cdot\min_{i=0,1}\{|S_i|\}. \tag{3.18}$$

Then the following is true:

$$E_k(f;S_1\bigcup S_2;L_q) \leqslant c\varepsilon^{-k+1}\sum_{i=0}^{1} E_k(f;S_i;L_q). \tag{3.19}$$

For the proof see, e.g., [BB, p. 123].

Lemma 3.7. *Let $\{S_j\}_{1\leqslant j\leqslant N}$ be a family of subsets in $\mathbb{R}^d$ of finite measure such that for some $\varepsilon > 0$*

$$|S_j \bigcap S_{j+1}| \geqslant \varepsilon \min\{|S_j|, |S_{j+1}|\},\ 1 \leqslant j < N. \tag{3.20}$$

Then it is true that

$$E_k(f; \bigcup_{j=1}^{N} S_j; L_q) \leqslant c \sum_{j=1}^{N} E_k(f, S_j; L_q) \tag{3.21}$$

where $c = (c(k,d)\varepsilon^{-k+1})^{N-1}$.

Proof (induction on N). For $N = 2$ the result follows from (3.19). Now assume that (3.21) holds for $N \geqslant 2$ and prove it for $N+1$.

Setting $S^M := \bigcup\limits_{j=1}^{M} S_j$ we get from (3.18)

$$|S^N \bigcap S_{N+1}| \ \geqslant\ \varepsilon |S_N \bigcap S_{N+1}| \ \geqslant\ \varepsilon \min\{|S_N|, |S_{N+1}|\} \ =\ \varepsilon \min\{|S^N|, |S_{N+1}|\}.$$

Further, Lemma 3.6 implies

$$E_k(f; \bigcup_{j=1}^{N+1} S_j; L_q) \leqslant c(k,d)\varepsilon^{-k+1}(E_k(f; S^N; L_q) + E_k(f; S_{N+1}; L_q))$$

while the induction hypothesis gives

$$E_k(f; S^N; L_q) \leqslant (c(k,d)\varepsilon^{-k+1})^{N-1} \sum_{j=1}^{N} E_k(f; S_j; L_q).$$

Combining these we get the result for $N+1$. ☐

Now we formulate the mentioned cover lemma proved in Appendix I.

Theorem I.1. *There exists a cover $\mathcal{K}$ of $Q^* \setminus Q$ by cubes such that the following holds:*

For every overlapping pair[2] $K_1, K_2 \in \mathcal{K}$

$$|K_1 \bigcap K_2| \geqslant \frac{1}{2} \min\{|K_1|, |K_2|\}, \tag{3.22}$$

and, moreover,

$$\operatorname{card} \mathcal{K} \leqslant 4(2^d - 1). \tag{3.23}$$

[2]i.e., intersecting by a set of positive measure

Now we complete the proof of Theorem 3.5.

By Lemma 3.7 and Theorem I.1 we have

$$E_k(f; Q^* \setminus Q; L_q) \leqslant \big(c(k,d)2^{k-1}\big)^{4(2^d-1)} \sum_{K \in \mathcal{K}} E_k(f; K; L_q).$$

Moreover, by the definition of (k,p)-variation, see (3.1),

$$E_k(f; K; L_q) \leqslant \operatorname{var}_p^k(f; Q^* \setminus Q; L_q)$$

for every $K \in \mathcal{K}$.

Together with the previous inequality this gets the required result

$$E_k(f; Q^* \setminus Q; L_q) \leqslant c(k,d) \operatorname{var}_p^k(f; Q^* \setminus Q; L_q).$$

□

§ 4. Proof of Theorem 2.1

4.1. We begin with part (a) of this result and then derive from (a) part (b).

Let $f \in V_{pq}^k (:= V_{pq}^k(Q^d))$ where

$$1 \leqslant p < q < \infty, \ d \geqslant 2 \text{ and } 0 < s := s(V_{pq}^k) \leqslant k. \tag{4.1}$$

Without loss of generality we assume that

$$|f|_{V_{pq}^k} = 1. \tag{4.2}$$

Under these assumptions given $N \in \mathbb{N}$ we prove existence of a cover Δ_N of Q^d by at most N dyadic cubes and a piecewise polynomial $g_N \in \mathcal{P}_{k-1}(\Delta_N)$ such that

$$\|f - g_N\|_q \leqslant c(k,d) N^{-s/d}. \tag{4.3}$$

First, let $C^\infty(\mathbb{R}^d) \cap V_{pq}^k$. The proof of (4.3) for this case begins with the construction of the cover Δ_N. This will be obtained by the algorithm presented now.

4.2. *Description of the algorithm.* The important ingredient of the algorithm is a *weight* W defined on the σ-algebra $A(\mathcal{D})$ generated by dyadic cubes of Q^d. This by definition is a function $W : A(\mathcal{D}) \to \mathbb{R}_+$ satisfying the conditions.

(*Subadditivity*) For a disjoint family $\{S_i\} \subset A(\mathcal{D})$

$$\sum W(S_i) \leqslant W\Big(\bigcup S_i\Big). \tag{4.4}$$

(*Absolute continuity*)

$$\lim_{|S| \to 0} W(S) = 0. \tag{4.5}$$

We *normalize* W by

$$W(Q^q) = 1. \tag{4.6}$$

To prove Theorem 2.1(a) for $f \in V^k_{pq} \bigcap C^\infty$ we define a weight W by

$$W(S) := \operatorname{var}^k_p(f; S; L_q)^p, \quad S \in A(\mathcal{D}). \tag{4.7}$$

Due to Propositions 3.1^a, 3.1^b and (4.2) W satisfies the required properties (4.4)–(4.6).

In the construction of the algorithm we essentially exploit the canonical *graph structure* of the set $\mathcal{D}$ regarding as the *vertex set* while the *edge set* consists of pairs $\{Q', Q\} \subset \mathcal{D}$ such that $Q' \subset Q$ and $|Q'| = 2^{-d}|Q|$. In this case, we use the notation $Q' \to Q$ and call Q' the *son* of Q and Q the *father* of Q'.

The set of all 2^d sons of Q is denoted by $\mathcal{D}_1(Q)$. This clearly is the uniform partition of Q into 2^d congruent subcubes.

Further, a *path* in the graph $\mathcal{D}$ is a sequence

$$P := \{Q_1 \to Q_2 \to \ldots \to Q_n\}. \tag{4.8}$$

The vertices (cubes) Q_1, Q_n are called the *tail* and the *head* of P, respectively. Moreover, we use the notations

$$P := [Q_1, Q_n],\ Q_1 =: T_P =: P^-,\ Q_n =: H_P =: P^+. \tag{4.9}$$

It is readily seen that the following is true.

Proposition 4.1. *If $Q' \subset Q$ are dyadic cubes of $\mathcal{D}$, there exists a unique path joining Q' and Q.*

In terms of Graph Theory, $\mathcal{D}$ is a *rooted tree* with the *root* Q^d.

More generally, the set $\mathcal{D}(Q)$ of all dyadic subcubes of $Q \in \mathcal{D}$ is a *rooted tree* with the *root* Q.

For $N \in \mathbb{N}$ and W given by (4.7) the subset of "bad" cubes of $\mathcal{D}$ is defined by

$$G_N := \{Q \in \mathcal{D};\ W(Q) \geqslant N^{-1}\}; \tag{4.10}$$

clearly, $Q^d \in G_N$, see (4.6), and G_N is finite, see (4.5).

The algorithm gives the following partition of G_N into the set of (basic) paths, see Proposition II.1 of Appendix II for the proof.

Proposition 4.2. *There exists partition $\mathcal{B}_N$ of the set $G_N \setminus \{Q^d\}$ into N paths such that*[3]

$$W(H_B \setminus T_B) < N^{-1}, \quad B \in \mathcal{B}_N, \tag{4.11}$$

and, moreover,

$$\operatorname{card} \mathcal{B}_N \leqslant 3N + 1. \tag{4.12}$$

Now we decompose the remaining part of $\mathcal{D}$

$$G^c_N := \mathcal{D} \setminus G_N. \tag{4.13}$$

[3]among basic paths may be singletons. In this case, $H_B = T_B$ and $W(H_B \setminus T_B) = 0$.

To this end we define the *boundary* of G_N denoted by ∂G_N that consists of all *maximal cubes* of G_N^c with respect to the set–inclusion order.

In other words, every $Q' \in \mathcal{D}$ containing $Q \in \partial G_N (\subset G_N^c)$ as a proper subset belongs to G_N. In particular, if Q^+ is the father of $Q \in \partial G_N$, then

$$W(Q) < N^{-1} \text{ and } W(Q^+) \geqslant N^{-1}. \tag{4.14}$$

Proposition 4.3. *(a) The family $\{\mathcal{D}(Q); Q \in \partial G_N\}$ is disjoint and*

$$G_N^c = \bigcup_{Q \in \partial G_N} \mathcal{D}(Q), \tag{4.15}$$

i.e., the family is a partition of G_N^c.

(b) The following is true

$$\operatorname{card}(\partial G_N) \leqslant 2^d N. \tag{4.16}$$

Proof. (a) Maximal cubes are pairwise disjoint. Hence, ∂G_N is a disjoint family.

Further, cubes $Q \in \partial G_N$ are roots of the trees $\mathcal{D}(Q)$ from (4.15). Since the roots are disjoint, the corresponding trees are as well, i.e., $\{\mathcal{D}(Q); Q \in \partial G_N\}$ is a disjoint family.

To prove that the family is a partition of G_N^c we check that every $Q' \in G_N^c$ belongs to some $\mathcal{D}(Q)$ where $Q \in \partial G_N$.

Let $Q' =: Q_1 \to Q_2 \to \ldots \to Q^d$ be the path joining Q' and Q^d, and Q_i, $i \geqslant 2$, be the smallest cube of the path belonging to G_N. Then its son Q_{i-1} belongs to G_N^c, i.e., Q_{i-1} is maximal, and $Q' \in \mathcal{D}(Q_{i-1})$ as required.

(b) Let Q^+ be the father of $Q \in \partial G_N$ and $(\partial G_N)^+ := \{Q^+; Q \in \partial G_N\}$. Since Q^+ is unique, the set $(\partial G_N)^+$ is disjoint.

