Adaptive multivariate piecewise polynomial approximation

Shai Dekel*, Dany Leviatan**

*RealTimeImage, 6 Hamasger St., Or-Yehuda 60408, Israel, shai.dekel@turboimage.com
**School of Mathematical Sciences, Tel-Aviv University, Tel-Aviv 69978, Israel, leviatan@math.tau.ac.il

ABSTRACT

We survey some recent results in the theory of multivariate piecewise polynomial approximation. In the univariate case this method is equivalent to Wavelet approximation, but in the multivariate case this is no longer true, since this form of approximation is more adaptive to the geometry of the singularities of the function to be approximated. The theory possibly has applications in image compression.

Keywords: Multivariate polynomial approximation, Piecewise polynomial approximation, nonlinear approximation.

1. INTRODUCTION

A problem that is motivated by multivariate signal processing, and in particular image compression, is the characterization of the degree of nonlinear approximation

$$s_{n,r}(f)_p := \inf_{S\in\Sigma_n^r} \|f-S\|_{L_p([0,1]^d)}, \tag{1.1}$$

where $f \in L_p\left([0,1]^d\right)$, and $\Sigma_n^r\left(\mathbb{R}^d\right)$ is the collection

$$\sum_{k=1}^{n} \mathbf{1}_{\Delta_k} P_k, \tag{1.2}$$

where $\{\Delta_k\}$ are $d-$simplices (triangles in the bivariate case) with disjoint interiors, such that $\bigcup_{k=1}^{n}\Delta_k = [0,1]^d$, and P_k, $1 \le k \le n$, are polynomials of total degree $r-1$.

In the univariate case there is certain equivalence between the two $n-$term approximation methods, Wavelets and piecewise polynomials. Namely, the so-called approximation spaces associated with the two methods are characterized by the same Besov spaces (see [7] for a survey on nonlinear approximation). The advantage of Wavelet approximation is the simplicity and efficiency with which one can implement it.

When $d \ge 2$, these two methods are no longer equivalent. Wavelet approximation is still characterized by the (linear) Besov spaces, while the approximation spaces associated with piecewise polynomials are known to be nonlinear spaces (see [6]) and their characterization remains an open problem. Furthermore, it is very difficult to find a near-best solution of (1.1). In fact, a discretized version of (1.1) is an NP-Hard problem (see [6] for details).

In this work we survey some recent results in the theory of piecewise polynomials approximation. We begin in Section 2 with an overview of some new Whitney-type estimates for polynomial approximation over multivariate domains. In Section 3 we present the Skinny B-spaces that were introduced by Karaivanov and Petrushev [11] for the purpose of characterizing piecewise polynomial approximation from a fixed triangulation of $\mathbb{R}^d$. In some sense these spaces are a natural generalization of Besov spaces. In Section 4 we present the bivariate $\tilde{B}$-spaces introduced in [6] for the purpose of investigating the way piecewise polynomial approximation can adapt to the geometry of the singularities of a given function.

It is interesting to note that recently R. Shukla, P. L. Dragotti, M. N. Do and M. Vetterli have published an image compression algorithm [16] that utilizes a form of adaptive piecewise polynomial approximation and outperforms JPEG2000 in the low-bit rate range on known test-images.

2. LOCAL POLYNOMIAL APPROXIMATION

Here and throughout the paper we assume that domains $\Omega \subset \mathbb{R}^d$ are bounded with a nonempty interior. Let $W_p^r(\Omega)$, $1 \le p \le \infty$, $r \in \mathbb{N}$, denote the ***Sobolev spaces***, equipped with the semi-norm $|g|_{r,p} := \sum_{|\alpha|=r} \|D^\alpha g\|_{L_p(\Omega)}$. The ***K-functional of order r*** of $f \in L_p(\Omega)$, $1 \le p \le \infty$ is defined by

$$K_r(f,t)_p := K\left(f, t, L_p(\Omega), W_p^r(\Omega)\right) := \inf_{g \in W_p^r(\Omega)} \left\{ \|f-g\|_p + t|g|_{r,p} \right\}. \tag{2.1}$$

