Near-degenerate Finite Element and Lacunary Multiresolution Methods of Approximation

Lubomir T. Dechevsky

Abstract. We study the effects of the use of near-degenerate elements in finite and boundary element methods, and their analogues with multiresolution methods. The main results include: an improved bound for the embedding constant in the Bramble–Hilbert lemma; characterization of the best N-term approximation of solutions of nonlinear operator equations; best N-term approximation by near-degenerate normal approximating families; atomic decomposition of Wiener amalgam spaces. In the context of these results, a brief comparison between finite element and wavelet methods is made.

Mathematics Subject Classification: Primary 41A17, 41A25, 41A65, 42C15, 46E35, 47H05, 47H06; Secondary 35A15, 35J65, 41A15, 45G05, 45L05, 45L10, 46M35, 47H17, 65J15

Key Words and Phrases: near-degenerate finite element, lacunary multiresolution method, Bramble–Hilbert lemma, best N-term approximation, nonlinear operator, monotone operator, accretive operator, Wiener amalgam space, atomic decomposition, biorthonormal wavelet

Resumé. Nous étudions les effets de l'utilisation des éléments presque dégénérés dans les méthodes des éléments finis et des éléments à la frontière et leurs analogues chez les méthodes de multiresolution. Les résultats principaux sont: une borne améliorée de la constante dans le lemme de Bramble–Hilbert; caractérisation de l'approximation la meilleure avec N termes des solutions des équations aux opérateurs nonlinéaires; caractérisation de l'approximation la meilleure avec N termes par des familles normales presque dégénérées; décomposition atomique des espaces d'amalgame de Wiener. Dans le contexte de ces résultats, on fait une comparaison brève entre les méthodes des éléments finis et des ondelettes.

Départent de mathématiques et statistique, Université de Montréal
C. P. 6128, Succursale A, Montréal, Québec H3C 3J7 Canada
dechevsk@dms.umontreal.ca

§1. Introduction

One important constraint in the theory of finite and boundary element methods (FEM and BEM) is that the respective partitions should not contain elements which tend to degenerate with the refinement of the step. However, there are important problems arising in industry, engineering and natural sciences where maintaining strict obeyance to this constraint in the mesh generation can only be achieved at the cost of a huge increase of execution time/computer memory during the mesh-refinement process needed to obtain a sufficiently precise numerical solution. Here are some examples: boundary problems over domains consisting of several "almost disconnected" parts and domains with cusps (see, e.g., [28], Fig. 2.2; [27], Fig. 3.2), or even when the domain is infinitely smooth (see [28], Fig. 3.14, 3.15, [25], Fig. 4.1-3); Maxwell's equations for rotor-stator electric generators and motors (see [27], Fig. 7.3) and for optimal shape design of electromagnetic devices (see [27], Fig. 14.17-20); stationary semiconductor equations (see [27], Fig. 10.4); parabolic problems over time-dependent domains; hyperbolic problems in the study of interference phenomena in wave propagation through narrow passages; solving singular boundary integral equations, etc. In order to reduce the volume of computation, such a problem is often being solved in practice by committing the so-called "variational crimes" (one of which is to admit near-degenerate elements of certain types during the mesh generation). So far, theoretical "legalization" of these "crimes" has been carried out successfully in a number of cases on a "case-by-case" principle. The aim of the present paper is to propose a general approach to FEM and BEM involving near-degenerate elements. As we shall see, the results obtained have their analogues for multiresolution methods. In fact, for wavelet methods the underlying ideas become much more transparent and easy to understand.

In the sequel we shall be considering FEM in detail; the results obtained can be transferred, mutatis mutandis, to BEM, but the case of BEM will not be discussed in details here.

§2. Near-degenerate Elements and the Bramble–Hilbert Lemma

Classical examples of approximation by near-degenerate elements are known for a long time. In 1957 Synge proposed the so-called maximum angle condition (MAC) for triangular elements (see [26]). Further results on this condition were obtained by Babuška and Aziz, Barnhill and Gregory, and Křížek [26]. We refer to [26], Remark 2.2, and [27], Remarks 4.19 and 4.35 for more details. The MAC admits near-degenerate triangles (see [26], Fig. 1) and improves upon Zlamal's minimum angle condition and the inscribed ball condition (see [26], (2.1) and (1.1)). The latter two conditions (which are the most commonly accepted ones at present) do not allow the use of near-degenerate triangles. These conditions are based on a simple, but rough, estimate of the Jacobian of the diffeomorphism between an arbitrary and a reference triangle, while the MAC is obtained by a more refined bound for this Jacobian (see [26]).

Here we propose a more general approach to this problem based on a strengthened version of the Bramble–Hilbert lemma, as follows.

Theorem A. *(Dechevski and Wendland [13].) Let $m,\ n \in \mathbb{N},\ 1 \le p \le \infty$, and let $\Omega \subset \mathbb{R}^n$ be a bounded domain with* $\operatorname{diam} \Omega = d \in (0, \infty)$. *Assume that Ω is star-shaped with respect to any point $x_0 \in D$, where D is a measurable subset of Ω. Let the sublinear functional l have the properties*

$$1)\ \ \exists M < \infty : |l(f)| \le M \sum_{k=0}^{m} d^k ||f||_{\dot{W}_p^k(\Omega)},\ \ \forall f \in W_p^k(\Omega), \tag{1}$$

where $||f||_{\dot{W}_p^0(\Omega)} = ||f||_{L_p(\Omega)},\ \ ||f||_{\dot{W}_p^k(\Omega)} = \sum_{|\alpha|=k} ||D^\alpha f||_{L_p(\Omega)}$ is the semi-norm of the homogeneous Sobolev space $\dot{W}_p^k(\Omega)$, and $W_p^k(\Omega) = \bigcap_{k=0}^m \dot{W}_p^k(\Omega)$ is the inhomogeneous Sobolev space with norm $||f||_{W_p^k(\Omega)} = \sum_{k=0}^m ||f||_{\dot{W}_p^k(\Omega)}$;

$$2)\ \ l(P) \equiv 0 \text{ on } \Omega,\ \ \forall P \in \Pi_n^{m-1}(\Omega),$$

where $\Pi_n^m(\Omega)$ is the space of all n-variate algebraic polynomials on Ω with total degree $\deg P \le m$.

Then,

$$\exists C < \infty : |l(f)| \le C d^m ||f||_{\dot{W}_p^m(\Omega)},\ \ \forall f \in W_p^m(\Omega), \tag{2}$$

$$C \le \overline{M} M \max[1, \Big(\frac{|D|_n}{|\Omega|_n}\Big)^{\frac{1}{n}-\frac{1}{p}}],\ \ p \ne n, \tag{3}$$

$$C \le \overline{M}_1 . \frac{1}{n} . \ln \frac{|\Omega|_n}{|D|_n} + M_2,\ \ p = n, \tag{4}$$

where $|.|_n$ denotes the n-dimensional Lebesgue measure, and $\overline{M} = \overline{M}(m, n, p) < \infty$, $\overline{M}_j = \overline{M}_j(m, n) < \infty,\ j = 1, 2$, are independent of Ω and D.

Corollary A. *Under the conditions of Theorem A, assume that Ω is convex. Then, the embedding constant C in (2-4) does not depend on the domain Ω for any $p : 1 \le p \le \infty$.*

A detailed account of the improvements obtained in Theorem A and Corollary A with respect to previous relevant results will be given in [13]. Here we note the most essential improvement: given that the constant M in (1) is independent of Ω, Theorem A shows that the embedding constant C is independent of Ω for $n < p \le \infty$, and for $1 \le p \le n$ the influence of C on the degeneracy of Ω (i.e., on the smallness of $|D|_n/|\Omega|_n$) is much weaker than previous results on the Bramble–Hilbert lemma (see [4,5,18,19,20]) have suggested. The main result of [12] coincides with the partial case of Theorem A when $n < p \le \infty$ and D is a one-point set, i.e., $|D|_n = 0$.