Further, every father has 2^d sons and therefore

$$\operatorname{card}(\partial G_N) \leqslant 2^d \operatorname{card}(\partial G_N)^+. \tag{4.17}$$

Finally, (4.14), subadditivity of W and (4.6) imply

$$N^{-1} \operatorname{card}(\partial G_N)^+ < \sum_{Q \in (\partial G_N)^+} W(Q) \leqslant W(Q^d) = 1. \tag{4.18}$$

This and (4.17) give (4.16). □

Finally, the required cover Δ_N is given by

$$\Delta_N := \{Q^d\} \bigcup \left(\bigcup_{B \in \mathcal{B}_N} \{T_B, H_B\} \bigcup \mathcal{D}_1(T_B) \right). \tag{4.19}$$

Due to (4.12)

$$\operatorname{card} \Delta_N \leqslant 1 + 2(3N+1) + 2^d(3N+1) =: c(d)N. \tag{4.20}$$

4.3. *Definition of N Term Approximation for C^∞ Functions.*

Now we define the required $g_N \in \mathcal{P}_{k-1}(\Delta_N)$ using to this end polynomials of best approximation determined by

$$\|f - m_S\|_{L_q(S)} = E_k(f; S; L_q). \tag{4.21}$$

Further, we use for brevity the following notations

$$M_Q := \sum_{Q' \in \mathcal{D}_1(Q)} m_{Q'} \cdot 1_{Q'} - m_Q \cdot 1_Q, \quad Q \in \mathcal{D}; \tag{4.22}$$

and, moreover,

$$B^+ := H_B, \quad B^0 := H_B \setminus T_B, \quad B^- := T_B. \tag{4.23}$$

Using this we write

$$g_N := m_{Q^d} + \sum_{B \in \mathcal{B}_N} [(m_{B^+} - m_{B^0}) \cdot 1_{B^+} + (m_{B^0} - m_{B^-}) \cdot 1_{B^-} + M_{B^-}]. \tag{4.24}$$

This clearly is a piecewise polynomial of degree $k-1$ over Δ_N, see (4.19).

Let us note that for B being a singleton $B^{\pm} = \{B\}$, $B^0 = \emptyset$, i.e., the corresponding terms in (4.24) and (4.19) equal M_B and $\{\{B\}, \mathcal{D}_1(\{B\})\}$, respectively.

Theorem 2.1(a) will be derived from the next key result.

Proposition 4.4. *Let $f \in V^k_{pq} \bigcap C^\infty(\mathbb{R}^d)$ where $d, p, q, s = s(V^k_{pq})$ satisfy (4.1) and (4.2). Given $N \in \mathbb{N}$ there exist a cover $\Delta_N \subset \mathcal{D}$ of Q^d and a piecewise polynomial $g_N \in \mathcal{P}_{k-1}(\Delta_N)$ such that*

$$\|f - g_N\|_q \leqslant c(k,d) N^{-s/d} \tag{4.25}$$

and, moreover,

$$\operatorname{card} \Delta_N \leqslant c(d) N. \tag{4.26}$$

The proof of the proposition is given in Section 4.4 below. It employs some auxiliary results presented now.

To introduce the family Δ_N we use the algorithm for the weight W given by (4.7).

Since W satisfies the assumptions of Proposition 4.2, see (4.4)–(4.6), it determines the *finite* set $G_N \subset \mathcal{D}$, and the algorithm gives the partition $\mathcal{B}_N$ of $G_N \setminus \{Q^d\}$ into the basic paths which in turn determines the required cover Δ_N, see (4.19) and (4.20).

To estimate $f - g_N$ we need a suitable presentation of this difference; the next lemmas are used for its derivation.

Lemma 4.5. *Let $f \in L_q(Q) \bigcap C(\mathbb{R}^d)$, $1 \leqslant q \leqslant \infty$, $Q \in \mathcal{D}$. Then the following holds*

$$f = m_Q + \sum_{Q' \in \mathcal{D}(Q)} M_{Q'} \tag{4.27}$$

with convergence in $L_q(Q)$.

Proof. Let $\mathcal{D}_j(Q)$, $j \in \mathbb{Z}$, be the partition of Q into 2^{jd} congruent (dyadic) cubes, e.g., $\mathcal{D}_0(Q) = \{Q\}$ and $\mathcal{D}_1(Q)$ is the set of sons for Q. Then $P_j \in \mathcal{P}_{k-1}(\mathcal{D}_j(Q))$ is defined by

$$P_j := \sum_{Q' \in \mathcal{D}_j(Q)} m_{Q'} \cdot 1_{Q'}. \tag{4.28}$$

We show that

$$f - m_Q = \sum_{j \geqslant 0} (P_{j+1} - P_j) \text{ (convergence in } L_q(Q)). \tag{4.29}$$

Let s_n be the n-th partial sum of the series (4.29). Then

$$f - m_Q - s_n = f - P_{n-1} = \sum_{Q' \in \mathcal{D}_n(Q)} (f - m_{Q'}) \cdot 1_{Q'}.$$

This and (4.21) imply that

$$\|f - m_Q - s_n\|_q = \left\{ \sum_{Q' \in \mathcal{D}_n(Q)} \|f - m_{Q'}\|^q_{L_q(Q')} \right\}^{1/q} = \left\{ \sum_{Q' \in \mathcal{D}_n(Q)} E_k(f; Q'; L_q) \right\}^{1/q}.$$

By Theorem 4 of [B-71, § 2] the right–hand side is bounded by $c(k,d)\omega_k \left(f; \frac{|Q|^{1/d}}{2^n}; L_q(Q) \right)$. Since this bound tends to 0 as $n \to \infty$ for $q < \infty$ and for $q = \infty$ and $f \in C(\mathbb{R}^d)$, (4.29) is proved. Using now notations (4.22) and (4.28) we obtain

$$P_{j+1} - P_j = \sum_{Q' \in \mathcal{D}_j(Q)} M_{Q'}.$$

Summing over $j \geq 0$ and using (4.29) we then have

$$f - m_Q = \sum_{j \geq 0} \sum_{Q' \in \mathcal{D}_j(Q)} M_{Q'} = \sum_{Q' \in \mathcal{D}(Q)} M_{Q'}.$$

The proof is complete. □

Now we apply (4.27) for $Q = Q^d$ and present $\mathcal{D} = \mathcal{D}(Q^d)$ as follows:

$$\mathcal{D} = \left(\sum_{B \in \mathcal{B}_N} \sum_{Q \in B} Q \right) \bigcup \left(\sum_{Q \in \partial G_N} \mathcal{D}(Q) \right),$$

see Proposition 4.2 and (4.15). This then implies the identity

$$f - m_{Q^d} = \sum_{B \in \mathcal{B}_N} \sum_{Q \in B} M_Q + \sum_{Q \in \partial G_N} \sum_{Q' \in \mathcal{D}(Q)} M_{Q'}.$$

Rewriting the second sum here by (4.27) we have

$$f - m_{Q^d} = \sum_{B \in \mathcal{B}_N} \sum_{Q \in B} M_Q + \sum_{Q \in \partial G_N} (f - m_Q) \cdot 1_Q.$$

Subtracting from here equality (4.24) for g_N we obtain the required presentation

$$f - g_N = \sum_{B \in \mathcal{B}_N} S_B + \sum_{Q \in \partial G_N} (f - m_Q) \cdot 1_Q; \tag{4.30}$$

here we set[4]

$$S_B = \left(\sum_{Q \in B \setminus \{B^-\}} M_Q \right) - [(m_{B^+} - m_{B^0}) \cdot 1_{B^+} + (m_{B^0} - m_{B^-}) \cdot 1_{B^-}]. \tag{4.31}$$

The next result gives the basic presentation of S_B.

Lemma 4.6. *The following is true*

$$S_B = \sum_{Q \in B \setminus \{B^-\}} \sum_{Q' \in \mathcal{D}_1(Q) \setminus B} (m_{B^0} - m_{Q'}) \cdot 1_{Q'}. \tag{4.32}$$

Proof. We begin with the identity

$$\sum_{Q \in B \setminus \{B^-\}} M_Q = \sum_{Q \in B \setminus \{B^-\}} \sum_{Q' \in \mathcal{D}_1(Q) \setminus B} [(m_{Q'} - m_{B^+}) \cdot 1_{Q'} + (m_{B^-} - m_{B^+}) \cdot 1_{B^-}] \tag{4.33}$$

proved by induction on card B.

Let $B := [Q_1, Q_n] = \{Q_1 \to Q_2 \to \ldots \to Q_n\}$, i.e., $B^- := Q_1$, $B^+ := Q_n$. Since $\mathcal{D}_1(Q) \setminus B$ for $Q \in B \setminus \{B^-\}$ consists of all sons of Q excluding the son belonging to B,

$$\mathcal{D}_1(Q_i) \setminus B = \mathcal{D}_1(Q_i) \setminus \{Q_{i-1}\}, \ i \geqslant 2.$$

Denoting the right–hand side by $\mathcal{D}^*(Q_i)$ we then rewrite (4.33) as follows.

$$\sum_{i=2}^{n} M_{Q_i} = \sum_{i=2}^{n} \sum_{Q \in \mathcal{D}_1^*(Q_i)} [(m_Q - m_{Q_n}) \cdot 1_Q] + (m_{Q_1} - m_{Q_n}) \cdot 1_{Q_1}. \tag{4.34}$$

For $n = 2$ the right–hand side of (4.34) equals

$$\sum_{Q \in \mathcal{D}_1^*(Q_2)} [(m_Q - m_{Q_2}) \cdot 1_Q] + (m_{Q_1} - m_{Q_2}) \cdot 1_{Q_1} := \sum_{Q \in \mathcal{D}_1(Q_2)} m_Q \cdot 1_Q - m_{Q_2} \left(\sum_{Q \in D_1^*(Q_2)} 1_Q + 1_{Q_1} \right).$$

Since $\mathcal{D}_1^*(Q_2)$ is a partition of $Q_2 \setminus Q_1$, the sum in the parentheses equals $1_{Q_2 \setminus Q_1} + 1_{Q_1} = 1_{Q_2}$. Hence, the right–hand side here equals M_{Q_2}, see (4.22), as required.

