Multilevel Frames and Riesz Bases in Sobolev Spaces

Peter Oswald

Bell Labs, Lucent Technologies, Rm. 2C403

600 Mountain Av.

Murray Hill, NJ 07974-0636

e-mail: poswald@research.bell-labs.com

www: http://cm.bell-labs.com/who/poswald

Abstract

This is the continuation of [1] where we specialize to multilevel space splittings for the solution of elliptic variational problems in Sobolev spaces. Starting from an increasing, dense sequence of subspaces

$$V_0 \subset V_1 \subset \ldots \subset V_j \subset \ldots$$

of $L_2(\Omega)$, where each V_j is equipped with a L_2-stable basis $\Phi_j = \{\phi_{j,i}\}$, we introduce multilevel systems by selecting bases $\Psi_j = \{\psi_{j,i}\}$ in certain 'detail' spaces W_j which complement V_{j-1} to the next space V_j in the sequence: $V_j = V_{j-1} + W_j$. Naturally, the system

$$F = \Phi_0 \cup \Psi_1 \cup \ldots \cup \Psi_j \cup \ldots$$

and its finite sections

$$F_J = \Phi_0 \cup \Psi_1 \cup \ldots \cup \Psi_J \ , \quad J \geq 0 \ ,$$

are candidates for multilevel frames resp. Riesz bases in function spaces on Ω and their restrictions to V_J, respectively. We derive the recursive strucutre of additive and multiplicative subspace correction methods associated with such multilevel systems, with emphasis on potential applications to the solution of PDE discretizations. Then we give a list of wavelet systems on $\mathbb{R}^1$ which are based on shifts and dyadic dilates of one or several ϕ (scaling functions) or ψ (wavelets) and discuss their properties with respect to the scale of Sobolev spaces. Finally, multilevel finite element frames and Riesz bases for Sobolev spaces on $\mathbb{R}^d$ and bounded polyhedral domains will be considered.

1 Introduction to multilevel systems

1.1 Definitions

We assume familiarity with Sobolev spaces $H^s(\Omega)$ on $\Omega = \mathbb{R}^d$ and on bounded Lipschitz domains $\Omega \subset \mathbb{R}^d$ [3, 67]. All our discussion is for a given sequence of subspaces $V_j \subset L_2(\Omega)$, $j = 0, 1, \ldots$, such that

$$\{0\} \neq V_0 \subset V_1 \subset \ldots \subset V_j \subset \ldots \ , \quad \mathrm{clos}_{L_2}(\cup_j V_j) = L_2(\Omega) \ , \tag{1}$$

We assume that there exist Riesz bases $\Phi_j = \{\phi_{j,i}\}$ for V_j with L_2-condition uniformly bounded with respect to $j \geq 0$. This means, that each $v_j \in V_j$ possesses a unique decomposition

$$v_j = \sum_i c_{j,i} \phi_{j,i} \tag{2}$$

such that

$$c_0 \|v_j\|_{L_2}^2 \leq \alpha_j^2 \sum_i c_{j,i}^2 \leq c_1 \|v_j\|_{L_2}^2 \tag{3}$$

for some scaling constants $\alpha_j > 0$ and constants $0 < c_0 \leq c_1 < \infty$ which are independent of v_j and (only the latter) of $j \geq 0$. If all these properties are satisfied, we will call $\{V_j\}$ a *multiresolution analysis* (MRA) in $L_2(\Omega)$ (this terminology goes back to Mallat and is often used in a more restrictive setting). We could have weakened the above assumptions by only requiring the frame property of Φ_j in $V_j \subset L_2(\Omega)$ but will not pursue this possibility further.

An immediate consequence of (1) and (3) is that all basis functions from V_{j-1} possess unique decompositions

$$\phi_{j-1,k} = \sum_i a_{j,i}^k \phi_{j,i} \tag{4}$$

with respect to Φ_j, $j \geq 1$. The relations (4) will be called *generalized refinement equations* associated with the MRA $\{V_j\}$, and the coefficient sequences $(a_{j,\cdot}^k)$ ϕ*-masks* associated with $\phi_{j-1,k}$.