Open problem. *For any $p : 1 \le p \le \infty$, characterize all sublinear functionals satisfying conditions 1 and 2 of Theorem A, such that the constant M in (1)* *does not depend on Ω. Especially important is the partial case of this problem when l is linear and/or Ω is convex.*

The solution of this problem is of key importance for the extension of the theory of direct (Jackson-type) inequalities for FEM and BEM to the case when the approximating family contains near-degenerate elements. We note that for some (but by far not for all) of the most frequently used finite elements on the simplex Ω in $\mathbb{R}^n$, the constant M in (1) is indeed independent of Ω, at least for $p = \infty$. The simplest such example is the linear interpolant at the vertices of the simplex. Another, much more general, example is provided by the Bernstein-Bézier family of interpolants on the simplex (see, e.g., [21], Chapt. 4 and 18). For these two examples, (1) takes the simple form $|l(f)| \leq M||f||_{L_\infty}$, where M is independent of Ω. It is possible to construct also finite elements based on Hermite interpolation, for which summands containing intermediate and/or highest-order derivatives are also present in (1). As for the case $p < \infty$, it can be seen that pointwise interpolants and, in particular, the two examples given above, *fail*: the constant M depends on Ω. In this case, one should look for "successfull candidates" among finite elements based on interpolants with respect to *integral, not pointwise,* functionals. Such types of finite elements (and finite difference schemes) are usually used in the search of *weak solutions* of operator equations by variational methods (see, e.g., [30], Chapt. III, section 1, subsection 1). All this suggests that the general solution of the open problem posed above must be closely related to the concept of *Wiener amalgam* spaces (see [15,22]). We shall return on this type of spaces later in the text.

Remark. The proof of Theorem A is based on the use of a regularized Taylor expansion and is rather technical and computationally involved. It is tempting to try to simplify this proof in the following way. Using Sobolev-type inequalities about the derivatives of intermediate order $k = 1, \ldots, m-1$, in the form (see [1,3,11,23])

$$\exists C_1(m,k,n,\Omega) < \infty : ||f||_{\dot{W}_p^k(\Omega)} \leq C_1(d^{-k}||f||_{L_p(\Omega)} + d^{m-k}||f||_{\dot{W}_p^m(\Omega)}),$$

for all $f \in W_p^k(\Omega)$, without loss of generality we may replace the RHS in (1) by the simpler $M_1(||f||_{L_p(\Omega)} + d^m||f||_{\dot{W}_p^m(\Omega)})$, where

$$M_1 = M(1 + \sum_{k=1}^{m-1} C_1(m,k,n,\Omega)). \tag{5}$$

Starting from this version of (1), the proof of Theorem A can be obtained in a simple way, e.g., by a much simpler version of our proof, or by applying the multidimensional Hardy inequality. Unfortunately, there is a hidden catch: Sobolev-type inequalities about intermediate derivatives are essentially equivalent to inverse (Markov-type) inequalities (cf., e.g., [3], Theorem 3.9) and, hence, the constants C_1 in (5) *depend strongly on the non-degeneracy of* Ω *for any* $p : 1 \leq p \leq \infty$! Therefore, in the context of near-degenerate Ω, this simple approach can be useful only for proving the partial case of Theorem A when l satisfies Condition 1 of the theorem without involving intermediate derivatives in the RHS of (1). Still, this partial case is quite important (for instance, it includes the examples of linear and Bernstein-Bézier interpolants considered above).

§3. Approximate Solutions of Nonlinear Operator Equations

This section is dedicated to the study of a very general class of nonlinear (including linear) operator equations, and to the study of the approximate solutions of these equations, obtained by iterative and projection methods.

Let X, Y be real Banach spaces, and let F be a (generally nonlinear) operator with $F(X) = Y$ and such that $\exists F^{-1} : F^{-1}(Y) = X$ with $F^{-1}F = I_X,\ FF^{-1} = I_Y$, i.e., F is bijective between X and Y, or, in other words, for the equation $F(x) = y,\ y \in Y$, there exists a unique solution x in X. Most of the results discussed below can be appropriately reformulated for complex Banach spaces, but in this short communication we shall consider only the case of real scalars.

In this context, we shall be especially interested in the case when F and F^{-1} are Lipschitz homeomorphisms between X and Y, and between Y and X, respectively, that is, $\exists C(F, X, Y) < \infty : ||F(x_1) - F(x_2)||_Y \leq C||x_1 - x_2||_X,\ \forall x_1 \forall x_2 \in X$, and analogously for F^{-1}. We shall denote this case for F and F^{-1} by $F \in LH(X, Y)$, for short. This definition of $LH(X, Y)$ continues to hold valid if, more generally, X and Y are quasi-Banach (see, e.g., [2,29,17,10]). Under natural restrictions on the Banach spaces X, Y, two very general subclasses of $LH(X, Y)$ are the class LSM of Lipschitz, strongly monotone operators and the class LSA of Lipschitz, strongly accretive operators (see Definitions 1 and 3 below).

Let H be a Hilbert space, such that $X \cap Y \subset H$ and $X \cap Y$ is dense on H, and let Y be the dual of X pivotal to H, i.e., the dual with respect to the duality functional defined by the scalar product of H. We shall denote this dual by $Y = X^* = X^*(H)$.

Definition 1. *Let $Y = X^*(H)$. The (generally nonlinear) operator $F : X \to X^*$ is called* Lipschitz, *if*

$$\exists C(F, X, H) < \infty : ||F(x_1) - F(x_2)||_{X^*} \leq C||x_1 - x_2||_X, \tag{6}$$

$\forall x_1 \forall x_2 \in X$, *and* strongly monotone, *if*

$$\exists c(F, X, H) < \infty : \langle F(x_1) - F(x_2), x_1 - x_2\rangle_H \geq c||x_1 - x_2||_X^2, \tag{7}$$

$\forall x_1 \forall x_2 \in X$. *The class $LSM = LSM(X, H)$ consists of exactly those $F : X \to X^*$ which satisfy (6,7).*

Definition 2. *Let $Y = X^*(H)$. The operator $U : X \to X^*$ is called* dual, *if*

$$\forall x \in X\ ||U(x)||_{X^*} = ||x||_X,\ \langle U(x), x\rangle_H = ||U(x)||_{X^*}||x||_X = ||x||_X^2.$$

It is known (see [35], Chapt. VII) that dual operators exist in every normed space, and that the dual operator (as a single-valued mapping) is unique if the space is strictly convex. The dual operator is homogeneous, but not additive. For example, if X has Gateau-differentiable norm, then $U(x) = ||x||_X.\mathrm{grad}||x||_X$, $x \neq 0$, and $U(0) = 0$. In particular, if $X = L_p(\Omega)$, $1 < p < \infty$, then $U(x,t) = ||x||^{2-p}_{L_p(\Omega)}|x(t)|^{p-2}x(t)$, $t \in \Omega$.

With the help of the dual operator $U : X \to X^*(H)$, it is possible to define a semiscalar product $[.,.]_X$ on X by $[x_1, x_2]_X := \langle x_1, U(x_2)\rangle_H$, for all x_1, $x_2 \in X$. For the properties of $[.,.]_X$, see [35], Chapt. VII.

Definition 3. *Let X be strictly convex, $Y = X^*(H)$, and $X \cap Y$ be dense on H. Let $U : X \to X^*$ be the (unique) dual operator. $F : X \to X$ is called* strongly accretive, *if*

$$\exists c(F,X,H) < \infty : [F(x_1) - F(x_2), x_1 - x_2]_X =$$

$$= \langle F(x_1) - F(x_2), U(x_1 - x_2)\rangle_H \geq c||x_1 - x_2||^2_X, \tag{8}$$

for all $x_1, x_2 \in X$. The class $LSA = LSA(X,H)$ consists of exactly those $F : X \to X$ which satisfy (8) and are Lipschitz on X.