[4] $S_B = 0$ if B is a singleton

Now let (4.33) hold for all paths of cardinality $n \geqslant 2$. To prove it for $n+1$ we write (4.34) for the n-term path $\{Q_2 \to \ldots \to Q_{n+1}\}$ and add to it (4.34) for $n=2$ written equivalently as follows:

$$m_{Q_2} = \sum_{Q \in \mathcal{D}_1^*(Q_2)} (m_Q - m_{Q_{n+1}}) \cdot 1_Q + (m_{Q_{n+1}} - m_{Q_2}) \cdot 1_{Q_2} + (m_{Q_1} - m_{Q_{n+1}}) \cdot 1_{Q_1}.$$

Together with the equality

$$\sum_{i=3}^{n+1} M_{Q_i} = \sum_{i=3}^{n+1} \sum_{Q \in \mathcal{D}_1^*(Q_i)} (m_Q - m_{Q_{n+1}}) \cdot 1_Q + (m_{Q_2} - m_{Q_{n+1}}) \cdot 1_{Q_2} \tag{4.35}$$

this gives

$$\sum_{i=2}^{n+1} M_{Q_i} = \sum_{i=2}^{n+1} \sum_{Q \in \mathcal{D}_1^*(Q_i)} (m_Q - m_{Q_{n+1}}) \cdot 1_Q + R \tag{4.36}$$

where we set

$$\begin{aligned} R := (m_{Q_{n+1}} - m_{Q_2}) \cdot 1_{Q_2} + (m_{Q_1} - m_{Q_{n+1}}) \cdot 1_{Q_1} + (m_{Q_2} - m_{Q_{n+1}}) \cdot 1_{Q_2} \\ = (m_{Q_1} - m_{Q_{n+1}}) \cdot 1_{Q_1}. \end{aligned}$$

Hence, (4.36) proves the required equality (4.33) for $n+1$.

Now we transform (4.34) by adding and subtracting $m_{B^0} (:= m_{Q_n \setminus Q_1})$. This gives

$$\begin{aligned} \sum_{i=2}^{n} M_{Q_i} = \sum_{i=2}^{n} \sum_{Q \in \mathcal{D}_1^*(Q_i)} (m_Q - m_{B^0}) \cdot 1_Q + (m_{B^0} - m_{Q_n}) \sum_{i=2}^{n} \sum_{Q \in \mathcal{D}_1^*(Q_i)} 1_Q \\ + (m_{Q_1} - m_{B^0}) \cdot 1_{Q_1} + (m_{B^0} - m_{Q_n}) \cdot 1_{Q_1}. \end{aligned}$$

Since the second sum here equals $\sum_{i=2}^{n} 1_{Q_i \setminus Q_{i-1}} = 1_{Q_n \setminus Q_1}$ and, in the chosen notations, see (4.31),

$$S_B := \sum_{i=2}^{n} M_{Q_i} - (m_{Q_n} - m_{B^0}) \cdot 1_{Q_n} - (m_{B^0} - m_{Q_1}) \cdot 1_{Q_1}, \tag{4.37}$$

these two equalities give

$$S_B = \sum_{i=2}^{n} \left[\sum_{Q \in \mathcal{D}_1^*(Q_i)} (m_Q - m_{B^0}) \cdot 1_Q \right] + R$$

where the remainder R equals

$$\begin{aligned} R := [(m_{B^0} - m_{Q_n}) \cdot 1_{Q_n \setminus Q_1} + (m_{B^0} - m_{Q_n}) \cdot 1_{Q_1} \\ + (m_{Q_1} - m_{B^0}) \cdot 1_{Q_1}] - [(m_{Q_n} - m_{B^0}) \cdot 1_{Q_n} + (m_{B^0} - m_{Q_1}) \cdot 1_{Q_1}]. \end{aligned} \tag{4.38}$$

Since the square parentheses here annihilate, $R = 0$.

The identity (4.32) is proved. $\square$

4.4. *Proof of Proposition 4.4.* We should prove that for $f \in V^k_{pq} \bigcap C^\infty(\mathbb{R}^d)$

$$\|f - g_N\|_q \leqslant c(k,d) N^{-s/d}. \tag{4.39}$$

Due to the presentation (4.30)

$$\|f - g_N\|_q \leqslant \left\| \sum_{B \in \mathcal{B}_N} S_B \right\|_q + \left\| \sum_{Q \in \partial G_N} (f - m_Q) \cdot 1_Q \right\|_q \tag{4.40}$$

and it remains to estimate each term of the sum.

Lemma 4.7. *(a) Supports of the functions S_B, $B \in \mathcal{B}_N$, are disjoint.*
(b) It is true that

$$\|S_B\|_q \leqslant c(k,d) \operatorname{var}^k_p(f; B^+ \setminus B^-; L_q). \tag{4.41}$$

Proof. (a) Since $supp\, S_B = B^+ \setminus B^-$, see Lemma 4.6, the supports of S_B and $S_{\widetilde{B}}$ are disjoint if their heads are. Otherwise, one of these (dyadic) cubes, say, $\widetilde{B}^+$, embeds into the other. Then $\widetilde{B}^+$ embeds into the tail B^- of the path B. Hence, $supp\, S_{\widetilde{B}}$ does not intersect $supp\, S_B = B^+ \setminus B^-$.

(b) By identity (4.32)

$$S_B = \sum_{Q \in B^*} \sum_{Q' \in \mathcal{D}^*_1(Q)} (m_{B^0} - f + f - m_{Q'}) \cdot 1_{Q'}$$

where for brevity we set $B^* := B \setminus \{B^-\}$.

Further, we have

$$S_B = (m_{B_0} - f) \sum_{Q \in B^*} \sum_{Q' \in \mathcal{D}^*_1(Q)} 1_Q + \sum_{Q \in B^*} \sum_{Q' \in \mathcal{D}^*_1(Q)} (f - m_{Q'}) \cdot 1_{Q'}.$$

Since the family $\bigcup\limits_{Q \in B^*} \mathcal{D}^*_1(Q)$ is a partition of $B^+ \setminus B^-$, the sum of indicators here equals $1_{B^+ \setminus B^-}$ and the equality implies

$$\|S_B\|_q \leqslant \|f - m_{B^0}\|_{L_q(B^+ \setminus B^-)} + \left(\sum_{Q \in B^*} \sum_{Q' \in \mathcal{D}^*_1(Q)} \|f - m_{Q'}\|^q \right)^{1/q}$$

$$= E_k(f; B^+ \setminus B^-; L_q) + \left(\sum_{Q \in B^*} \sum_{Q' \in \mathcal{D}^*_1(Q)} E_k(f; Q'; L_q)^q \right)^{1/q}.$$

By the Jenssen inequality the second term is bounded by

$$\left\{ \sum_{Q \in B^*} \sum_{Q' \in \mathcal{D}^*_1(Q)} E_k(f; Q'; L_q)^p \right\}^{1/p}.$$

Since the family $\bigcup\limits_{Q\in B^*} \mathcal{D}_1^*(Q)$ is a partition of $B^+ \setminus B^-$, this sum is bounded by $\operatorname{var}_p^k(f; B^+ \setminus B^-; L_q)$, see the definition of (k,p)–variation in (3.1).

Moreover, by Theorem 3.5

$$\|f - m_{B^0}\|_{L_q(B^0)} := E_k(f; B^+ \setminus B^-) \leqslant c(k,d) \cdot \operatorname{var}_p^k(f; B^+ \setminus B^-; L_q).$$

Combining this with the previous inequality we obtain (4.41). □

Now we use Lemma 4.7 to estimate the first term in (4.40). We have

$$\left\| \sum_{B\in\mathcal{B}_N} S_B \right\|_q \leqslant \left\{ \sum_{B\in\mathcal{B}_N} \|S_B\|_q^q \right\}^{1/q}$$

$$\leqslant c(k,d) \left\{ \sum_{B\in\mathcal{B}_N} \operatorname{var}_p^k(f; B^+ \setminus B^-; L_q)^q \right\}^{1/q}.$$

Moreover, by the definition of the weight W, see (4.7), and the inequality (4.11) of Proposition 4.2

$$\operatorname{var}_p^k(f; B^+ \setminus B^-; L_q) := W(B^+ \setminus B^-)^{1/p} \leqslant N^{-1/p}.$$

Combining with the previous inequality and using (4.12) we finally have the required estimate

$$\left\| \sum_{B\in\mathcal{B}_N} S_B \right\|_q \leqslant c(k,d) \left(N^{-q/p} \operatorname{card} \mathcal{B}_N \right)^{1/q} \leqslant c(k,d) \left(N^{-q/p}(3N+1) \right)^{1/q}$$

$$\leqslant c_1(k,d) N^{-1/p+1/q} := c_1(k,d) N^{-s/d}.$$

It remains to obtain the similar bound for the sum over boundary ∂G_N in (4.40).

Due to Proposition 4.3 and (4.14) ∂G_N is disjoint and, moreover,

$$\operatorname{var}_p^k(f; Q; L_q)^p =: W(Q) < N^{-1}$$

for every $Q \in \partial G_N$.

This immediately implies

$$\left\| \sum_{Q\in\partial G_N} (f - m_Q) \cdot 1_Q \right\|_q = \left\{ \sum \|f - m_Q\|_{L_q(Q)}^q \right\}^{1/q} := \left\{ \sum_{Q\in\partial G_N} E_k(f; Q; L_q)^q \right\}^{1/q}$$

$$\leqslant \left\{ \sum_{Q\in\partial G_N} \operatorname{var}_p^k(f; Q; L_q)^{q/p} \right\}^{1/p} \leqslant N^{-1/p} \left(\operatorname{card} G_N \right)^{1/q}.$$

Since $\operatorname{card} G_N \leqslant 2^d N$, see (4.16), this finally gives

$$\left\| \sum_{Q\in\partial G_N} (f - m_Q) \cdot 1_Q \right\|_q \leqslant 2^{d/q} N^{-s/d}$$

as required.

Proposition 4.4 is proved. □

4.5. *Proof of Theorem 2.1(a).* We derive the result from Theorem 3.2 and Proposition 4.4.

Let $Q := [1-\delta, \delta)$, $\delta > 0$, and $f \in V^k_{pq}$ if $q < \infty$ and $f \in V^k_{pq} \bigcap C(\mathbb{R}^d)$ if $q = \infty$.

Given $\varepsilon > 0$ Theorem 3.2 then yields a function $f_\varepsilon \in C^\infty(\mathbb{R}^d)$ such that

$$\|f - f_\varepsilon\|_{L_q(Q)} \leqslant \varepsilon \tag{4.42}$$

and, moreover,

$$\operatorname{var}^k_p(f_\varepsilon; Q; L_q) \leqslant |f|_{V^k_{pq}}. \tag{4.43}$$

Since Proposition 4.4 is homothety–invariant, it remains true for Q substituted for Q^d. Hence, given $N \in \mathbb{N}$ there exist a cover $\widetilde{\Delta}_N \subset \mathcal{D}(Q)$ of Q and a piecewise polynomial $\widetilde{g}_N \in \mathcal{P}_{k-1}(\widetilde{\Delta}_N)$ such that

$$\|f_\varepsilon - \widetilde{g}_N\|_{L_q(Q)} \leqslant c(k,d) N^{-s/d} \operatorname{var}^k_p(f; Q; L_q) \tag{4.44}$$

and, moreover,

$$\operatorname{card} \widetilde{\Delta}_N \leqslant c(d) N. \tag{4.45}$$

Now let h be a homothety mapping Q onto Q^d, i.e.,

$$h(x) := \frac{x - \delta e}{1 - 2\delta}, \quad x \in \mathbb{R}^d,$$

where $e := (1, \dots, 1)$.

