

Nonlinear Approximation and Adaptive Techniques for Solving Elliptic Operator Equations

Stephan Dahlke and Wolfgang Dahmen
Institut für Geometrie
und Praktische Mathematik
RWTH Aachen
Templergraben 55
52056 Aachen
Germany

Ronald A. DeVore
Department of Mathematics
University of South Carolina
Columbia, S.C. 29208
USA

Abstract

This survey article is concerned with two basic approximation concepts and their interrelation with the numerical solution of elliptic operator equations, namely nonlinear and adaptive approximation. On one hand, for nonlinear approximation based on wavelet expansions the best possible approximation rate, which a function can have for a given number of degrees of freedom, is characterized in terms of its regularity in a certain scale of Besov spaces. Therefore, after demonstrating the gain of nonlinear approximation over linear approximation measured in a Sobolev scale, we review some recent results on the Sobolev and Besov regularity of solutions to elliptic boundary value

problems. On the other hand, nonlinear approximation requires information that is generally not available in practice. Instead one has to resort to the concept of adaptive approximation. We briefly summarize some recent results on wavelet based adaptive schemes for elliptic operator equations. In contrast to more conventional approaches one can show that these schemes converge without prior assumptions on the solution such as the saturation property. One central objective of this paper is to contribute to interrelating nonlinear approximation and adaptive methods in the context of elliptic operator equations.

Key Words: Nonlinear approximation, adaptive methods, elliptic equations, wavelets, characterization of function spaces.

AMS subject classification: 35B65, 41A46, 46E35, 65N12, 65N30, 65N55.

1 Introduction

Adaptive methods are an important tool for numerically solving elliptic equations. Their origins appear in the adaptive grid refinements in Finite Element Methods. Heuristically, adaptive methods are effective when trying to recover solutions u which exhibit singularities. Much impetus to the theory of adaptive finite element methods was provided by the introduction and analysis of **p** and **h-p** methods by Babuska and his collaborators (see e.g. [2] and [3]). A lot of further results on this subject have been developed in the last years, see e.g. [4, 33, 52]. For an overview on the theory of adaptive finite elements, the reader is referred to [28] and [53]. On the other hand, for most adaptive algorithms, there exist no proofs of convergence.

The purpose of this paper is to phrase the problem of designing and analyzing adaptive methods in the context of approximation of functions. In this way, we shall introduce some analytical tools which may prove useful for constructing and analyzing adaptive algorithms. In particular, we shall utilize heavily the theory of multilevel methods and wavelets.

An outline of this paper is as follows. In §2, we introduce the elliptic problems that we shall consider. They take the form

$$Au = f \tag{1.1}$$

where A is a symmetric positive definite operator which is boundedly invertible on some Sobolev space. Thus, (1.1) includes both integral equations and boundary value problems.

In §3, we briefly recall the theory of Sobolev and Besov spaces which we shall need for this paper. While these spaces have their classical definitions in terms of derivatives and smoothness, we introduce them from the viewpoint of wavelet decompositions. This gives a simple criterion for membership in these spaces in terms of certain sequence norms applied to the wavelet coefficients. This also gives us an opportunity to introduce wavelet decompositions for various types of domains which will be important for both our numerical and analytic considerations. We conclude this section with summarizing some relevant facts about wavelet discretizations of (1.1).

In §4,5, we discuss some approximation concepts which are relevant for numerically approximating the solution u to (1.1). We draw distinctions between two cases:

- *Linear methods* where the approximation comes from a linear space.
- *Nonlinear methods* where the approximation takes place from a nonlinear set.

Adaptive methods are a form of nonlinear approximation. In these sections, we shall discuss the smoothness required of a function v in order that it can be approximated with a certain efficiency by linear or nonlinear methods. This is important, vis a vis (1.1), since it tells us the smoothness (regularity) the solution u must have in order for it to be approximated with a given efficiency.

There will be two scales of regularity: one for linear methods, the other for nonlinear methods. The regularity for linear methods is given by the usual scale H^α of Sobolev spaces. This regularity is well-known and often used in error estimates for Finite Element Methods. The corresponding regularity for nonlinear methods takes place in a certain scale of Besov spaces. This scale of spaces does not seem to be as fully understood in the Finite Element literature. This type of regularity needs to be used (in place of Sobolev regularity) when analyzing nonlinear methods such as adaptive finite elements. In particular, this type of regularity needs to be kept in mind when constructing adaptive numerical methods and in analyzing their performance (establishing error estimates).

In §6, we discuss the regularity of solutions to (1.1) from the viewpoint of the Sobolev and Besov scales noted above. The regularity of the solution tells the maximum efficiency a numerical method can achieve.

However, the problem still exists to construct such numerical methods and establish their convergence and error estimates in specific settings. In §7, we discuss how this might be accomplished in the context of wavelet decompositions.

2 A class of elliptic problems

In this section, we shall introduce the generic elliptic problems which we wish to analyze. We shall consider the model case of a linear operator equation

$$Au = f \tag{2.2}$$

where $A : H \rightarrow H^*$ is a boundedly invertible operator from some Hilbert space H into its dual H^* , i.e.,

$$\|Av\|_{H^*} \asymp \|v\|_H, \quad v \in H. \tag{2.3}$$

We use the notation $A \asymp B$ to express that $A \lesssim B$ and $B \lesssim A$. Here $A \lesssim B$ means that $A \leq CB$ for some constant C . In the case that A and B depend on parameters or variables the constant C is to be independent of these parameters or variables unless explicitly stated otherwise.

To simplify the exposition we will confine the discussion to selfadjoint operators A , i.e.,

$$a(u, v) := \langle Au, v \rangle, \tag{2.4}$$

is a bilinear symmetric form and (2.3) means that the energy norm

$$\|u\| = a(u, u)^{\frac{1}{2}} \tag{2.5}$$

satisfies

$$\|\cdot\| \asymp \|\cdot\|_H. \tag{2.6}$$

In variational (weak) form the solution u to (2.2) is a function $u \in H$ which satisfies

$$a(u, v) = \langle f, v \rangle, \quad v \in H. \tag{2.7}$$

The following examples indicate the scope of problems we have in mind.

Suppose that $\Omega \subset \mathbb{R}^d$ is a domain. We shall always assume that Ω is a bounded, open, and connected Lipschitz domain. This covers all domains of practical interest. If k is positive integer, the Sobolev space $H^k(\Omega) := W^k(L_2(\Omega))$ consists of all functions $f \in L_2(\Omega)$ whose distributional derivatives $D^\nu f$, $|\nu| = k$, satisfy

$$|f|_{H^k(\Omega)}^2 := \sum_{|\nu|=k} \|D^\nu f\|_{L_2(\Omega)}^2 \quad (2.8)$$

is finite. The square root of (2.8) is the semi-norm for $H^k(\Omega)$ and adding to it $\|f\|_{L_2(\Omega)}$ gives the norm $\|f\|_{H^k(\Omega)}$ in $H^k(\Omega)$. The spaces $H^\alpha(\Omega)$ can then be defined for non-integer values $\alpha \geq 0$ in several equivalent ways, for example, using either interpolation theory, or viewing them as special cases of Besov spaces. The Besov spaces on domains are defined using moduli of smoothness (see e.g. [25]). For negative α one can employ duality.

Simple examples for A are in this case $Au = -\Delta u$ or $Au = -\Delta u + cu$ where $\Delta = \sum_{j=1}^d \frac{\partial^2}{\partial x_j^2}$ is the Laplacian and $c > 0$. Here $H = H_0^1(\Omega)$ or $H = H^1(\Omega)$, respectively, where $H_0^m(\Omega)$ is the closure of $C_0^\infty(\Omega)$ with respect to the $\|\cdot\|_{H^m}$ -norm. Of course, when A is the biharmonic operator one has $H = H_0^2(\Omega)$ or similar spaces incorporating mixed boundary conditions.

The second example represents a different type of problems which still fits into the present framework. In order to solve an exterior boundary value problem

$$\Delta U = 0 \quad \text{in } \mathbb{R}^3 \setminus \Omega, \quad U = f \quad \text{on } \partial\Omega, \quad (2.9)$$

it is tempting to transform this problem into a boundary integral equation. For instance, the *indirect method* leads to the equation on the boundary

$$Au = f, \quad A = I + 2K \quad (2.10)$$

where K involves the *double layer potential* operator

$$Ku(x) := [1/2 - \theta_\Omega(x)]u(x) + \frac{1}{4\pi} \int_{\partial\Omega} \frac{n_y \cdot (x - y)}{|y - x|^3} u(y) dy \quad (2.11)$$

and $\theta_\Omega(x)$ denotes the interior angle at $x \in \partial\Omega$ located on an edge of $\partial\Omega$. In this case it is known that $H = L_2(\Gamma)$, $\Gamma := \partial\Omega$. Solving (2.10) for u leads to the solution of (2.9) by evaluating a singular integral.

Alternatively, the direct method yields an integral equation of the first kind

$$Au = Vu = \left(\frac{1}{2} + K\right)f,$$

where V is the *single layer potential* operator

$$Vu(x) = \int_{\partial\Omega} \frac{u(y)}{4\pi|x-y|} dy.$$

It is known that this fits into the above framework with $H = H^{-1/2}(\Gamma)$, $H^* = H^{1/2}(\Gamma)$. Of course the latter context requires a proper definition of Sobolev spaces on surfaces or manifolds. This is usually done by “lifting” Sobolev spaces from domains in \mathbb{R}^d with the aid of an atlas and a partition of unity (see e.g. [9, 19]). We shall comment on this issue later in more detail.

Thus in what follows H typically stands for a Sobolev space $H^t = H^t(\Omega)$ (or some subspace determined by boundary conditions) while for $t \geq 0$, H^{-t} is to be understood as the dual of H^t .

We conclude pointing out one more property of the class of operators under consideration. Note that in the above examples A has support of measure zero or has a Schwarz kernel with certain asymptotic properties. More precisely, we will assume that

$$(Av)(x) = \int_{\Omega} \mathcal{K}(x, y)v(y)dy,$$

where $\mathcal{K}(x, y)$ is smooth off the diagonal $x = y$ and where we require that whenever $d + \rho + |\mu| + |\nu| > 0$

$$|\partial_x^\mu \partial_y^\nu \mathcal{K}(x, y)| \leq c_{\mu, \nu} \text{dist}(x, y)^{-(d+\rho+|\mu|+|\nu|)} \quad (2.12)$$

holds with constants $c_{\mu, \nu}$ depending only on the multi-indices $\mu, \nu \in \mathbb{Z}_+^d$. Estimates of the type (2.12) are known to hold for a wide range of cases including classical pseudo-differential operators and Calderón-Zygmund operators (see e.g. [18, 50]). Thus the single and double layer potential operator above as well as classical differential operators fall into this category.

3 Smoothness spaces and wavelet decompositions

We wish to treat the above type of problems by means of wavelet methods. To this end, we have to explain first what is meant by wavelets defined on the various types of domains appearing in the previous section and how wavelet expansions are related to smoothness spaces on such domains. The simplest setting is $\Omega = \mathbb{R}^d$. Although this is of limited use for problems of the above nature it is particularly well suited for bringing out the essentials of wavelet analysis and serves as a core ingredient for the construction of wavelets on other domains.

3.1 Wavelets on Euclidean space

We begin by discussing smoothness spaces on \mathbb{R}^d . As above, if k is positive integer, the Sobolev space $H^k(\mathbb{R}^d) := W^k(L_2(\mathbb{R}^d))$ is defined as before with $\Omega = \mathbb{R}^d$, compare with (2.8). The Sobolev spaces $H^\alpha(\mathbb{R}^d)$ for other values of $\alpha \in \mathbb{R}$ are usually defined by Fourier transforms. In particular, $H^0(\mathbb{R}^d) = L_2(\mathbb{R}^d)$ and $H^{-\alpha}(\mathbb{R}^d)$ is the dual space of $H^\alpha(\mathbb{R}^d)$.

There is an equivalent definition of the Sobolev spaces in terms of wavelet decompositions which is of primary importance in the present context and which we now describe. Let \mathcal{D} denote the set of all dyadic cubes in \mathbb{R}^d and let \mathcal{D}_j be the collection of all dyadic cubes at level j . Then, $I \in \mathcal{D}_j$ if and only if $I = 2^{-j}(k + [0, 1]^d)$, for some $j \in \mathbb{Z}$, $k \in \mathbb{Z}^d$. If $\eta \in L_2(\mathbb{R}^d)$, we define

$$\eta_I(x) := |I|^{-1/2} \eta(2^j x - k), \quad I = 2^{-j}(k + [0, 1]^d). \quad (3.1.13)$$

Then η_I is a scaled, shifted, dilate of η and $\|\eta_I\|_{L_2(\mathbb{R}^d)} = \|\eta\|_{L_2(\mathbb{R}^d)}$ for all $I \in \mathcal{D}$.

We begin our discussion of wavelet bases with the biorthogonal bases of compactly supported basis functions. We recall briefly how such bases are constructed from multiresolution analysis. A function η is said to satisfy a *refinement equation* (sometimes called a *two scale relation*) if

$$\eta(x) = \sum_{k \in \mathbb{Z}^d} a_k \eta(2x - k). \quad (3.1.14)$$

We shall only deal with compactly supported functions η . In this case, only a finite number of the coefficients a_k are nonzero.

The starting point for constructing biorthogonal wavelets is a pair of univariate functions φ and $\tilde{\varphi}$ of compact support which each satisfy a refinement equation and are in duality

$$\int_{\mathbb{R}} \varphi(x)\tilde{\varphi}(x-k) = \delta(k), \quad k \in \mathbb{Z}, \quad (3.1.15)$$

with δ the Kronecker delta function on \mathbb{Z} . From φ and $\tilde{\varphi}$ we construct the multivariate functions $\phi(x) := \varphi(x_1)\cdots\varphi(x_d)$ and $\tilde{\phi}(x) := \tilde{\varphi}(x_1)\cdots\tilde{\varphi}(x_d)$ which are also in duality:

$$\int_{\mathbb{R}^d} \phi(x)\tilde{\phi}(x-k) = \delta(k), \quad k \in \mathbb{Z}^d. \quad (3.1.16)$$

Let V_0 be the $L_2(\mathbb{R}^d)$ closure of the linear span of the shifts $\phi(\cdot - k)$, $k \in \mathbb{Z}^d$, of ϕ . From the existence of a compactly supported dual function $\tilde{\phi}$ it follows that the functions $\phi(\cdot - k)$, $k \in \mathbb{Z}^d$, are a Riesz basis for V_0 and each element $v \in V_0$ has the representation

$$v = \sum_{k \in \mathbb{Z}^d} \langle v, \tilde{\phi}(\cdot - k) \rangle \phi(\cdot - k), \quad (3.1.17)$$

with

$$\langle f, g \rangle = \langle f, g \rangle_{L_2(\mathbb{R}^d)} := \int_{\mathbb{R}^d} f(x)\bar{g}(x) dx$$

the inner product for $L_2(\mathbb{R}^d)$.

By dilation, we obtain the spaces

$$V_j := \{v(2^j \cdot) : v \in V_0\}. \quad (3.1.18)$$

Then V_j is spanned by the functions ϕ_I , $I \in \mathcal{D}_j$, and each $v \in V_j$ has the representation

$$v = \sum_{I \in \mathcal{D}_j} \langle v, \tilde{\phi}_I \rangle \phi_I. \quad (3.1.19)$$

Because ϕ satisfies a refinement equation, the spaces V_j are nested

$$V_j \subset V_{j+1}, \quad j \in \mathbb{Z}. \quad (3.1.20)$$

It follows [7] that $\cup_j V_j$ is dense in $L_2(\mathbb{R}^d)$.