Examples of dyadic MRA which are generated by a single scaling function $\phi \in L_2(\mathbb{R}^d)$ satisfying the dyadic refinement equation

$$\phi(x) = \sum_{\alpha \in \mathbf{Z}^d} a_\alpha \phi(2x - \alpha) \ , \tag{5}$$

are extensively discussed in connection with wavelet analysis on $\mathbb{R}^d$, see [30, 14]. The extension to the torus $\mathbb{T}^d$ is straightforward. The case of general domains is by no means trivial, already the construction of non-periodic wavelet systems on the interval $[0, 1]$ has triggered a significant number of investigations. Another important model case which covers the more traditional finite element applications is described by the following

Example 8. Let $\Omega \subset \mathbb{R}^d$ be a polyhedral domain equipped with an initial partition $\mathcal{T}_0$ of Ω into simplices (triangles ($d = 2$), tetrahedra ($d = 3$), and so on). Other types of *elementary cells* replacing or complementing simplices may be allowed, e.g., mixtures of triangles and quadrilaterals are popular in some applications. From $\mathcal{T}_0$, a sequence of partitions $\mathcal{T}_j$ is created by some global refinement process. We will assume the following

- *Monotonicity*: $\mathcal{T}_0 \subset \mathcal{T}_1 \subset \ldots \subset \mathcal{T}_j \subset \ldots$.
- *Regularity*: Simplices are uniformly shape-regular, intersections of different simplices in $\mathcal{T}_j$ can only be empty, a common vertex, a common edge, etc.
- *Quasi-uniformity*: The diameters of simplices in $\mathcal{T}_j$ are of size $\asymp 2^{-j}$, uniformly in $j \geq 0$.

Given such a sequence $\{\mathcal{T}_j\}$, spaces V_j of piecewise polynomial functions can be introduced in various ways. *Finite element spaces* are usually defined by local interpolation problems (i.e., triples $\{\Delta, P_\Delta, I_\Delta\}$ where $\Delta \in \mathcal{T}_j$ is a simplex, P_Δ a specified class of polynomials, and I_Δ an associated projector of interpolation type which is well-defined on certain classes of functions on Δ (including P_Δ), maps into P_Δ and satisfies $I_\Delta^2 = I_\Delta$). Global smoothness properties for functions in V_j are achieved by sharing interpolation conditions with support on the intersection of neighbored simplices, compare [18] for examples. *Spline spaces* are characterized by fixing a space of polynomials P and a global smoothness parameter $r = -1, 0, \ldots$, and defined as

$$S_P^r(\mathcal{T}) = \{u \in C^r(\Omega) \,:\, u|_\Delta \in P \quad \forall\, \Delta \in \mathcal{T}\}\,,$$

where $C^{-1}(\Omega) := L_\infty(\Omega)$. For $\Omega = \mathbb{R}^d$ and $\mathcal{T}_j$ which are invariant under 2^{-j}-shifts, an alternative are *box spline spaces* where certain basis systems Φ_j of piecewise polynomial functions are used to create V_j (see [8]). These possibilities cover a variety of numerical discretization schemes for partial differential equations. Whether the spaces V_j created in this way satisfy the above assumptions of a MRA depends on the choices made for P, I, and r, respectively. The crucial properties to be verified are the monotonicity in (1) which is often violated, and the existence of suitable Riesz bases Φ_j. For computer implementations, the case of locally supported functions $\phi_{j,i}$, with only a few non-zero terms in the generalized refinement equations (4), is of interest.

The simplest case in which a MRA is obtained is that of *linear finite element spaces*. It is covered by either approach: Take $P = P_\Delta = \mathbb{P}_1$ (the space of all polynomials of total degree ≤ 1 in $\mathbb{R}^d$), $r = 0$, and consider linear interpolation at the vertices of a simplex Δ as I_Δ. The basis functions $\phi_{j,i}$, the so-called *hat functions*, are associated with the vertices of $\mathcal{T}_j$. More precisely, if $\{P_{j,i}\}$ is the set of all vertices in $\mathcal{T}_j$ then $\phi_{j,i} \in V_j$ is characterized by the requirements $\phi_{j,i}(P_{j,i}) = 1$ and $\phi_{j,i}(P_{j,k}) = 0$ for all remaining vertices with $k \neq j$. Clearly, the support $K_{j,i}$ of $\phi_{j,i}$ coincides with the union of all simplices $\Delta \in \mathcal{T}_j$ that have $P_{j,i}$ as a vertex. Formula (3) holds since

$$\int_\Delta |p|^2 \asymp |\Delta| \sum_{P_{j,i} \in \Delta} |p(P_{j,i})|^2$$

for any $p \in \mathbb{P}_1$ and any $\Delta \in \mathcal{T}_j$, and the definition of the $\phi_{j,i}$ yields $c_{j,i} = v_j(P_{j,i})$ in (2). In this application, $\alpha_j \asymp 2^{-jd/2}$ would be a correct choice of the scaling constants in (3). We will come back to this example later on.