Then $\Delta_N := h(\widetilde{\Delta}_N) \subset \mathcal{D}(:= \mathcal{D}(Q^d))$ is a cover of Q^d satisfying

$$\operatorname{card} \Delta_N = \operatorname{card} \widetilde{\Delta}_N \leqslant c(d) N; \tag{4.46}$$

moreover, $g_N := \widetilde{g}_N \circ h^{-1}$ is a piecewise polynomial from $\mathcal{P}_{k-1}(\Delta_N)$.

We will show that for $f \in V^k_{pq}$ with $q < \infty$ and for $f \in V^k_{p\infty} \bigcap C(\mathbb{R}^d)$

$$\|f - g_N\|_q \leqslant c(k,d) N^{-s/d} |f|_{V^k_{pq}}; \tag{4.47}$$

this clearly implies Theorem 1.2(a) for $N \geqslant c(d)$, see (4.46).

Let $h^* g := g \circ h^{-1}$, $g \in L_q(Q^d)$. Then $h^* : L_q(Q^d) \to L_q(Q)$ and $\|h^*\| = (1-2\delta)^{d/q}$.

Further, we write

$$\begin{aligned} \|f - g_N\|_q &\leqslant \|(f \circ h - f_\varepsilon) \circ h^{-1}\|_q + \|(f_\varepsilon - \widetilde{g}_N) \circ h^{-1}\|_q \\ &\leqslant (1-2\delta)^{d/q} \left(\|f - f_\varepsilon\|_{L_q(Q)} + \|f - f \circ h\|_{L_q(Q)} + \|f_\varepsilon - \widetilde{g}_N\|_{L_q(Q)} \right). \end{aligned}$$

By (4.44) and (4.43) the third term in the parentheses is bounded by $c(k,d) N^{-s/d} |f|_{V^k_{pq}}$ while the first tends to 0 as $\varepsilon \to 0$, see (4.42), and the second does as $\delta \to 0$ for $q < \infty$, and also for $q = \infty$, if f is uniformly continuous on Q.

This proves (4.47) and Theorem 1.2(a) for $N \geqslant c(d)$.

To obtain the result for $1 \leqslant N < c(d)$ we simply set $\Delta_N := \{Q^d\}$ and $g_N := m_{Q^d}$. Then

$$\|f - g_N\|_q = E_k(f; Q^d; L_q) < c(d)^{s/d} N^{-s/d} |f|_{V^k_{pq}}$$

and, moreover, card $\Delta_N = 1 \leqslant N$.

This gives Theorem 2.1(a) for all $N \geqslant 1$. $\square$

4.6. *Proof of Theorem 2.1(b).* We establish the analog of Theorem 2.1(a) with a partition of Q^d into d-rings of cardinality at most $c(d)N$.

To this end we write the piecewise polynomial g_N of Theorem 2.1(a), see (4.24), in the form

$$g_N := m_{Q^d} + \sum_{Q \in \Delta_N} P_Q \cdot 1_Q \tag{4.48}$$

where $P_Q \in \mathcal{P}_{k-1}$ and

$$\Delta_N := \bigcup_{B \in \mathcal{B}_N} (\{H_B, T_B\} \bigcup \mathcal{D}_1(B^-)),$$

see (4.19).

First, we assume that Δ_N *covers* Q^d. If Δ_N is not a partition (otherwise, the result is clear), it contains at least one *tower*

$$T := \{Q_1 \subsetneqq \dots \subsetneqq Q_n\} \subset \Delta_N.$$

This means that for every $0 \leqslant i \leqslant n$ there is no $Q \in \Delta_N$ such that $Q_i \subsetneqq Q \subsetneqq Q_{i+1}$; here $Q_0 := \emptyset$, $Q_{n+1} := Q^d$; hence, the *bottom* $Q_1 \neq \emptyset$ and the *top* Q_n are, respectively, minimal and maximal cubes of T closest to Q^d.

According to this definition $(\Delta_N \setminus T) \bigcup \{Q_n\}$ still covers Q^d. Moreover, the tops of different towers do not intersect.

These, in particular, imply that if T_j, $1 \leqslant j \leqslant m$, are all towers of Δ_N and $Q(T_j)$ are their tops, then

$$\left(\Delta_N \setminus \bigcup_{j=1}^m T_j\right) \bigcup \left(\bigcup_{j=1}^m Q(T_j)\right)$$

is a partition of Q^d.

Hence, it suffices to subdivide each $Q(T_j)$ into a set of d-rings whose cardinality equals card T_j. We do this for $m = 1$ and then repeat the procedure for the remaining towers.

Now let $T := \{Q_1 \subsetneqq \dots \subsetneqq Q_n\}$ be the single tower of Δ_N. Setting

$$R_i := Q_i \setminus Q_{i-1}, \quad 1 \leqslant i \leqslant n,$$

where $R_1 = Q_1$ as $Q_0 := \emptyset$, we obtain the partition $\mathcal{R}_n := \{R_i\}_{1 \leqslant i \leqslant n}$ of $Q_n := Q(T)$ into d-rings.

Further, we define the family of polynomials $\{P_{R_i}\} \subset \mathcal{P}_{k-1}$ given by $P_{R_i} := \left(\sum_{j=i}^n P_{Q_j}\right) \cdot 1_{R_j}$, $1 \leqslant i \leqslant n$. These definitions imply the identity

$$\sum_{i=1}^n P_{Q_i} \cdot 1_{Q_i} = \sum_{R \in \mathcal{R}_n} P_R \cdot 1_R. \tag{4.49}$$

Moreover, the T is single in Δ_N, hence, $\mathcal{R}_n \bigcup (\Delta_N \setminus T)$ is a partition of Q^d into $\leqslant n + (N-n) = N$ d-rings while the piecewise polynomial

$$\widetilde{g}_N := m_{Q^d} + \sum_{R \in \mathcal{R}_n} P_R \cdot 1_R + \sum_{Q \in \Delta_N \setminus T} P_Q \cdot 1_Q$$

belongs to $\mathcal{P}_{k-1}(\mathcal{R}_n \bigcup (\Delta_N \setminus T))$ and equals g_N by (4.49) and (4.48).

This gives the result for Δ_N being a cover of Q^d.

Now suppose that Δ_N is not a cover of Q^d. Then $\mathcal{D}_1(Q^d) \bigcap G_N^c \neq \emptyset$, since otherwise $\mathcal{D}_1(Q^d) \subset G_N$, i.e., every son of Q^d is the head of a basic path. By the definition of Δ_N, see (4.19), this implies that $\mathcal{D}_1(Q) \subset \Delta_N$, i.e., Δ_N is a cover of Q^d, a contradiction.

Further, the set of heads $\mathcal{D}_1(Q^d) \bigcap G_N$ is contained in $\Delta_N \subset \mathcal{D} \setminus \{Q^d\}$, and, moreover, it is nonempty as for otherwise $\Delta_N = \{Q^d\}$.

Hence, the set

$$\widetilde{\Delta}_N := \Delta_N \bigcup (\mathcal{D}_1(Q^d) \bigcap G_N^c)$$

is a cover of Q^d and its cardinality is bounded by

$$N + \operatorname{card} \mathcal{D}_1(Q^d) - 1 = N + 2^d - 1 \leqslant 2^d N.$$

To complete the proof it suffices to modify the g_N to obtain $\widetilde{g}_N \in \mathcal{P}_{k-1}(\Delta_N \bigcup (\mathcal{D}_1(Q^d) \bigcap G_N^c))$ such that

$$\|f - \widetilde{g}_N\|_q \leqslant c(k,d) N^{-s/d} |f|_{V_{pq}^k}. \tag{4.50}$$

We define $\widetilde{g}_N$ by

$$\widetilde{g}_N := g_N + \sum_{Q \in \mathcal{D}_1(Q^d) \bigcap G_N^c} (m_Q - m_{Q^d}) \cdot 1_Q$$

and then prove (4.50).

Substituting here g_N by the right–hand side of (4.48) and using the notations

$$S := \bigcup_{Q \in \Delta_N} Q, \quad \widetilde{\Delta} := \mathcal{D}_1(Q^d) \bigcap G_N^c$$

we have

$$\widetilde{g}_N = g_N \cdot 1_S + \left(\sum_{Q \in \widetilde{\Delta}} m_Q \cdot 1_Q \right) \cdot 1_{Q^d \setminus S}.$$

This, in turn, implies

$$\|f - \widetilde{g}_N\|_q \leqslant \|f - g_N\|_q + \left\| \sum_{Q \in \widetilde{\Delta}} (f - m_Q) \cdot 1_Q \right\|_q . \tag{4.51}$$

The first summand is clearly bounded by the right–hand side of (4.50).

Moreover, the second one equals

$$\left(\sum_{Q\in\widetilde{\Delta}} E_k(f;Q;L_q)^q\right)^{1/q} \leqslant \left\{\sum_{Q\in\widetilde{\Delta}} E_k(f;Q;L_q)^p\right\}^{1/p}$$

$$\leqslant \left\{\sum_{Q\in\widetilde{\Delta}} \operatorname{var}_p^k(f;Q;L_q)^p\right\}^{1/p} =: \left\{\sum_{Q\in\widetilde{\Delta}} W(Q)\right\}^{1/p}.$$

Since $\widetilde{\Delta} \subset G_N^c$, every $W(Q) < N^{-1}$.

Hence, the second summand in (4.51) is bounded by $(2^d N^{-1})^{1/p}$ that clearly also majorates by the right–hand side of (4.50).

This proves (4.50) and Theorem 2.1(b). $\square$

§ 5. Proofs of Corollaries

5.1. *Proof of Theorem 1.2.* We obtain this result from Theorem 2.1 with $s(V_{pq}^k) = k$ and $q < \infty$. It asserts in this case that under the assumptions

$$d \geqslant 2, \quad 1 \leqslant p < q < \infty \text{ and } \frac{k}{d} = \frac{1}{p} - \frac{1}{q} \tag{5.1}$$

there exist a cover $\Delta_N \subset \mathcal{D}$ of Q^d of at most N cubes and $g_N \in \mathcal{P}_{k-1}(\Delta_N)$ such that

$$\|f - g_N\|_q \leqslant c(k,d) N^{-k/d} |f|_{V_{pq}^k}. \tag{5.2}$$

It remains to replace here $|f|_{V_{pq}^k}$ by the Sobolev seminorm $|f|_{W_p^k(Q^d)}$ if $p > 1$ and by the $BV^k(Q^d)$ seminorm if $p = 1$. This substitution is justified by the two–sided inequality

$$|f|_{V_{pq}^k} \approx \begin{cases} |f|_{W_p^k} & \text{if } \ p > 1, \\ |f|_{BV^k} & \text{if } \ p = 1, \end{cases} \tag{5.3}$$

where constants are independent of f, see Theorems 4 and 12 from [B-71, § 4].