Let P_j be the projector from $L_2(\mathbb{R}^d)$ onto V_j given by

$$P_j f = \sum_{I \in \mathcal{D}_j} \langle f, \tilde{\phi}_I \rangle \phi_I. \quad (3.1.21)$$

The projectors P_j are uniformly bounded on $L_2(\mathbb{R}^d)$ and for each $f \in L_2(\mathbb{R}^d)$,

$$\|f - P_j f\|_{L_2(\mathbb{R}^d)} \rightarrow 0. \quad (3.1.22)$$

The operators $Q_j := P_{j+1} - P_j$ are also uniformly bounded on $L_2(\mathbb{R}^d)$ and their range $W_j \subset L_2(\mathbb{R}^d)$ is called a *wavelet space*. The spaces W_j inherit the same structure as the V_j . For example, each W_j is the dilate (by 2^j) of W_0 . Also, W_0 is a shift invariant space generated by a set Ψ° of $2^d - 1$ functions ψ . That is, W_0 is the closed linear span of the functions $\psi(\cdot - k)$, $k \in \mathbb{Z}^d$, $\psi \in \Psi^\circ$. Moreover, there is a dual set $\tilde{\Psi}^\circ$ of $2^d - 1$ functions $\tilde{\psi}$ such that

$$\langle \psi(\cdot - k), \tilde{\eta}(\cdot - j) \rangle = \delta(k - j) \delta(\psi - \eta), \quad j, k \in \mathbb{Z}^d, \psi, \eta \in \Psi^\circ. \quad (3.1.23)$$

Now suppose that $\phi, \tilde{\phi}$ are *exact* of order N, \tilde{N} , respectively, i.e.,

$$x^\alpha = \sum_{k \in \mathbb{Z}^d} \langle (\cdot)^\alpha, \tilde{\phi}(\cdot - k) \rangle \phi(x - k), \quad x \in \mathbb{R}^d, |\alpha| < N, \quad (3.1.24)$$

$$x^\alpha = \sum_{k \in \mathbb{Z}^d} \langle (\cdot)^\alpha, \phi(\cdot - k) \rangle \tilde{\phi}(x - k), \quad x \in \mathbb{R}^d, |\alpha| < \tilde{N},$$

where $|\alpha|$ denotes the sum of the components of the multi-indices $\alpha \in \mathbb{Z}_+^d$. An immediate consequence of (3.1.23) and (3.1.24) is that the $\psi \in \Psi^\circ$ have *vanishing moments*, i.e.,

$$\int_{\mathbb{R}^d} P(x) \psi(x) dx = 0, \quad \psi \in \Psi^\circ, \quad (3.1.25)$$

for all polynomials P of coordinate degree less than \tilde{N} . An analogous statement holds for $\tilde{\psi}$ and \tilde{N} replaced by ψ and N , respectively. It is known that for any $N, \tilde{N} \in \mathbb{N}$ such that $N + \tilde{N}$ is even there exist compactly supported dual pairs $\varphi, \tilde{\varphi}$ whose order of polynomial exactness is N, \tilde{N} , respectively [11]. Thus the biorthogonal setting not only offers more flexibility in constructing compactly supported wavelets, where all filters have finite support and therefore give rise to fast reconstruction and decomposition algorithms,

but also allows one to construct wavelets with an arbitrarily high number of vanishing moments which is crucial for treating integral equations.

It follows that each function in $L_2(\mathbb{R}^d)$ has the wavelet decomposition

$$f = \sum_{I \in \mathcal{D}} \sum_{\psi \in \Psi^\circ} \langle f, \tilde{\psi}_I \rangle \psi_I. \quad (3.1.26)$$

An alternative wavelet decomposition and the one preferred in numerical considerations starts at a finite dyadic level (which we, for notational convenience, will take as level 0). Then, for each $f \in L_2(\mathbb{R}^d)$, we have

$$\begin{aligned} f &= P_0(f) + \sum_{I \in \mathcal{D}^+} \sum_{\psi \in \Psi^\circ} \langle f, \tilde{\psi}_I \rangle \psi_I \\ &= \sum_{k \in \mathbb{Z}^d} \langle f, \tilde{\phi}(\cdot - k) \rangle \phi(\cdot - k) + \sum_{I \in \mathcal{D}^+} \sum_{\psi \in \Psi^\circ} \langle f, \tilde{\psi}_I \rangle \psi_I \end{aligned} \quad (3.1.27)$$

with \mathcal{D}^+ the set of dyadic cubes with measure ≤ 1 .

The set of functions $\{\psi_I\}_{I \in \mathcal{D}, \psi \in \Psi^\circ}$ is a Riesz basis for $L_2(\mathbb{R}^d)$ (sometimes called a stable basis). This means that

$$\|f\|_{L_2(\mathbb{R}^d)}^2 \asymp \sum_{I \in \mathcal{D}} \sum_{\psi \in \Psi^\circ} |\langle f, \tilde{\psi}_I \rangle|^2. \quad (3.1.28)$$

As mentioned earlier the constants in (3.1.28) are independent of f and we will use similar notation throughout this paper.

The set $\{\psi_I\}_{I \in \mathcal{D}, \psi \in \Psi^\circ}$ is also a Riesz basis for H^α , and

$$\|f\|_{H^\alpha(\mathbb{R}^d)}^2 \asymp \sum_{I \in \mathcal{D}} \sum_{\psi \in \Psi^\circ} |I|^{-2\alpha/d} |\langle f, \tilde{\psi}_I \rangle|^2. \quad (3.1.29)$$

for a certain range $-\tilde{\gamma} < \alpha < \gamma$ that depends on the sets Ψ° and $\tilde{\Psi}^\circ$.

The equivalence (3.1.29) gives us a simple way to compute equivalent H^α norms in terms of wavelet coefficients. Also note that in this form, we see that we scale up or scale down the Sobolev spaces by simply multiplying wavelet coefficients by $|I|^{s/d}$. Namely, let

$$\mathcal{I}_s(f) := \sum_{I \in \mathcal{D}} \sum_{\psi \in \Psi^\circ} |I|^{s/d} \langle f, \tilde{\psi}_I \rangle \psi_I. \quad (3.1.30)$$

Then $\mathcal{I}_s(f) \in H^{\alpha+s}(\mathbb{R}^d)$ if and only if $f \in H^\alpha(\mathbb{R}^d)$, $s + \alpha \in (-\tilde{\gamma}, \gamma)$ and

$$\|\mathcal{I}_s f\|_{H^{\alpha+s}(\mathbb{R}^d)} \asymp \|f\|_{H^\alpha(\mathbb{R}^d)}. \quad (3.1.31)$$

The Besov spaces $B_q^\alpha(L_p(\mathbb{R}^d))$, $\alpha > 0$, $0 < p, q \leq \infty$ are smoothness spaces in $L_p(\mathbb{R}^d)$. The index α is the primary index and gives the order of smoothness (analogous to the number of derivatives). The second parameter q gives a finer scaling. For example, the space $B_2^\alpha(L_2(\mathbb{R}^d)) = H^\alpha(\mathbb{R}^d)$ and $B_\infty^\alpha(L_p(\mathbb{R}^d))$ is the Lipschitz space of order α in $L_p(\mathbb{R}^d)$ provided $\alpha > 0$ is not an integer.

Besov spaces are usually defined by Fourier transforms or moduli of smoothness (see e.g. [25]). There is, however, an equivalent definition in terms of wavelet decompositions which we shall employ. In fact, for the above range of α one has

$$\|f\|_{B_q^\alpha(L_p(\mathbb{R}^d))}^q \asymp \sum_{I \in \mathcal{D}} \sum_{\psi \in \Psi^\circ} |I|^{-q(\alpha/d + 1/2 - 1/p)} |\langle f, \tilde{\psi}_I \rangle|^q < \infty, \quad (3.1.32)$$

whenever $f \in B_q^\alpha(L_p(\mathbb{R}^d))$. We shall mainly be concerned with a particular scale of these spaces which will replace the role of the Sobolev spaces when treating nonlinear approximation. If $\alpha \geq 0$, we let

$$\tau := (\alpha/d + 1/2)^{-1} \quad (3.1.33)$$

so that $\tau \leq 2$. Then the Besov space $B^\alpha(\mathbb{R}^d) := B_\tau^\alpha(L_\tau(\mathbb{R}^d))$ is the set of all function in $L_2(\mathbb{R}^d)$ which have a wavelet decomposition (3.1.26) and the wavelet coefficients of f satisfy

$$\|f\|_{B^\alpha(\mathbb{R}^d)}^\tau \asymp \sum_{I \in \mathcal{D}} \sum_{\psi \in \Psi^\circ} |\langle f, \tilde{\psi}_I \rangle|^\tau < \infty. \quad (3.1.34)$$

Then, (3.1.34) gives an equivalent quasi-norm for B^α . Note that $B^0(\mathbb{R}^d) = L_2(\mathbb{R}^d)$ with equivalent norms. As α gets larger, the spaces $B^\alpha(\mathbb{R}^d)$ get smaller: $B^\alpha(\mathbb{R}^d) \subset B^\beta(\mathbb{R}^d)$, $\alpha \geq \beta$.

Wavelets of the above type are still of limited use for the numerical treatment of operator equations. Below we will briefly indicate how to obtain wavelet bases with the above properties (3.1.28), (3.1.29), (3.1.34) in several other cases of practical relevance.

3.2 Wavelets on the interval

As the simplest example of a bounded domain let us consider first $\Omega = [0, 1]$. This case deserves particular attention because it will also serve as a core ingredient of constructions for more complex situations.

The common starting point (see e.g. [1, 8, 12, 17]) is to construct collections $\Phi_k = \{\phi_{k,m} : m \in \Delta_k\} \subset L_2([0, 1])$ such that the spaces $S_k := \text{span } \Phi_k$ are nested and contain all polynomials up to a certain desired degree. Taking some dual pair $\varphi, \tilde{\varphi}$ as in (3.1.15) and fixing ℓ such that for $k > k_0$ we have $\text{supp } \varphi(2^k \cdot -m), \text{supp } \tilde{\varphi}(2^k \cdot -m) \subset (0, 1)$, $m = \ell, \dots, 2^k - \ell$, the collections Φ_k are comprised of these *interior* translates $2^{k/2}\varphi(2^k \cdot -m)$ together with certain *boundary* functions which are needed to preserve the desired degree of polynomial exactness. If φ has exactness order N these boundary functions are simply obtained by truncating the expansions (3.1.24). For instance, for the left end of the interval one adds the N functions

$$\varphi_{k,\ell-n+r}^L(x) := \sum_{m=-\infty}^{\ell-1} \langle (\cdot)^r, \tilde{\varphi}(\cdot - m) \rangle 2^{k/2} \varphi(2^k \cdot -m) |_{[0,1]}, \quad (3.2.35)$$

$r = 0, \dots, N - 1$, and analogously at the right end. One easily infers from (3.1.24) that these boundary functions together with the interior translates reconstruct all polynomials up to degree $N - 1$ on $[0, 1]$. Thus the resulting spaces inherit the approximation properties of their shift-invariant counterparts defined on all of \mathbb{R} . Moreover, since $\varphi_{k,\ell-n+r}^L(x)$ behaves near 0 like $2^{k/2}x^r$, it is relatively easy to incorporate homogeneous boundary conditions.

The construction of dual collections $\tilde{\Phi}_k$ differs somewhat in the above mentioned papers. In [17], it is shown that also $\tilde{\Phi}_k$ can be made to be exact of order \tilde{N} in an analogous fashion and that the resulting sets can indeed be biorthogonalized, compare with (3.1.23) and (3.1.24). Let us denote the biorthogonalized sets again by $\Phi_k, \tilde{\Phi}_k$ with elements $\phi_{k,m}, \tilde{\phi}_{k,m}$. Except for a finite number of boundary functions the $\phi_{k,m}, \tilde{\phi}_{k,m}$ still have the form $2^{k/2}\eta(2^k \cdot -m)$, $\eta \in \{\varphi, \tilde{\varphi}\}$, respectively. Moreover, compactly supported biorthogonal wavelets $\psi_{k,m}, \tilde{\psi}_{k,m}$, $m = 1, \dots, 2^k$, are constructed which form Riesz bases for $L_2([0, 1])$.

Due to the modifications of the basis functions near the end points of the interval the simple recipe from (3.1.13) of taking translates of dilated functions is no longer applicable. Nevertheless, it will be convenient to accept the slight abuse of notation and still write ϕ_I, ψ_I . In fact, setting in this case $\mathcal{D}_c := \{k_0\} \times \{\ell - n, \dots, 2^{k_0} - \ell + n\}$, $\mathcal{D}_k := \{k\} \times \{1, \dots, 2^{k-1}\}$, $k > k_0$ we can still identify the indices (k, m) with dyadic cubes $I = 2^{-k}(m + [0, 1])$. Defining $\mathcal{D}^+ := \cup_{k > k_0} \mathcal{D}_k$ and $\mathcal{D} := \mathcal{D}_c \cup \mathcal{D}^+$, we obtain essentially the same

format as above:

$$f = \sum_{I \in \mathcal{D}_c} \langle f, \tilde{\phi}_I \rangle \phi_I + \sum_{(\psi, I) \in \Psi^\circ \times \mathcal{D}^+} \langle f, \tilde{\psi}_I \rangle \psi_I, \quad (3.2.36)$$

as well as

$$\|f\|_{H^\alpha([0,1])}^2 \asymp \sum_{I \in \mathcal{D}_c} |\langle f, \tilde{\phi}_I \rangle|^2 + \sum_{(\psi, I) \in \Psi^\circ \times \mathcal{D}^+} |I|^{-2\alpha} |\langle f, \tilde{\psi}_I \rangle|^2, \quad (3.2.37)$$

or

$$\|f\|_{B^\alpha([0,1])}^\tau \asymp \sum_{I \in \mathcal{D}_c} |\langle f, \tilde{\phi}_I \rangle|^\tau + \sum_{(\psi, I) \in \Psi^\circ \times \mathcal{D}^+} |\langle f, \tilde{\psi}_I \rangle|^\tau, \quad (3.2.38)$$

where $\tau := (\alpha + 1/2)^{-1}$. Of course, in this case one has $\#\Psi^\circ = 1$ but in anticipation of the tensor product case below this redundancy is useful. Also one should note, that for notational simplicity we have suppressed the fact that, due to boundary modifications, Ψ° actually depends on I . Again the range for which (3.2.37) is valid is $(-\tilde{\gamma}, \gamma)$ where

$$\gamma := \sup \{ \alpha : \varphi \in H^\alpha(\mathbb{R}) \}, \quad \tilde{\gamma} := \sup \{ \alpha : \tilde{\varphi} \in H^\alpha(\mathbb{R}) \}$$

(see [17]). The case $\alpha = 0$, of course, recovers the Riesz basis property.

Since by construction the spaces $\tilde{S}_k := \text{span } \tilde{\Phi}_k$ are exact of order \tilde{N} one has

$$\int_0^1 x^r \psi_I(x) dx = 0, \quad I \in \mathcal{D}^+, \quad r = 0, \dots, \tilde{N} - 1. \quad (3.2.39)$$

3.3 The isoparametric case

Taking tensor products of wavelets on $[0, 1]$ immediately yields biorthogonal wavelet bases on the unit d -cube $\square := [0, 1]^d$ with analogues of (3.2.37), (3.2.38), (3.2.39). One can push this line a little further in the following direction. Suppose that for some $d' \geq d$, κ is a regular mapping from $\mathbb{R}^{d'}$ into \mathbb{R}^d , i.e., κ is smooth and its Jacobian is bounded away from 0, and let $\Omega := \kappa(\square)$. Sobolev spaces or Besov spaces on Ω can be defined by *lifting* corresponding spaces from \square with the aid of κ . In fact,

$$\langle f, g \rangle := \int_{\square} f(\kappa(x)) g(\kappa(x)) |\det \kappa'(\kappa^{-1}(x))| dx$$

is a natural inner product which can be used to define Sobolev norms. On the other hand,

$$(f, g) := \int_{\square} f(\kappa(x))g(\kappa(x))dx \quad (3.3.40)$$

induces equivalent norms whenever κ is sufficiently regular. Taking tensor products of the above mentioned wavelets on the interval, readily yields biorthogonal wavelet bases $\Psi = \{\phi_I : I \in \mathcal{D}_c\} \cup \{\psi_I : \psi \in \Psi^\circ, I \in \mathcal{D}^+\}$, $\tilde{\Psi} = \{\tilde{\psi}_I : \psi_I \in \Psi\}$ on \square . Here we have used the convention $\psi_I := \phi_I$ for $I \in \mathcal{D}_c$. Of course, in this case one has $\#\Psi^\circ = 2^d - 1$ and the structure of the sets $\mathcal{D}_c, \mathcal{D}^+$ is clear from the tensor product construction. Then the collections

$$\Psi_\Omega := \{\psi_I^\Omega := \psi_I \circ \kappa^{-1} : \psi_I \in \Psi\}, \quad (3.3.41)$$

$$\tilde{\Psi}_\Omega := \{\tilde{\psi}_I^\Omega := \tilde{\psi}_I \circ \kappa^{-1} : \tilde{\psi}_I \in \tilde{\Psi}\},$$

are obviously biorthogonal relative to the inner product (\cdot, \cdot) in (3.3.40) which again satisfy (3.2.37) and (3.2.38). The moment conditions take the form

$$(P, \psi_I^\Omega) = 0, \quad (\psi, I) \in \Psi^\circ \times \mathcal{D}^+, \quad (3.3.42)$$

whenever $P \circ \kappa^{-1}$ is a polynomial of coordinate degree less than \tilde{N} . Here and in the following we reserve the notation \mathcal{D}_c for those dyadic cubes associated to the scaling functions on the coarsest level. The importance of this case will become clearer below.