It can be shown that the constants C and c in (6,7) are related by $c \leq C$. The same is true for the Lipschitz constant C in Definition 3 and c in (8).

It should be noted that if X is a Hilbert space (and, therefore, $X^*(H)$ is such, too), and if $X = H = X^*$, then the concepts of a monotone and an accretive operator coincide. However, if X and X^* are Hilbert spaces, but $X \neq H$, then monotonicity and accretivity are already diverse concepts. The typical case here is $X \hookrightarrow H \hookrightarrow X^*$ or $X \hookleftarrow H \hookleftarrow X^*$, where, as ususal, $A \hookrightarrow B$ or $B \hookleftarrow A$ denotes continuous embedding: $A \subset B$ and $||.||_B \leq C||.||_A$.

Theorem 1. *(Generalization and strengthening of Theorem 18.5 in [35] for the case of Lipschitz operators.) Let X be H-reflexive, i.e., $[X^*(H)]^*(H) = X$. Then, $LSM(X,H) \subset LH(X, X^*(H))$.*

Theorem 2. *(Generalization of Theorem 21.1 in [35].) Let the norm in X be uniformly Frechet-differentiable. Then, $LSA(X,H) \subset LH(X,X)$.*

Theorem B. *(See [28], Theorem 18.5.) Assume that $X = X^* = H$. Then, $LSM(H,H) = LSA(H,H) \subset LH(H,H)$ and the operator $T_{\varepsilon,y}(x) = x - \varepsilon[F(x) - y]$, $x \in H$, is contractive in H for $0 < \varepsilon < \frac{2c}{C^2}$, uniformly in $y \in H$, where C and c are defined in (6,7). The best contraction factor is $1 - c^2/C^2$ and is achieved for $\varepsilon = c/C^2$.*

In the remaining part of this section we shall consider methods for approximate solution of the equation $F(x) = y$, $x \in X$, $y \in Y$, X, Y - quasi-Banach spaces.

Definition 4. *(See [29].) Let X be a quasi-Banach space. $G \subset X$ is called an* existence set *for X, if $\forall x \in X\ \exists g_x \in G : ||x - g_x||_X = \min_{g\in G} ||x - g||_X = E_G(x)_X$. (The best approximation g_x need not be necessarily unique.) The* *sequence* $\{G_N\}_{N=1}^{\infty}$, $G_N \subset X$ *is called a* normal approximation family in X, *if for any* $N \in \mathbb{N}$, G_N *is an existence set, with* $G_N \subset G_{N+1}$ *and* $G_N - G_{N-1} \subset G_{2N}$.

Obviously, an existence set in X is closed in X (typical example: any finite-dimensional subspace of X).

Definition 5. *Let* X *be a quasi-Banach space. The sequence* $\{G_N\}_{N=1}^{\infty}$: $G_N \subset X$ *is called to* have the strong (weak) approximation property *(SAP, resp. WAP, for short) if* $\bigcup_{N=1}^{\infty} G_N$ *is dense on* X *in the quasinorm topology of* X *(resp., in the weak topology on* X*, which in this case is assumed to be a Banach space).*

Lemma 1. *Let* X, Y *be quasi-Banach spaces,* $F \in LH(X,Y)$, $X \supset G \neq \emptyset$. *Consider* $y^* : ||y-y^*||_Y \leq 2\inf_{\eta\in F(G)} ||y-\eta||_Y$. *(Obviously, such* y^* *exists.) Then, there exists a unique solution* $\tilde{x}$ *of the equation* $F(x) = y^*$, *and*

$$\exists C(F,X,Y) < \infty : \frac{1}{2}\inf_{\xi\in G} ||x-\xi||_X \leq ||x-\tilde{x}||_X \leq C\inf_{\xi\in G} ||x-\xi||_X.$$

Proposition 1. *Let* X, Y *be quasi-Banach,* $F \in LH(X,Y)$, $\{G_N\}_{N=1}^{\infty}$: $G_N \subset X$, $G_N \subset G_{N+1}$. *Assume that* G_N *is an existence set in* X, $\Gamma_N = F(G_N)$ *is an existence set in* Y*, and* $\{\Gamma_N\}_{N=1}^{\infty}$ *has the SAP in* Y*. Then,*

(i) the equation $F(x) = y$ has a unique solution in G_N for any $y \in \Gamma_N$;
(ii) $\{G_N\}$ has the SAP in X;
(iii) if y_N^* is a best approximation to y in $F(G_N)$, and if $\tilde{x}_N : F(\tilde{x}_N) = y_N^*$, then

$$\exists C(F,X,Y) < \infty : E_N(x)_X := E_{G_N}(x)_X \leq ||x-\tilde{x}_N||_X \leq CE_N(x)_X.$$

Let us consider now the *Galerkin–Petrov projection methods.* Let $P_N : X \to X$, $Q_N : X \to X$ be projectors with $\dim P_N(x) = \dim Q_N(X) = N$, and $P_N P_{N+1} = P_{N+1}P_N = P_N$, $Q_N Q_{N+1} = Q_{N+1}Q_N = Q_N$.

Example 1. (*Galerkin–Petrov method for monotone operators.*) Let $Y = X^*(H)$. The equation $F(x) = y$, $x \in X$, $y \in X^*$, is replaced by $Q_N^* F(P_N x) = Q_N^* y$, where $Q_N^* : X^* \to X^*$, $\dim Q_N^*(X^*) = N$, is the Banach adjoint of Q_N. The $N \times N$ nonlinear system is determined by

$$\langle Q_N^* F(P_N x), Q_N h\rangle_H = \langle Q_N^* y, Q_N h\rangle_H. \tag{9}$$

Example 2. (*Galerkin–Petrov method for accretive operators.*) Let $Y = X$ and let $X^*(H)$ be strictly convex. Let $U : X \to X^*(H)$ be the dual operator. Then, by Lemma 22.4 in [35], $UQ_N = Q_N^* U Q_N$, which implies that the equation $F(x) = y$, $x \in X$, $y \in X$ is now replaced by $Q_N F(P_N x) = Q_N y$. The $N \times N$ nonlinear system is determined by

$$[Q_N F(P_N x), Q_N h]_X = [Q_N y, Q_N h]_X. \tag{10}$$

By Lemma 23.1 in [35], it follows that if $F \in LMS(X,H)$, where X and H are separable, then (9) has a unique solution for N large enough. By Lemma 23.2 and Remark 23.1 in [35], it follows also that if $F \in LMA(X,H)$, where X and H are separable, if X satisfies the conditions of Definition 22.1 in [35], and if X^* is strictly convex, then (10) has a unique solution for N large enough. In the Hilbert case $X = X^* = H$, if N is large enough, so that $Q_N^* F(P_N H) = Q_N^* H$ holds, then, by Theorem B, (9) can be computed by quickly converging contractive iterations. For small N the condition $Q_N^* F(P_N H) = Q_N^* H$ may fail even if F is linear and $P_N = Q_N$ (see [6], Theorem 10.1.1).

Returning to the general real non-Hilbertian case of Example 1, we note that, if F is twice Gateau-differentiable, then Newton's method can be used where the inverse matrix involved in each iteration is usually sparse. In general, this method needs an appropriate initial approximation x_0 to the solution of $F(x) = y$, but if F is strongly monotone and potential, that is, if there exists a real functional $f : X \to \mathbb{R}$ such that $F = \operatorname{grad} f$, then, by Theorem 5.1 in [35], f is strictly convex and the solution of (9) is equivalent to minimizing the three times Gateau-differentiable functional f, hence, Newton's method converges to the solution of (9) for any $x_0 \in X$, the rate of convergence depending on the constant c in (7). This technique is still numerically efficient if F is only Lipschitz and Newton's method or its various modifications be replaced by the respective variants of the more general F. Clarke's subdifferential method. In the case of potential F, this type of technique remains efficient even if only f is Lipschitz and Clarke's subgradient method is applied to the optimization of f. In the case of potential F, the *Bubnov–Galerkin method* ($P_N = Q_N$) coincides with the *Ritz method for minimization of* f.