We denote

$$K_r(f,\Omega)_p := K_r\left(f, diam(\Omega)^r\right)_p. \tag{2.2}$$

We note that the classical Peetre K-functional (2.1) is unsuitable as a measure of smoothness whenever $p<1$ (see [8]). For $f \in L_p(\Omega)$, $0 < p \le \infty$, $h \in \mathbb{R}^d$ and $r \in \mathbb{N}$ we recall the r th order difference operator $\Delta_h^r(f): \Omega \to \mathbb{R}$

$$\Delta_h^r(f,x) := \Delta_h^r(f,\Omega,x) := \begin{cases} \sum_{k=0}^{r} (-1)^{r+k} \binom{r}{k} f(x+kh), & [x, x+rh] \subset \Omega, \\ 0, & otherwise, \end{cases}$$

where $[x,y]$ denotes the line segment connecting the two points $x, y \in \mathbb{R}^n$. The ***modulus of smoothness of order r*** is defined by

$$w_r(f,t)_p := \sup_{|h| \le t} \left\| \Delta_h^r(f,\Omega,\cdot) \right\|_{L_p(\Omega)}, \quad t>0, \tag{2.3}$$

where for $h \in \mathbb{R}^d$, $|h|$ denotes the norm of h. As above we also denote

$$w_r(f,\Omega)_p := w_r\left(f, diam(\Omega)\right)_p. \tag{2.4}$$

Let $\Pi_{r-1} := \Pi_{r-1}\left(\mathbb{R}^d\right)$ denote the multivariate polynomials of total degree $r-1$ (order r) in d variables. Given $\Omega \subset \mathbb{R}^d$, our initial goal is to estimate the degree of approximation of a function $f \in L_p(\Omega)$, $0 < p \le \infty$

$$E_{r-1}(f,\Omega)_p := \inf_{P \in \Pi_{r-1}} \|f-P\|_{L_p(\Omega)}.$$

By combining several classical results, one can obtain for $1 \le p \le \infty$ and domains with the finite cone property (see definition in [1])

$$E_{r-1}(f,\Omega)_p \approx K_r(f,\Omega)_p \approx w_r(f,\Omega)_p, \tag{2.5}$$

where the constants depend on the shape of the domain Ω.

When $d \ge 2$, the main drawback of (2.5) is that the constants of equivalence may 'blow-up' for example in the case of a sequence of triangles of equivalent diameter that become thinner and thinner. This dependence is too restrictive to be applied when trying to estimate the degree of nonlinear approximation by piecewise polynomials of the type (1.1). In order to apply (2.5), we would need to assume that all of the simplices $\{\Delta_k\}$ in (1.2) have some sort of uniform geometric properties. In the finite-element community (see [2]) one says that the mesh $\{\Delta_k\}$ is required to be 'quasi-uniform'. This limitation is in contradiction to the main idea behind piecewise polynomial approximation which is to adaptively place the simplices $\{\Delta_k\}$ over sub-domains where the function is smooth. These sub-domains may be long and narrow and thus under the constraint of a 'quasi-uniform' mesh it may take many simplices to cover them.

In [5] we proved the following

Proposition 2.1 For all bounded convex domains $\Omega \subset \mathbb{R}^d$ and functions $f \in L_p(\Omega)$, $1 \le p \le \infty$, we have the equivalence

$$E_{r-1}(f,\Omega)_p \approx K_r(f,\Omega)_p \approx w_r(f,\Omega)_p , \tag{2.6}$$

and for $0 < p < 1$ we have the equivalence

$$E_{r-1}(f,\Omega)_p \approx w_r(f,\Omega)_p , \tag{2.7}$$

where the constants of equivalency depend only on d , r and p.

We would like to highlight a simple yet powerful tool that was used to prove Proposition 2.1, namely John's theorem [10]. This classical result implies that for any bounded convex domain $\Omega \subset \mathbb{R}^d$ there exists a nonsingular affine transform A such that

$$B(0,1) \subseteq A^{-1}(\Omega) \subseteq B(0,d) , \tag{2.8}$$

where $B(x,r)$ is the ball of radius r with center at $x \in \mathbb{R}^d$.