Multilevel systems associated with a MRA are given by at most countable collections

$$\Psi_j = \{\psi_{j,i}\} \subset V_j \ , \quad j \geq 1 \ , \tag{6}$$

of functions $0 \not\equiv \psi_{j,i} \in V_j$ such that $\Phi_{j-1} \cup \Psi_j$ is a *representation system* in V_j, i.e., for each $v_j \in V_j$, there exists at least one (L_2-convergent) representation

$$v_j = \sum_i c_{j-1,i}\phi_{j-1,j} + \sum_k d_{j,k}\psi_{j,k} \ .$$

In particular, this is satisfied if Ψ_j is a Riesz basis in its L_2-closed span $W_j = [\Psi_j]$ and if each $v_j \in V_j$ admits at least one representation $v_j = v_{j-1} + w_j$, where $v_{j-1} \in V_{j-1}$, $w_j \in W_j$. If the latter representation is unique then V_j is the *direct sum* of V_{j-1} and W_j which will be denoted by $V_j = V_{j-1} \dot{+} W_j$. In any case, since $\Psi_j \subset V_j$ we have unique representations

$$\psi_{j,k} = \sum_i b^k_{j,i}\phi_{j,i} \tag{7}$$

for all $\psi_{j,k} \in \Psi_j$, the coefficient sequences $(b^k_{j,\cdot})$ are called *ψ-masks.*

The *multilevel system*

$$F = \Phi_0 \cup \Psi_1 \cup \ldots \cup \Psi_j \cup \ldots \tag{8}$$

resp. its sections

$$F_J = \Phi_0 \cup \Psi_1 \cup \ldots \cup \Psi_J \ , \qquad J \geq 0 \ ,$$

and their scaled versions will be investigated as candidates for frames in function spaces on Ω resp. in V_J in connection with possible applications to the efficient iterative solution of variational problems arising from PDE. In the remainder of this subsection we mention some simple results concerning the frame property of multilevel systems in Sobolev spaces which happen to be the natural *energy spaces* used in the variational approach to elliptic (and parabolic) PDEs. The corresponding algorithms resulting along the lines of [1, Section 4], especially, their arithmetical complexity and factors that influence it, are discussed in some detail in the next subsection 1.2.

The following result appears to be (after some transformations) a special case of [1, Example 7].

Theorem 1 *Let be given a MRA, and set $\Psi_j = \Phi_j$, $j \geq 0$. Assume that the pair $(L_2(\Omega), H^\gamma(\Omega))$ satisfies Jackson-Bernstein inequalities with respect to $\{V_j\}$ and with constants $\gamma > 0$ and $a > 1$, i.e.,*

$$e_j(u)_{L_2} \equiv \inf_{v_j \in V_j} \|u - v_j\|_{L_2} \leq Ca^{-\gamma j}\|u\|_{H^\gamma} \qquad \forall\ u \in H^\gamma(\Omega) \ ,$$

and

$$\|v_j\|_{H^\gamma} \le C a^{\gamma j} \|v_j\|_{L_2} \qquad \forall\, v_j \in V_j$$

hold with constants C independent of $j \ge 0$. Then, for $0 < s < \gamma$, the scaled system

$$F' = \{(a^{js}\alpha_j)^{-1}\phi_{j,i} \,:\, \phi_{j,i} \in \Phi_j\,,\; j \ge 0\}\,,$$

where α_j is defined in (3), is a frame in $H^s(\Omega)$. Its sections F'_J form frames in V_J (the latter considered as subspaces of $H^s(\Omega)$), with a possible choice of frame bounds that are independent of $J \ge 0$.