For projection methods, (see Examples 1 and 2) the strong approximation property can be written as $\lim_{N\to\infty} ||(I_X - P_N)x||_X = 0$. A typical example when $G_N = P_N X$ forms a NAF having the SAP is when P_N is obtained by multiresolution.

By Theorem 23.3 in [35], if X is a real reflexive Banach space with unconditional basis, then $G_N = P_N X$, as defined in Example 1, has the SAP; by Lemma 23.1 in [35] and in view of $F \in LMS(X,H)$, the solution x_N of (9) exists for N large enough and $||x_N - x||_X \to 0$, where x is the solution of $F(x) = y,\ y \in X^*$. By Theorem 23.2 in [35], if X is a real H-reflexive Banach space satisfying the conditions of Definition 22.1 in [35], if $X^*(H)$ is strictly convex, and if the dual operator $U : X \to X^*$ is continuous and sequentially weakly continuous, then, $G_N = P_N X$, as defined in Example 2, has the WAP; by Lemma 23.2 in [35], and since $F \in LMA(X,H)$, the solution x_N of (10) exists for N large enough, and the weak limit of any convergent subsequence of $\{x_N\}$ is x, the solution of $F(x) = y,\ y \in X$.

For the formulation of Céa's lemma in the linear and nonlinear case for F, in Hilbert spaces, see [31], Lemma 2.8; [28], Theorems 4.1 and 18.8.

Theorem 3. *(Generalization of Céa's lemma for Banach spaces, and strengthening of Proposition 1 in the case of Example 1.) Under the assumptions of Example 1, let* $x = F^{-1}(y) \in X$ *be the solution of* $F(x) = y \in X^*(H)$, *and*

$x_n \in X$ *be the solution of (9). Then,*

$$\exists C(F, X, H) < \infty : E_N(F^{-1}(y))_X \leq ||F^{-1}(y) - x_N||_X \leq C E_N(F^{-1}(y))_X.$$

The idea of the proof of Céa's lemma does not admit generalization for accretive operators in Banach spaces (Example 2), because the dual operator $U : X \to X^*$ is not additive.

§4. Best N-term Approximation

For the general paradigm of best N-term approximation (BNTAP) we refer to [29], section 3.5, and [16].

Definition 6. *Let* $X_j,\ Y_j,\ j = 0, 1$ *be quasi-Banach spaces,* $X_1 \hookrightarrow X_0,\ Y_1 \hookrightarrow Y_0$, *and let* $F \in LH(X_0, Y_0) \cap LH(X_1, Y_1)$. *The NAF* $\{G_N\}_{N=1}^{\infty} :\ G_N \subset X_1$, *is called* near-degenerate of order $(\lambda; \alpha, \beta)$, $\lambda > 0,\ \alpha \geq 0,\ \beta \geq 0$, *if it satisfies a* direct inequality *of the type*

$$\exists C < \infty :\ E_N(F^{-1}(y))_{X_0} \leq C \frac{||F^{-1}(y)||_{X_1}}{N^{\lambda}},\ \ \forall y \in Y_1, \tag{11}$$

where $C = C(N)$, *with* $C \asymp N^{\alpha}$*; and an* inverse inequality *of the type*

$$\exists D < \infty :\ ||x||_{X_1} \leq D N^{\lambda} ||x||_{X_0},\ x \in G_N, \tag{12}$$

where $D = D(N)$, *with* $D \asymp N^{\beta}$. *The partial case* $\alpha = \beta = 0$ *corresponds to a* non-degenerate (regular) *NAF.*

Remark. The inequalities (11) are also called Jackson-type inequalities. The inverse results (12) have many names depending on the context (Bernstein, Markov, Babuška, etc.). In our opinion, when X_1 is a function space over a domain Ω and the domain is *open*, (12) should be called Bernstein-type inequalities; when Ω is *closed*, (12) should be called Markov-type inequalities, in correspondence with the original results of Bernstein and Markov. Thus, in the case of FEM and BEM we are dealing with Markov-type inequalities. Our results in section 2 show that it is possible for $\Omega = \Omega_N$ to tend to degenerate as $N \to \infty$ while in the Jackson-type inequality α remains fixed at zero. However, the sharp results available for the Markov-type inequality (see [3]) show that if Ω_N tends to degenerate then always $D = D_N \to \infty$ holds in (12), and if Ω_N is *polynomially cuspidal*, then, necessarily, $\beta > 0$ holds.

Consider the approximation space $A_q^s(X_0) :=$

$$= \{f \in X_0 : ||f||_{A_q^s(X_0)} = \big(||f||_{X_0}^q + \sum_{j=0}^{\infty} [2^{js} E_{2^j}(f)_{X_0}]^q\big)^{1/q} < \infty\} \tag{15}$$

and the real interpolation space $(Y_0, Y_1)_{\theta,q} := \{f \in X_0 :$

$$||f||_{(Y_0,Y_1)_{\theta,q}} = \big(||f||_{X_0}^q + \sum_{j=0}^{\infty} [2^{j\theta} K(2^{-j}, f; Y_0, Y_1)]^q\big)^{1/q} < \infty\}, \tag{16}$$

where $K(t, f; Y_0, Y_1)$ is Peetre's K-functional (see [2,17,29,11,10]), $0 < t < \infty$, $s > 0$, $0 < \theta < 1$, $0 < q \leq \infty$ (with the usual sup-modification in (15,16) for $q = \infty$). (Recall that $X_1 \hookrightarrow X_0$, which explains the presence of the saturation term $||f||_{X_0}^q$ in (15,16).)

Theorem 4. *(Characterization of the best N-term approximation of solutions of nonlinear operator equations by near-degenerate NAF.) Assume that the conditions of Definition 6 hold. Let $0 < q \leq \infty$. Then,*

(i) if $0 \leq \alpha < \lambda$ and $s : 0 < s < \lambda - \alpha$, then, $\exists C_1 < \infty$:

$$||F^{-1}(y)||_{A_q^s(X_0)} \leq C_1[||F^{-1}(0)||_{X_1} + ||y||_{(Y_0,Y_1)_{\frac{s}{\lambda-\alpha},q}}]; \tag{13}$$

(ii) if $\beta \geq 0$ and $0 < s < \lambda + \beta$, then $\exists C_2 < \infty$:

$$||y||_{(Y_0,Y_1)_{\frac{s}{\lambda+\beta},q}} \leq C_2[||F(0)||_{Y_1} + ||F^{-1}(y)||_{A_q^s(X_0)}]. \tag{14}$$

Corollary 1. *Under the conditions of Theorem 4, let $0 < s < \lambda - \alpha$, and assume that $F(0) = 0_{Y_1},\ F^{-1}(0) = 0_{X_1}$. Then,*

$$(Y_0, Y_1)_{\frac{s}{\lambda-\alpha},q} \hookrightarrow A_q^s(X_0) \hookrightarrow (Y_0, Y_1)_{\frac{s}{\lambda+\beta},q}. \tag{17}$$

In particular, if $\alpha = \beta = 0$, then

$$(Y_0, Y_1)_{\frac{s}{\lambda},q} = A_q^s(X_0) \tag{18}$$

(isomorphism of the spaces, equivalence of the quasinorms).

Remark. If the dependence of C in (11) and/or D in (12) on N is weaker than polynomial, e.g., logarithmic, then the left-hand and/or right-hand embedding in (17) can be sharpened by setting $\alpha = 0$ and/or $\beta = 0$ and modifying the index q. We omit the details.