Given a bounded convex domain $\Omega \subset \mathbb{R}^d$ and a function $f \in L_p(\Omega)$ we apply the corresponding affine transform A, and approximate the function $\tilde{f} := f(A\cdot)$ on the domain $\tilde{\Omega} := A^{-1}(\Omega)$. For example, if $f \in W_p^r(\Omega)$, $1 \le p \le \infty$, then for $\tilde{f} \in W_p^r(\tilde{\Omega})$ one can obtain by means of the classical Bramble-Hilbert Lemma (see [2, Chapter 4] or [4])

$$E_{r-1}(\tilde{f},\tilde{\Omega})_p \le C(r,d)\left|\tilde{f}\right|_{r,p} ,$$

which implies

$$E_{r-1}(f,\Omega)_p \le C(r,d)\,diam(\Omega)^r \left|f\right|_{r,p} . \tag{2.9}$$

From (2.9) it is easy to obtain the left-hand side equivalence in (2.6) for $f \in L_p(\Omega)$ (see [4] for more details).

Proposition 2.1 allows us to analyze adaptive multivariate piecewise polynomial approximation schemes that employ convex elements, in particular $d-$simplices (triangles in the bivariate case). One can also obtain the following result for arbitrary domains

Proposition 2.2 Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with non-empty interior and $f \in L_p(\Omega)$, $1 \le p < \infty$. Then for any set of interior disjoint convex domains $\{\Omega_n\}_{n=1}^N$, $\Omega_n \subseteq \Omega$, we have

$$\sum_{n=1}^N E_{r-1}(f,\Omega_n)_p^p \le C(r,d) K_r(f,\Omega)_p^p .$$

The proof of Proposition 2.2 for the bivariate case where $\{\Omega_n\}_{n=1}^N$ are triangles can be found in [6]. The general case follows in similar manner using (2.9).

3. MULTIVARIATE SKINNY B-SPACES

Karaivanov and Petrushev [11] introduced bivariate smoothness spaces, the so-called bivariate Skinny B-spaces, for the purpose of characterizing the approximation spaces corresponding to nonlinear $n-$term piecewise polynomials over nested triangulations (see also [12] for a survey on such spaces). As anticipated by Karaivanov and Petrushev, we were able in [5] to generalize their approach to the general case $d \geq 2$.

A set $\mathcal{T}$ of $d-$simplices is called a ***weak locally regular (WLR-) triangulation*** of $\mathbb{R}^d$ with levels $\{\mathcal{T}_m\}_{m\in\mathbb{Z}}$ if $\mathcal{T} = \bigcup_{m\in\mathbb{Z}} \mathcal{T}_m$ satisfies the following conditions:

(i) Every level $\mathcal{T}_m$ is a set of $d-$simplices with disjoint interiors such that

$$\mathbb{R}^d = \bigcup_{\Delta\in\mathcal{T}_m} \Delta .$$

Note that, by definition, simplices are compact and convex.

(ii) The levels $\mathcal{T}_m$ are nested, that is, for every $\Delta \in \mathcal{T}_m$,

$$\Delta = \bigcup_{\substack{\Delta'\in\mathcal{T}_{m+1}\\ \Delta'\subset\Delta}} \Delta' .$$

Therefore, any two simplices in $\mathcal{T}$ either have disjoint interiors or one of them contains the other. We shall call Δ' a ***child*** of $\Delta \in \mathcal{T}_m$ if $\Delta' \in \mathcal{T}_{m+1}$ and $\Delta' \subset \Delta$.

(iii) There exist constants $0 < r_1 < r_2 < 1$ $(r_1 \leq 1/4)$ such that for each $\Delta \in \mathcal{T}$ and any child Δ' of Δ

$$r_1 |\Delta| \leq |\Delta'| \leq r_2 |\Delta| .$$

In particular, the number of children of any $\Delta \in \mathcal{T}$ satisfies

$$1 < \lfloor r_2^{-1} \rfloor \leq \# child(\Delta) \leq \lceil r_1^{-1} \rceil .$$