The proof is complete. $\square$

5.2. *Proof of Theorem 2.3.* We should prove the analog of the previous result for the homogeneous Besov $\dot{B}_p^\lambda(Q)$, $\lambda > 0$, $Q \subset \mathbb{R}^d$, whose associated seminorm is given by

$$|f|_{B_p^\lambda(Q)} := \left\{ \int_0^{|Q|^{1/d}} \left(\frac{\omega_k(f;t;L_p(Q))}{t^\lambda}\right)^\lambda \frac{dt}{t} \right\}^{1/p} \tag{5.4}$$

where $k = k(\lambda) := \min\{n \in \mathbb{N}; n > \lambda\}$.

We derive this from Theorem 2.1 with $s(V_{pq}^k) = \lambda$, $k = k(\lambda)$ and $q < \infty$. Hence, in this case,

$$1 \leqslant p < q < \infty, \quad \frac{\lambda}{d} = \frac{1}{p} - \frac{1}{q}, \tag{5.5}$$

and Theorem 2.1 gives under these assumptions the inequality

$$\|f - g_N\|_q \leqslant c(k,d) N^{-\lambda/d} |f|_{V^k_{pq}} \tag{5.6}$$

with the corresponding $g_N \in \mathcal{P}_{k(\lambda)-1}(\Delta_N)$ and Δ_N.

It remains to replace here $|f|_{V^k_{pq}}$ by $|f|_{B^\lambda_p(Q^d)}$.

To this end we use the classical embedding theorem that under the assumptions (5.5) gives the inequality

$$E_k(f; Q; L_q) \leqslant c(d, \lambda, q) |f|_{B^\lambda_p(Q)}, \tag{5.7}$$

see Remark 5.1 below for details.

Now let $\Delta := \{Q\}$ be a disjoint family of cubes from Q^d. Then (5.7) implies

$$\left(\sum_{Q \in \Delta} E_k(f; Q; L_q)^p \right)^{1/p} \leqslant c(d, \lambda, q) \left(\sum_{Q \in \Delta} (|f|_{B^\lambda_p(Q)})^p \right)^{1/p}.$$

Due to Lemma 2 from [B-94, § 5] the sum in the right–hand side is bounded by $c(k,d)|f|_{B^\lambda_p(Q^d))}$. Taking supremum over Δ we then obtain the required inequality

$$|f|_{V^k_{pq}} \leqslant c(k, \lambda, q) |f|_{B^\lambda_p(Q^d)} \tag{5.8}$$

and prove Theorem 2.3. □

5.3. *Proof of Theorem 2.4.* Now we deal with the homogeneous space $\dot{B}^{\lambda 1}_p(Q)$ whose associated seminorm is given by

$$|f|_{B^{\lambda 1}_p(Q)} := \int_0^{|Q|^{1/d}} \frac{\omega_k(f; t; L_p(Q))}{t^{\lambda+1}} dt \tag{5.9}$$

where $k = k(\lambda)$.

We prove, under the conditions

$$1 \leqslant p < q = \infty, \quad d \geqslant 2 \text{ and } \frac{\lambda}{d} = \frac{1}{p}, \tag{5.10}$$

existence of the corresponding Δ_N and $g_N \in \mathcal{P}_k(\Delta_N)$ such that the next inequality is true:

$$\|f - g_N\|_\infty \leqslant c(d, \lambda, p) N^{-\lambda/d} |f|_{B^{\lambda 1}_p(Q^d)}; \tag{5.11}$$

here $k = k(\lambda)$.

Due to (5.10) $\lambda = \frac{d}{p} \leqslant d$ and therefore $k(\lambda) \leqslant d + 1$. Since norms $\|f\|_{B^{\lambda 1}_p(Q)} := \|f\|_{L_p(Q)} + |f|_{B^{\lambda 1}_p(Q)}$ with different $k \geqslant k(\lambda)$ are equivalent, it suffices to prove (5.11) for $k := d + 1$ instead of $k(\lambda)$.

We derive (5.11) from Theorem 2.1(a) with $s(V^k_{pq}) = \lambda$ and $q = \infty$. This requires the embedding

$$\dot{B}^{\lambda 1}_p(Q^d) \subset V^k_{p\infty}(Q^d) \bigcap C(\mathbb{R}^d), \tag{5.12}$$

because Theorem 2.1 with $q = \infty$ holds only for $f \in V^k_{p\infty} \bigcap C(\mathbb{R}^d)$. But $C(\mathbb{R}^d)$ in (5.12) can be removed as condition (5.10) implies that $B^{\lambda 1}_p(Q^d) \subset C(\mathbb{R}^d)|_{Q^d}$, see, e.g., [BL, Thm. 6.8.9(a)].

By a reason explained later we begin with the case

$$\dot{B}^{\lambda 1}_p(\mathbb{R}^d) \subset V^k_{p\infty}(\mathbb{R}^d), \quad k = d+1, \ \lambda = \frac{d}{p}. \tag{5.13}$$

This will be proved for $p = 1$ and ∞ while the general case will be then derived from those by the method of real interpolation.

If $p = 1$, then (5.10) implies $\lambda = d$ and $k(\lambda) = d+1$; moreover, by definition $\dot{B}^{\lambda 1}_1 = \dot{B}^{\lambda}_1$. In this case (5.7) is still true, i.e., we have

$$E_{d+1}(f; Q; L_\infty) \leqslant c(k,d)|f|_{B^d_1(Q)}, \tag{5.14}$$

see Remark 5.1 below.

Using the argument used in the proof of (5.8) we obtain from (5.14) the required inequality

$$|f|_{V^{d+1}_{1\infty}(\mathbb{R}^d)} \leqslant c(d)|f|_{B^d_1(\mathbb{R}^d)}. \tag{5.15}$$

This proves (5.13) for $p = 1$.

Now let $p = \infty$, hence, $\lambda = \frac{d}{p} = 0$. The arising space $\dot{B}^{01}_\infty(\mathbb{R}^d)$ is defined by the seminorm

$$|f|_{B^{01}_\infty(\mathbb{R}^d)} := \sum_{j\in\mathbb{Z}} \|f * \varphi_j\|_{L_\infty(\mathbb{R}^d)}$$

where $\{\varphi_j\}$ is a sequence of test functions, satisfying, in particular, the condition

$$f = \sum_j f * \varphi_j$$

with convergence in the distributional sense, see, e.g., [BL, sec. 6.3].

This implies

$$|f|_{V^{d+1}_{\infty\infty}(\mathbb{R}^d)} := \sup_{Q\subset\mathbb{R}^d} E_d(f; Q; L_\infty) \leqslant \|f\|_{L_\infty(\mathbb{R}^d)} \leqslant \sum_j \|f * \varphi_j\|_{L_\infty(\mathbb{R}^d)} = |f|_{B^{01}_\infty(\mathbb{R}^d)}.$$

Hence, we prove (5.13) for $p = \infty$ as well.

Interpolating the embeddings obtained we then have

$$(\dot{B}^{01}_\infty, \dot{B}^{d1}_1)_{\theta p} \subset (V^{d+1}_{\infty\infty}, V^{d+1}_{1\infty})_{\theta p}; \tag{5.16}$$

hereafter $\mathbb{R}^d$ is omitted for brevity.

Taking $\theta := \frac{\lambda}{d} = \frac{1}{p}$ we obtain for the left–hand side the embedding

$$\dot{B}^{\lambda 1}_p \subset (\dot{B}^{01}_\infty, \dot{B}^{d1}_1)_{\theta p} \tag{5.17}$$

see [P, Ch. 5, Thm. 6(9)].

Now we show that the right–hand side is contained in $V^{d+1}_{p\infty}(\mathbb{R}^d)$ with $p := \frac{d}{\lambda}$.

Let $\mathcal{E} : L_\infty(\mathbb{R}^d) \to l_\infty(\Delta)$ be a map given by

$$\mathcal{E} : f \mapsto (E_d(f; Q; L_\infty))_{Q\in\Delta};$$

here Δ is a disjoint family of cubes $Q \subset \mathbb{R}^d$.

By definition

$$\|\mathcal{E}(f)\|_{l_p(\Delta)} := \left(\sum_{Q\in\Delta} E_d(f; Q; L_\infty)^p\right)^{1/p} \leqslant |f|_{V^{d+1}_{p\infty}(\mathbb{R}^d)},$$

i.e., $\mathcal{E}$ maps $V^{d+1}_{p\infty}(\mathbb{R}^d)$ into $l_p(\Delta)$ and $\|\mathcal{E}\| \leqslant 1$, $1 \leqslant p \leqslant \infty$.

Interpolating this sublinear operator[5] by the real method we obtain

$$\|\mathcal{E}(f)\|_{(l_\infty(\Delta), l_1(\Delta))_{\theta p}} \leqslant |f|_{(V^{d+1}_{\infty\infty}, V^{d+1}_{1\infty})_{\theta p}},$$

see, e.g., [BK, 4.1.5(c)] for validity of the interpolation result for sublinear operators.

Moreover, $(l_\infty(\Delta), l_1(\Delta))_{\theta p}$ with $\theta = \frac{1}{p}$ equals $l_p(\Delta)$, see, e.g., [BL, Thm. 5.6.1]. Together with the previous this implies

$$\left(\sum_{Q\in\Delta} E_d(f; Q; L_\infty)^p\right)^{1/p} \leqslant |f|_{(V^{d+1}_{\infty\infty}, V^{d+1}_{1\infty})_{\theta p}}$$

where $\theta = \frac{1}{p} = \frac{\lambda}{d}$.

Taking here supremum over all Δ we obtain the embedding

$$(V^{d+1}_{\infty\infty}, V^{d+1}_{\infty 1})_{\theta p} \subset V^{d+1}_{p\infty},$$

implying the required embedding (5.13).