Multiresolution Galerkin-Petrov methods for monotone and accretive operators (Examples 1 and 2) are included as partial cases in Theorem 4 and Corollary 1. For monotone operators, we have $Y_0 = X_0^*(H_0),\ Y_1 = X_1^*(H_1)$, where $H_0,\ H_1$ are Hilbert spaces with $H_0 \hookleftarrow H_1$ which are *sufficiently far away from each other so that $X_0 \hookleftarrow X_1$ and $X_0^*(H_0) \hookleftarrow X_1^*(H_1)$ hold simultaneously.* For accretive operators, $Y_0 = X_0,\ Y_1 = X_1$ holds, and the above requirement about $H_0,\ H_1$ being "far enough" from each other is not needed. (It is possible to have even the inverse embedding $H_0 \hookrightarrow H_1$.) In the cases of Example 1 and 2, $X_1 \cap Y_1$ is assumed to be dense in H_0 and H_1. The projectors P_N and Q_N in Examples 1 and 2 are assumed generated by multiresolution, which ensures that $G_N - G_{N-1} \subset G_{2N}$.

In the rest of this section and in the next section we shall discuss how to reduce the rates α and β in Theorem 4 and Corollary 1 to zero in the presence of near-degeneracy. To this end, we shall study the analogue of the phenomenon of near-degeneracy with multiresolution methods based on biorthogonal wavelets.

One equivalent quasinorm in the *inhomogeneous* Besov spaces B_{pq}^s (cf., e.g., [33,8]) is given by

$$||f||_{B_{pq}^s} \asymp \{(\sum_{k\in\mathbb{Z}^n} |\alpha_{0k}|^p)^{q/p} +$$

$$+\sum_{j=0}^{\infty}[2^{j[s+n(1/2-1/p)]}(\sum_{k\in\mathbb{Z}^n}\sum_{l=1}^{2^n-1}|\beta_{jk}^{[l]}|^p)^{1/p}]^q\}^{1/q}, \tag{19}$$

with $0<p\le\infty,\ 0<q\le\infty,\ n\max(\frac{1}{p}-1,0)<s<r$, where r is the Lipschitz regularity of the compactly supported scaling functions $\varphi\in B^r_{\infty\infty},\ \tilde{\varphi}\in B^r_{\infty\infty}$ and wavelets $\psi^{[l]}\in B^r_{\infty\infty},\ \tilde{\psi}^{[l]}\in B^r_{\infty\infty}$ of the biorthonormal wavelet bases, with respect to which $f\in B^s_{pq}$ can be expanded as follows:

$$f(x)=\sum_{k\in\mathbb{Z}^n}\alpha_{0k}\varphi_{0k}(x)+\sum_{j=0}^{\infty}\sum_{k\in\mathbb{Z}^n}\sum_{l=1}^{2^n-1}\beta_{jk}^{[l]}\psi_{jk}^{[l]}(x),\ \text{ a. e. } x, \tag{20}$$

where $\alpha_{0k}=\langle f,\tilde{\varphi}_{0k}\rangle_{L_2},\ \beta_{jk}^{[l]}=\langle f,\tilde{\psi}_{jk}^{[l]}\rangle_{L_2}$. Each hypercube in the Calderon-Zygmund decomposition of $\mathbb{R}^n$ and Stein's construction of Whitney-type extension operators (see [34]) corresponds to $2^{n-1}-1$ basis functions $\tilde{\psi}_{jk}^{[l]},\ \psi_{jk}^{[l]}$, $l=1,\dots,2^n-1$, in each of the two biorthonormal bases. The convergence in (20) is in the quasinorm topology of B^s_{pq}, but also Lebesgue a. e. on the domain Ω of the functions, because of the lower bound on s. Ω may be $\mathbb{R}^n$, hyperrectangle, correspond to the periodic case, or even general Lipschitz-graph domain. We refer to the currently most advanced work on this topic [7], as well as to the extensive account [8] (for the case of *homogeneous* Besov spaces, see [14]).

Definition 7. *Let $j_1\in\mathbb{N}$. A* non-degenerate wavelet-based projector *(NWP) is denoted by P_{j_1} and defined by*

$$P_{j_1}f(x)=\sum_k\alpha_{0k}\varphi_{0k}(x)+\sum_{j=0}^{j_1-1}\sum_k\sum_{l=1}^{2^n-1}\beta_{jk}^{[l]}\psi_{jk}^{[l]}(x),\ \ x\in\Omega, \tag{21}$$

cf. (20). A near-degenerate wavelet-based projector *(NDWP) is denoted by $\tilde{P}_{j_1}$ and defined by*

$$\tilde{P}_{j_1}f(x)=\sum_k\alpha_{0k}\varphi_{0k}(x)+\sum_k\sum_{j=0}^{J(j_1,k)-1}\sum_{l=1}^{2^n-1}\beta_{jk}^{[l]}\psi_{jk}^{[l]}(x),\ \ x\in\Omega, \tag{22}$$

$$\forall k\ J(j_1,k)>J(j_1-1,k),\ J(j_1,k)\ge j_1;\ \ \exists k_{j_1}:J(j_1,k_{j_1})=j_1. \tag{23}$$

In other words, for a NWP $J(j_1,k)=j_1=\text{const}$, uniformly in k. Thus NDWP's are a specific partial case of *lacunary wavelet-based projectors* (see the concluding remarks in [8], subsection 6.2), lacunarity being with respect to the NWP corresponding to $J_1:=\max_k J(j_1,k)$.

Example 3. A typical (and very general) example when near-degenerate FEM or lacunary wavelet-based projectors of NDWP type are needed, is the error analysis of numerical solutions in the immediate neighbourhood of the boundary $\partial\Omega$ (see, e.g., [28], Fig. 3.14, 3.15, 6.14, 6.15, 8.12). Then it is desirable to ensure that the local approximation rates near and on $\partial\Omega$ *do not deteriorate* compared to the local approximation rates in the interior of Ω. Indeed, assume that $\partial\Omega$ is regular enough (Lipschitz or smoother). Then, by the trace theorem (see, e.g., [2,24]), if $f \in B^s_{pq}(\Omega)$, $\Omega \subset \mathbb{R}^n$, then the restriction of f on $\partial\Omega$ is less regular for $p < \infty$, namely, $f_{|\partial\Omega} \in B^{s-1/p}_{pq}(\partial\Omega)$ holds. Let, for simplicity, $q = \infty$. Then, the local approximation rate achieved via NWP, given in (21), is $O(2^{-j_1[s+n(1/2-1/p)]})$ in the interior of Ω and only $O(2^{-j_1[s-1/p+(n-1)(1/2-1/p)]})$ near $\partial\Omega$. To achieve the desired uniform distribution of the error in the interior and near the boundary when f is smooth enough ($s > \max(0, \frac{n}{p} - \frac{n-1}{2})$), NDWP given in (22,23) should be employed, with $J(j_1, k) \approx j_1$ for k corresponding to the interior of Ω, and with $J(j_1, k) \asymp C_1 j_1 + C_2$ otherwise, where

$$C_2 \geq 0 \;\; C_1 = 1 + \frac{1}{2(s + \frac{n-1}{2} - \frac{n}{p})} > 1. \tag{24}$$

If the Lipschitz graph $\partial\Omega$ contains edges or corners of dimensions lower than $n-1$, then, by the same type of argument, C_1 in (24) should be increased. If $\partial\Omega$ contains *polynomial cusps* (see [3]) or is a *fractal d-set* (see [24]), then C_1 in (24) must be increased, too. Finally, in the neighbourhood of an *exponential cusp* (see [3]) $C_1 = C_1(N) \to \infty$ holds, so that $J(j_1, k)/j_1 \to \infty$ as $j_1 \to \infty$.