For a given WLR-triangulation $\mathcal{T}$ of $\mathbb{R}^d$ we define the ***Skinny B-space*** $\mathcal{B}_{q}^{\alpha,r}(\mathcal{T})$, $\alpha > 0$, $0 < t < \infty$, $r \in \mathbb{N}$, as the set of functions $f \in L_t(\mathbb{R}^d)$ for which

$$\|f\|_{\mathcal{B}_t^{\alpha,r}(\mathcal{T})} := \left(\sum_{\Delta\in\mathcal{T}} \left(|\Delta|^{-\alpha} w_r(f,\Delta)_t \right)^t \right)^{1/t}, \tag{3.1}$$

is finite. In some sense these spaces generalize the classical Besov spaces $B_t^{d\alpha,r}(L_t)$ whose norm is equivalent to

$$\|f\|_{B_t^{d\alpha,r}(L_t)} \sim \left(\sum_{Q} \left(|Q|^{-\alpha} w_r(f,Q)_t \right)^t \right)^{1/t}, \tag{3.2}$$

where $\{Q\}$ is the collection of all dyadic cubes in $\mathbb{R}^d$ (for a proof of this equivalence in the bivariate case see [11, Lemma 2.15]. The proof of the general case is similar). Comparing (3.1) with (3.2) we see that if the triangulation $\mathcal{T}$ is adapted to the geometry of the singularities of f, then the Skinny B-norm of f can be smaller than the Besov norm of f, that is governed by partitions of dyadic cubes.

Let $\Sigma_n^r(\mathcal{T})$ be the collection

$$\sum_{k=1}^{n} \mathbf{1}_{\Delta_k} P_k \ ,$$

where $\Delta_k \in \mathcal{T}$ and $P_k \in \Pi_{r-1}(\mathbb{R}^d)$, $1 \le k \le n$, and let

$$s_{n,r}(f,\mathcal{T})_p := \inf_{S \in \Sigma_n^r(\mathcal{T})} \|f - S\|_p ,$$

denote the degree of nonlinear approximation from $\Sigma_n^r(\mathcal{T})$. For a WLR-triangulation $\mathcal{T}$, we denote the ***approximation space*** $A_q^{g,r}(L_p,\mathcal{T})$, $g > 0$, $0 < q \le \infty$, to be the set of functions $f \in L_p(\mathbb{R}^d)$ for which

$$|f|_{A_q^{g,r}(L_p,\mathcal{T})} := \begin{cases} \left(\sum_{m=0}^{\infty} \left(2^{mg} s_{2^m,r}(f,\mathcal{T})_p \right)^q \right)^{1/q}, & 0 < q < \infty, \\ \sup_{m \ge 0} \left(2^{mg} s_{2^m,r}(f,\mathcal{T})_p \right), & q = \infty, \end{cases}$$

is finite. One sees that the space $A_q^{g,r}(L_p,\mathcal{T})$ is the collection of functions for which the degree of $n-$term approximation from $\mathcal{T}$ decays 'like' n^{-g}.

The proofs of the following Jackson and Bernstein estimates were established in [11] for the bivariate case (see [5] for the general case $d \ge 2$).

Proposition 3.1 [Jackson estimate] Let $\mathcal{T}$ be a WLR-triangulation, $0 < p < \infty$, $a > 0$ and $r \in \mathbb{N}$. If $f \in \mathcal{B}_t^{a,r}(\mathcal{T})$, $1/t = a + 1/p$, then

$$s_{n,r}(f,\mathcal{T})_p \le C n^{-a} \|f\|_{\mathcal{B}_t^{a,r}(\mathcal{T})} ,$$

with $C := C(a,d,r,p,r_1,r_2)$.

We note that a crucial ingredient in the proof of the Jackson estimate is the equivalence

$$\|f\|_{\mathcal{B}_t^{a,r}(\mathcal{T})} \sim \left(\sum_{\Delta \in \mathcal{T}} \left(|\Delta|^{-a} E_{r-1}(f,\Delta)_t \right)^t \right)^{1/t} ,$$

that is guaranteed by Proposition 2.1 with constants that do not depend on the minimal angles of the simplices of $\mathcal{T}$.

Proposition 3.2 [Bernstein estimate] Let $\mathcal{T}$ be a WLR-triangulation and let $S \in \Sigma_n^r(\mathcal{T})$. Then for $0 < p < \infty$, $a > 0$ and $1/t = a + 1/p$,

$$\|S\|_{\mathcal{B}_t^{a,r}(\mathcal{T})} \le C n^a \|S\|_p ,$$

with $C := C(a,d,r,p,r_1,r_2)$.

By using the Jackson and Bernstein estimates and applying the real interpolation method (see [7] for a survey of this technique) one is able to characterize the spaces $A_q^{g,r}(L_p,\mathcal{T})$.