3.4 Wavelets on manifolds

When $d = d' = 2$ the above construction yields for instance wavelet bases on various planar domains. However the case $d' > d$ is important too. In fact, the examples in Section 2 show that one needs wavelets defined on manifolds which are embedded in some higher dimensional Euclidean space.

The simplest case is the d -torus. Functions defined on the d -torus correspond in a one-to-one way to 1-periodic functions $f(x+m) = f(x)$, $m \in \mathbb{Z}^d$. Clearly every compactly supported function η in $L_2(\mathbb{R}^d)$ is easily periodized by

$$[\eta](x) := \sum_{k \in \mathbb{Z}^d} \eta(x+k). \quad (3.4.43)$$

Moreover this is easily seen to preserve orthogonality relations, i.e.,

$$\int_{\mathbb{R}^d} g(x)f(x)dx = 0 \quad \Longrightarrow \quad \int_{\square} [f](x)[g](x)dx = 0 .$$

Thus, given wavelets $\psi_I, \tilde{\psi}_I$ on \mathbb{R}^d , the functions $[\psi_I], [\tilde{\psi}_I]$ form corresponding wavelet bases on the d -torus. The ease of this construction is exploited in many papers.

Again the case $d = 1$, the circle, deserves special attention. Suppose \mathcal{C} is any smooth closed curve (without selfintersection) in \mathbb{R}^2 . Then \mathcal{C} can be written as a parametric image $\mathcal{C} = \kappa([0, 1])$ of a smooth 1-periodic mapping κ . Thus combining periodization with the isoparametric approach from Section 3.3 immediately provides wavelet bases on \mathcal{C} giving rise to analogous norm equivalences and moment conditions. These wavelets can be used to discretize for instance boundary integral equations of the type mentioned in Section 2 arising from exterior boundary value problems for planar domains with smooth boundaries.

When the curve is not smooth but has corners it may have to be subdivided into smooth sections and wavelet bases can be obtained by piecing together parametric images of wavelets on the interval. This gives stable bases for L_2 . However, for the characterization of smoothness spaces, this is not sufficient. Here the transition between adjacent segments requires special care. We will briefly indicate a systematic approach to this problem below in the context of a more general situation.

Note that example (2.11) requires wavelets defined on two-dimensional closed surfaces in \mathbb{R}^3 . In such a case periodization does not help. Instead one can use the tools developed in Computer Aided Geometric Design where such surfaces are modeled as a union of parametric patches. Thus assume that Γ is a piecewise smooth d -dimensional manifold of the form

$$\Gamma = \bigcup_{i=1}^M \bar{\Gamma}_i, \quad \Gamma_i \cap \Gamma_l = \emptyset, \quad i \neq l, \quad (3.4.44)$$

where $\Gamma_i = \kappa_i(\square)$ and κ_i are regular sufficiently smooth parametrizations. Again one can consider function spaces $\mathcal{F}(\Gamma)$ where $\mathcal{F}(\Gamma) = H^\alpha(\Gamma)$ or $\mathcal{F}(\Gamma) = B_q^\alpha(L_p(\Gamma))$ and the range of α depends on the global regularity of Γ . For instance, when Γ is at least Lipschitz it makes sense to consider Sobolev spaces

with index $\alpha < 1$. For practical purposes and for the sake of constructing wavelets on Γ the characterization of \mathcal{F} via an atlas and a partition of unity is rather useless. An interesting alternative was offered in [9] where a characterization of $\mathcal{F}(\Gamma)$ is directly based on a decomposition of Γ into patches Γ_i . The following brief indication of the basic ideas is taken from [19] where an attempt is made to make the existence statements from [9] constructive and where the details of the following comments are given. First one orders the patches Γ_i in a certain fashion. If $\bar{\Gamma}_i \cap \bar{\Gamma}_l := \epsilon_{i,l}$ is a common face and $i < l$, then $\epsilon_{i,l}$ is called an *outflow* (*inflow*) face for Γ_i (Γ_l). $\partial\Gamma_i^\uparrow, \partial\Gamma_i^\downarrow$ are called the *outflow* and *inflow* boundary of the patch Γ_i . Let Γ_i^\uparrow denote an extension of Γ_i in Γ which contains the outflow boundary $\partial\Gamma_i^\uparrow$ in its relative interior and whose boundary contains the inflow boundary $\partial\Gamma_i^\downarrow$ of Γ_i . Now suppose that E_i is an extension operator from the domain Γ_i to Γ_i^\uparrow such that

$$\|E_i f\|_{\mathcal{F}(\Gamma_i^\uparrow)} \lesssim \|f\|_{\mathcal{F}(\Gamma_i)}, \quad \|(E_i^* f)^\uparrow\|_{\mathcal{F}(\Gamma_i)^\uparrow} \lesssim \|f\|_{\mathcal{F}(\Gamma_i^\uparrow)}, \quad (3.4.45)$$

where

$$f^\uparrow(x) := \begin{cases} f(x), & x \in \Gamma_i, \\ 0, & x \in \Gamma_i^\uparrow \setminus \Gamma_i, \end{cases}$$

and

$$\mathcal{F}(\Gamma_i)^\uparrow := \{f \in \mathcal{F}(\Gamma_i) : f^\uparrow \in \mathcal{F}(\Gamma_i^\uparrow)\}$$

consists of those elements in the local space $\mathcal{F}(\Gamma_i)$ whose trace vanishes on the outflow boundary $\partial\Gamma_i^\downarrow$. Such extensions can be constructed explicitly as tensor products of Hestenes-type extensions [9, 19]. Then, denoting by χ_Ω the characteristic function of Ω and defining

$$\mathcal{P}_1 f := E_1(\chi_{\Gamma_1} f), \quad \mathcal{P}_i f := E_i(\chi_{\Gamma_i}(f - \sum_{l < i} \mathcal{P}_l f)), \quad i = 2, \dots, N,$$

one can prove that the mapping

$$T : f \mapsto \{\chi_{\Gamma_i} \mathcal{P}_i f\}_{i=1}^N \quad (3.4.46)$$

defines a topological isomorphism acting from $\mathcal{F}(\Gamma)$ onto the product space $\prod_{i=1}^N \mathcal{F}(\Gamma_i)^\downarrow$, where the spaces $\mathcal{F}(\Gamma_i)^\downarrow$ are defined analogously to $\mathcal{F}(\Gamma_i)^\uparrow$.

Since in view of (3.4.45), an analogous statement holds for the mapping

$$R : f \mapsto \{\chi_{\Gamma_i} \mathcal{P}_i^* f\}_{i=1}^N,$$

which takes $\mathcal{F}(\Gamma)$ onto the product space $\prod_{i=1}^N \mathcal{F}(\Gamma_i)^\dagger$, T is also isomorphic for the dual $\mathcal{F}^*(\Gamma)$, i.e.,

$$\|f\|_{\mathcal{F}(\Gamma)} \asymp \sum_{i=1}^N \|\mathcal{P}_i f\|_{\mathcal{F}(\Gamma_i)^\dagger}, \quad (3.4.47)$$

and likewise for $\mathcal{F}(\Gamma)$, $\mathcal{F}(\Gamma_i)^\dagger$ replaced by the duals $\mathcal{F}^*(\Gamma)$, $\mathcal{F}^*(\Gamma_i)^\dagger$, respectively.

Recall that the component spaces $\mathcal{F}(\Gamma_i)^\dagger$ are Sobolev or Besov spaces with certain boundary conditions (while their duals satisfy complementary boundary conditions) which can be viewed as *liftings* of analogous spaces defined on the unit cube \square as described in Section 3.3. Biorthogonal wavelet bases for these spaces, in turn, can be constructed via tensor products of suitable wavelet bases on the interval $[0, 1]$ satisfying certain boundary conditions [19]. Lifting such bases for each patch Γ_i and then applying T^{-1} produces biorthogonal wavelet bases on Γ which due to (3.4.47) gives again rise to norm equivalences of the type (3.2.37) but this time for Γ . Note that the latter step is only done for the sake of analysis. In practical calculations one would avoid executing T^{-1} but rather transfer all the computations to the component spaces and thereby to functions defined on the unit cube. Likewise moment conditions are formulated as in (3.3.42) ultimately on \square . For corresponding consequences with regard to domain decomposition see again [19].

3.5 Lipschitz domains

The above techniques are of limited use when dealing with bounded domains in \mathbb{R}^d with complicated boundaries. We shall always assume that Ω is an open and connected Lipschitz domain. This covers all domains of practical interest. In the following we shall briefly recall the results from [10].

One can also realize the Sobolev and Besov spaces by extension operators. The conditions we assume on Ω guarantee that there is an extension operator E which simultaneously extends Sobolev and Besov spaces. For example, in the cases of interest to us, if r is any positive real number, there is an extension operator $E = E_r$ such that

$$E : H^\alpha(\Omega) \rightarrow H^\alpha(\mathbb{R}^d), \quad 0 \leq \alpha \leq r$$

$$\begin{aligned}
E : B^\alpha(\Omega) &\rightarrow B^\alpha(\mathbb{R}^d), & 0 \leq \alpha \leq r & \\
Ef|_\Omega = f, & & f \in L_2(\Omega), & \|Ef\|_{\mathcal{F}(\mathbb{R}^d)} \lesssim \|f\|_{\mathcal{F}(\Omega)},
\end{aligned} \tag{3.5.48}$$

where $\mathcal{F} = H^\alpha$ or $\mathcal{F} = B^\alpha$.

In principle, the extension E can be used to generate a wavelet basis for Ω from a wavelet basis on \mathbb{R}^d . To this end, suppose that $\Psi = \{\phi_I : I \in \mathcal{D}_c\} \cup \{\psi_I : \psi \in \Psi^\circ, I \in \mathcal{D}^+\}$ and analogously $\tilde{\Psi}$ are biorthogonal wavelet bases for $L_2(\mathbb{R}^d)$. Given $f \in L_2(\Omega)$ it follows that Ef has a wavelet expansion (3.1.26). Let P_0 be the projector defined in (3.1.21). Then, for each $f \in L_2(\Omega)$, we have on Ω ,

$$f = P_0(Ef) + \sum_{I \in \mathcal{D}^+} \sum_{\psi \in \Psi^\circ} \langle Ef, \tilde{\psi}_I \rangle \psi_I \tag{3.5.49}$$

with \mathcal{D}^+ the set of dyadic cubes with measure ≤ 1 . The sum in (3.5.49) can be restricted to those ψ_I whose support nontrivially intersects Ω . The function f is in $H^\alpha(\Omega)$ (respectively $B^\alpha(\Omega)$) if and only if for the series in (3.5.49) the expressions (3.1.29) (respectively (3.1.34)) are finite.

Denoting again by E^* the adjoint of E this can be interpreted as

$$\begin{aligned}
\|f\|_{H^\alpha(\Omega)}^2 &\asymp \sum_{I \in \mathcal{D}_c} |\langle f, E^* \tilde{\phi}_I \rangle_\Omega|^2 \\
&+ \sum_{I \in \mathcal{D}^+} \sum_{\psi \in \Psi^\circ} |I|^{2\alpha/d} |\langle f, E^* \tilde{\psi}_I \rangle_\Omega|^2, \quad f \in H^\alpha(\Omega),
\end{aligned} \tag{3.5.50}$$

where $\langle f, g \rangle_\Omega := \int_\Omega f(x)g(x)dx$. This requires evaluating E^* which is in general not feasible numerically. Moreover, the above procedure does not necessarily preserve biorthogonality. Therefore this approach is useful as an analytical tool, but not for practical purposes.

For somewhat more specialized domains, it is possible to develop an extension strategy which is more numerically accessible. Indeed, we can utilize multiresolution analysis to construct E and E^* . For this, we shall use the approach described in [10]. The class of suitable domains is specified there. Roughly speaking, these domains are coordinatewise Lipschitz. Given such a domain and a dual pair $\phi, \tilde{\phi}$ as in (3.1.16), biorthogonal collections $\Phi_k = \{\phi_{k,m} : m \in \Delta_k\}$, $\tilde{\Phi}_k = \{\tilde{\phi}_{k,m} : m \in \Delta_k\}$ on Ω were constructed where the $\phi_{k,m}$ are adapted to the boundary so as to ensure polynomial exactness

while the $\tilde{\phi}_{k,m}$ involve only translates $\tilde{\phi}(2^k \cdot -m)$ which are fully supported in Ω . Now, we define

$$P_k f := \sum_{m \in \Delta_k} \langle f, \tilde{\phi}_{k,m} \rangle_{\Omega} \phi_{k,m}.$$

It was shown in [10] that for $0 \leq \alpha < \gamma$

$$\|f\|_{H^\alpha(\Omega)}^2 \asymp \|P_{k_0} f\|_{L_2(\Omega)}^2 + \sum_{k=k_0+1}^{\infty} 2^{2\alpha k} \|(P_k - P_{k-1})f\|_{L_2(\Omega)}^2, \quad (3.5.51)$$

where again $\gamma = \sup \{s > 0 : \phi \in H^s(\mathbb{R}^d)\}$. Each $\phi_{k,m}$ is either of the form $2^{dk/2} \phi(2^k \cdot -m)$ or is a linear combination of such functions restricted to Ω . Therefore the functions $\phi_{k,m}$ possess a canonical extension $\phi_{k,m}^e$ to \mathbb{R}^d . Note that since the $\tilde{\phi}_{k,m}$ are supported in Ω the collections $\Phi^e, \tilde{\Phi}_k$ are still biorthogonal. Thus, the operators

$$P_k^e f := \sum_{m \in \Delta_k} \langle f, \tilde{\phi}_{k,m} \rangle_{\Omega} \phi_{k,m}^e$$

take $L_2(\Omega)$ into $L_2(\mathbb{R}^d)$ and its adjoint

$$(P_k^e)^* f := \sum_{m \in \Delta_k} \langle f, \phi_{k,m}^e \rangle \tilde{\phi}_{k,m}$$

takes $L_2(\mathbb{R}^d)$ into $L_2(\Omega)$. Moreover, the mapping

$$E f := P_{k_0}^e f + \sum_{k=k_0}^{\infty} (P_{k+1}^e - P_k^e) f \quad (3.5.52)$$

is an extension satisfying (3.5.48) for any $r < \gamma$.

Due to biorthogonality, evaluating $E^* \tilde{\psi}_I$ requires computing the inner products $\langle \tilde{\psi}_I, \phi_{k,m}^e \rangle_{\mathbb{R}^d}$ for levels k larger than the level of I . In numerical implementations, by using decay properties, this can in turn be restricted to finitely many levels depending on the required accuracy.

In view of the above comments, we shall assume in the following that we always have a pair of biorthogonal bases $\Psi = \{\psi_I : \psi \in \Psi_I^\circ, I \in \mathcal{D}_c \cup \mathcal{D}^+\}$ and $\tilde{\Psi} = \{\tilde{\psi}_I : \tilde{\psi} \in \tilde{\Psi}_I^\circ, I \in \mathcal{D}_c \cup \mathcal{D}^+\}$ where \mathcal{D}_c corresponds to functions on the coarsest level while for $I \in \mathcal{D}^+$ the $\psi_I, \tilde{\psi}_I$ play the role of wavelets.

Again the sets Ψ_I° will generally depend on I but will always contain at most finitely many functions. Setting as above $\mathcal{D} := \mathcal{D}_c \cup \mathcal{D}^+$, on one hand moment conditions of the form (3.2.39) or (3.3.42) hold, while on the other hand relations like

$$\|f\|_{B_q^\alpha(L_p(\Omega))}^q \asymp \sum_{I \in \mathcal{D}} \sum_{\psi \in \Psi_I^\circ} |I|^{-q(\alpha/d+1/2-1/p)} |\langle f, \tilde{\psi}_I \rangle|^q, \quad (3.5.53)$$

and

$$\|f\|_{H^\alpha(\Omega)}^2 \asymp \sum_{I \in \mathcal{D}} \sum_{\psi \in \Psi_I^\circ} |I|^{-2\alpha/d} |\langle f, \tilde{\psi}_I \rangle|^2, \quad \alpha \in (-\tilde{\gamma}, \gamma), \quad (3.5.54)$$

are valid when Ω is a domain or manifold of dimension d as discussed above. As in all the above examples it will be convenient to identify always the indices I with dyadic cubes of volume $|I|$.

So far, we have outlined several principles to construct wavelets on various types of domains and manifolds.