In the context of Theorem 4 and Corollary 1, if $X_j = B^{s_j}_{p_j q_j}$, $j = 0, 1$, and the Sobolev embedding holds, with $s_0 - \frac{n}{p_0} \leq s_1 - \frac{n}{p_1}$, $0 < q_1 \leq q_0 \leq \infty$, so that $X_1 \hookrightarrow X_0$ is fulfilled, then it can be verified that for NWP both the direct inequality $||f - P_{j_1} f||_{X_0} \leq C_1 2^{-j_1\lambda} ||f||_{X_1}$ and the inverse inequality $||P_{j_1} f||_{X_1} \leq C_2 2^{j_1\lambda} ||P_{j_1} f||_{X_0}$ hold when $p_0 = p_1$, $q_0 = q_1$, with $\lambda = s_1 - n/p_1 - (s_0 - n/p_0)$, and with C_1, C_2 independent of j_1, hence, in Definition 6 $\alpha = \beta = 0$ is attained. On the contrary, for NDWP satisfying (24) the constants C_1 and C_2 depend on j_1 and $\alpha > 0$, $\beta > 0$ holds.

In the case of NDWP, is it possible to somehow reduce α and β to zero, thereby achieving isomorphism in (17)? It turns out that *the answer is positive*, and below we shall propose a general method how to achieve this. Another approach will be discussed in the next section.

Our first approach is to try to save the entire BNTAP intact, by considering *more general spaces* X_0, X_1 *than Besov spaces*, so that, for the new X_0 and X_1, $\alpha = \beta = 0$. Consider the quasi-Banach space B^W_{pq} with quasinorm

$$||f||_{B^W_{pq}} \asymp [(\sum_k |\alpha_{0k}|^p)^{q/p} + \sum_{j=0}^{\infty} (\sum_k 2^{pW(j,k)} \sum_{l=1}^{2^n-1} |\beta^{[l]}_{jk}|^p)^{q/p}]^{1/q}.$$

This space scale is very general: it contains not only Besov spaces (for $W(j,k) = j(s + \frac{n}{2} - \frac{n}{p})$, but also Nikol'skii-Besov spaces (when $W(j,k) = W(j)$). The spaces from this scale still admit atomic decomposition via the same Riesz bases of biorthonormal wavelets as the Besov spaces. We now specify the choice of $W(j,k)$, as follows: $W(j,k) = w(j,k)[s + n(\frac{1}{2} - \frac{1}{p})]$, where s is subject to the same constraints as in (19). The weight $w(j,k)$ is positive, monotonously increasing function in j for each fixed k, and depends on the choice of $J(j_1,k)$ in (23). The definition of $w(j,k)$ is

$$w(J(j_1,k),k) = j_1, \tag{25}$$

$$w(j,k) = j_1 - 1, \ \ j = J(j_1-1,k), \ J(j_1-1,k)+1, \ \ldots, \ J(j_1,k)-1, \tag{26}$$

$\forall j_1 \in \mathbb{N}\ \forall k \in \mathbb{Z}^n$. For this choice of $W(j,k)$ we shall write $B^W_{pq} = B^{s,w}_{pq}$.

Now, take $X_j = B^{s_j,w}_{p_j q_j}$, $j=0,1$, with $s_0 - \frac{n}{p_0} < s_1 - \frac{n}{p_1}$, $0 < q_1 \le q_0 \le \infty$. It can be seen that $X_1 \hookrightarrow X_0$ holds, and we can consider this pair of spaces in the context of Theorem 4 and Corollary 1. It can be verified that the bounds

$$||f - \tilde{P}_{j_1} f||_{B^{s_0,w}_{p_0 q_0}} \le$$

$$\le \Big[\sum_{j=j_1}^{\infty} \Big(\sum_{k: w(j,k) \ge j_1} 2^{p_0 w(j,k)(s_0+n/2-n/p_0)} \sum_l |\beta^{[l]}_{jk}|^{p_0} \Big)^{q_0/p_0} \Big]^{1/q_0}, \tag{27}$$

$$||\tilde{P}_{j_1} f||_{B^{s_1,w}_{p_1 q_1}} \le \Big[\Big(|\alpha_{0k}|^{p_1} \Big)^{q_1/p_1} +$$

$$+ \Big[\sum_{j=0}^{J_1} \Big(\sum_{k: w(j,k) \le j_1} 2^{p_1 w(j,k)(s_1+n/2-n/p_1)} \sum_l |\beta^{[l]}_{jk}|^{p_1} \Big)^{q_1/p_1} \Big]^{1/q_1}, \tag{28}$$

hold. (Recall that $j_1 = \min_k J(j_1,k)$, $J_1 = \max_k J(j_1,k)$.) After some computations, (27,28) imply

$$||f - \tilde{P}_{j_1} f||_{X_0} \le C_1 2^{-j_1 \lambda} ||f||_{X_1}, \ \ \forall f \in X_1, \tag{29}$$

$$||\tilde{P}_{j_1} f||_{X_1} \le C_2 2^{j_1 \lambda} ||\tilde{P}_{j_1} f||_{X_0}, \ \ \forall f \in X_0, \tag{30}$$

(where (30) holds only when $p_0 = p_1$, $q_0 = q_1$), with $\lambda = s_1 - n/p_1 - (s_0 - n/p_0)$, and the constants C_1 and C_2 in (29,30) *do not depend on* j_1, i.e., for this choice of the spaces X_0, X_1 in Definition 6 $\alpha = \beta = 0$ holds. Thus, we have solved the problem about characterization of the best approximation spaces induced by NDWP defined in (22,23). This solution is given by

Corollary 2. *Under the conditions of Corollary 1, assume that* $X_j = B^{s_j,w}_{p_j q_j}$, $j = 0,1$, *where* $w = w(j,k)$ *is the left inverse of* $J(j,k)$ *as defined in (23). Assume also that* $N = 2^{j_1}$ *and* $G_N = \tilde{P}_{j_1} X_0$, *where the NDWP* $\tilde{P}_{j_1}$ *is defined in (22), with the same* $J(j,k)$ *in (23). Let* $s : 0 < s < \lambda = s_1 - n/p_1 - (s_0 - n/p_0)$. *If* $p_0 = p_1$, $q_0 = q_1$, *then,*

$$A^s_q(X_0) = (Y_0, Y_1)_{\frac{s}{\lambda},q}.$$

§5. Atomic Decomposition of Wiener Amalgam Spaces

The second approach we suggest to deal with reducing the size of α and β in Definition 6 is to go beyond the general BNTAP, by abandoning one part of it, namely, the use of the real interpolation functor. The main object of interest here are *Wiener amalgam spaces* (WAS) whose relevance to the problem about near-degenerate partitions was already noted in section 2. More precisely, Theorem A tells us that L_p-metrics with higher value of p are in a certain sense more "tolerant" to the use of near-degenerate elements. This suggests to use two metrics in the error estimates - one, with p large, for local estimates on each separate element - and a second one, with smaller p, to obtain the global averaging of the local results. This type of error estimates has been already in consideration, for over 30 years, when applying the so called *average moduli of smoothness (τ-moduli)* (see [32,9,10,11]). On the other hand, Feichtinger (see, e.g., [22]) and others have developed an elegant general theory of Wiener amalgam spaces. This theory is oriented towards applications other than error analysis. Due to this, little attention has been paid in this theory to the "window size" of the local space. However, in error estimation this size is related to the approximation step and is, therefore, of crucial importance. This aspect of the theory of Wiener amalgam spaces has been studied in detail in [9,10,11]. In [10,11] we indicated that *Wiener amalgam spaces cannot be interpolated by the real (as well as the complex) interpolation method*, because in the case of WAS the spaces involved in Peetre's K-functional *depend themselves on the step of the K-functional.* Taking this in consideration, in [10,11] a different approach was proposed to interpolate operators acting between WAS, and the idea of this approach will be essential for the construction proposed below. For simplicity, we shall consider here only the case when the domain of the functions is $\Omega = \mathbb{R}^n$. In the definition (19) of the quasinorm of $B^s_q(L_p) := B^s_{pq}$ we shall replace L_p by the WAS $W(L_\rho, L_p)$, where $p \le \rho \le \infty$, and (cf. [10,11])