For a Skinny B-space, B, we introduce the K-functional corresponding to the pair L_p and B

$$K(f,t) := K(f,t,L_p,B) := \inf_{g \in B} \left\{ \|f - g\|_p + t|g|_B \right\}, \qquad t > 0 .$$

The ***interpolation space*** $(L_p,B)_{l,q}$, $l > 0$, $0 < q \le \infty$, is defined as the set of all $f \in L_p(\mathbb{R}^d)$ such that

$$|f|_{(L_p,B)_{l,q}} := \begin{cases} \left(\sum_{m=0}^{\infty} \left(2^{ml} K\left(f, 2^{-m}\right) \right)^q \right)^{1/q}, & 0 < q < \infty, \\ \sup_{m \ge 0} 2^{ml} K\left(f, 2^{-m}\right), & q = \infty, \end{cases}$$

is finite. The norm in $\left(L_p, B\right)_{l,q}$ is defined by $\|f\|_{(L_p,B)_{l,q}} := \|f\|_p + |f|_{(L_p,B)_{l,q}}$.

We get (see [11] for the case $d = 2$ and [5] for $d \ge 2$)

Proposition 3.3 If $\mathcal{T}$ is a WLR-triangulation, $0 < g < a$, $0 < q \le \infty$, $0 < p < \infty$ and $r \in \mathbb{N}$, then

$$A_q^{g,r}\left(L_p, \mathcal{T}\right) \approx \left(L_p, \mathcal{B}_t^{a,r}\left(\mathcal{T}\right)\right)_{\frac{g}{a},q},$$

where $t = \left(a + 1/p\right)^{-1}$.

4. BIVARIATE $\widetilde{B}-$SPACES

In [6] we proposed that in the multivariate setting, a measure of smoothness that incorporates **measures of smoothness in several dimensions**, is more appropriate. Our efforts to proceed in this direction rely on recent attempts (e.g. [3], [9], [13], [15], [16]) to find compact representations by computing the geometry of the singularities of functions, i.e. segmentation in the case of images. Indeed, one of the goals of [6] is to try and understand on which 'class' of functions these approaches outperform wavelets. Thus, we try to quantify, in an approximation theoretical sense, the amount of 'structure' present in the signal.

Definition 4.1 For $t > 0$ we define $\Lambda(t)$ as the set of partitions Λ of a bounded domain Ω with the following properties:

(i) As in Figure 4.1, the partition Λ is defined by non-intersecting curves $b_j : [0,1] \to \Omega$, $j = 1, \ldots n_E\left(\Lambda\right)$, each of finite length, denoted by $len\left(b_j\right)$. The curves may intersect only at the endpoints and a subset of the curves should compose the boundary of Ω. Observe that we allow the curves to have 'crack-tips' (see for example [12]), that is, an end of a curve possibly does not touch the end of any of the other curves.

(ii) To each curve b_j we associate a parameter $0 < t_j \le 1$ such that $\sum_{j=1}^{n_E(\Lambda)} t_j^{-1} \le t^{-1}$ (in particular this implies that $n_E\left(\Lambda\right) \le t^{-1}$).

(iii) The curves partition Ω into open connected sub-domains Ω_k, $k = 1, \ldots, n_F\left(\Lambda\right)$.

The smoothness of a continuous planar curve $b : [0,1] \to \mathbb{R}^2$ is measured by the $K-$functional

$$K_r\left(b,t\right)_{\infty,1} := K\left(b, t, C[0,1], W^{r-1}\left(BV[0,1]\right)\right) := \inf_{g \in W^{r-1}(BV[0,1])} \left\{ \left\| \left| b - g \right| \right\|_\infty + t \left| g^{(r-1)} \right|_{BV} \right\}, \tag{4.1}$$

where for a planar curve $j : [0,1] \to \mathbb{R}^2$

$$\left| j \right|_{BV} := \sup_{0 = t_0 < \cdots < t_n = 1} \sum_{i=0}^{n-1} \left| j\left(t_{i+1}\right) - j\left(t_i\right) \right|.$$

In some sense (4.1) is the weakest form of curve smoothness that we found to be appropriate for this type of application.