To derive from (5.13) the similar embedding for Q^d we use a bounded linear extension operator

$$Ext : \dot{B}^{\lambda 1}_p(Q^d) \to \dot{B}^{\lambda 1}_p(\mathbb{R}^d),$$

with $\|Ext\| \leqslant c(\lambda, d)$, see, e.g., [BB, Thm. 2.7.2], and the restriction operator

$$Res : V^{d+1}_{p\infty}(\mathbb{R}^d) \to V^{d+1}_{p\infty}(Q^d).$$

Denoting the embedding operator in (5.13) by U and composing it with the now introduced ones we obtain the operator $U_{Q^d} := Ext \circ U \circ Res$ that embeds $\dot{B}^{\lambda 1}_p(Q^d)$ into $V^{d+1}_{p\infty}(Q^d)$ with the embedding constant $\|Ext\| \leqslant c(d, \lambda)$.

This proves the required inequality (5.12) and, therefore, Theorem 2.4. □

Remark 5.1. We prove inequalities (5.7) and (5.14).

Let $f \in L_p(Q)$, $1 \leqslant p < q \leqslant \infty$, and $m_Q \in \mathcal{P}_{k-1}$ be the best approximation of f in $L_p(Q)$. Setting for brevity

$$\omega(t) := \omega_k(f; t; L_p(Q)), \quad t > 0,$$

[5]i.e., $\mathcal{E}(f+g) \leqslant \mathcal{E}(f) + \mathcal{E}(g)$ and $\mathcal{E}(\lambda f) = |\lambda|\mathcal{E}(f)$, $\lambda \in \mathbb{R}$

we estimate the nonincreasing rearrangement of $f - m_Q$ as follows

$$(f - m_Q)^*(t) \leqslant c(k,d) \int_{t/2}^{|Q|} \frac{\omega(u^{1/d})}{u^{1+1/p}} du, \quad t \leqslant |Q|, \tag{5.18}$$

see [B-94, Appendix II, Cor. 2'].

Taking L_q–norm and applying the Hardy inequality we have

$$\begin{aligned} \|f - m_Q\|_{L_q(Q)} &= \|(f - m_Q)^*\|_{L_q(0,|Q|)} \leqslant \\ &\leqslant c(k,d)\|\mathcal{H}_{1/q}\| \left(\int_0^{|Q|} \left(\frac{\omega(u^{1/d})}{u^{1/p-1/q}} \right)^q du \right)^{1/q} \end{aligned} \tag{5.19}$$

where $\mathcal{H}_\mu$, $\mu > 0$, is the Hardy operator given by

$$\mathcal{H}_\mu g(t) := t^\mu \int_t^{|Q|} \frac{g(u)}{u^\mu} \frac{du}{u}.$$

Since $\|\mathcal{H}_\mu\| < \infty$ for $\mu > 0$, inequality (5.19) is true for $1/q > 0$, i.e., for $q < \infty$.

Since $\frac{1}{p} - \frac{1}{q} = \frac{\lambda}{d}$, the integral in (5.19) is bounded by

$$d^{1/q} \left(\int_0^{|Q|^{1/d}} \left(\frac{\omega(t)}{t^\lambda} \right)^q \frac{dt}{t} \right)^{1/q} \leqslant c(k,\lambda) d^{1/q} \left(\int_0^{|Q|^{1/d}} \left(\frac{\omega(t)}{t} \right)^p \frac{dt}{t} \right)^{1/p}$$
$$= c(k,d,\lambda)|f|_{B_p^\lambda(Q)}.$$

Hence, for $q < \infty$

$$\|f - m_Q\|_{L_q(Q)} \leqslant c(k,d,\lambda)|f|_{B_p^\lambda(Q)}$$

which implies (5.7) as the left–hand side is clearly bigger than $E_k(f;Q;L_q)$.

For $q = \infty$ we pass in (5.18) to the limit as $t \to 0^+$ to obtain

$$\|f - m_Q\|_{L_\infty(Q)} = \lim_{t\to 0}(f - m_Q)^*(t) \leqslant c(k,d) \int_0^{|Q|} \frac{\omega(u^{1/d})}{u^{1/p}} \frac{du}{u} = d \cdot c(k,d)|f|_{B_p^{\lambda 1}(Q)}.$$

Hence, (5.14) follows. □

Appendix I Covering Dyadic Rings

First, let Q be a dyadic subcube of Q^* such that

$$dist(Q, \mathbb{R}^d \setminus Q^*) > 0. \tag{I.1}$$

The general result will be reduced to this case, see Remark I.1.

Theorem I.1. *Let (I.1) hold. There exists a cover $\mathcal{K}$ of $Q^* \setminus Q$ by cubes*[6] *such that for every overlapping*[7] *pair $\{K_1, K_2\} \subset \mathcal{K}$*

$$|K_1 \bigcap K_2| \geqslant \frac{1}{2} \min_{i=1,2} |K_i| \tag{I.2}$$

and, moreover,

$$\operatorname{card} \mathcal{K} = 4(2^d - 1). \tag{I.3}$$

Proof. Without loss of generality we assume that $Q^* = Q^d := [0,1)^d$. By (I.1) the dyadic cube Q is contained in one of sons of Q^d, say, in $[1/2e, e) := \prod_{i=1}^{d} [1/2, 1)$, $e := (1, 1, \ldots, 1)$. Denoting $Q := \prod_{i=1}^{d} [a_i, b_i)$ we, in particular, have

$$0 < 1 - a_i \leqslant 1/2, \quad 1 \leqslant i \leqslant d. \tag{I.4}$$

Now let π denote a partition of Q^d by hyperplanes passing through the vertex $a \in Q$ and parallel to the coordinate hyperplanes. It consists of 2^d parallelotopes every of which consists a single vertex $\varepsilon \in \{0,1\}^d$ of Q^d. We enumerate elements of π by these vertices, so that $\pi := \{\Pi_\varepsilon\}_{\varepsilon \in \{0,1\}^d}$ and ε is contained in the closure of $Q^d \bigcap \Pi_\varepsilon$. Then Π_ε and $\Pi_{\varepsilon'}$ have a (unique) common face whenever ε, ε' differ by a single coordinate. Moreover, the edge $[\varepsilon, \varepsilon')$ of Q^d is orthogonal to this face and intersects Π_ε and $\Pi_{\varepsilon'}$.

Let $G(\pi)$ denote a graph with the vertex set $\pi = \{\Pi_\varepsilon\}$ and the edges consisting of pairs $\{\Pi_\varepsilon, \Pi_{\varepsilon'}\}$ with a common face. The bijection $\varphi : \Pi_\varepsilon \longleftrightarrow \varepsilon$ is an isomorphism of $G(\pi)$ onto the *hypercube graph* Γ_d whose vertices and edges are those of the cube Q^d.

In fact, $\varepsilon = \varphi(\Pi_\varepsilon)$ and $\varepsilon' = \varphi(\Pi_{\varepsilon'})$ are joined by an edge in Γ_d whenever ε differs from ε' by a single coordinate, i.e., whenever Π_ε and $\Pi_{\varepsilon'}$ have a common face and therefore are joined by an edge in $G(\pi)$.

Further, the graph Γ_d has a *Hamiltonian cycle*, i.e., a cycle that visits each vertex of Γ_d exactly once, see, e.g., [HHW]. Therefore, $G(\pi)$ also has such a cycle denoted by $\mathcal{C}(\pi)$.

Now we apply this construction to the parallelotope $\Pi_e := \prod_{i=1}^{d} [a_i, 1)$ containing $Q = [a, b) := \prod_{i=1}^{d} [a_i, b_i)$ and the vertex b substituting for that of a. This gives a partition $\widehat{\pi}$ of Π_e into 2^d parallelotopes one of which is Q. Then we enumerate them by the vertex set V of Π_e such that $\widehat{\pi} = \{\Pi_v\}_{v \in V}$ and v belong to the closure of $\Pi_v \bigcap \Pi_e$, e.g., $\Pi_a = Q$.

Using the partition $\widehat{\pi}$ we, as above, define the graph $G(\widehat{\pi})$ isomorphic to Γ_d and denote by $\mathcal{C}(\widehat{\pi})$ the corresponding Hamiltonian cycle. Hence, Π_v, $\Pi_{v'}$ are neighbours in $\mathcal{C}(\widehat{\pi})$ if they have a common face orthogonal to $[v, v']$.

[6] let us recall, that all cubes have a form $\prod_{i=1}^{d} [a_i, b_i)$

[7] i.e., $|K_1 \bigcap K_2| > 0$.

Now we define a new graph G with the vertex set

$$V(G) := (\pi \setminus \{\Pi_e\}) \bigcup (\widehat{\pi} \setminus \{\Pi_a\})$$

where $\Pi_a = Q$, and with the edge set $E(G)$ of two parts.

The first consists of edges from $G(\pi)$ and $G(\widehat{\pi})$ such that both of their endpoints belong to either $\pi \setminus \{\Pi_e\}$ or $\widehat{\pi} \setminus \{\Pi_a\}$.

The second part is as follows.

Let Π_ε, $\Pi_{\varepsilon'}$ from $\mathcal{C}(\pi)$ have common faces with $\Pi_e (\in \mathcal{C}(\pi))$. Since $\Pi_e := [a, e)$, the vertex $a \in Q$ belongs to Π_ε and to $\Pi_{\varepsilon'}$. Therefore there exist parallelotopes Π_v and $\Pi_{v'}$ from $\widehat{\pi}$ each having one of faces common with that of Q and another containing in Π_ε and $\Pi_{\varepsilon'}$, respectively.

Then the pairs $\{\Pi_\varepsilon, \Pi_v\}$, $\{\Pi_{\varepsilon'}, \Pi_{v'}\}$ from $V(G)$ form the remaining part of edges from $E(G)$.

It is now the matter of definition to check that

$$\mathcal{C} = \mathcal{C}_1 \bigcup \mathcal{C}_2 := (\mathcal{C}(\pi) \setminus \{\Pi_e\}) \bigcup (\mathcal{C}(\widehat{\pi}) \setminus \{Q\})$$

is a Hamiltonian cycle in G.

Now we construct the desired cover $\mathcal{K}$ of $Q^d \setminus Q$ beginning first with extension of each parallelotope of $\mathcal{C}_i$, $i = 1, 2$, to a cube contained in $Q^d \setminus Q$.

We begin with the set

$$\mathcal{C}_1 := \{\Pi_\varepsilon; \varepsilon \in \{0,1\}^d \setminus \{e\}\}$$

containing $2^d - 1$ elements.