If one wants to employ such wavelet bases for solving an operator equation, the issue of boundary conditions is, of course, important. When dealing with boundary integral equations on a closed manifold this problem does not arise. It may also not be severe in connection with natural boundary conditions for elliptic problems on bounded domains. Appending essential boundary conditions is, in principle, a possibility to avoid incorporating boundary conditions in the trial spaces and to preserve possibly many favorable properties of wavelets defined on simple domains [41]. For domains which can be represented as a union of parametric images of a cube the approach outlined above also facilitates incorporating essential boundary conditions in the wavelet spaces. We dispense here with elaborating further on this issue and refer to [19, 36] for details of corresponding recent progress in this problem.

3.6 Wavelet discretization of operator equations

We return now to the operator equation (2.2) where in the following $H = H^t$ and $H^* = H^{-t}$ where either $H^t = H^t(\Omega)$ when Ω is a closed surface or when natural boundary conditions are assumed or H^t is a closed subspace of $H^t(\Omega)$ determined by boundary conditions so that A is injective on H^t . In fact, we will assume that (2.3) holds with $H = H^t$.

The standard Galerkin method for approximating the solution u of (2.2) begins with a finite dimensional space $S \subset H^t$ and finds the function $u_S \in S$ such that

$$\langle Au_S, s \rangle = \langle f, s \rangle, \quad s \in S. \quad (3.6.55)$$

By choosing a basis $\{s_k\}$ for S , (3.6.55) becomes a system of linear equations

$$\langle a(s_i, s_j) \rangle_{i,j} \mathbf{c} = \mathbf{f}, \quad (3.6.56)$$

with $\mathbf{f} := (f_i)$ and $f_i := \langle f, s_i \rangle$, \mathbf{c} the vector of coefficients of u_S with respect to this basis and the matrix $\langle a(s_i, s_j) \rangle_{i,j}$ the stiffness matrix.

In the sections that follow, we shall be interested in the efficiency in which u_S approximates the exact solution u of (2.2). The typical choices for S in the standard Finite Element Theory are spaces of piecewise polynomials on some partition associated to Ω . An analogous choice in the context of wavelets are spaces $S = S_j := \text{span}\{\psi_I : \psi \in \Psi_I^\circ, |I| \leq 2^{-jd}\}$ or more generally $S = S_\Lambda := \text{span}\{\psi_I : (\psi, I) \in \Lambda\}$, where Λ is some finite subset of $\nabla := \{(\psi, I) : \psi \in \Psi_I^\circ, I \in \mathcal{D}_c \cup \mathcal{D}^+\}$.

The efficiency of Galerkin methods depends on:

- (i) the approximation power of the spaces S ;
- (ii) properties of the stiffness matrix (condition number and sparsity).

We shall see in the following sections how the accuracy of the approximation of u_S to u depends on the regularity of u .

The properties of the stiffness matrix including its amenability to preconditioning is a central theme in Finite Element Methods amply reported on e.g. in [16, 18, 48, 49]. Wavelet discretizations offer the following advantages with regard to (ii). To describe this let for $\Lambda \subset \nabla$ as above

$$P_\Lambda f := \sum_{(\psi, I) \in \Lambda} \langle f, \tilde{\psi}_I \rangle \psi_I. \quad (3.6.57)$$

Note that under the assumption (2.3), which we will quantify as,

$$c'_1 \|Av\|_{H^{-t}} \leq \|v\|_{H^t} \leq c'_2 \|Av\|_{H^{-t}}, \quad v \in H^t, \quad (3.6.58)$$

and the selfadjointness of A , the Galerkin schemes are stable. In this case, this means in terms of the projectors P_Λ that

$$\|P_\Lambda^* Av\|_{H^{-t}} \asymp \|v\|_{H^t}, \quad v \in S_\Lambda. \quad (3.6.59)$$

A first important relation between the wavelet bases and A is

$$|t| < \gamma, \tilde{\gamma}. \quad (3.6.60)$$

To explain its relevance for preconditioning consider the general version of the scale shift (3.1.30)

$$\mathcal{I}_s f := \sum_{(\psi, I) \in \nabla} |I|^{s/d} \langle f, \tilde{\psi}_I \rangle \psi_I, \quad (3.6.61)$$

so that by (3.5.54)

$$\|\mathcal{I}_{-s} f\|_{H^\alpha} \asymp \|f\|_{H^{s+\alpha}}, \quad s + \alpha \in (-\tilde{\gamma}, \gamma). \quad (3.6.62)$$

Thus considering $w := \mathcal{I}_{-t} v$ for $v \in S_\Lambda$, (3.6.62) and (3.6.59) yield under the assumption (3.6.60)

$$\|w\|_{L_2} \asymp \|v\|_{H^t} \asymp \|P_\Lambda^* A v\|_{H^{-t}} \asymp \|\mathcal{I}_t^* P_\Lambda^* A P_\Lambda \mathcal{I}_t w\|_{L_2}, \quad (3.6.63)$$

where we have used that

$$\mathcal{I}_s^{-1} = \mathcal{I}_{-s},$$

since $\Psi, \tilde{\Psi}$ are biorthogonal. Clearly (3.6.63) means that the operator

$$\mathcal{B}_\Lambda := \mathcal{I}_t^* P_\Lambda^* A P_\Lambda \mathcal{I}_t$$

satisfies

$$\text{cond}_2(\mathcal{B}_\Lambda) := \|\mathcal{B}_\Lambda\|_{\ell_2} \|\mathcal{B}_\Lambda^{-1}\|_{\ell_2} = \mathcal{O}(1), \quad \#\Lambda \rightarrow \infty, \quad (3.6.64)$$

where $\|\cdot\|_{\ell_2}$ denotes the spectral norm. It is not hard to verify that the matrix representation of \mathcal{B}_Λ is given by

$$\mathcal{A}_\Lambda = \langle \mathcal{A}((I, \psi), (J, \eta)) \rangle_{(I, \psi), (J, \eta) \in \Lambda}, \quad \mathcal{A}(I, J) = |I|^{t/d} \langle A \eta_J, \psi_I \rangle |J|^{t/d}. \quad (3.6.65)$$

Thus, a suitable diagonal scaling applied to the stiffness matrix relative to the wavelet basis results in a matrix with uniformly bounded condition numbers [18].

This suggests to reformulate the equation (2.2) as an infinite discrete system by representing u and f with respect to the primal and dual wavelet basis, respectively, i.e.,

$$u = \sum_{(\psi, I) \in \nabla} \langle u, \tilde{\psi}_I \rangle \psi_I, \quad f = \sum_{(\psi, I) \in \nabla} \langle f, \psi_I \rangle \tilde{\psi}_I.$$

In view of (3.6.64) and (3.6.65), it is useful to introduce the following rescaling. Let $\bar{\mathbf{f}} := (\bar{f}_{\psi,I})$ with $\bar{f}_{\psi,I} := |I|^{t/d} \langle f, \psi_I \rangle$ and let $\bar{\mathbf{u}} := (\bar{u}_{\psi,I})$ with $\bar{u}_{\psi,I} := |I|^{-t/d} \langle u, \psi_I \rangle$ and let \mathcal{A} be the infinite matrix with entries $\mathcal{A}((I, \psi), (J, \eta)) := |I|^{t/d} a(\psi_I, \eta_J) |J|^{t/d}$. Then, (2.2) becomes

$$\mathcal{A}\bar{\mathbf{u}} = \bar{\mathbf{f}}. \quad (3.6.66)$$

Clearly \mathcal{A}_Λ is a finite submatrix of \mathcal{A} and \mathcal{A} is boundedly invertible on $\ell_2(\nabla)$.

Another important advantage of the wavelet basis is that the matrix \mathcal{A} is almost diagonal in the sense that

$$\mathcal{A}((\psi, I), (\eta, J)) \lesssim \left(1 + \frac{|\xi_I - \xi_J|}{\max(|I|, |J|)^{1/d}}\right)^{d+2\tilde{N}+2t} \left(\min\left(\frac{|I|}{|J|}, \frac{|J|}{|I|}\right)\right)^r, \quad (3.6.67)$$

where ξ_I is a point in the support of ψ_I and where $r > d/2$ depends on the regularity of the wavelets. The parameter \tilde{N} again denotes the number of vanishing moments of the wavelets $\psi_I, I \in \mathcal{D}^+$. Thus in the present biorthogonal framework r and \tilde{N} can be made as large as one wishes by proper choice of the wavelet basis.

We shall make a distinction in what follows between the cases of linear methods in which the space S is chosen independent of u and nonlinear or adaptive methods in which S depends on u and previous choices for S . We wish in particular to understand what, if any, are the advantages of adaptive methods.

4 Linear approximation

For simplicity we confine the following discussion to the case that $\Omega \subset \mathbb{R}^d$ is a bounded and connected Lipschitz domain. We have noted above that in standard Finite Element Theory one seeks an approximation to the solution u of (2.2) from a finite dimensional linear space S . It is well understood in approximation theory that for standard spaces S , consisting for example of polynomials, splines, or wavelets, the efficiency of the approximation to u by elements of S is related to the regularity of u in the scale of Sobolev spaces H^α . To briefly recall this theory, we shall restrict ourselves to the case where S is chosen from a sequence (S_n) of spaces of the following two general types: (i) S_n is a linear space of piecewise polynomials on a partition related to Ω ;

(ii) S_n is a subspace of a multiresolution space V_j with the V_j as described in §3. We assume that S_n has dimension $\sim n^d$. In case (i), this corresponds to partitions with cell size $h \sim n^{-1}$ and in case (ii) this corresponds to taking $S_n = V_m$, $n = 2^m$, or a subspace of V_m reflecting boundary conditions. We let

$$E_n(f) := E(f, S_n) := \inf_{s \in S_n} \|f - s\|_{L_2(\Omega)}. \quad (4.68)$$

The following is a generic theorem for approximation by the elements of S_n . It requires additional conditions on S_n which we shall discuss after the theorem is stated.

Theorem 1 *For $\alpha > 0$, and a function $g \in L_2(\Omega)$, the following are equivalent*

- (i) $g \in H^\alpha(\Omega)$;
- (ii) $\sum_{n=1}^{\infty} [n^\alpha E_n(g)]^2 \frac{1}{n}$ is finite.

The sum in (ii) is equivalent to the semi-norm for $H^\alpha(\Omega)$. Moreover, a similar result holds for approximation in $H^t(\Omega)$ provided $E_n(g)$ is replaced by the error in approximation by functions from S_n in the norm of H^t and H^α is replaced by $H^{\alpha+t}$ in (ii).

Remark 1 *The condition (i) is slightly stronger than requiring $E_n(g) = \mathcal{O}(n^{-\alpha})$. The class of functions which satisfy the latter condition is precisely the Besov space $B_\infty^\alpha(L_2(\Omega))$ in the case of approximation in $L_2(\Omega)$. A similar result holds for approximation in $H^t(\Omega)$.*

We are purposefully not being precise about the conditions needed on the spaces S_n so that Theorem 1 is valid. In the case that S_n is a wavelet space V_m with $n = 2^m$ then there is a real number r such that Theorem 1 holds for $0 < \alpha < r$. The number r is related to the smoothness and polynomial exactness of the wavelet basis. It is the same number r such that the Sobolev spaces H^α are characterized by the wavelet coefficients as in (3.1.29) for $0 < \alpha < r$. To see precise conditions under which Theorem 1 holds for wavelets, we refer the reader to any of the standard treatments on wavelets such as Meyer [47], DeVore and Lucier [23], or Frazier and Jawerth [29]. In the case of spline approximation, necessary and sufficient conditions for the validity of Theorem 1 can be quite subtle (see e.g. Jia [34]). We refer the reader to one of the standard texts on Finite Elements, e.g. Oswald [48]. We should also mention that if boundary conditions are to be incorporated

in H^α then these boundary conditions must be incorporated into S_n and this must be incorporated into the analysis.

We can use Theorem 1 to infer the potential accuracy of numerical methods for solving (2.2) based on Galerkin solutions. Since the numerical solution u_{S_n} comes from the space S_n , it will provide efficiency of approximation of order $\mathcal{O}(n^{-\alpha})$ in the L_2 norm in the sense of (i) of Theorem 1 only if the solution u has smoothness of order α in the scale of Sobolev spaces. Thus the maximum regularity of u in this scale determines the maximum efficiency a linear numerical scheme can have.

We shall see in the next section that using nonlinear methods changes the scale of smoothness spaces in the generic theorem.

5 Nonlinear wavelet approximation in $L_2(\Omega)$

We are ultimately interested in adaptive methods for solving (2.2). Adaptive methods are a form of nonlinear approximation. For the purposes of orientation, it will be useful to consider the following simpler (but related) form of nonlinear wavelet approximation called n -term approximation.

Recall that $\nabla = \{(\psi, I) : \psi \in \Psi_I^\circ, I \in \mathcal{D}_c \cup \mathcal{D}^+\}$. Let Σ_n denote the set of all function

$$S = \sum_{(\psi, I) \in \Lambda} c_{\psi_I} \psi_I$$

where $\Lambda \subset \nabla$ and $\#\Lambda \leq n$. Then Σ_n is a nonlinear space. We let

$$\sigma_n(f) := \inf_{S \in \Sigma_n} \|f - S\|_{L_2(\Omega)} \quad (5.69)$$

be the error in approximating f by the elements of Σ_n .

We are interested in characterizing the functions $f \in L_2(\Omega)$ for which $\sigma_n(f)$ tends to zero like $\mathcal{O}(n^{-\alpha/d})$ for some $\alpha > 0$. It is not difficult (see [24]) to prove the following theorem.

Theorem 2 *For each $0 \leq \alpha < r$, we have that the following two statements are equivalent*

- (i) $f \in B^\alpha$,
- (ii) $\sum_{n=1}^{\infty} [n^{\alpha/d} \sigma_n(f)]^r \frac{1}{n} < \infty$.

Condition (ii) in Theorem 2.1 is the analogue of (ii) of Theorem 1. Notice that in Theorem 1 the dimension of S_n is $\mathcal{O}(n^d)$ but in Theorem 2 Σ_n is of dimension n which explains the difference in the form of (ii) in these two theorems. Thus both theorems talk about the same approximation rate in terms of space dimension. As noted before (ii) of Theorem 2 is close to $\sigma_n(f) = \mathcal{O}(n^{-\alpha/d})$.

We shall now make several remarks which will bring out the differences between Theorem 1 for linear approximation and Theorem 2 for nonlinear approximation. Both theorems characterize functions with a prescribed accuracy of approximation by smoothness conditions. But these smoothness conditions are of quite a different nature. In Theorem 1 the function g should be in H^α and thus have α orders of smoothness in L_2 . In contrast, Theorem 2 requires (for the same approximation rate) only that $g \in B^\alpha$. Recall that B^α measures smoothness of order α in a space L_τ , with $\tau := (\alpha/d + 1/2)^{-1}$. Since τ is generally much smaller than 2 this is a much weaker smoothness condition. Another view of the spaces B^α comes from the Sobolev embedding theorem. These are in some sense the smallest spaces of smoothness α which are embedded in L_2 ; for example for $\mu < \tau$, $B_\mu^\alpha(L_\mu)$ is not embedded in L_2 .

The spaces B^α contain functions which are very unsmooth in the classical sense. For example when $d = 1$, any piecewise analytic function is contained in all of the spaces B^α but only in H^α if $\alpha < 2$.

Since the wavelet basis is a Riesz basis, the following simple algorithm asymptotically realizes the best n -term approximation.

Remark 2 *We take Λ to be a set of n pairs (I, ψ) for which $|\langle f, \tilde{\psi}_I \rangle|$ is largest. Then,*

$$S_n =: P_\Lambda f = \sum_{(I, \psi) \in \Lambda} \langle f, \tilde{\psi}_I \rangle \psi_I$$

is in Σ_n and

$$\|f - S_n\|_{L_2(\Omega)} \asymp \sigma_n(f)$$

with constants depending only on the constants in (3.1.28). It follows that that

(iii) The property

$$\sum_{n=1}^{\infty} [n^{\alpha/d} \|f - S_n\|_{L_2(\Omega)}]^\tau \frac{1}{n} < \infty$$

is equivalent to (i) and (ii) in Theorem 2.

Note that this algorithm requires knowledge of all of the wavelet coefficients of f .

The n -term approximation is not directly applicable to numerical methods for operator equations since the wavelet coefficients of the solution u are not available to us. Instead, one constructs nonlinear approximations to the solution u using adaptive algorithms. An adaptive wavelet method for approximating the solution u of (2.2) would select the wavelet terms to be retained in the approximation from prior computations combined with any additional information that may be available. We discuss specific adaptive methods later in this paper. For the present, we want only to draw out the distinction in the form of the approximation between these adaptive methods and the n -term approximation just described.