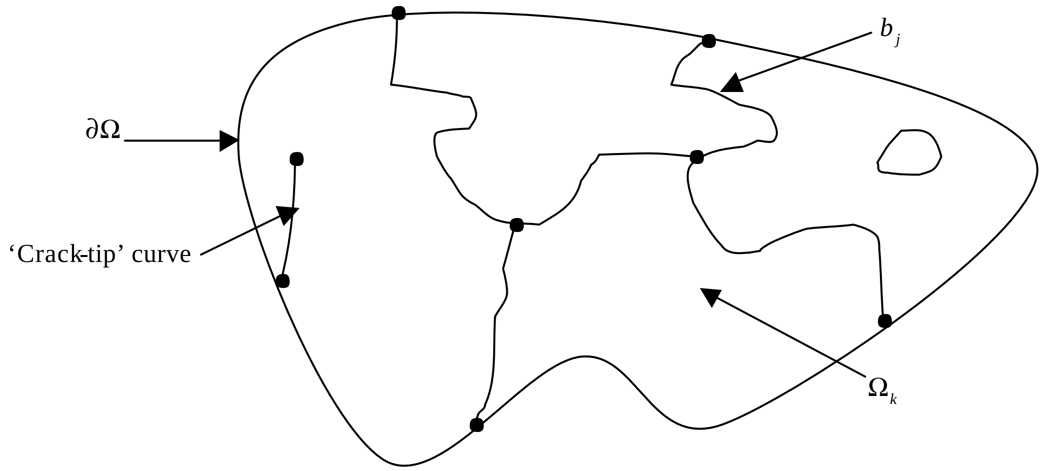

Figure 4.1 A partition Λ of the domain Ω

The following notion of smoothness combines measures of smoothness at several dimensions.

Definition 4.2 For $1 \le p < \infty$, $t > 0$, $r_1, r_2 \in \mathbb{N}$ and $f \in L_p(\Omega)$ we define the $\widetilde{K}$ **- functional**

$$\widetilde{K}_{r_1,r_2}(f,t)_p := \inf_{\Lambda\in\Lambda(t)} \left(\sum_{j=1}^{n_E(\Lambda)} len(b_j) K_{r_1}(b_j, t_j^{r_1})_{\infty,1} + \sum_{k=1}^{n_F(\Lambda)} K_{r_2}(f,\Omega_k)_p^p \right)^{1/p}. \tag{4.2}$$

One can see that the $\widetilde{K}-$functional is defined using sums of 'curve' and 'surface' smoothness terms. The $\widetilde{K}-$functional is highly nonlinear in the following sense

(i) $\widetilde{K}$ is non-decreasing as a function of t, but in general is not continuous.

(ii) $\widetilde{K}$ in general is not sub-linear, that is, there does not exist any constant C such that for any $f, g \in L_p(\Omega)$ we have $\widetilde{K}_{r_1,r_2}(f+g,t)_p \le C\left(\widetilde{K}_{r_1,r_2}(f,t)_p + \widetilde{K}_{r_1,r_2}(g,t)_p\right)$.

In [6] we discuss the relationships between the $\tilde{K}-$functional and the Mumford-Shah functional of [12]. Using (4.2) we define the following bivariate smoothness spaces.

Definition 4.3 We say that a function $f \in L_p(\mathbb{R}^2)$, $1 \le p < \infty$, is in the $\tilde{\boldsymbol{B}}$ **- space** $\tilde{\boldsymbol{B}}_{\boldsymbol{q}}^{\alpha,r_1,r_2}(\boldsymbol{L_p}(\Omega))$ if

$$(f)_{\tilde{B}_q^{\alpha,r_1,r_2}(L_p(\Omega))} := \begin{cases} \sum_{m=0}^{\infty} \left(2^{m\alpha}\, \widetilde{K}_{r_1,r_2}(f,2^{-m})_p\right)^q, & 0<q<\infty, \\ \sup_{m\ge 0} 2^{m\alpha}\, \widetilde{K}_{r_1,r_2}(f,2^{-m})_p, & q=\infty, \end{cases}$$

is finite. Observe that $(\cdot)_{\tilde{B}_q^{\alpha,r_1,r_2}(L_p(\Omega))}$ serves as a measure of smoothness, but it is not a semi-norm because in general the triangle inequality is not fulfilled. This should be compared with the following equivalent form of the Besov semi-norm of the function

$$|f|_{B_q^\alpha(L_p(\Omega))} \sim \begin{cases} \sum_{m=0}^{\infty} \left(2^{m\alpha} K_r(f,2^{-mr})_p\right)^q, & 0<q<\infty, \\ \sup_{m\ge 0} 2^{m\alpha} K_r(f,2^{-mr})_p, & q=\infty, \end{cases}$$

where $r := \lfloor \alpha \rfloor + 1$. Denoting $\tilde{B}_q^{\alpha,r}(L_p(\Omega)) := \tilde{B}_q^{\alpha,2,r}(L_p(\Omega))$, we have the following relationship between the Besov spaces that characterize nonlinear Wavelet approximation and the $\tilde{B}$ spaces.