Proof. We sketch the main steps. According to [1, Example 7], the following space splitting is stable:

$$[L_2(\Omega), H^\gamma(\Omega)]_{s/\gamma,2} = \sum_{j=0}^{\infty} \{V_j; a^{2js}(\cdot,\cdot)_{L_2}\}\,, \quad 0 < s < \gamma\,.$$

To this end, set $V = H^\gamma(\Omega)$ and $H = L_2(\Omega)$ there. For domains satisfying the extension property (e.g., Lipschitz domains), we have

$$H^s(\Omega) = [L_2(\Omega), H^\gamma(\Omega)]_{s/\gamma,2}\,, \quad 0 < s < \gamma\,,$$

with spectrally equivalent Hilbert space structure. This gives the norm equivalence

$$\|u\|_{H^s}^2 \asymp \inf_{v_j \in V_j\,:\, u = \sum_{j\ge 0} v_j} \sum_{j \ge 0} a^{2sj} \|v_j\|_{L_2}^2 \quad \forall\, u \in H^s(\Omega)\,.$$

Using the L_2-stability (3) of the basis Φ_j in V_j, this implies

$$\begin{aligned}
\|u\|_{H^s}^2 &\asymp \inf_{u = \sum_{j\ge0}\sum_i c_{j,i}\phi_{j,i}} \sum_{j\ge 0}\sum_i a^{2sj}\alpha_j^2 c_{j,i}^2 \\
&= \inf_{u = \sum_{j\ge0}\sum_i c'_{j,i}\phi'_{j,i}} \sum_{j\ge 0}\sum_i (c'_{j,i})^2\,,
\end{aligned}$$

where we have denoted $\phi'_{j,i} = (a^{js}\alpha_j)^{-1}\phi_{j,i}$ (note that this step can be viewed as an example of the refinement technique for stable splittings mentioned in [1, Section 5]).

But this is equivalent to the frame property of F' with respect to $H^s(\Omega)$, compare [1, Theorem 2]. The result for the finite sections F'_J is obtained in complete analogy, if one starts from the uniform stability of the subsplittings

$$\{V_J; (\cdot,\cdot)_{H^s}\} = \sum_{j=0}^{J} \{V_j; a^{2js}(\cdot,\cdot)_{L_2}\}\,,$$

which was already established in [1, Example 7] in the abstract setting.

Theorem 1 generates an initial multilevel frame for Sobolev spaces from basic approximation-theoretic properties of the underlying MRA $\{V_j\}$. A classical way to check Jackson-Bernstein inequalities (such as needed in Theorem 1) for functions on (bounded) domains, is to use *moduli of smoothness* and *Besov space techniques*, see [56, 27]. An example of such techniques is given in Section 3.1 below. For the case of MRAs on $\mathbb{R}^d$ which are obtained by translation and dilation, Fourier transform techniques are usually prefered. Analogously, Fourier series are the tool for MRA on $\mathbb{T}^d$ and on smooth closed curves (a case which is similar to $\mathbb{T}^1$).

Note that by interpolation from the Bernstein inequality, and from (3), we have

$$\|\phi_{j,i}\|_{H^s} \leq C a^{sj} \|\phi_{j,i}\|_{L_2} \asymp a^{sj}\alpha_j$$

for all basis functions in F. Thus, the scaling constants $a^{sj}\alpha_j$ which make F into F' could be replaced by $a^{sj}\|\phi_{j,i}\|_{L_2}$ or, if also the opposite inequality

$$a^{sj}\|\phi_{j,i}\|_{L_2} \leq C\|\phi_{j,i}\|_{H^s}$$

holds true, by $\|\phi_{j,i}\|_{H^s}$, without changing the result of Theorem 1. However, such modifications have impact on the size of the frame bounds, with consequences for the practical performance of iterative solvers associated with the multilevel systems.