We shall discuss adaptive wavelet approximation based on the wavelet decomposition (3.1.27). It is notationally convenient to combine all wavelet terms that correspond to a fixed dyadic cube. Therefore, for $I \in \mathcal{D}_c \cup \mathcal{D}^+$, we define

$$f_I := \sum_{\psi \in \Psi^\circ} \langle f, \tilde{\psi}_I \rangle \psi_I \quad (5.70)$$

and

$$c_I(f) := \left(\sum_{\psi \in \Psi^\circ} |\langle f, \tilde{\psi}_I \rangle|^2 \right)^{1/2}. \quad (5.71)$$

If we assume to work with wavelets adapted in some way to the underlying domain the sets $\mathcal{D}_c, \mathcal{D}^+$ of dyadic cubes is to be understood in the above sense. In what follows we could likewise employ wavelets defined on all of \mathbb{R}^d when Ω is a bounded domain in \mathbb{R}^d . In this case one can always assume that there is a fixed dyadic cube Q such that $\text{supp } \psi_I \cap \Omega = \emptyset$ unless $I \subseteq Q$. For notational convenience, we assume that $Q = [0, 1]^d$. We shall adopt the following notational conventions. We let $\mathcal{D}_c \cup \mathcal{D}^+$ denote the set of dyadic cubes $I \subseteq Q$. If $I \subset \mathcal{D}_c \cup \mathcal{D}^+$, and $\text{supp } \psi_I \cap \Omega = \emptyset$, then define $f_I = 0$ and $c_I(f) = 0$. In this way, we can always consider the wavelet decomposition in (3.1.27) to be indexed on all the cubes of $\mathcal{D} := \mathcal{D}_c \cup \mathcal{D}^+$.

An adaptive procedure usually creates approximations of the form

$$P_0(f) + \sum_{I \in \mathcal{J}} f_I \quad (5.72)$$

where \mathcal{J} is a set of dyadic cubes which have a certain tree structure. If $I \in \mathcal{D}$, we say that J is a *child* of I (and I is a *parent* of J), iff $|J| = 2^{-d}|I|$ and $J \subset I$. We denote the set of all children of I by $\mathcal{C}(I)$. We say J is a *sibling* of I if they are both children of the same parent.

Trees typically arise in an adaptive algorithm where at each inductive stage cubes are refined by adding all their children to the tree. The trees $\mathcal{J} \subset \mathcal{D}$ that arise in adaptive algorithms have the following two properties:

P1 If $I \in \mathcal{J}$ with $|I| < 1$, then its parent is in \mathcal{J} .

P2 If $I \in \mathcal{J}$, then all of its siblings are in \mathcal{J} .

We call a set $\mathcal{J} \subset \mathcal{D}_+$ which satisfies **P1**, **P2** an admissible tree. For an admissible tree \mathcal{J} , we let $\mathcal{F}(\mathcal{J}) \subset \mathcal{J}$ denote the set of final leaves of \mathcal{J} , i.e., the set of those $I \in \mathcal{J}$ such that I has no children in \mathcal{J} .

It is of interest to know what is the overhead (in efficiency when compared to n -term approximation) in forcing such a tree structure in the approximant. In [10] a simple adaptive wavelet algorithm was given that shows that this cost is quite minimal. To describe this algorithm, we define for any $\mathcal{J} \subset \mathcal{D}$,

$$P_{\mathcal{J}}f := \sum_{I \in \mathcal{J}} f_I$$

where f_I is defined by (5.70). It follows that

$$\|f - P_{\mathcal{J}}f\|_{L_2(\Omega)}^2 \lesssim \sum_{I \in \mathcal{D} \setminus \mathcal{J}} |c_I(f)|^2. \quad (5.73)$$

For a dyadic cube $I \in \mathcal{D}$, we let $T(I)$ be the *tower* of I which is the collection of all $J \in \mathcal{D}$ such that $J \subset I$, $J \neq I$ and we let

$$R(I) := \left(\sum_{J \in T(I)} |c_J|^2 \right)^{\frac{1}{2}}. \quad (5.74)$$

Algorithm 1 Fix $\varepsilon > 0$ and choose an initial admissible tree \mathcal{J}_0 . Set $\mathcal{B}_\varepsilon = \{J \in \mathcal{F}(\mathcal{J}_0) : R(J) > \varepsilon\}$, $\mathcal{J}_\varepsilon = \mathcal{J}_0$. If $\mathcal{B}_\varepsilon = \emptyset$ stop. Otherwise, for $I \in \mathcal{B}_\varepsilon$ do:

- replace \mathcal{J}_ε by $\mathcal{J}_\varepsilon \cup \mathcal{C}(I)$
- replace \mathcal{B}_ε by $(\mathcal{B}_\varepsilon \setminus \{I\}) \cup \{J \in \mathcal{C}(I) : R(J) > \varepsilon\}$

Since $\|P_{\mathcal{J}_n} f - f\|_{L_2(\Omega)} \rightarrow 0$, $n \rightarrow \infty$, $\mathcal{J}_n = \{I \in \mathcal{D} : |I| \leq 2^{-nd}\}$ the above algorithm terminates for every $\varepsilon > 0$ after finitely many refinement steps, i.e., eventually one obtains $\mathcal{B}_\varepsilon = \emptyset$ and the resulting tree \mathcal{J}_ε has the property that

$$R(J) \leq \varepsilon \text{ for } J \in \mathcal{F}(\mathcal{J}_\varepsilon). \quad (5.75)$$

The following theorem from [10] estimates the approximation error of the adaptive algorithm.

Theorem 3 *Let $\alpha > 0$ and $\tau := (\alpha/d + 1/2)^{-1}$ be as in Theorem 2. If $g \in B_\mu^\beta(L_\mu(\Omega))$ for any $\beta > \alpha$ and $\mu > \tau$, then*

$$\|g - P_{\mathcal{J}_\varepsilon} f\|_{L_2(\Omega)} \lesssim \|g\|_{B_\mu^\beta(L_\mu(\Omega))} (\#\mathcal{J}_\varepsilon)^{-\frac{\alpha}{d}}. \quad (5.76)$$

with a constant depending only on d and α .

When compared to Theorem 2, this theorem shows that with only a slightly stronger assumption on g , we obtain the same approximation order as in n -term approximation.

There is an analysis, similar to the above, for adaptive approximation based on piecewise polynomials. In [26], this was carried out for adaptive algorithms which use partitions into cubes. It should be possible to carry over the arguments in [26] to more general adaptive partitions, for example, to triangulations, provided the refining triangulation are always done in the same manner and lead to shape preserving triangulations.

The above analysis shows that it is the regularity of the solution u in the Besov scale B^α which determines its approximability by nonlinear methods. Adaptive methods therefore should be evaluated against the optimal value that is theoretically possible. In this context the Besov spaces B^α replace the role of the Sobolev spaces H^α when analyzing adaptive numerical methods.

6 Regularity of solutions to PDE's and approximation order

In the proceeding sections, we have already seen that the maximal possible efficiency that a numerical method to recover the solution of (2.2) can have is determined by the regularity of the exact solution of (2.2) in specific

smoothness spaces. It was emphasized that the approximation order of linear methods is related to the Sobolev scale $H^\alpha(\Omega)$, compare with Theorem 1, whereas the efficiency of nonlinear and adaptive methods is determined by the Besov scale B^α , compare with Theorem 2 and Theorem 3. Therefore, in this section, we shall give a short survey on the regularity theorems for partial differential equation for both kinds of smoothness spaces. Let L be an elliptic differential operator of order $2m$ on a bounded and connected Lipschitz domain Ω ,

$$L = \sum_{|k| \leq m} \sum_{|l| \leq m} (-1)^{|l|} D^l a_{k,l}(x) D^k, \quad a_{k,l} \in L_\infty(\overline{\Omega}). \quad (6.77)$$

For simplicity, we shall restrict ourselves to Dirichlet boundary conditions, i.e., we consider the problem: find $u \in H_0^m(\Omega)$ such that

$$a(u, v) = \sum_{|k|, |l| \leq m} \int_{\Omega} a_{k,l}(x) (D^k u)(D^l v) dx = \int_{\Omega} f(x) v(x) dx \quad (6.78)$$

holds for all $v \in H_0^m(\Omega)$.

Let us start with the usual Sobolev scale H^α . We want to investigate how the regularity of the solution u of (6.78) depends on the coefficients $a_{k,l}$, the right-hand side f and on the shape of the domain Ω . Most of the time, this question is formulated in the following form. Suppose that f is contained in $H^{\alpha-m}(\Omega)$ for some $\alpha \geq 0$. What are the conditions which imply that the variational solution $u \in H_0^m(\Omega)$ is in fact in $H^{m+\alpha}(\Omega)$ and satisfies

$$\|u\|_{H^{m+\alpha}} \lesssim \|f\|_{H^{\alpha-m}} + \|u\|_{H^m} ? \quad (6.79)$$

A boundary value problem with these properties is called α -regular. A first answer is that α -regularity holds if the coefficients and the domain are sufficiently smooth.

Theorem 4 *Let $\Omega \in C^{s+m}$ for some $s \geq 0$. Let $\alpha \geq 0$ satisfy*

$$\alpha + 1/2 \notin \{1, 2, \dots, m\}; \quad 0 \leq \alpha \leq s, \text{ if } s \in \mathbb{N}; \quad 0 \leq \alpha < s, \text{ if } s \notin \mathbb{N}.$$

For the coefficients let the following hold:

$$D^n a_{k,l} \in L_\infty(\Omega) \text{ for all } k, l, n \text{ with } |n| \leq \max(0, s + |l| - m), \text{ if } s \in \mathbb{N},$$

$a_{k,l} \in C^{s+|l|-m}(\overline{\Omega})$ for $|l| = m$, $a_{k,l} \in L_\infty(\Omega)$ otherwise, if $s \notin \mathbb{N}$.

Then the solution $u \in H_0^m(\Omega)$ of (6.78) with $f \in H^{-m+\alpha}(\Omega)$ belongs to $H^{m+\alpha}(\Omega) \cap H_0^m(\Omega)$ and satisfies

$$\|u\|_{H^{m+\alpha}} \lesssim \|f\|_{H^{\alpha-m}} + \|u\|_{H^m}.$$

Results of this form were e.g. obtained by Lions and Magenes [42], see also Hackbusch [32]. Theorem 4 implies that for problems satisfying the conditions of that theorem linear methods are sufficient in the sense that they can provide a convergence rate of order $\mathcal{O}(n^{-(m+\alpha)})$, compare with Theorem 1 and Remark 1. Such a result does not hold for nonsmooth domains, e.g. for domains with edges and corners, even if the coefficients are arbitrarily smooth. In this case, the regularity is only preserved strictly in the interior [32].

Theorem 5 *Let Ω be a Lipschitz domain and let $\Omega_0 \subset\subset \Omega_1 \subset \Omega$ and $\alpha \geq 0$. Let us assume that the coefficients satisfy the conditions of Theorem 4 with Ω replaced by Ω_1 and with $s \geq \alpha$, where $\alpha \in \mathbb{N}$, $s > \alpha$, where $\alpha \notin \mathbb{N}$, respectively. Suppose that the restriction f_{Ω_1} belongs to $H^{-m+\alpha}(\Omega_1)$. Then the restriction of u to Ω_0 belongs to $H^{m+\alpha}(\Omega_0)$ and satisfies*

$$\|u\|_{H^{m+\alpha}(\Omega_0)} \lesssim \|f\|_{H^{-m+\alpha}(\Omega_1)} + \|u\|_{H^m(\Omega)}.$$

In general the smoothness of the solution u will decrease significantly as one approaches the boundary. Therefore, the estimation of the Sobolev regularity on nonsmooth domains is a delicate task. Roughly speaking, the results can be divided into three types: results on specific operators and specific domains, see e.g. Grisvard [30, 31], results on specific operators and general domains, see e.g. Jerison and Kenig [35], and results on general operators on general domains, see e.g. Dauge [21], Kondrat'ev [37, 38, 39, 40] and Maz'ja and Plamenevskii [43, 44, 45, 46]. Among other things, Grisvard has intensively studied the Poisson equation, i.e., $L = -\Delta$ on polyhedral domains in \mathbb{R}^2 and \mathbb{R}^3 , respectively, primarily with $f \in L_2$. Let us first describe a typical result in \mathbb{R}^2 .

Let Ω be a polygonal domain with vertices ζ_j , $j = 1, 2, \dots$ and let ω_j denote the measure of the interior angle at ζ_j . We introduce polar coordinates r_j, θ_j in the vicinity of each vertex ζ_j . Furthermore, let ϑ_j denote a suitable truncation function which depends only on the distance r_j to ζ_j . Then the following holds.

Theorem 6 Assume that Ω is a bounded polygonal open subset of \mathbb{R}^2 . For $f \in L_2(\Omega)$ let u denote the unique solution of

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx$$

for every $v \in H_0^1(\Omega)$. Then there exist unique numbers c_j such that

$$u - \sum_{\omega_j > \pi} c_j \vartheta_j(r_j) r_j^{\pi/\omega_j} \sin(\pi\theta_j/\omega_j) \in H^2(\Omega). \quad (6.80)$$

The explicit expression (6.80) enables one to determine exactly the Sobolev regularity of the solution u . Since only the nonconvex corners contribute, we obtain that $u \in H^2$ for convex polygonal domains. In the worst case, u will only be contained in $H^{3/2+\epsilon}$ for some $\epsilon > 0$ which can be arbitrary small. This result implies that for general polygonal domains linear method can only provide an approximation of order $\mathcal{O}(n^{-3/2})$.

A similar result holds for polyhedral domains in \mathbb{R}^3 . However, the treatment of this case is much more complicated since one has to deal with different types of singularities according to edges and vertices. The regularity along edges can in principle be reduced to a two-dimensional problem with parameters, see again [30] for details. To keep the technical difficulties at a reasonable level, we shall not discuss this case here and confine ourselves to a typical result concerning the behavior of the solution at one particular vertex. We need some further notation. Let Γ_i, Γ_l be two faces of Ω and let $\epsilon_{i,l}$ denote the edge between Γ_i and Γ_l whenever $\bar{\Gamma}_i$ and $\bar{\Gamma}_l$ intersect. The measure of the interior angle of the edge $\epsilon_{i,l}$ is denoted by $\omega_{i,l}$. We set

$$\Upsilon := \inf\{m\pi/\omega_{i,l} \in]0, 1[; \bar{\Gamma}_i \cap \bar{\Gamma}_l \neq \emptyset, m \geq 1\}.$$

For convenience, we translate the typical vertex to zero. Thus, in a neighborhood V of 0, Ω coincides with a cone C whose intersection with the unit sphere S^2 is denoted by G . Thus G is an open subset of the unit sphere whose boundary is the union of a finite number of arcs of great circles. We introduce spherical coordinates ϱ, σ and denote by Δ' the Laplace–Beltrami operator on S^2 . It can be shown that the spectrum of Δ' is an infinite sequence of real numbers $-\zeta_l$, $l = 1, 2, \dots$ where $\zeta_l \geq 0$, with no limit points. We denote by v_l , $l = 1, 2, \dots$ the orthonormalized sequence of related eigenfunctions, i.e.,

$$-\Delta' v_l = \zeta_l v_l.$$

The following theorem was shown in [30].

Theorem 7 *Let Ω be a bounded polyhedral open subset of \mathbb{R}^3 . For $f \in L_2(\Omega)$ let u denote the solution of*

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx$$

for all $v \in H_0^1(\Omega)$. Then there exist unique numbers c_l such that

$$u - \sum_l c_l \varrho^{-1/2 + \sqrt{(\zeta_l + 1/4)}} v_l(\sigma) \in H^\alpha(V) \quad (6.81)$$

for every $\alpha \leq 2$ such that $\alpha < \Upsilon + 1$, where the sum is over the l such that $\zeta_l \leq \alpha^2 - 2\alpha + 3/4$.

It is an easy consequence of (6.81) that again the critical value for regularity is $3/2$. Similar results hold for the biharmonic problem and for elasticity systems in 2 and 3D, see again [30] for details.

Jerison and Kenig [35] have studied the Poisson equation on arbitrary Lipschitz domains. Their work can be interpreted as the continuation of Grisvard's study. One of their results is the famous " $H^{3/2}$ -Theorem".

Theorem 8 *Let Ω be a bounded Lipschitz domain in \mathbb{R}^d . If $f \in L_2(\Omega)$, then $u \in H^{3/2}(\Omega)$.*

The value $3/2$ in Theorem 8 is best possible. For example, for $d = 2, 3$, Grisvard's results stated above give examples of domains where regularity bigger than $3/2$ cannot be obtained.