Proposition 4.4 For $\alpha > 0$, $1 < p < \infty$, $\tau = (\alpha/2 + 1/p)^{-1}$ and $r > \alpha + 1 - 1/p$ one has

$$B_\tau^\alpha\left(L_\tau\left([0,1]^2\right)\right) \subseteq \tilde{B}_q^{\alpha/2,r}\left(L_p\left([0,1]^2\right)\right).$$

The $\tilde{B}$ notion of smoothness is designed to capture singularities along smooth curves as the following example shows.

Example 4.5 If $\tilde{\Omega} \subset \Omega$ and $\partial\tilde{\Omega} \in C^\infty$ then $\mathbf{1}_{\tilde{\Omega}} \in \tilde{B}_q^{\alpha,r_1,r_2}(L_p(\Omega))$, $1 \le p < \infty$, $0 < q \le \infty$, whenever $r_1 \ge \lfloor \alpha \rfloor + 1$. On the other hand we have that $\mathbf{1}_{\tilde{\Omega}} \notin B_q^\alpha(L_p(\Omega))$ if $\alpha > 1/p$.

Let $S_n^{r_1,r_2}(\Omega)$ denote the collection of piecewise polynomials of type $\sum_{k=1}^n \mathbf{1}_{\Omega_k} P_k$, where $\Omega_k \subset \Omega$ are domains with disjoint interiors whose boundary is composed of a fixed number of non-intersecting piecewise polynomial segments of degree $r_1 - 1$ and P_k are bivariate polynomials of degree $r_2 - 1$. In the special case where $r_1 = 2$, the approximation takes the form of piecewise polynomials over polygonal domains. By triangulating these polygonal domains, we may consider $S_n^{2,r}(\Omega)$ to be the collection of functions of type $\sum_{k=1}^n \mathbf{1}_{\Delta_k} P_k$, where Δ_k are triangles with disjoint interiors. The parameters r_1, r_2 allow us to 'tune' the approximation method to the lower or higher dimensional smoothness of the approximated functions. For $f \in L_p(\Omega)$ we define the degree of approximation

$$s_{n,r_1,r_2}(f)_p := \inf_{\tilde{f} \in S_n^{r_1,r_2}} \|f - \tilde{f}\|_{L_p(\Omega)}.$$

Denoting $s_{n,r}(f)_p := s_{n,2,r}(f)_p$, we have the following Jackson-type inequality for the approximation by piecewise polynomials over triangles.

Proposition 4.6 Let Ω be a bounded domain with a piecewise $Lip^*(2)$ boundary and let $f \in L_\infty(\Omega)$. Then for $1 \le p < \infty$ and each $n \ge 1$, we have that

$$s_{n,r}(f)_p \le C_1(p,r)\max\left(\|f\|_{L_\infty(\Omega)}, 1\right)\tilde{K}_{2,r}\left(f, C_2 n^{-1}\right)_p. \tag{4.3}$$

Remarks

1. The estimate (4.3) is good for scenarios where f is 'approximately' a piecewise smooth function with piecewise smooth curve singularities. Indeed, in such a case the 'smoothness quantity' $\tilde{K}_{2,r}(f,t)_p$ is small for sufficiently small values of t.
2. We note that the proof of Proposition 4.6 uses Proposition 2.2.
3. A typical application of (4.3) is in image processing, where images are always in L_∞ and errors are traditionally measured in the L_1 or L_2 norms.