As was already discussed in [1], the result does not extend to L_2 ($s = 0$) neither to $s < 0$ (in this section, we take $H^{-s}(\Omega) = H^s(\Omega)'$, $s > 0$, as the definition of Sobolev spaces of negative order by duality which is not exactly the standard if Ω is a bounded domain, see Section 3.1 for some explaining words). A possible repair for this drawback is as follows. Assume that

$$\Psi_j = \{\psi_{j,i}\} \subset W_j = V_j \ominus_{L_2(\Omega)} V_{j-1} \tag{9}$$

is a Riesz basis in the orthogonal complement spaces W_j, $j \geq 1$. I.e., in analogy to (3) we assume that each $w_j \in W_j$ possesses a unique decomposition with respect to Ψ_j such that

$$w_j = \sum_i d_{j,i}\psi_{j,i} \ : \ \|w_j\|_{L_2}^2 \asymp \beta_j^2 \sum_i d_{j,i}^2 \,, \tag{10}$$

where the constants in the two-sided inequality can be chosen independently of $j \geq 1$ (again, we could have assumed only the frame property of Ψ_j in W_j). The resulting system F as defined by (8) will be called *L_2-semiorthogonal multilevel system* associated with the MRA $\{V_j\}$.

Corollary 2 *Let the MRA satisfy the assumptions of Theorem 1. Then, for any $-\gamma < s < \gamma$, the scaled semiorthogonal multilevel system*

$$F' = \{\alpha_0^{-1}\phi_{0,i}\,;\, \phi_{0,i} \in \Phi_0\} \cup \{(a^{sj}\beta_j)^{-1}\psi_{j,i} \ : \ \psi_{j,i} \in \Psi_j,\ j \geq 1\}$$

forms a Riesz basis in $H^s(\Omega)$. Moreover, its sections F'_J form Riesz bases in V_J (considered as subspaces of $H^s(\Omega)$), with a choice of Riesz bounds that is independent of $J \geq 0$.

Proof. The case $s \geq 0$ has essentially been considered in [1, Example 7]. The following argument covers the case of negative Sobolev exponents. Let $s \in (0, \gamma)$. Since $L_2(\Omega)$ is dense in $H^{-s}(\Omega)$, it is enough to consider $u \in L_2(\Omega)$ in the sequel. Recall that by duality

$$\|u\|_{H^{-s}} = \sup_{0 \neq v \in H^s(\Omega)} \frac{(u,v)_{L_2}}{\|v\|_{H^s}}$$

for all $u \in L_2(\Omega)$. By construction of the W_j, we have a unique L_2-orthogonal decompositions

$$u = \sum_{j \geq 0} w_j \ , \quad v = \sum_{j \geq 0} w'_j \ , \qquad w_j, w'_j \in W_j \quad (W_0 = V_0) \ .$$

By the (already established) result for $H^s(\Omega)$, we have

$$\|v\|^2_{H^s} \asymp \sum_{j \geq 0} a^{2sj} \|w'_j\|^2_{L_2} \ , \quad v \in H^s(\Omega)$$

(the reader may derive this as an exercise along the lines of [1, Example 7]). Thus, for any such u and v,

$$\begin{aligned} |(u,v)_{L_2}| &\leq \sum_j |(w_j, w'_j)_{L_2}| \leq \sum_j a^{-js} \|w_j\|_{L_2} \cdot a^{js} \|w'_j\|_{L_2} \\ &\leq (\sum_j a^{-2sj} \|w_j\|^2)^{1/2} (\sum_j a^{2sj} \|w'_j\|^2_{L_2})^{1/2} \end{aligned}$$

Equality is achieved if we define v from the decomposition of u by setting $w'_j = a^{-2sj} w_j$ (check that this particular v belongs to $H^s(\Omega)$). From these remarks we have

$$\|u\|^2_{H^{-s}} \asymp \sum_j a^{-2sj} \|w_j\|^2_{L_2}$$

for any $u \in L_2(\Omega)$, and by the above mentioned density argument, for all $u \in H^{-s}(\Omega)$. Thus,

$$H^{-s}(\Omega) = \sum_{j \geq 0} \{W_j; a^{-2sj} (\cdot, \cdot)_{L_2}\}$$

is stable. The remaining steps of the proof are as for Theorem 1. The result for the sections F'_J follows immediately since subsplittings of a splitting into a direct sum of subspaces are again stable (with the same or better Riesz bounds).

Examples of semiorthogonal multilevel systems are given below, in some sense, they are still hard to construct in situations that deviate from model cases (although the task is simpler than to construct CONS). A brief discussion of the broader concept of biorthogonal multilevel systems will also be given later. It turns out that the basic result of Theorem 1 is the key to most of the investigations on multilevel systems in Sobolev spaces.