For general operators, the technical and notational difficulties increase alarmingly. Therefore, we shall only give a theorem for domains with regular cones. For the general case, the reader is referred to [21] from which also the following result is quoted. Again we translate the vertex of the cone C to zero and introduce spherical coordinates ϱ, σ . The set G is defined as above. The cone is called *regular* if G has a smooth boundary and $C = \{x \in \mathbb{R}^d \mid \frac{x}{|x|} \in G\}$. Suppose that the coefficients of L are in $C^\infty(\overline{\Omega})$ and let L_x be the principal part of L , frozen in x . We define

$$\mathcal{S}^\zeta(C) := \{v \mid v = \varrho^\zeta \sum_{0 \leq q \leq Q} v_q(\sigma) \log^q \varrho, v_q \in H_0^m(G)\}. \quad (6.82)$$

We say that L_x is *injective modulo polynomials on* $\mathcal{S}^\zeta(C)$ if whenever $v \in \mathcal{S}^\zeta(C)$ is such that $L_x v$ is polynomial, v is a polynomial itself. For regular conical domains, the following theorem holds.

Theorem 9 *Let $\alpha \geq 0, \alpha \notin \{1/2, \dots, m - 1/2\}$. Then (6.78) is α -regular if and only if for all $\zeta \in \mathbf{C}$ satisfying*

$$\Re \zeta \in [m - d/2, \alpha + m - d/2], \quad (6.83)$$

L_x is injective modulo polynomials on $\mathcal{S}^\zeta(C)$.

There exist also a lot of regularity theorems for Besov and non-Hilbertian Sobolev spaces. They are concerned with questions of the form: Given $f \in B_p^{\alpha-m}(L_p(\Omega))$, what are the conditions that imply that $u \in B_p^{\alpha+m}(L_p(\Omega))$? Consequently, these theorems provide us with information concerning the approximation order of linear methods as measured in L_p . We shall state two typical results for the spaces $W^\alpha(L_p(\Omega))$ and $B_p^\alpha(L_p(\Omega))$, respectively. For instance, Theorem 6 has the following extension to non-Hilbertian Sobolev spaces.

Theorem 10 *Let Ω be a bounded polygonal open subset of \mathbb{R}^2 . For each $f \in L_p(\Omega), 1 < p < \infty$, there exists a unique solution u of*

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx$$

for every $v \in H_0^1(\Omega)$ and in addition there exist numbers c_j such that

$$u - \sum_{\pi/\omega_j < 2-2/p} c_j \vartheta_j(r_j) r_j^{\pi/\omega_j} \sin(\pi\theta_j/\omega_j) \in W^2(L_p(\Omega))$$

provided that none of the numbers π/ω_j is equal to $2 - 2/p$.

For the case of Lipschitz domains, the most general results were again obtained by Jerison and Kenig [35]. We shall restrict ourselves here to the case $d \geq 3$, a similar result holds for $d = 2$.

Theorem 11 *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^d, d \geq 3$. There exists $\varepsilon, 0 < \varepsilon \leq 1$, depending only on the Lipschitz constant of Ω such that for every $f \in W^{\alpha-1}(L_p(\Omega))$ there is a unique solution $u \in W^{\alpha+1}(L_p(\Omega))$ to*

$$\begin{aligned} -\Delta u &= f & \text{on } \Omega, \\ u &= 0 & \text{on } \partial\Omega, \end{aligned} \quad (6.84)$$

provided one of the following holds:

- (a) $p_0 < p < p'_0$ and $\frac{1}{p} - 1 < \alpha < \frac{1}{p}$
- (b) $1 < p \leq p_0$ and $\frac{3}{p} - 2 - \varepsilon < \alpha < \frac{1}{p}$
- (c) $p'_0 \leq p < \infty$ and $\frac{1}{p} - 1 < \alpha < \frac{3}{p} - 1 + \varepsilon$

where $1/p_0 = 1/2 + \varepsilon/2$ and $1/p'_0 = 1/2 - \varepsilon/2$. Moreover, we have the estimate

$$\|u\|_{W^{\alpha+1}(L_p(\Omega))} \lesssim \|f\|_{W^{\alpha-1}(L_p(\Omega))}$$

for all $f \in W^{\alpha-1}(L_p(\Omega))$.

So far, we have seen that linear methods are only suitable for smooth and convex domains, even if the coefficients are arbitrarily smooth. Therefore, the hope is to gain efficiency by employing an adaptive numerical scheme. According to Theorem 2 the use on nonlinear methods is justified if the weak solution u is contained in the scale B^α , $0 \leq \alpha < \alpha^*$, where the maximal index α^* is significantly higher than the one for usual Sobolev scale H^α . Therefore, the first step of a systematic study of adaptive schemes should consist in the derivation of regularity theorems for u with respect to B^α . It seems that this kind of study is still in its infancy. First regularity theorems for the specific scale $B_\tau^\alpha(L_\tau(\Omega))$, $\frac{1}{\tau} = (\frac{\alpha}{d} + \frac{1}{2})$ were given for certain model problems by Dahlke and DeVore [15]. We give the following example for the Poisson equation taken from [13].

Theorem 12 *Let Ω be a bounded Lipschitz domain in R^d . Let u denote the solution of (6.84) with $f \in B_2^{\alpha-1}(L_2(\Omega))$ for some $\alpha > -1/2$. Then the following holds:*

$$u \in B_\tau^s(L_\tau(\Omega)), \quad \frac{1}{\tau} = \left(\frac{s}{d} + \frac{1}{2} \right), \quad 0 < s < \min \left\{ \frac{3d}{2(d-1)}, \alpha + 1 \right\}.$$

We observe that for a large range of the parameter α we have a jump of two for the smoothness of the solution in the special scale $B_\tau^s(L_\tau(\Omega))$, $\frac{1}{\tau} = (\frac{s}{d} + \frac{1}{2})$. For instance, for $d = 2$, we obtain the condition $\alpha < 2$, whereas for the usual scale $H^\alpha = B_2^\alpha(L_2(\Omega))$ we have the jump of two only for $\alpha < 1/2$, compare with Theorem 11. Therefore, the maximal index for the spaces B^α is in general much larger than the one for H^α . Consequently, Theorem 12

can be interpreted as a justification for adaptive and nonlinear methods. Indeed, this theorem implies that adaptive methods on Lipschitz domains can perform as good as linear methods on smooth domains, provided that the right-hand side f is contained in a suitable smoothness space.

7 An adaptive scheme for elliptic equations

In this section, we discuss possible connections between the above concept of nonlinear approximation and certain adaptive schemes for elliptic problems of the type considered above. We adhere to the assumptions made in Sections 2 and 3. While nonlinear approximation uses idealized information about the approximand any adaptive solver has to contend itself with information acquired during the computation combined perhaps with some information of the given data. In fact, the basic idea of adaptive schemes is to refine step by step the discretization only at those places where the behaviour of the searched object requires a higher resolution so that the error is more or less balanced throughout the domain. In the presence of singularities this results in highly nonuniform meshes. In the present context we do not have to think though in terms of refined meshes but rather in terms of refined spaces. By this we mean the following. How can one find possibly few further wavelets which when added to the current trial space guarantee a prescribed decay of the error of the corresponding Galerkin approximation. One can therefore perhaps not expect to obtain equally strong theoretically founded results in this latter context. Nevertheless, an issue of central importance will be to interrelate both concepts.

7.1 Preliminary remarks

In the context of elliptic problems it is natural to measure errors in the energy norm $\|\cdot\|$ defined in (2.5) or equivalently in $\|\cdot\|_{H^1}$ while the concept of nonlinear approximation has so far been formulated for the L_2 -norm $\|\cdot\|_{L_2}$. It is easy though to carry this over to measuring errors in Sobolev norms which will be indicated first. To this end, it will be convenient to economize our notation a little further by subsuming all information on a wavelet in

one index λ containing its type (if applicable), location and scale. For $\nabla = \{(\psi, I) : \psi \in \Psi_I^\circ, I \in \mathcal{D}_c \cup \mathcal{D}^+\}$ as before let

$$\lambda := (\psi, I), \quad |\lambda| := |I|^{1/d},$$

so that the biorthogonal bases are briefly denoted by $\Psi = \{\psi_\lambda : \lambda \in \nabla\}$, $\tilde{\Psi} = \{\tilde{\psi}_\lambda : \lambda \in \nabla\}$ and \mathcal{I}_s takes the form

$$\mathcal{I}_s g = \sum_{\lambda \in \nabla} |\lambda|^s \langle g, \tilde{\psi}_\lambda \rangle \psi_\lambda.$$

In analogy to (5.69) let

$$\sigma_{n,t}(g) := \inf \left\{ \left\| g - \sum_{\lambda \in \Lambda} d_\lambda \psi_\lambda \right\|_{H^t} : d_\lambda \in \mathbb{R}, \lambda \in \Lambda \subset \nabla, \#\Lambda = n \right\}.$$

We have the following analog to Remark 2.

Remark 3 *Let $g \in H^t$. We take Λ_n to be a set of n indices λ for which $|\lambda|^{-t} |\langle g, \tilde{\psi}_\lambda \rangle|$ is largest. Then, for P_Λ defined by (3.6.57) one has*

$$\sigma_{n,t}(g) \asymp \|g - P_{\Lambda_n}(g)\|_{H^t}, \quad n \in \mathbb{N}. \quad (7.1.85)$$

Thus picking the n first largest weighted coefficients realizes asymptotically the best n -term approximation relative to the norm $\|\cdot\|_{H^t}$ and hence, by (2.6), also relative to the energy norm $\|\cdot\|$.

Proof: Under the assumption (3.6.60) the assertion is an immediate consequence of the norm equivalence (3.5.54) which implies that

$$\sigma_{n,t}(g) \sim \sigma_{n,0}(\mathcal{I}_{-t}g) := \sigma_n(\mathcal{I}_{-t}g); \quad (7.1.86)$$

compare with (3.6.62). \square

Again the Besov regularity of a function g can be characterized in terms of its best n -term approximation relative to $\|\cdot\|_{H^t}$.

Proposition 1 *Assume that $\alpha - t < \gamma$ and let for $t \leq \alpha$*

$$\frac{1}{\tau^*} := \frac{\alpha - t}{d} + \frac{1}{2}. \quad (7.1.87)$$

Then one has

$$\sum_{n=1}^{\infty} \left(n^{(\alpha-t)/d} \sigma_{n,t}(g) \right)^{\tau^*} < \infty \quad (7.1.88)$$

if and only if $g \in B_{\tau^}^\alpha(L_{\tau^*}(\Omega))$.*

Proof: Combining Theorem 2 with (7.1.86) ensures that $\mathcal{I}_{-t}g$ belongs to $B_{\tau^*}^{\alpha-t}(L_{\tau^*}(\Omega))$ if and only if

$$\sum_{n=1}^{\infty} \left(n^{(\alpha-t)/d} \sigma_{n,t}(g) \right)^{\tau^*} < \infty.$$

So it remains to verify that \mathcal{I}_s not only shifts between Sobolev but also between Besov scales. In fact, one easily infers from (3.1.32) that the statements $\mathcal{I}_{-t}g \in B_{\tau^*}^{\alpha-t}(L_{\tau^*}(\Omega))$, $g \in B_{\tau^*}^{\alpha}(L_{\tau^*}(\Omega))$, and

$$\sum_{\lambda \in \nabla} |\lambda|^{-t\tau^*} |\langle g, \tilde{\psi}_{\lambda} \rangle|^{\tau^*} < \infty$$

are equivalent. \square

Proposition 1 has an interesting application to the Poisson equation (6.84). We have already discussed the efficiency of the best n -term approximation when applied to the solution of (6.84); compare with Theorem 12. However, these results were formulated with respect to approximation in $L_2(\Omega)$. For elliptic equations, the energy norm is slightly more natural. A combination of Proposition 1 and Theorem 12 provides us with the following result concerning approximation relative to $\|\cdot\|_{H^1}$.

Proposition 2 *Let u denote the solution of (6.84) with $f \in B_2^{\alpha-1}(L_2(\Omega))$, $\alpha \geq 1$. Then*

$$\sum_{n=1}^{\infty} \left(n^{s/d} \sigma_{n,1}(u) \right)^{\tau} < \infty \quad \text{for all } 0 < s < s^*/3, \quad (7.1.89)$$

where $s^* = \min\{\frac{3d}{2(d-1)}, \alpha + 1\}$ and $\tau = (s - 1)/d + 1/2$.

Proof: Since $\alpha \geq 1$, the right-hand side f is contained in $L_2(\Omega)$. Therefore, Theorem 8 implies that $u \in H^{3/2}(\Omega) = B_2^{3/2}(L_2(\Omega))$. On the other hand, we know from Theorem 12 that

$$u \in B_{\mu}^{\beta}(L_{\mu}(\Omega)), \quad \frac{1}{\mu} = \left(\frac{\beta}{d} + \frac{1}{2} \right), \quad 0 < \beta < s^*.$$

By interpolation and embeddings of Besov spaces, we can conclude that u is in a family of Besov spaces $B_q^s(L_q(\Omega))$ for a certain range of parameters q and

s , i.e., $u \in B_q^s(L_q(\Omega))$ whenever $(1/q, s)$ is in the interior of the quadrilateral with vertices $(1/2, 0)$, $(1/2, 3/2)$, $(s^*/d + 1/2, 0)$, $(s^*/d + 1/2, s^*)$. Therefore, to compute the range of parameters s for which u is contained in $B_\tau^s(L_\tau(\Omega))$, $\tau = (s - 1)/d + 1/2$, we have to determine the intersection of the lines

$$\frac{1}{q} - \frac{1}{2} = \left(\frac{2s^*}{d(2s^* - 3)} \right) \left(s - \frac{3}{2} \right)$$

and

$$\frac{1}{q} - \frac{1}{2} = \frac{s - 1}{d}$$

which is the point $(s^*/(3d) + 1/2, s^*/3 + 1)$. An application of Proposition 1 with $t = 1$ yields the result. \square

To illustrate this result, we consider the example where $d = 2$. If $\alpha \geq 2$, then $s^* = 3$. Hence, in this case, the nonlinear method gives an H^1 -approximation to u of order up to $n^{-1/d}$, whereas a linear method using n terms could only give $n^{-1/2d}$ in the worst case.

In general, a-priori knowledge about the Besov regularity of the solution u to (2.2) would give lower bounds for the errors produced by any adaptive method. Conversely, if we knew that a particular adaptive scheme is asymptotically as efficient as best n -term approximation in $\|\cdot\|_{H^t}$ its performance would allow us to infer the regularity of u . Of course, since the wavelet coefficients of the solution u are not known a-priori, one cannot apply Remark 3 directly. There are several possible ways of dealing with this problem.

Let $\mathbf{d}_\Psi(g)$ denote the sequence of wavelet coefficients of g relative to Ψ , i.e., $d_{\Psi,\lambda}(g) = \langle g, \tilde{\psi}_\lambda \rangle$, $\lambda \in \nabla$, and analogously $\mathbf{d}_{\tilde{\Psi}}(g)$. By (3.6.66) the solution u of (2.2) is determined by

$$\mathbf{d}_\Psi(\mathcal{I}_{-t}u) = \mathcal{A}^{-1} \mathbf{d}_{\tilde{\Psi}}(\mathcal{I}_t^* f). \quad (7.1.90)$$

Recall from (7.1.86) that the best n -term approximation of u in $\|\cdot\|_{H^t}$ corresponds to the best n -term approximation of $\mathcal{I}_{-t}u$ in the L_2 -norm $\|\cdot\|_{L_2}$ which, by Remark 2, corresponds to selecting the n largest terms $|d_{\Psi,\lambda}(\mathcal{I}_{-t}u)| = |\lambda|^{-t} |d_{\Psi,\lambda}(u)|$. It is known that in certain cases the decay properties of (3.6.67) of the infinite matrix \mathcal{A} imply similar decay properties for \mathcal{A}^{-1} , perhaps with different parameters, see e.g. [51]. In such a case the largest coefficients of $\mathcal{I}_{-t}u$ are expected to appear in a ‘neighborhood’ of the (accessible) largest coefficients $\tilde{f}_\lambda = |\lambda|^t d_{\tilde{\Psi},\lambda}(f)$. The effect of the smearing caused

by the application of \mathcal{A}^{-1} can in principle be estimated by the same methods as used in connection with matrix compression [18]. However, this assumes that the singular behavior of u is completely governed by the right hand side f .