For $\alpha > 0$, $0 < q \le \infty$ we denote the approximation space corresponding to piecewise polynomials of degree $r-1$ over triangles by $A_q^{\alpha,r}(L_p(\Omega), \Delta)$, i.e., the set of functions $f \in L_p(\Omega)$ for which

$$(f)_{A_q^{a,r}(L_p,\Delta)} := \begin{cases} \left(\sum_{m=0}^{\infty} \left(2^{ma} s_{2^m,r}(f)_p \right)^q \right)^{1/q}, & 0<q<\infty, \\ \sup_{m\ge 0} 2^{ma} s_{2^m,r}(f)_p, & q=\infty, \end{cases}$$

is finite. The $\tilde{B}-$spaces can 'almost' characterize nonlinear piecewise polynomial approximation in the following way.

Proposition 4.7 Let Ω be a bounded domain with a piecewise linear boundary. Then, for any $a>0$, $1\le p<\infty$, $0<q\le\infty$ and $r\in\mathbb{N}$, the set $A_q^{a,r}(L_p(\Omega),\Delta)$ is contained in $\tilde{B}_q^{a,r}(L_p(\Omega))$, moreover $(f)_{\tilde{B}} \le C(f)_A$. On the other hand,

$$f\in \tilde{B}_q^{a,r}(L_p(\Omega)) \cap L_\infty(\Omega) \Rightarrow f\in A_q^{a,r}(L_p(\Omega),\Delta).$$

Perhaps our discussion so far quantifies the following 'intuition'. Wavelet approximation cannot well capture singularities along curves while nonlinear piecewise polynomial approximation does. Thus piecewise polynomials over triangles outperform wavelets on images which 'approximately' fit the model of a piecewise smooth function with piecewise smooth edges. On the other hand, if a function is a typical Besov-type function that is smooth only in a weak-sense with oscillations 'randomly' distributed over the time and frequency domains, then we should not expect $n-$term piecewise polynomials approximation to perform any better than the $n-$term wavelet approximation.

REFERENCES

1. R. Adams, *Sobolev Spaces*, Academic Press, 1975.
2. S. C. Brenner and L. R. Scott, *The mathematical theory of finite elements methods*, Springer-Verlag, 1994.
3. A. Cohen and B. Matei, "Compact representation of images by edge adapted multiscale transforms", *preprint*, 2001.
4. S. Dekel and D. Leviatan, "The Bramble-Hilbert lemma for convex domains", *preprint*, 2002.
5. S. Dekel and D. Leviatan, "On local and piecewise multivariate polynomial approximation", *preprint*, 2003.
6. S. Dekel, D. Leviatan and M. Sharir, "On bivariate smoothness spaces associated with nonlinear approximation", *Constructive Approximation*, to appear.
7. R. DeVore, "Nonlinear approximation", *Acta Numerica* 7 (1998), 51-150.
8. Z. Ditzian, V. H. Hristov and G. Ivanov, "Moduli of smoothness and K-functionals in L_p, $0<p<1$", *Constructive Approximation* 11 (1995), 67-83.
9. M. N. Do, P. L. Dragotti, R. Shukla and M. Vetterli, "On the compression of two-dimensional piecewise smooth functions", *Proc. of IEEE International Conference on Image Processing (ICIP)*, Greece, 2001.
10. Fritz John, "Extremum problems with inequalities as subsidiary conditions", in *Studies and Essays Presented to R. Courant on his 60th Birthday*, Interscience Publishers (1948), 187-204.
11. B. Karaivanov and P. Petrushev, "Nonlinear piecewise polynomial approximation beyond Besov spaces", *Industrial Mathematics Institute*, University of South Carolina, technical report 01:13, 2001.
12. D. Mumford and J. Shah, "Optimal Approximations of Piecewise Smooth Functions and Associated Variational Problems", *Comm. in Pure and Appl. Math.* 42 (1989), 577-685.
13. E. Pennec and S. Mallat, "Image compression with geometrical wavelets", *preprint*, 2001.
14. P. Petrushev, "Nonlinear $n-$term approximation from hierarchical spline bases", *Industrial Mathematics Institute*, University of South Carolina, technical report 02:21, 2002.
15. P. Prandoni and M. Vetterli. "Approximation and compression of piecewise-smooth functions", *Phil. Trans. R. Soc. Lond. A.* 357 (1999), 2573-2591.
16. R. Shukla, P. L. Dragotti, M. N. Do and M. Vetterli, "Rate-distortion optimized tree structured compression algorithms for piecewise smooth images", *preprint*, 2003.