1.2 Multilevel algorithms

For the discussion in this subsection we will accept some further, reasonable restrictions. First of all, we assume that all spaces V_j are *finite-dimensional*, and that the collections Φ_j (obvious) and Ψ_j are finite as well. For simplicity, we will also assume that the numbers

$$n_j = \#\Phi_j = \dim\, V_j \ , \quad m_j = \#\Psi_j$$

grow exponentially, i.e., $n_j/n_{j-1} \geq b > 1$ for all $j \geq 1$, analogously for m_j. Adaptivity applications, where this assumption might be violated, are discussed in later parts. In case that the multilevel system F (resp. its scaled versions and finite sections) is minimal, then necessarily

$$m_j = \dim\, W_j = n_j - n_{j-1} \ , \quad j \geq 1 \ .$$

All this implies that the subspaces $W_j = [\Psi_j] \subset V_j$ are well-defined and finite-dimensional ($\dim\, W_j \leq m_j$), and that the generating system Ψ_j is at least a frame in W_j if we equip W_j with an appropriate scalar product. In most cases, this scalar product will be a scaled L_2-scalar product, in order to cover possible modifications we will identify this scalar product with a spd bilinear form $b_j(\cdot,\cdot)$. Moreover, we will exclusively deal with the finite sections F_J resp. F'_J of a multilevel system such that only $0 \leq j \leq J$ is of interest.

We concentrate on the solution of the restriction of a $H^s(\Omega)$-elliptic variational problem to V_J with $J \geq 0$ temporarily fixed: Find $u_J \in V_J$ such that

$$a(u_J, v_J) = \Phi(v_J) \qquad \forall\ v_J \in V_J \ . \tag{11}$$

To match the notation used throughout [1], we will assume that, for a finite section F_J resp. F'_J, we have defined the one-dimensional subspaces

$$\begin{aligned} V_{j,i} &= [\phi_{j,i}] \qquad & i = 1, \ldots, n_j \ , \ \ j \geq 0 \ , \\ W_{j,i} &= [\psi_{j,i}] \qquad & i = 1, \ldots, m_j \ , \ \ j \geq 1 \ , \end{aligned}$$

(for our convenience, set $W_{0,i} = V_{0,i},\ m_0 = n_0$) and bilinear forms $b_{j,i}$ on $W_{j,i}$ such that the assumed frame property (with respect to $\{V_J; a(\cdot,\cdot)\}$) of the system F'_J is equivalent to the stability of the splitting

$$\{V_J; a\} = \sum_{j=0}^{J} \{W_j; b_j\} = \sum_{j=0}^{J} \sum_{i=1}^{m_j} \{W_{j,i}; b_{j,i}\} \ . \tag{12}$$

The inclusion of the intermediate splitting (involving $\{W_j; b_j\}$) does not restrict the generality (as we have seen so far, it naturally appears in the derivation of multilevel frames and Riesz bases). That operators R_j are not explicitly indicated means that the natural embedding operators of the subspaces into V_J are used.

There are two generic ways to use the splitting (12) for purposes of discretizing and solving (11). They have some common parts but have different potential and bottlenecks.

- *Standard discretization and multilevel preconditioning.* In the standard setup of, say, finite element spaces V_J and a PDE-based variational problem, the ansatz

$$u_J = \sum_{i=1}^{n_J} x_{J,i}\phi_{J,i} \tag{13}$$

substituted into (11), and tested with the set Φ_J leads to a linear spd system

$$A_J x_J = f_J \ , \quad A_J = (a(\phi_{J,i}, \phi_{J,i'}))_{i',i=1,\ldots,n_J} \ , \tag{14}$$

(with right-hand side $f_J = (\Phi(\phi_{J,i'}))^T_{i'=1,\ldots,n_J}$). Typical features of A_J are *sparsity* and *ill-conditioning* (exponential growth of condition numbers if $J \to \infty$). With all the engineering codes in mind that produce and store A_J in large-scale applications, the issue seems to be just the efficient *preconditioning* of A_J for the stable and fast solution of the linear system (14). Looking at the theory outlined in [1], we guess that the splitting (12) will provide suitable preconditioners C_J. In this section we will have some look at their structure, and the connection with standard multigrid algorithms for the same type of problem.