Here we shall describe next a somewhat different approach from [14]. To motivate this let us briefly recall first a basic strategy employed by many adaptive finite element schemes. A key observation is the equivalence between the validity of two-sided error estimates and the so called *saturation* property. In the context of finite element methods this issue is discussed in [5]. The basic reasoning can be sketched as follows. Suppose that $S \subset V \subset H^t$ are two trial spaces with respective Galerkin solutions u_S, u_V . By orthogonality one has

$$\|u_V - u_S\| \leq \|u - u_S\|,$$

where $\|\cdot\|$ denotes again the energy norm. Moreover, one easily checks that one has

$$\|u - u_V\| \leq \beta \|u - u_S\| \tag{7.1.91}$$

for some $\beta < 1$ if and only if

$$(1 - \beta^2)^{1/2} \|u - u_S\| \leq \|u_V - u_S\|. \tag{7.1.92}$$

Thus if the refined solution u_V captures a sufficiently large portion of the remainder (7.1.92) the global energy error is guaranteed to decrease by a factor β when passing to the refined solution u_V and one has the bounds

$$\|u_V - u_S\| \leq \|u - u_S\| \leq (1 - \beta^2)^{-1/2} \|u_V - u_S\|, \tag{7.1.93}$$

which are computable. In practice one controls the local behavior of $u_V - u_S$ and refines the mesh at places where (an estimate for) this difference is largest. This results in balancing the error bounds. Although this has been observed to work well in many cases, the principal problem remains that something like (7.1.92) has to be *assumed* to prove convergence of the overall adaptive algorithm.

It is perhaps worth stressing that wavelet analysis allows us to remedy this conceptual deficiency and derive much stronger information about remainders. In fact, we shall see below that the assumption (7.1.92) about the *unknown* solution u can be replaced (quite in the spirit of the previous comments) by some (rather weak) information on the *accessible* data f .

To this end, let us first relate the type of estimates (7.1.93) to n -term approximation. Instead of minimizing the error for a given allowance of n terms one can minimize the number of terms needed to meet a given error tolerance. Specifically, given any strictly decreasing sequence $\{\tau_i\}_{i \in \mathbb{N}}$, we can look for a sequence $\{\Lambda(\tau_i)\}_{i \in \mathbb{N}}$ of index sets $\Lambda(\tau_i) \subset \nabla$ such that

$$\sigma_{\#\Lambda(\tau_i), t}(u) \asymp \tau_i, \quad i \in \mathbb{N}. \quad (7.1.94)$$

The following observation is an immediate consequence of Remark 3.

Remark 4 *One has*

$$\|u - P_{\Lambda(\tau_i)}u\| \asymp \tau_i, \quad i \in \mathbb{N}, \quad (7.1.95)$$

and the sets $\Lambda(\tau_i)$ can be chosen to be nested, i.e.,

$$\Lambda(\tau_i) \subset \Lambda(\tau_{i+1}), \quad i \in \mathbb{N}. \quad (7.1.96)$$

Let, with a slight abuse of notation, u_Λ denote the solution of Galerkin problem (3.6.55) with

$$S := S_\Lambda := \text{span} \{\psi_\lambda : \lambda \in \Lambda\}.$$

If $\Lambda \subset \tilde{\Lambda}$ we have

$$\|u - u_{\tilde{\Lambda}}\|^2 = \|u - u_\Lambda\|^2 - \|u_\Lambda - u_{\tilde{\Lambda}}\|^2,$$

since the Galerkin approximation is an orthogonal projection relative to the energy inner product. Therefore, we obtain

Remark 5 *Consider the following sequence $\{\Lambda^i\}_{i \in \mathbb{N}}$:*

(I) *Fix some $\Lambda^1 \subset \nabla$ and $\kappa < 1$. Define $\tau_1 := \|u - u_{\Lambda^1}\|$.*

(II) *Given Λ^i choose $\Lambda^{i+1} \subset \nabla$, $\Lambda^i \subset \Lambda^{i+1}$ such that*

$$\|u_{\Lambda^i} - u_{\Lambda^{i+1}}\| \geq \kappa \|u - u_{\Lambda^i}\| \quad (7.1.97)$$

while for any $\Lambda \subset \nabla$ with $\Lambda^i \subseteq \Lambda$ and $\#(\Lambda \setminus \Lambda^i) \leq \#(\Lambda^{i+1} \setminus \Lambda^i)$ one has

$$\|u_{\Lambda^i} - u_\Lambda\| < \kappa \|u - u_{\Lambda^i}\|. \quad (7.1.98)$$

Set $\tau_{i+1} := (1 - \kappa^2)^{1/2} \|u - u_{\Lambda^i}\|$.

(III) Replace $i + 1$ by i and go to (II).

Then one has

$$\|u - u_{\Lambda^n}\| \asymp \sigma_{\#\Lambda^n+c,t}(u), \quad n \in \mathbb{N}, \quad (7.1.99)$$

where c is some constant.

In practice, it will generally not be possible to realize the above strategy of capturing a significant portion of the remainder by a possibly small set of additional indices since the exact estimation required in (7.1.97) and (7.1.98) is generally not possible. However, it will be possible to bound quantities of the form $\|u_{\Lambda} - u_{\tilde{\Lambda}}\|$, for $\Lambda \subset \tilde{\Lambda}$, from below and above by computable local quantities times constants which are independent of the sets $\Lambda, \tilde{\Lambda}$ but different from one.

7.2 A-posteriori error estimates

Suppose that for some $\Lambda \subset \nabla$, S_{Λ} is the current trial space and that we have computed the solution u_{Λ} of (3.6.55) (within some appropriate tolerance). According to Remark 5, the next step is to estimate the error $\|u - u_{\Lambda}\|$ in the energy norm in a way that indicates how to select next a bigger set $\tilde{\Lambda} \subset \nabla$, $\Lambda \subset \tilde{\Lambda}$, of wavelet indices so that, on one hand, $\tilde{\Lambda}$ stays still possibly small while, on the other hand, the error $\|u - u_{\tilde{\Lambda}}\|$ is guaranteed to decrease by a certain amount. As mentioned before, selecting the index sets Λ implicitly corresponds to creating possibly nonuniform meshes. In fact, the spaces $S_n = \text{span}\{\psi_{\lambda} : |\lambda| \leq 2^{-n}\}$ correspond to uniformly refined meshes and taking only subsets of the complement bases $\{\psi_{\lambda} : |\lambda| = \ell\}$ corresponds to a nonuniform refinement.

To this end, we exploit the commonly used fact that, the error in the energy norm can be estimated by the residual in a dual norm which, at least in principle, can be evaluated. In fact, since

$$r_{\Lambda} := Au_{\Lambda} - f = A(u_{\Lambda} - u),$$

by (3.6.58) and (2.6), one has

$$c_1 \|r_{\Lambda}\|_{H^{-t}} \leq \|u - u_{\Lambda}\| \leq c_2 \|r_{\Lambda}\|_{H^{-t}}. \quad (7.2.100)$$

Expanding the residual r_{Λ} by the dual basis $\tilde{\Psi}$ and taking the Galerkin conditions

$$P_{\Lambda}^* Au_{\Lambda} = P_{\Lambda}^* f \quad (7.2.101)$$

into account, yields

$$r_\Lambda = \sum_{\lambda \in \nabla} \langle r_\Lambda, \psi_\lambda \rangle \tilde{\psi}_\lambda = \sum_{\lambda \in \nabla \setminus \Lambda} \langle r_\Lambda, \psi_\lambda \rangle \tilde{\psi}_\lambda.$$

Bearing (3.6.60) in mind, and quantifying the constants in (3.5.54), ensures the existence of finite positive constants c_3, c_4 such that

$$c_3 \left(\sum_{\lambda \in \nabla \setminus \Lambda} |\lambda|^{2t} |\langle r_\Lambda, \psi_\lambda \rangle|^2 \right)^{\frac{1}{2}} \leq \|r_\Lambda\|_{H^{-t}} \leq c_4 \left(\sum_{\lambda \in \nabla \setminus \Lambda} |\lambda|^{2t} |\langle r_\Lambda, \psi_\lambda \rangle|^2 \right)^{\frac{1}{2}}. \quad (7.2.102)$$

Thus, in principle, the nonnegative quantities

$$\delta_\lambda = \delta_\lambda(\Lambda) := |\lambda|^t |\langle r_\Lambda, \psi_\lambda \rangle|, \quad \lambda \in \nabla \setminus \Lambda,$$

are in some sense the desired local quantities bounding the error $\|u - u_\Lambda\|$ from below and above. However, in the present form, (7.2.102) is still useless since the bounds involve generally infinitely many terms δ_λ .

To understand these bounds a little better, suppose that $u_{\lambda'}$ denote the wavelet coefficients of the current solution

$$u_\Lambda = \sum_{\lambda' \in \Lambda} u_{\lambda'} \psi_{\lambda'}.$$

Straightforward calculations then yield

$$\delta_\lambda = |\lambda|^t \left| f_\lambda - \sum_{\lambda' \in \Lambda} \langle A\psi_{\lambda'}, \psi_\lambda \rangle u_{\lambda'} \right|, \quad (7.2.103)$$

where as above $f_\lambda := \langle f, \psi_\lambda \rangle$ denote the wavelet coefficients of the right hand side f relative to the dual basis $\tilde{\Psi}$. (7.2.103) shows that the size of δ_λ is influenced by two quantities. First, if the right hand side f itself has singularities this will result in large wavelet coefficients f_λ . Second, the sum $\sum_{\lambda' \in \Lambda} \langle A\psi_{\lambda'}, \psi_\lambda \rangle u_{\lambda'}$ gives the contribution of the current solution which, for instance, could reflect the influence of the boundary.

Thus replacing the bounds in (7.2.102) by finitely many computable but still sufficiently accurate terms requires

- a) estimating the smearing effect of A as well as

b) some a-priori knowledge about f .

So far we have only used the ellipticity (2.3) or (2.6) of A and the norm equivalence (3.5.54). To deal with problem a) one has to make essential use of the decay estimates (3.6.67). These estimates are usually deduced from (2.12) with the aid of moment conditions (see (3.1.25)). We describe now how they can be utilized. Let again \tilde{N} denote the order of vanishing moments of the wavelets ψ_λ and let $\delta < r - d/2$, where r is the constant in (3.6.67). Choose for any $\varepsilon > 0$ positive numbers $\varepsilon_1, \varepsilon_2$ such that

$$\varepsilon_1^{2\tilde{N}+2t} + 2^{-\frac{\delta}{\varepsilon_2}} \leq \varepsilon. \quad (7.2.104)$$

For each $\lambda \in \nabla$, we define the *influence sets*

$$\begin{aligned} \nabla_{\lambda,\varepsilon} := \{ \lambda' \in \nabla : & \left| |\log \lambda| - |\log \lambda'| \right| \leq \log 2 \varepsilon_2^{-1} \text{ and} \\ & \min\{|\lambda|^{-1}, |\lambda'|^{-1}\} \text{dist}(\Omega_\lambda, \Omega_{\lambda'}) \leq \varepsilon_1^{-1} \}, \end{aligned}$$

where Ω_λ denotes the support of ψ_λ . The sets $\nabla_{\lambda,\varepsilon}$ describe that portion of the sum

$$\Sigma_{\lambda,\Lambda} = \sum_{\lambda' \in \Lambda} \langle A\psi_{\lambda'}, \psi_\lambda \rangle u_{\lambda'},$$

appearing in the residual weights δ_λ (7.2.103), which is significant. In fact, the remainder

$$e_\lambda := \sum_{\lambda' \in \Lambda \setminus \nabla_{\lambda,\varepsilon}} \langle A\psi_{\lambda'}, \psi_\lambda \rangle u_{\lambda'}$$

can be estimated as follows [14].

Proposition 3 *For e_λ and $\nabla_{\lambda,\varepsilon}$ as above there exists a constant c_5 independent of f and Λ such that*

$$\left(\sum_{\lambda \in \nabla \setminus \Lambda} |\lambda|^{2t} |e_\lambda|^2 \right)^{\frac{1}{2}} \leq c_5 \varepsilon \|u_\Lambda\|. \quad (7.2.105)$$

Note that, again by (3.5.54),

$$\|u_\Lambda\| \sim \|u_\Lambda\|_{H^t} \sim \left(\sum_{\lambda \in \Lambda} |\lambda|^{-2t} |u_\lambda|^2 \right)^{\frac{1}{2}},$$

so that the right hand side in (7.2.105) can be evaluated by means of the wavelet coefficients of the current solution u_Λ . Moreover, one can even give an a-priori bound. In fact, the stability of the Galerkin scheme assured by (3.6.59) says, on account of the uniform boundedness of the P_Λ^* in H^{-t} , that

$$\|u_\Lambda\| \lesssim \|P_\Lambda^* f\|_{H^{-t}} \lesssim \|f\|_{H^{-t}}. \quad (7.2.106)$$

As for b) above, by construction, the *significant neighborhood* of Λ in $\nabla \setminus \Lambda$

$$N_{\Lambda, \varepsilon} := \{\lambda \in \nabla \setminus \Lambda : \Lambda \cap \nabla_{\lambda, \varepsilon} \neq \emptyset\} \quad (7.2.107)$$

is finite

$$\#N_{\Lambda, \varepsilon} < \infty.$$

Outside $N_{\Lambda, \varepsilon}$ the quantities δ_λ in (7.2.103) are essentially influenced by wavelet coefficients of f . But this portion is essentially a remainder of f . In fact, by (3.5.54),

$$\begin{aligned} \left(\sum_{\lambda \in \nabla \setminus (\Lambda \cup N_{\Lambda, \varepsilon})} |\lambda|^{2t} |f_\lambda|^2 \right)^{\frac{1}{2}} &\leq c_2 \|f - P_{\Lambda \cup N_{\Lambda, \varepsilon}}^* f\|_{H^{-t}} \\ &\leq c_6 \inf_{v \in \tilde{S}_{\Lambda \cup N_{\Lambda, \varepsilon}}} \|f - v\|_{H^{-t}} \leq c_6 \inf_{v \in \tilde{S}_\Lambda} \|f - v\|_{H^{-t}}, \end{aligned}$$

for some $c_6 < \infty$. This suggests defining

$$d_\lambda(\Lambda, \varepsilon) := |\lambda|^t \left| \sum_{\lambda' \in \Lambda \cap \nabla_{\lambda, \varepsilon}} \langle A\psi_{\lambda'}, \psi_\lambda \rangle u_{\lambda'} \right|, \quad \lambda \in \nabla \setminus \Lambda.$$

Note that, in view of (7.2.107),

$$d_\lambda(\Lambda, \varepsilon) = 0, \quad \lambda \in \nabla \setminus \Lambda, \quad \lambda \notin N_{\Lambda, \varepsilon}. \quad (7.2.108)$$

The main result can now be formulated as follows [14].

Theorem 13 *Under the above assumptions, one has*

$$\|u - u_\Lambda\| \leq c_2 c_4 \left(\left(\sum_{\lambda \in N_{\Lambda, \varepsilon}} d_\lambda(\Lambda, \varepsilon)^2 \right)^{\frac{1}{2}} + c'_5 \varepsilon \|f\|_{H^{-t}} + c_6 \inf_{v \in \tilde{S}_\Lambda} \|f - v\|_{H^{-t}} \right)$$

as well as,

$$\left(\sum_{\lambda \in N_{\Lambda, \varepsilon}} d_\lambda(\Lambda, \varepsilon)^2 \right)^{\frac{1}{2}} \leq \frac{1}{c_1 c_3} \|u - u_\Lambda\| + c'_5 \varepsilon \|f\|_{H^{-t}} + c_6 \inf_{v \in \tilde{S}_\Lambda} \|f - v\|_{H^{-t}}.$$

Moreover, for any $\tilde{\Lambda} \subset \nabla$, $\Lambda \subset \tilde{\Lambda}$, one has

$$\left(\sum_{\lambda \in \tilde{\Lambda} \cap N_{\Lambda, \varepsilon}} d_{\lambda}(\Lambda, \varepsilon)^2 \right)^{\frac{1}{2}} \leq \frac{1}{c_1 c_3} \|u_{\tilde{\Lambda}} - u_{\Lambda}\| + c'_5 \varepsilon \|f\|_{H^{-t}} + c_6 \inf_{v \in \tilde{S}_{\Lambda}} \|f - v\|_{H^{-t}}.$$

This result provides, up to the controllable tolerance

$$\tau(\Lambda, \varepsilon) := c'_5 \varepsilon \|f\|_{H^{-t}} + c_6 \inf_{v \in \tilde{S}_{\Lambda}} \|f - v\|_{H^{-t}},$$

computable lower and upper bounds for the error $\|u - u_{\Lambda}\|$. Usually under more specialized assumptions results of similar nature have been obtained also in the finite element context (see e.g. [27]). Furthermore, one expects that nonlinear problems can be handled by combining such estimates with known abstract results.

7.3 Convergence of an adaptive refinement scheme

In the present setting, it can be shown with the aid of Theorem 13 that under mild assumptions on the right hand side f a suitable adaptive choice of $\tilde{\Lambda}$ enforces the validity of the saturation property (7.1.92). We continue with the notation of Section 7.2. The following theorem was proved in [14].