$$||f||_{W(L_\rho,L_p;2^{-j})} = \Big(\sum_{\nu\in\mathbb{Z}^n} 2^{-jn(1-p/\rho)}||\chi_{2^{-j},\nu}.f||^p_{L_\rho}\Big)^{1/p}, \tag{31}$$

where $\chi_{\delta,\nu},\ \delta > 0$, are the indicator functions of sets forming a $(\delta; C, c)$-quasiuniform disjoint cover of $\mathbb{R}^n$. (The disjoint cover $\{\Delta_\nu\}_{\nu\in\mathbb{Z}^n}$ of $\mathbb{R}^n$ (hence, also of $\mathbb{Z}^n$) is called $(\delta; C, c)$-quasiuniform, $\delta > 0,\ 0 < c \le C < \infty$, if diam $\Delta_\nu \le 2C\delta$ and Δ_ν contains a ball of radius $\ge c\delta$, uniformly in $\nu \in \mathbb{Z}^n$.) Note that the space $A_{p,\delta}$, defined in [10,11], is isomorphic to $W(L_\infty, L_p)$, $0 < p \le \infty$ with the quasinorm $||.||_{W(L_\infty,L_p;\delta)}$ as of (31), and with constants of quasinorm equivalence which are independent of δ.

Using one of the versions of the discrete Hardy inequality (see [24], Chapt. V, section 2.1, Lemma 3), we can prove the following.

Lemma 2. *Fix $\sigma > -(s + \frac{n}{2} - \frac{n}{p})$. Then, for any admissible value of p, q and s, for which (19) holds, there is also the following equivalent quasinorm* *in* $B^s_q(L_p)$*:*

$$||f||_{B^s_q(L_p)} \asymp \{(\sum_{k\in\mathbb{Z}^n} |\alpha_{0k}|^p)^{q/p}+$$

$$+\sum_{j=0}^{\infty} 2^{jq[s+n(1/2-1/p)+\sigma]}[\sum_{\mu=j}^{\infty} 2^{-\sigma\mu}(\sum_{k\in\mathbb{Z}^n}\sum_{l=1}^{2^n-1}|\beta^{[l]}_{\mu k}|^p)^{1/p}]^q\}^{1/q}. \qquad (32)$$

The analogue of (32) with $B^s_q(W(L_\rho,L_p))$ is

$$||f||_{B^s_q(W(L_\rho,L_p))} \asymp \{(\sum_{k\in\mathbb{Z}^n} |\alpha_{0k}|^p)^{q/p}+$$

$$+\sum_{j=0}^{\infty} 2^{jq[s+n(1/2-1/p)+\sigma]}[\sum_{\mu=j}^{\infty} 2^{-\sigma\mu}(\sum_{\nu\in\mathbb{Z}^n}(\sum_{k\in\Delta_{j\nu}}\sum_{l=1}^{2^n-1}|\beta^{[l]}_{\mu k}|^\rho)^{p/\rho})^{1/p}]^q\}^{1/q}, \quad (33)$$

where $0<p\le\rho\le\infty$, and there exist C and c with $0<c\le C<\infty$, such that $\{\Delta_{j\nu}\}_{\nu\in\mathbb{Z}^n}$ is a $(2^{-j};C,c)$-quasiuniform disjoint cover of $\mathbb{R}^n$ and $\mathbb{Z}^n$, uniformly in $j=0,1,\dots$.

Formula (33) shows that the space $B^s_q(W(L_\rho,L_p))$ admits atomic decomposition via biorthonormal wavelet bases. This opens a way to study best N-term approximation in WAS, in a situation where the standard BNTAP does not work.

In the end of this section, we note that the new spaces $B^{s,w}_{pq}$, considered in section 4, and formula (33) are of considerable interest also for *penalized wavelet estimation and smoothing in statistics* (for this approach to wavelet estimation and smoothing, see [15]).

§6. Comparison between Finite Element and Wavelet Methods

In this concluding section we make a short and inexhaustive comparison between FEM and wavelet methods, in the light of the results obtained in the previous sections.

As we have seen in Theorem 3 and Corollary 1, (18), both FEM and wavelet methods achieve the best N-term approximation rates. In the event of near-degeneracy, we have been able to prove stronger results in the case of wavelets, since for lacunary multiresolution methods we found a way to reduce the values of α and β in the direct and inverse inequalities to zero, while in the case of FEM we were able to achieve this only for α.

FEM seem to have advantage over wavelet methods in the case when additional information about properties of the solution (e.g., edges of discontinuity) is available a priori and can be taken into account in the mesh generation. This is more economical to do with FEM. For the same reason, FEM has also some advantages when the boundary-value problem is over a domain with piecewise smooth boundary which has “difficult geometry” (e.g., a cuspidal domain). Nevertheless, the recent multilevel implementation of Stein’s construction of Whitney-type extensions (see [7]) on Lipschitz-graph domains is a remarkable achievement with great conceptual importance, in particular, because it opens the way to developing a numerically implementable wavelet-based analogue of the convolution/local-Taylor-expansion method of [24] for constructing Whitney-type extension operators in the very general situation when $\partial\Omega$, or even Ω itself, is a fractal subset of $\mathbb{R}^n$ with possibly fractional Hausdorff dimension. This would result in achieving biorthogonal wavelet atomic decomposition of function spaces over fractal subsets of $\mathbb{R}^n$. (Envisaged applications are, e.g., in solving boundary problems for stochastic differential equations of diffusion type, with potential contributions to the solution of problems arising in financial mathematics.)

In the interior of a Lipschitz-graph domain Ω, adaptive wavelet methods seem to do better than FEM (even multilevel FEM) in unbiased situations when only a minimal amount of additional information is available about the solution of the operator equation. One general reason for this to happen is that in unbiased situations multiresolution seems to be the most flexible and universal approach to mesh generation.

In sections 4 and 5 we have seen that the wavelet approach offers also a somehow deeper insight into the theoretical consequences of the use of near-degenerate element partitions.

Let us summarize the conclusions of our comparison:

1) without replacing existing finite element/boundary element methodology, wavelet methods are a valuable complement to existing FEM/BEM, with ready application to operator equations with spatially inhomogeneous solutions;

2) wavelet methods are a powerful tool for theoretical analysis, and for efficient numerical implementation of so-far theoretical constructions;

3) rather than supplanting each other, FEM and wavelet methods have the potential to harmoniously coexist and complement each other.

Example. One common trick helping to reduce the time and computer memory needed for solving a boundary-value problem is to use simplectic and special finite elements near $\partial\Omega$, and rectangular elements in the interior of Ω. In this case, the domain $\Omega_{1,h} \subset \Omega$ (eventually depending on the approximation step h of the method), in which rectangular elements are used, is a disjoint union of cubes or rectangles. It is exactly for boundary problems on this type of domains that wavelet methods (involving boundary-adjusted wavelets near $\partial\Omega_{1,h}$) are at their best. In our opinion, when the solution of the boundary problem is expected to change relatively sharply within the interior of Ω, the rectangular finite elements should be replaced by an adaptive wavelet method on $\Omega_{1,h}$. In this construction, the non-rectangular finite elements on $\Omega\backslash\Omega_{1,h}$ play a double role: they serve both for computation of the approximate solution on $\Omega \setminus \Omega_{1,h}$, and for "transferring" the boundary-value problem from $\partial\Omega$ to $\partial\Omega_{1,h}$. Such a composite method is expected to enjoy the advantages of both FEM and wavelet methods: a more economical and flexible fitting of $\partial\Omega$ and a better spatial adaptivity in the interior of Ω, respectively. Besides, it can be used not only for Lipschitz-graph, but also for cuspidal domains. In this context, the use of near-degenerate elements in $\Omega \setminus \Omega_{1,h}$ and lacunary wavelet expansions in $\Omega_{1,h}$ can prove to be very helpful during the mesh refinement.