- *Multilevel discretization and standard solution.* At least, if F'_J is minimal (i.e., the splittings in (12) are into a direct sum of subspaces, and F'_J is a Riesz basis not only a frame in V_J) then the following alternative is available: One can discretize directly with respect to F_J,

$$u_J = \sum_{j=0}^{J}\sum_{i=1}^{m_j} \hat{x}_{j,i}\psi_{j,i} \ , \tag{15}$$

which leads to another linear spd system

$$\hat{A}_J \hat{x}_J = \hat{f}_J \ , \tag{16}$$

where the matrix $\hat{A}_J$ can be viewed as $(J+1)\times(J+1)$ block matrix with blocks

$$\hat{A}_{j,j'} = (a(\psi_{j',i'}, \psi_{j,i}))_{i'=1,\ldots,m_{j'},i=1,\ldots,m_j} \ , \quad j',j = 0,\ldots,J \ .$$

The vectors $\hat{x}_J$, $\hat{f}_J$, are analogously composed from sections corresponding to the different levels $j = 0,\ldots,J$. This matrix is *well-conditioned after diagonal scaling* which is nothing but the application of the scaling factors for deriving F'_J from F_J (this diagonal scaling can be avoided if the discretization is performed directly with respect to F'_J). Thus, any *standard iterative method* such as Richardson- or cg-iteration would do. The drawback is that in cases where A_J is extremely sparse, the new matrix $\hat{A}_J$ is less sparse. There is a close connection between A_J and a natural factorization of the multilevel preconditioner C_J, on the one hand, and $\hat{A}_J$, on the other, which would result in a fast implementation of matrix-vector products involving $\hat{A}_J$, and avoid the explicit computation and storage of $\hat{A}_J$.

However, this would be just a reformulation of the above method based on (14) and the multilevel preconditioning introduced through (12).

The reduced sparsity of $\hat{A}_J$ is no longer a disadvantage if the matrix A_J itself is not extremely sparse. E.g., *integral equations* but also a number of PDE problems with global low-order terms or some applications to data analysis and denoising typically lead to *dense matrices* when the standard basis representation (13) is used. Although this does not change if we turn to (15) but a special design of the detail spaces W_j resp. of the systems Ψ_j can increase the 'diagonal dominance' in the resulting $\hat{A}_J$: it has often much less significant entries and *matrix compression* can sparsify it dramatically. In other words, the system (16) can often be replaced by another linear spd system

$$A_J^{comp} x_J^{comp} = f_J^{comp} \tag{17}$$

such that the error $\|u_J - u_J^{comp}\|$ remains within the order of the discretization error achievable from V_J which is sparse and well-conditioned at the same time. Speaking in terms of algorithms, this approach compares with replacing $A_J x_J$ by an *approximate matrix-vector multiplication* of lower complexity (see [37, 11]) which has been tried before (and after) the availability of multiscale and wavelet analysis. On the other hand, the fact that whole classes of operators (such as operators of Calderon-Zygmund type) can be essentially diagonalized when represented in wavelet bases [52] will have further impact in scientific computing. See the survey [25], where this approach is emphasized. Clearly, the approach can be carried out also for situations where A_J is sparse although it is less obvious why to do it.

In these notes, we will concentrate on the first approach and emphasize the intimate *connections with* other, more classical methods in the field of scientific computing such as *multigrid and domain decomposition.* From now on, let us talk about solving (14) by some of the abstract Schwarz methods (see [1, Section 4]) associated with the splitting (12). We start with the structure of the Schwarz operator $\mathcal{P}_J$. To prepare the statement of the next theorem, introduce the following matrices:

I_j: These are the $n_j \times n_{j-1}$ matrices representing the natural embedding $V_{j-1} \to V_j$ in the bases Φ_{j-1} resp. Φ_j. The $k-th$ column ($k = 1, \ldots, n_{j-1}$) is given by the coefficients $a_{j,i}^k$, $i = 1, \ldots, n_j$, of the ϕ-mask associated with $\phi_{j-1,k}$, see (4).

$\hat{I}_j$: These $n_j \times m_j$ matrices describe the transformation of a Ψ_j-representation into the basis representation in V_j, i.e., with respect to Φ_j. Clearly, the columns of $\hat{I}_j$ contain the coefficients of the ψ-masks given by (7).