Theorem 14 *Let $\text{tol} > 0$ be a given tolerance and fix $\theta \in (0, 1)$. Define*

$$C^* := \left(\frac{1}{c_1 c_3} + \frac{1 - \theta}{2c_2 c_4} \right), \quad (7.3.109)$$

choose $\mu > 0$ such that

$$\mu C^* \leq \frac{1 - \theta}{2(2 - \theta)c_2 c_4}, \quad (7.3.110)$$

and set

$$\varepsilon := \frac{\mu \text{tol}}{2c'_5 \|f\|_{H^{-t}}}. \quad (7.3.111)$$

Suppose that for $\Lambda \subset \nabla$, one has

$$c_6 \inf_{v \in \tilde{S}_{\Lambda}} \|f - v\|_{H^{-t}} < \frac{1}{2} \mu \text{tol}.$$

Then, whenever $\tilde{\Lambda} \subset \nabla$, $\Lambda \subset \tilde{\Lambda}$ is chosen so that

$$\left(\sum_{\lambda \in \tilde{\Lambda} \cap N_{\Lambda, \varepsilon}} d_{\lambda}(\Lambda, \varepsilon)^2 \right)^{\frac{1}{2}} \geq (1 - \theta) \left(\sum_{\lambda \in N_{\Lambda, \varepsilon}} d_{\lambda}(\Lambda, \varepsilon)^2 \right)^{\frac{1}{2}},$$

there exists a constant $\kappa \in (0, 1)$ depending only on the constants μ , θ , c_i , $i = 1, \dots, 6$, such that either

$$\|u - u_{\tilde{\Lambda}}\| \leq \kappa \|u - u_{\Lambda}\|$$

or

$$\left(\sum_{\lambda \in N_{\Lambda, \varepsilon}} d_{\lambda}(\Lambda, \varepsilon)^2 \right)^{\frac{1}{2}} = \left(\sum_{\lambda \in \nabla \setminus \Lambda} d_{\lambda}(\Lambda, \varepsilon)^2 \right)^{\frac{1}{2}} < \text{tol}.$$

Of course, the idea is to choose $\tilde{\Lambda} \supset \Lambda$ as small as possible, i.e., in any case $\tilde{\Lambda} \setminus \Lambda \subset N_{\Lambda, \varepsilon}$. This leads to the following

Algorithm 2 *0. Choose $\text{tol} > 0$, $\theta \in (0, 1)$ and compute C^* , μ according to (7.3.109), (7.3.110).*

1. *Compute $\varepsilon = \varepsilon(\mu, \text{tol})$ by (7.3.111).*

2. *Determine $\Lambda \subset \nabla$ such that*

$$c_6 \inf_{v \in \tilde{S}_{\Lambda}} \|f - v\|_{H^{-t}} < \frac{1}{2} \mu \text{ tol}.$$

3. *Solve*

$$\langle Au_{\Lambda}, v \rangle = \langle f, v \rangle, \quad \forall v \in S_{\Lambda}.$$

4. *Compute*

$$\eta_{\Lambda, \varepsilon} := \left(\sum_{\lambda \in N_{\Lambda, \varepsilon}} d_{\lambda}(\Lambda, \varepsilon)^2 \right)^{\frac{1}{2}}.$$

If $\eta_{\Lambda, \varepsilon} < \text{tol}$ stop, accept u_{Λ} as solution.

5. *Determine $\tilde{\Lambda}$ with $\Lambda \subset \tilde{\Lambda} \subset \Lambda \cup N_{\Lambda, \varepsilon}$ such that*

$$\left(\sum_{\lambda \in \tilde{\Lambda}} d_{\lambda}(\Lambda, \varepsilon)^2 \right)^{\frac{1}{2}} \geq (1 - \theta) \eta_{\Lambda, \varepsilon}.$$

Set $\tilde{\Lambda} \rightarrow \Lambda$ and go to (3).

Although quite different with regard to its technical ingredients the above algorithm is very similar in spirit to the adaptive scheme proposed in [27] for bivariate piecewise linear finite element discretizations of Poisson's equation. As above the coarsest grid is chosen in [27] in such a way that all errors stemming from data are kept below any desired tolerance. In that sense the approach in [27] has motivated part of the developments described above and in [14].

We wish to add a few more comments on the above scheme. It may not be practically efficient to shoot for the final accuracy in the first step. One would rather select a sequence $\text{tol}_\ell = \frac{1}{2}\varepsilon_{\ell-1}$, $\ell = 1, \dots, N$, say where $\text{tol}_N =: \text{Tol}$ is the final accuracy. One would then proceed as follows:

Algorithm 3 0. Choose $\text{Tol} = 2^{-N}\text{tol}_0$. Set $\text{tol} = \text{tol}_0$.

1. Apply Algorithm 2 with tol .
2. If $\text{tol} \leq \text{Tol}$ stop, accept u_Λ as solution. Otherwise set $\frac{\text{tol}}{2} \rightarrow \text{tol}$ and go to (1).

A brief comment on Step 3 in Algorithm 2 is in order. By (3.6.64), the principal sections of the matrix \mathcal{A} are well-conditioned. This can be used to update a current Galerkin approximation u_Λ as follows. Let $\mathbf{u}_\Lambda := \mathbf{d}_\Psi(u_\Lambda)$ be the vector of wavelet coefficients of u_Λ . To compute the coefficient vector $\mathbf{u}_{\tilde{\Lambda}}$ of $u_{\tilde{\Lambda}}$ we choose an initial approximation \mathbf{v} according to

$$v_\lambda = \begin{cases} u_\lambda, & \lambda \in \Lambda \\ w_\lambda, & \lambda \in \tilde{\Lambda} \setminus \Lambda \end{cases}, \quad (7.3.112)$$

where $\mathbf{w}_{\tilde{\Lambda} \setminus \Lambda} = \mathbf{d}_\Psi(w_{\tilde{\Lambda} \setminus \Lambda})$ are the coefficients of the Galerkin solution $w_{\tilde{\Lambda} \setminus \Lambda}$ of the complement system

$$\langle Aw_{\tilde{\Lambda} \setminus \Lambda}, v \rangle = \langle f, v \rangle, \quad v \in S_{\tilde{\Lambda} \setminus \Lambda},$$

where $S_{\tilde{\Lambda} \setminus \Lambda} := \text{span}\{\psi_\lambda : \lambda \in \tilde{\Lambda} \setminus \Lambda\}$. The corresponding matrix entries have to be determined anyway for the adaptive refinement. Since by (3.6.64), the corresponding section $\mathcal{A}_{\tilde{\Lambda} \setminus \Lambda}$ of \mathcal{A} is well-conditioned only a few conjugate gradient iterations are expected to be necessary to approximate $\mathbf{w}_{\tilde{\Lambda} \setminus \Lambda}$ well enough to provide a good starting approximation of the form (7.3.112) which will then have to be improved by (a few) further iterations on the system matrix $\mathcal{A}_{\tilde{\Lambda}}$.

References

- [1] Andersson, L., N. Hall, B. Jawerth and G. Peters, Wavelets on closed subsets of the real line, in *Topics in the Theory and Applications of Wavelets*, L. L. Schumaker and G. Webb (eds.), Academic Press, Boston, 1994, 1–61.
- [2] Babuska, I., Advances in the p and h-p versions of the finite element method, in *Proc. Singapore Conf. ISNM 86*, Birkhäuser, Basel, 1988, 31–46.
- [3] Babuska, I. and W. C. Rheinboldt, A posteriori error estimates for finite element computations, *SIAM J. Numer. Anal.* **15** (1978), 736–754.
- [4] Bank, R. E. and A. Weiser, Some a posteriori error estimators for elliptic partial differential equations, *Math. Comp.* **44** (1985), 283–301.
- [5] Bornemann, F., B. Erdmann and R. Kornhuber, A posteriori error estimates for elliptic problems in two and three space dimensions, *SIAM J. Numer. Anal.* **33** (1996), 1188–1204.
- [6] Bertoluzza, S., A posteriori error estimates for wavelet Galerkin methods, Preprint Nr. 935, Istituto di Analisi Numerica, Pavia, 1994.
- [7] de Boor, C., R. DeVore and A. Ron, On the construction of multivariate (pre-) wavelets, *Constr. Approx.* **9(2)** (1993), 123–166.
- [8] Chui, C. K. and E. Quak, Wavelets on a bounded interval, in *Numerical Methods of Approximation Theory*, D. Braess and L. L. Schumaker (eds.), Birkhäuser, 1992, 1–24.
- [9] Ciesielski, Z. and T. Figiel, Spline bases in classical function spaces on compact C^∞ manifolds, part I, *Studia Math.* **76** (1983), 1–58.
- [10] Cohen, A., W. Dahmen and R. DeVore, Multiscale decompositions on bounded domains, IGPM-Bericht Nr. 113, RWTH Aachen, 1995.
- [11] Cohen, A., I. Daubechies and J.-C. Feauveau, Biorthogonal bases of compactly supported wavelets, *Comm. Pure and Appl. Math.* **45** (1992), 485–560.

- [12] Cohen, A., I. Daubechies and P. Vial, Wavelets on the interval and fast wavelet transforms, *Appl. Comput. Harmonic Anal.* **1** (1993), 54–81.
- [13] Dahlke, S., Wavelets: Construction principles and applications to the numerical treatment of operator equations, habilitation thesis, RWTH Aachen, 1996.
- [14] Dahlke, S., W. Dahmen, R. Hochmuth and R. Schneider, Stable multiscale bases and local error estimation for elliptic problems, IGPM–Bericht Nr. 124, RWTH Aachen, 1996.
- [15] Dahlke, S. and R. DeVore, Besov regularity for elliptic boundary value problems, IGPM-Bericht Nr. 116, RWTH Aachen, 1995.
- [16] Dahmen, W. and A. Kunoth, Multilevel preconditioning, *Numer. Math.* **63** (1992), 315–344.
- [17] Dahmen, W., A. Kunoth and K. Urban, Biorthogonal spline-wavelets on the interval: Stability and moment conditions, RWTH Aachen, 1996.
- [18] Dahmen, W., S. Pröbldorf and R. Schneider, Multiscale methods for pseudo-differential equations on smooth manifolds, in *Proceedings of the International Conference on Wavelets: Theory, Algorithms, and Applications*, C. K. Chui, L. Montefusco, and L. Puccio (eds.), Academic Press, 1994, 385–424.
- [19] Dahmen, W. and R. Schneider, Wavelets on manifolds, in preparation.
- [20] Daubechies, I., *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Math. **61**, SIAM, Philadelphia, 1992.
- [21] Dauge, M., *Elliptic boundary value problems on corner domains*, Lecture Notes in Mathematics **1341**, Springer, Berlin, 1988.
- [22] DeVore, R., B. Jawerth and V. Popov, Compression of wavelet decompositions, *Amer. J. Math.* **114** (1992), 737–785.
- [23] DeVore, R. and B. Lucier, Wavelets, *Acta Numerica* (1991), 1–56.

- [24] DeVore, R. and V. Popov, Interpolation spaces and nonlinear approximation, in *Function Spaces and Approximation*, M. Cwikel, J. Peetre, Y. Sagher, and H. Wallin (eds.), Springer Lecture Notes in Math., 1988, 191–205.
- [25] DeVore, R. and B. Sharpley, Besov spaces on domains in \mathbb{R}^d , *Trans. Amer. Math. Soc.* **335** (1993), 843–864.
- [26] DeVore, R. and X. M. Yu, Degree of adaptive approximation, *Math. Comp.* **55** (1990), 625–635.
- [27] Dörfler, W., A convergent adaptive algorithm for Poisson’s equation, *SIAM J. Numer. Anal.* **33** (1996), 1106–1124.
- [28] Eriksson, K., P. Hansbo and C. Johnson, Adaptive finite element methods, manuscript, 1994.
- [29] Frazier, M. and B. Jawerth, A discrete transform and decompositions of distribution spaces, *J. of Funct. Anal.* **93** (1990), 34–170.
- [30] Grisvard, P., *Singularities in Boundary Value Problems*, Research Notes in Applied Mathematics **22**, Springer, Berlin, 1992.
- [31] Grisvard, P., Behavior of the solutions of elliptic boundary value problems in a polygonal or polyhedral domain, in *Symposium on Numerical Solutions of Partial Differential Equations III*, B. Hubbard (ed.), Academic Press, New York, 1975, 207–274.
- [32] Hackbusch, W., *Elliptic Differential Equations*, Springer, Berlin, 1992.
- [33] Hansbo, P. and C. Johnson, Adaptive finite element methods in computational mechanics, *Comp. Methods Appl. Mech. Eng.* **101** (1992), 143–181.
- [34] Jia, R. Q., Approximation by smooth bivariate splines on a three direction mesh, in *Approximation Theory IV*, C. K. Chui, L. L. Schumaker, and J. D. Ward, (eds.), Academic Press, 1983, 539–545.
- [35] Jerison, D. and C. E. Kenig, The inhomogeneous Dirichlet problem in Lipschitz domains, *J. of Funct. Anal.* **130** (1995), 161–219.

- [36] Jouini, A. and P. G. Lemarié-Rieusset, Ondelettes sur un ouvert borné du plan, Preprint Nr. 92–46, Université de Paris-Sud, 1992.
- [37] Kondrat'ev, V. A., Boundary–value problems for elliptic equations in domains with conical or angular points, *Trans. Moscow Math. Soc.* **16** (1967), 227–313.
- [38] Kondrat'ev, V. A., On the smoothness of the solutions of the Dirichlet problems for second order elliptic equations in a piecewise–smooth domain, *Differential Equations* **6**, 1392–1401.
- [39] Kondrat'ev, V. A., Singularities of a solution of Dirichlet's problem for a second order elliptic equation in a neighborhood of an edge, *Differential Equations* **13**, 1411–1415.
- [40] Kondrat'ev, V. A. and O. A. Oleinik, Boundary–value problems for partial differential equations in non-smooth domains, *Russian Math. Surveys* **38** (1983), 1–86.
- [41] Kunoth, A. Multilevel preconditioning – Appending boundary conditions by Lagrange multipliers, *Adv. Comp. Math.* **4** (1995), 145–170.
- [42] Lions, J. and E. Magenes, *Problèmes aux limites non homogènes et applications*, Dunod, Paris, 1968.
- [43] Maz'ja, V. G. and B. A. Plamenevskii, Estimates in L_p and in Hölder classes and the Miranda–Agmon maximum principle for solutions of elliptic boundary value problems in domains with singular points on the boundary, *Amer. Math. Soc. Transl. (2)* **123** (1984), 1–56.
- [44] Maz'ja, V. G. and B. A. Plamenevskii, Weighted spaces with nonhomogeneous norms and boundary value problems in domains with conical points, *Amer. Math. Soc. Transl. (2)* **123** (1984), 89–107.
- [45] Maz'ja, V. G. and B. A. Plamenevskii, L_p estimates of solutions of elliptic boundary value problems in a domain with edges, *Trans. Moscow Math. Soc.* **1** (1980), 49–97.
- [46] Maz'ja, V. G. and B. A. Plamenevskii, Boundary value problems on manifolds with singularities, *Problems of Math. Anal.* **6** (1977), 85–142.

- [47] Meyer, Y., *Wavelets and Operators*, Cambridge Studies in Advanced Mathematics, **vol. 37**, Cambridge, 1992.
- [48] Oswald, P., *Multilevel finite element approximation*, Teubner Skripten zur Numerik, Stuttgart, 1994.
- [49] von Petersdorff, T. and C. Schwab, Fully discrete multiscale Galerkin BEM, Research Rep. No. 95-08, Seminar Angew. Mathematik, ETH Zürich.
- [50] Stein, E. M., *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, 1970.
- [51] Tchamitchian, P., Wavelets, functions, and operators, in *Wavelets: Theory and Application*, G. Erlebacher, M. Y. Hussaini, and L. Jameson, (eds.), ICASE/LaRC Series in Computational Science and Engineering, Oxford University Press, New York, 1996.
- [52] Verfürth, R., A posteriori error estimation and adaptive mesh refinement techniques, *Comput. and Appl. Math.* **50** (1994), 67–83.
- [53] Verfürth, R., A review of a posteriori error estimation and adaptive mesh-refinement techniques, manuscript, 1993.

Stephan Dahlke

Institut für Geometrie und Praktische Mathematik
 RWTH Aachen
 Templergraben 55
 52056 Aachen
 Germany
 dahlke@igpm.rwth-aachen.de

Wolfgang Dahmen

Institut für Geometrie und Praktische Mathematik
 RWTH Aachen
 Templergraben 55
 52056 Aachen
 Germany
 dahmen@igpm.rwth-aachen.de

Ronald A. DeVore
Department of Mathematics
University of South Carolina
Columbia, S.C. 29208
U.S.A.
devore@math.sc.edu