Let $\Pi_\varepsilon := \prod_{i=1}^{d} [a_i^\varepsilon, b_i^\varepsilon)$ and l^ε be the maximal edgelength of Π_ε. Since by the definition of Π_ε every edge $[a_i^\varepsilon, b_i^\varepsilon)$ equals either $A_i := [0, a_i)$ or $B_i := [a_i, 1)$ and $|A_i| \geqslant |B_i|$, see (I.4), the maximal edge of Π_ε, say, $[a_{i_0}^\varepsilon, b_{i_0}^\varepsilon)$, has the form

$$[a_{i_0}^\varepsilon, b_{i_0}^\varepsilon) = A_{i_0} = [0, a_{i_0}). \tag{I.5}$$

Now we extend Π_ε to a cube replacing every edge $[a_i^\varepsilon, b_i^\varepsilon) = A_i$ by $[\widehat{a}_i^\varepsilon, \widehat{b}_i^\varepsilon) := [0, a_{i_0})$ and every edge equal to B_i by $[\widehat{a}_i^\varepsilon, \widehat{b}_i^\varepsilon) := [1 - a_{i_0}, 1)$.

In this way, we obtain the cube

$$Q_\varepsilon := \prod_{i=1}^{d} [\widehat{a}_i^\varepsilon, \widehat{b}_i^\varepsilon) \subset Q^d$$

of edgelength a_{i_0} that contains Π_ε and, moreover, is contained in $Q^d \setminus \Pi_e$.

In fact, the projections of Q_ε and Π_e on the x_{i_0} – axis are $[\widehat{a}_i^\varepsilon, \widehat{b}_i^\varepsilon) = [0, a_{i_0})$, see (I.5), and $[a_{i_0}, 1)$, respectively, that do not intersect.

Thus, we have

$$\bigcup \mathcal{C}_1 = \bigcup_{\varepsilon \neq e} Q_\varepsilon, \quad Q_\varepsilon \supset \Pi_\varepsilon, \quad \varepsilon \neq e,$$

where $\bigcup \mathcal{C}_1 := \bigcup \{\Pi; \Pi \in \mathcal{C}_1\}$.

Further, we cover $\bigcup \mathcal{C}_2$ similarly. By definition

$$\mathcal{C}_2 = \left\{ \Pi_v := \prod_{i=1}^{d} [a_i^v, b_i^v);\ v \in V \setminus \{a\} \right\}$$

where $[a_i^v, b_i^v)$ equals either $A_i := [b_i, 1)$ or $B_i := [a_i, b_i)$.

Let us show that $|A_i| \geqslant |B_i|$. In fact, Q is a dyadic cube, say, $Q := 2^{-n}(\alpha + Q^d)$, $\alpha \in \mathbb{Z}_+^d$, and therefore $|B_i| = 2^{-n}$ while $|A_i| = 1 - b_i = 2^{-n}(2^n - \alpha_i - 1) \geqslant 2^{-n}$ as $b_i < 1$.

Then the maximal edge of Π_v, say, $[a_{i_0}^v, b_{i_0}^v)$ has the form

$$[a_{i_0}^v, b_{i_0}^v) = A_{i_0} = [b_{i_0}, 1). \tag{I.6}$$

Now we extend Π_v replacing every $[a_i^v, b_i^v) = A_i$ by $[\widehat{a}_i^v, \widehat{b}_i^v) := [1 - l^v, 1)$ and every $[a_i^v, b_i^v) = B_i$ by $[b_i, b_i - l^v)$; here $l^v = 1 - b_{i_0}$ is the maximal edgelength of Π_v.

In this way, we obtain the cube

$$Q_v := \prod_{i=1}^{d} [\widehat{a}_i^v, \widehat{b}_i^v) \subset Q^d$$

of volume $(l^v)^d$ that contains Π_v and, moreover, is contained in $Q^d \setminus Q$.

In fact, the embedding $\Pi_v \subset Q_v$ follows from the inequality $a_i \geqslant \widehat{a}_i^v := b_i - l^v$ equivalent to

$$|B_i| = b_i - a_i \leqslant |A_i| \leqslant l^v.$$

Further, $Q_v \bigcap Q = \emptyset$, as the projections on the x_{i_0} – axis of these cubes $[a_{i_0}^v, b_{i_0}^v) = [b_{i_0}, 1)$ and $[a_{i_0}, b_{i_0})$, respectively, do not intersect.

Thus, we have

$$\bigcup \mathcal{C}_2 \subset \prod_{v \neq a} Q_v \subset Q \setminus Q^d \ \text{ and } \ \Pi_v \subset Q_v.$$

This gives the family $\mathcal{F} := \{Q_\varepsilon\} \bigcup \{Q_v\}$ of $2(2^d - 1)$ cubes that cover the d-ring $Q^d \setminus Q$ such that Q_ε, Q_v are uniquely defined by the corresponding $\Pi_\varepsilon \supset Q_\varepsilon$, $\Pi_v \supset Q_v$ from the Hamiltonian cycle $\mathcal{C}$.

Further, we enumerate the cycle $\mathcal{C}$ by integers to obtain

$$\mathcal{C} = \{\Pi_i; 1 \leqslant i \leqslant 2 \cdot 2^d - 1\}$$

where $\Pi_i := \Pi_1$ for $i = 2 \cdot 2^d - 1$, such that Π_i, Π_{i+1} are neighbours in $\mathcal{C}$. Hence, they adjoint to some edge of Q^d denoted by $[v_i, v_{i+1})$ such that a small shift along this edge of the smaller parallelotope remains in $\Pi_i \bigcup \Pi_{i+1} \subset Q^d \setminus Q$.

Now let $\{Q_i; 1 \leqslant i \leqslant 2 \cdot 2^d - 1\}$ where $Q_i := Q_1$ for $i = 2 \cdot 2^d - 1$ and the numeration of the family $\mathcal{F}$ is induced by that of $\mathcal{C}$.

Then by the definition of cubes from $\mathcal{F}$ the following is true.

(a) $\bigcup\limits_i Q_i$ covers $Q^d \setminus Q$;

(b) cubes $Q_i \supset \Pi_i$, $Q_{i+1} \supset \Pi_{i+1}$ adjoint to the edge $[v_i, v_{i+1})$ and the shift along this edge of the smaller one, say Q_i, by its length remains in $Q_{i+1} \subset Q^d \setminus Q$.

Let then $Q_{i+1/2}$ denote the image of Q_i under such a shift by the one–half of its length. Then the cover $\mathcal{K} := \{Q_i, Q_{i+1/2}\}$ of $Q^d \setminus Q$ consists of $4(2^d - 1)$ cubes satisfying the inequality

$$|Q_j \bigcap Q_{i+1/2}| \geqslant 1/2 \min\{|Q_j|, |Q_{i+1/2}|\}$$

for $j = i, i+1$.

Hence, Theorem I.1 is proved for Q contained in the interior of Q^d, see (I.1). □

Remark I.1. Now let (I.1) do not hold. Then some face of Q is contained in a face of Q^d with the vertex e. To reduce this case to the previous one we introduce the pair Q, $Q^* := [0, 2)^d$. Clearly, Q is a dyadic subcube of Q^* and (I.1) holds for this pair.

Applying to $\{Q, Q^*\}$ the first step of the previous procedure we obtain the families of parallelotopes $\{\Pi_\varepsilon\}_{\{0,2\}^d \setminus \{2e\}}$ and $\{\Pi_v\}_{V \setminus \{a\}}$ that cover $Q^* \setminus \Pi_{2e}$ and $\Pi_{2e} \setminus Q$, respectively (here V is the vertex set of Π_{2e}).

Then we use the second step of the procedure extending all of the parallelotopes $\Pi_\varepsilon \bigcap Q^d$, $\varepsilon \neq 2e$, $\Pi_v \bigcap Q^d$, $v \neq a$, to the cubes. In some cases, the extension is fictional as the corresponding intersections are empty. E.g., if the vertex b of Q coincides with e, then $\Pi_v \bigcap Q^d = \emptyset$ if $v \neq a$.

At the third step we as above shift the obtained cubes forming the pairs $\{Q_i, Q_{i+1/2}\}$, $1 \leqslant i \leqslant 2(2^d - 1)$; here $Q_i = Q_1$ for $i = 2 \cdot 2^d - 1$ and the numeration of the family of cubes $\{Q_i\}$ is induced by the natural numeration of the Hamiltonian cycle $\mathcal{C}$ generated by the pair $\{Q, Q^*\}$.

Discarding empty pairs $\{Q_i, Q_{i+1/2}\}$ we obtain the required cover $\mathcal{K}$ of $Q^d \setminus Q$ satisfying (I.2). In this case, $2(2^d - 1) \leqslant \operatorname{card} \mathcal{K} < 4(2^d - 1)$.

Hence, Theorem I.1 is true also in this case.

Appendix II Approximation Algorithm

We describe the algorithm giving as output the cover Δ_N in Theorem 2.1. In what follows, we freely use terms and definitions of section 4.2, e.g., weight, dyadic tree $\mathcal{D} := \mathcal{D}(Q^d)$, paths etc. Proofs of some statements below will be left to the reader (all of them are presented in details in [B-2004, § 6]).

Let $W : A(\mathcal{D}) \to \mathbb{R}_+$ be a subadditive absolutely continuous weight normed by the condition

(II.1) $$W(Q^d) = 1.$$

Then the set

(II.2) $$G_N := \{Q \in \mathcal{D}; W(Q) \geqslant N^{-1}\}, \quad N \in \mathbb{N},$$

is a *finite rooted subtree* of $\mathcal{D}$ with the root Q^d. Hence, every path connecting $Q \in G_N$ and Q^d is unique and belongs to G_N.

Further, let $G_N^{\min}$ be the set of minimal elements of G_N with respect to the set–inclusion order.

Hence, every $Q \in G_N$ contains properly some minimal cube and a son Q' of such a cube satisfies

$$W(Q') < N^{-1}.$$

In particular, $G_N^{\min}$ is disjoint and as every disjoint subset of G_N has at most N elements.

Somehow enumerating $G_N^{\min}$, say,

$$G_{\min}^N := \{Q_i\}_{1 \leqslant i \leqslant m_N}$$

where

$$m_N := \operatorname{card} G_N^{\min} \leqslant N, \tag{II.3}$$

we then denote by L_i a (unique) path in G_N joining Q_i and Q^d.

By the definition of $G_N^{\min}$

$$G_N = \bigcup_{i=1}^{m_N} L_i. \tag{II.4}$$

We divide each L_i into more small paths

$$P_i := L_i \setminus \bigcup_{j=0}^{i-1} L_j, \quad 1 \leqslant i \leqslant m_N$$

where $L_0 := \{Q^d\}$.