In this example, FEM and wavelet methods are used in their "pure" form, and the results are "glued together" at $\partial\Omega_{1,h}$. It is possible instead to try to use one of the two types of methods as a fundament for a hybrid method having also some of the properties of the other type. Examples of this approach, when FEM are chosen as the fundament, are discussed in [8], subsection 6.6. An interesting and far going example of this approach, when multiresolution methods are chosen as the fundament, and which also takes near-degeneracy into account, is provided by the work of Donoho and Candès on curvelets.

Acknowledgments. Supported in part by the Natural Sciences and Engineering Research Council of Canada and by the Priority Programme "Boundary Element Methods" of the German Research Foundation.

First and most of all, I would like to acknowledge the crucial importance of my collaboration with Wolfgang Wendland, to whom much of the credit for writing this paper is due. I also thank him for his great patience to wait until I was working on my part of our joint paper [13], which is now to be finished soon. My only excuse for the delay is that the problem turned out to be not only very important for applications, but also very hard to study in full depth.

I had the chance to benefit from the proficient expertise of Michal Křížek, in a valuable discussion and from his magnificent books [27,28]. From him I learned much about the use of FEM for solving nonlinear equations, and about the challenges of the utilization of near-degenerate finite elements.

Upon my request, Wolfgang Dahmen kindly sent me some of his recent papers on multiresolution methods, and they proved to be of key importance for the understanding of near-degeneracy "from the wavelet side". His moral support and understanding of the seriousness of the topic are very much appreciated.

Ron DeVore who, together with Vasil Popov and Pentcho Petrushev, has been the founder of the theory of best N-term approximation, has also had the most important personal contribution to the further development and applications of this theory. Several recent papers authored and co-authored by him, of which I explicitly emphasize here on the work of Cohen, Dahmen and DeVore [7], have contributed in a very essential way to my knowledge about wavelet approximation, and have thus helped me to successfully complete my present work.

The kind interest of Larry Schumaker in our results on the Bramble–Hilbert lemma and his valuable comments concerning the Bernstein–Bézier interpolants are highly appreciated.

I also appreciate very much the attention, sympathy and encouragement of Vidar Thomée and Ian Sloan.

To all the above-named people, I express my deep and cordial gratitude.

References

1. Adams, R. A., *Sobolev Spaces*, Academic Press, New York, 1975.
2. Bergh, J., and J. Löfström, *Interpolation Spaces. An Introduction*, Springer, Berlin, 1976.
3. Bos, L. P., and P. D. Milman, On Markov and Sobolev type inequalities on compact sets in $\mathbb{R}^n$, in *Topics in Polynomials of One and Several Variables and their Applications*, Th. Rassias (ed.), World Scientific Publishing, River Edge, N. J., 1993.
4. Bramble, J., and S. R. Hilbert, Estimation of linear functionals on Sobolev spaces with applications to Fourier transforms and spline interpolation, SIAM J. Numer. Anal. **8** (1970), 112–124.
5. Bramble, J., and S. R. Hilbert, Bounds for a class of linear functionals with application to Hermite interpolation, Numer. Math. **16** (1971), 362–369.
6. Chen, G., and J. Zhou, *Boundary Element Methods*, Academic Press, London, 1992.
7. Cohen, A., W. Dahmen, and R. A. DeVore, Multiscale decompositions on bounded domains, Trans. Amer. Math. Soc., to appear.
8. Dahmen, W., Wavelet and multiscale methods for operator equations, Acta Numer. (1997), 55–228.
9. Dechevski, L. T., Network-norm error estimates of the numerical solution of evolutionary equations, Serdica **12** (1986), 53–64.
10. Dechevski, L. T., τ-moduli and interpolation, in *Theory of Function Spaces and Applications*, M. Cwikel, J. Peetre, Y. Sagher and H. Wallin (eds.), Springer, Berlin, 1988, 177–190.
11. Dechevski, L. T., *Ph. D. Dissertation*, V. A. Popov (advisor), Sofia University, Sofia, 1988.
12. Dechevski, L. T., and E. Quak, On the Bramble–Hilbert lemma, Numer. Func. Anal. Optim. **11** (1990), 485–495.
13. Dechevski, L. T., and W. L. Wendland, On the Bramble–Hilbert lemma II, in preparation.
14. Dechevsky, L. T., Atomic decomposition of function spaces and fractional integral and differential operators, preprint.
15. Dechevsky, L. T., and J. O. Ramsay, Penalized wavelet estimation with Besov regularity constraints. Research report CRM-2602, Université de Montréal, 1999.
16. DeVore, R. A., Nonlinear approximation, Acta Numer. (1998), 51–150.
17. DeVore, R. A., and G. G. Lorentz, *Constructive Approximation*, Springer, Berlin, 1993.

18. Dupont, T., and R. Scott, Constructive polynomial approximation in Sobolev spaces, in *Recent Advances in Numerical Analysis*, C. de Boor and G. Golub (eds.), Academic Press, New York, 1978.
19. Dupont, T., and R. Scott, Polynomial approximation of functions in Sobolev spaces, Math. Comp. **34** (1980), 441–463.
20. Durán, R., On polynomial approximation in Sobolev spaces, SIAM J. Numer. Anal. **10** (1983), 985–988.
21. Farin, G., *Curves and Surfaces for Computer Aided Geometric Design*, Academic Press, San Diego, 1993.
22. Feichtinger, H.-G., Amalgam spaces and generalized harmonic analysis, Proc. Sympos. Appl. Math. **52** (1997), 141–150.
23. Johnen, H., and K. Scherer, On the equivalence of the K-functional and moduli of continuity and some applications, in *Constructive Theory of Functions of Several Variables*, W. Schempp and K. Zeller (eds.), Springer, Berlin, 1977, 119–140.
24. Jonsson, A., and H. Wallin, *Function Spaces on Subsets of* $\mathbb{R}^n$, Harwood, London, 1984.
25. Korotov, S., and M. Křížek, Finite element analysis of variational crimes for a quasilinear elliptic problem in 3D, Numer. Math., to appear.
26. Křížek, M., On semiregular families of triangulations and linear interpolation, Applic. Math. **36** (1991), 223–232.
27. Křížek, M. and P. Neittaanmäki, *Mathematical and Numerical Modelling in Electrical Engineering Theory and Applications*, Kluwer, Dordrecht, 1996.
28. Křížek, M. and P. Neittaanmäki, *Finite Element Approximation of Variational Problems and Applications*, Longman, New York, 1990.
29. Petrushev, P. P., and V. A. Popov, *Rational Approximation of Real Functions*, Cambridge University Press, Cambridge, 1987.
30. Samarskii, A. A., R. D. Lazarov, and V. L. Makarov, *Difference Schemes for Differential Equations with Generalized Solutions*, Vysshaya Shkola, Moscow, 1987, in Russian.
31. Schatz, A. H., V. Thomée, and W. L. Wendland, *Mathematical Theory of Finite and Boundary Element Methods*, Birkhäuser, Basel, 1990.
32. Sendov, Bl., and V. A. Popov, *The Averaged Moduli of Smoothness. Applications in Numerical Methods and Approximation*, Wiley, Chichester, 1988.
33. Sickel, W., Spline representations of functions in Besov–Triebel–Lizorkin spaces on $\mathbb{R}^n$, Forum Math. **2** (1990), 451–475.
34. Stein, E., *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, N. J., 1970.
35. Vainberg, M. M., *Variational Method and Method of Monotone Operators in the Theory of Nonlinear Equations*, Halsted Press, New York – Toronto, 1973.