Lemma II.1. ([B-2004, p. 164]) *(a) Family $\{P_i\}_{1 \leqslant i \leqslant m_N}$ is a partition of $G_N \setminus \{Q^d\}$.*

(b) Every P_i is of the form

$$P_i := [Q_i, Q_i^c) := [Q_i, Q_i^c] \setminus \{Q_i^c\} \tag{II.5}$$

where Q_i^c is the tail of a path $L_i \bigcap L_j$ with $j < i$.

The set

$$\mathcal{C}_N := \{Q^d\} \bigcup \{Q_i^c\}_{1 \leqslant i \leqslant m_N}$$

contains at most $m_N + 1$ elements called *contact cubes.*

Now we refine G_N subdividing each P_i by contact cubes from $P_i \bigcap \mathcal{C}_N$. In this way, we define a set of subpaths $[Q', Q'')$ where Q' is either a minimal cube or a contact cube, and Q'' is a contact cube.

Denoting the set of these subpaths by $\mathcal{P}_N$ we obtain from (II.3)

$$\operatorname{card} \mathcal{P}_N \leqslant 2m_N + (m_N + 1) = 3m_N + 1. \tag{II.6}$$

Finally, we divide each path $P \in \mathcal{P}_N$ in the required *basic paths.* To this end, we use an auxiliary weight defined on paths $P = [T_P, H_P]$ of $\mathcal{D}$ by

$$\widetilde{W}(P) := W(H_P \setminus T_P). \tag{II.7}$$

Now we define for each $P \in \mathcal{P}_N$ a family of vertices (cubes) $\{Q_i(P) \in P; 1 \leqslant i \leqslant i_p\}$ using induction on i.

We begin with $Q_1(P) := T_P$ and then having $Q_i(P)$ define $Q_{i+1}(P)$ as a vertex in the half–open from the left path

$$(Q_i(P), H_P] := [Q_i(P), H_P] \setminus \{Q_i(P)\}$$

satisfying the conditions

$$\widetilde{W}([Q_i(P), Q_{i+1}(P)]) \geqslant N^{-1},$$

$$\widetilde{W}([Q_i(P), Q_{i+1}(P))) < N^{-1}.$$

Then we define the i-th basic path $B_i(P)$ by setting

$$B_i(P) := [Q_i(P), Q_{i+1}(P)). \tag{II.8}$$

The vertex $Q_{i+1}(P)$ may be undetermined, if

$$\widetilde{W}([Q_i(P), H_p]) < N^{-1}.$$

In this case, we complete induction setting $i_P := i$ and defining $B_i(P)$ to be equal to $[Q_i(P), H_P]$. However, to preserve formula (II.8) for this case, we define $Q_{i+1}(P)$ as the *father* of H_P. Denoting it, say, $H_P^+(\in P)$ we define $B_i(P)$ for this case by (II.8) with $Q_{i+1}(P) := H_P^+$ and $i := i_P$.

Hence, the induction has been completed with $Q_{i+1}(P) = H_P$ or $Q_{i+1}(P) = H_P^+$ for $i = i_P$. In this way, we obtain a partition of P by subpaths $B_i(P) := [Q_i(P), Q_{i+1}(P))$, $1 \leqslant i \leqslant i_P$. Let us single out that if $Q_{i+1} = H_p^+$, then $B_i(P)$ may be a singleton $\{H_p\}$.

By definition these subpaths satisfy

$$\begin{aligned} &\widetilde{W}([Q_i(P), Q_{i+1}(P))) < N^{-1},\\ &\widetilde{W}([Q_i(P_i), Q_{i+1}(P))) \geqslant N^{-1} \end{aligned} \tag{II.9}$$

for $1 \leqslant i \leqslant i_P - \varepsilon_P$ where $\varepsilon_P := 0$ if $Q_{i_P+1}(P) = H_P^+$ and $\varepsilon_P := 1$ otherwise; in the first case, only the first of inequalities (II.9) holds.

Collecting all the basic paths we obtain the refinement of $\mathcal{P}_N$ given by

$$\mathcal{B}_N := \{B_i(P);\ 1 \leqslant i \leqslant i_P,\ P \in \mathcal{P}_N\}. \tag{II.10}$$

The next result (=Proposition 4.2) gives the output of the algorithm.

Proposition II.1. *(a) $\mathcal{B}_N$ is a partition of $G_N \setminus \{Q^d\}$.*

(b) For every $B = [T_B, H_B] \in \mathcal{B}_N$

$$W(H_B \setminus T_B) := \widetilde{W}(B) < N^{-1}. \tag{II.11}$$

(c) The following is true

$$\operatorname{card} \mathcal{B}_N \leqslant 3N + 1. \tag{II.12}$$

Proof. (a) $\mathcal{B}_N$ is a refinement of the partition $\mathcal{P}_N$, hence, it is also a partition.

(b) is given by the first inequality in (II.9) and the definition of $B_i(P)$.

(c) Let $\{P_i\}$ be a strictly monotone sequence of subpaths in a path P, i.e.. the head of P_i is a *proper* subset of the tail of P_{i+1}. Then by the definition of $\widetilde{W}$, see (II.7),

$$\sum_i \widetilde{W}(P_i) \leqslant W(H_P \setminus T_P).$$

Now let $B_i(P) := [Q_i(P), Q_{i+1}(P))$, $1 \leqslant i \leqslant i_P$, be the partition of $P \in \mathcal{P}_N$ into the basic paths. By the second inequality (II.9)

$$(i_P - \varepsilon_P)N^{-1} \leqslant \sum_{i=1}^{i_P-\varepsilon_P} \widetilde{W}([Q_i, Q_{i+1}]). \tag{II.13}$$

Since the sequence $\{[Q_i, Q_{i+1}]\}_{1 \leqslant i \leqslant i_P - \varepsilon_P}$ has multiplicity 2, it can be divided into two strictly monotone subsequences. Hence, the right–hand side of (II.13) is bounded by $2W(H_P - T_P)$. This implies

$$\operatorname{card} \mathcal{B}_N = \sum_{P \in \mathcal{P}_N} i_P \leqslant 2N \sum_{P \in \mathcal{P}_N} W(H_P) + \sum_{P \in \mathcal{P}_N} \varepsilon_P.$$

Since the set $\{H_P\}_{P \in \mathcal{P}_N}$ is disjoint, $\sum\limits_{P \in \mathcal{P}_N} W(H_P) \leqslant W(Q^d) = 1$.

Further, $\varepsilon_P = 1$ if and only if the endpoint of $B_i(P)$ with $i = i_P$ is H_P^+. By the definition of $\mathcal{P}_N$ every head H_P of $P \in \mathcal{P}_N$ is a contact cube. Hence,

$$\sum_{P \in \mathcal{P}_N} \varepsilon_P \leqslant \operatorname{card} \mathcal{C}_N \leqslant m_N + 1 \leqslant N + 1,$$

see (II.3).

Combining this with the previous estimates we finally get

$$\operatorname{card} \mathcal{B}_N \leqslant 2N + N + 1 = 3N + 1. \quad \square$$

References

[BB] *A. Brudnyi and Yu. Brudnyi*, Methods of Geometric Analysis in Extension and Trace Problems. Vol. I. Birkhäuser. 2011.

[BDKPW] *P. Bechler, R. Devore, A. Kamont, G. Petrova and P. Wojtaszczyk*, Greedy wavelet projections are bounded on BV, Trans. AMS, 359 (2007), 619–635.

[BI-87] *Yu. Brudnyi and I. Irodova*, Nonlinear spline–approximation and B-spaces. Proc. Intern. Conf. in Approx. Theory (Kiev, 1987). Nauka. M. 71–75.

[BI-92] *Yu. Brudnyi and J. Irodova*, Nonlinear spline approximation of functions of several variables and B-spaces, Algebra and Analiz, 4 (1992); English Trans. in St. Petersburg Math. J., 4 (1993), 667–694.

[BK] *Yu. Brudnyi and N. Kruglyak*, Interpolation Functions and Interpolation Methods. I. North–Holland. 1991.

[BL] *J. Bergh and J. Löfström*, Interpolation Spaces. An Introduction. Springer–Verlag. 1976.

[BS] *M. Birman and M. Solomyak*, Piecewise polynomial approximation of functions of classes W_p^α, Math. Sb. 73 (1967), 331–355; English transl. in Math. USSR – Sb., 2 (1967), 295–317.

[B-70] *Yu. Brudnyi*, A multidimensional analog of a theorem of Whitney. Mat. Sb. 82 (1970), 169–191; English transl. in Math. USSR – Sb. 11 (1970).

[B-71] *Yu. Brudnyi*, Spaces defined by local approximation, Trudy Moscov. Mat. Obshch. 24 (1971), 69–132; English transl. in Trans. Moscow Math. Soc. 24 (1971).

[B-1994] *Yu. Brudnyi*, Adaptive approximation of functions with singularities, Trudy Moscov. Mat. Obshch., 55 (1994), 149–245; English transl. in Trans. Moscow Math. Soc. (1995), 23–186.

[B-2002] *Yu. Brudnyi*, Nonlinear approximation of refinable functions and piecewise polynomials. Appr. Theory X: Splines, Wavelets and Applications, Vanderbilt Univ. Press, Nashville. 2002.

[B-2004] *Yu. Brudnyi*, Nonlinear N-term approximation by refinable functions. Algebra i Analiz, 16 (2004); English transl. in St. Petersburg Math. J. 16 (2005), 143–179.

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

[DJP] *R. DeVore, B. Jawerth and V. Popov*, Compression of wavelet decompositions, Amer. J. Math. 114 (1992), 737–785.

[DL] *R. DeVore and G.G. Lorentz*, Constructive Approximation. Springer. 1993.

[DPY] *R. DeVore, P. Petrushev and X.M. Yu*, Nonlinear wavelet approximation in the space $C(\mathbb{R}^d)$. Progress in Appr. Theory. Springer–Verlag. 1992. pp. 261–283.

[HHW] *F. Harary, J. Hayers and H.-J. Wu*, A survey of the theory of hypercube graph, Comp. and Math. with Applications, 15 (4), (1988), 277–289.

[N] *S. Nikolskii*, Approximation of Functions of Several variables and Embedding Theorems. Springer. 1975.

[P] *J. Peetre*, New Thoughts on Besov spaces. Duke Univ. Math. Ser. I. 1976.

[W] *P. Wojtaszczyk*, Projections and nonlinear approximation in the space $BV(\mathbb{R}^d)$, Proc. London Math. Soc. 87 (2003). 471–497.

[Z] *W. Ziemer*, Weakly Differentiable Functions. Springer–Verlag. 1989.