S_j: Define these diagonal $m_j \times m_j$ matrices as $S_j = \mathrm{diag}(b_{j,i}(\psi_{j,i}, \psi_{j,i}) : i = 1, \ldots, m_j)^{-1}$, $j \geq 0$ (for $j = 0$, set $\psi_{0,i} = \phi_{0,i}$).

Theorem 3 *The Schwarz operator $\mathcal{P}_J$ of the space splitting (12) associated with the section F'_J of a multilevel frame takes the form*

$$\mathcal{P}_J u_J = \sum_{j=0}^{J} \sum_{i=1}^{m_j} \frac{a(u_J, \psi_{j,i})}{b_{j,i}(\psi_{j,i}, \psi_{j,i})} \psi_{j,i} \qquad \forall\, u_J \in V_J \; . \tag{18}$$

The matrix representation of $\mathcal{P}_J : V_J \to V_J$ with respect to the basis Φ_J of V_J is given by $C_J A_J$ where C_J is spd and recursively given by

$$C_0 = S_0 \; , \qquad C_j = I_j C_{j-1} I_j^T + \hat{I}_j S_j \hat{I}_j^T \; , \; j = 1, \ldots, J \; . \tag{19}$$

Proof. Solving the one-dimensional auxiliary problems associated with $\{W_{j,i}; b_{j,i}\}$ (and, thus, determining $R^*_{j,i} u_J$) means to find the constant c in $R^*_{j,i} u_J = c\psi_{j,i}$ such that $c b_{j,i}(\psi_{j,i}, \psi_{j,i}) = a(u, \psi_{j,i})$. This gives (18).

Consequently turning to matrix representations (note that the iterated matrix $I_J \ldots I_{j+1}$ corresponds to the natural embedding $V_j \subset V_J$) we have

$$\begin{aligned}
(a(u_J, \phi_{J,i}))^T_{i=1,\ldots,n_J} &= A_J x_J \; , \\
(a(u_J, \phi_{j,i}))^T_{i=1,\ldots,n_j} &= I^T_{j+1} \ldots I^T_J A_J x_J \; , \; j < J \\
(a(u_J, \psi_{j,i}))^T_{i=1,\ldots,m_j} &= \hat{I}^T_j I^T_{j+1} \ldots I^T_J A_J x_J \; , \; j \le J \; , \\
\left(\frac{a(u_J, \psi_{j,i})}{b_{j,i}(\psi_{j,i}, \psi_{j,i})}\right)^T_{i=1,\ldots,m_j} &= S_j \hat{I}^T_j I^T_{j+1} \ldots I^T_J A_J x_J \; , \; j \le J \; .
\end{aligned}$$

This gives the coefficients in front of $\psi_{j,i}$ in the sum representation (18) of $\mathcal{P}_J u_J$. Now it remains to recursively use the matrix representations $I_j, \hat{I}_j$, $j \ge 1$, of the embeddings to arrive at the coefficients of $\mathcal{P}_J u_J$ with respect to the basis Φ_J. Altogether, we have (19).

The symmetry of C_J is obvious from the recursion (19), the positive definitness follows from

$$0 < a(\mathcal{P}_J u_J, u_J) = (A_J C_J A_J x_J, x_J) = (C_J y_J, y_J) \quad \forall\, y_J = A_J x_J \ne 0 \; ,$$

since A_J is invertible, $y_J \ne 0$ implies $x_J \ne 0$ and $u_J \ne 0$, and $\mathcal{P}_J$ was spd with respect to $a(\cdot, \cdot)$.

Before we come to the multiplicative algorithms, let us recall that the additive Schwarz method **(AS)** defined in [1, Section 4] is given by the stationary linear iteration

$$x_J^{(n+1)} = x_J^{(n)} + \tilde{C}_J r_J^{(n)} \; , \quad r^{(n)} = f^{(n)} - A_J x_J^{(n)} \tag{20}$$

where one has to put $\tilde{C}_J = \omega C_J$ with some relaxation parameter $\omega > 0$ (note that the pcg-method **(SCG)** is quite similar but nonstationary, the only change is that the relaxation parameter is now computed from $x^{(n)}$: $\omega = \lambda^{(n)}$). If the matrix-vector multiplication with C_J is implemented according to the recursion (19) than the 'parallelism' of visiting the auxiliary subspace problems is obviously lost, and the preconditioning step reminds us more of a multigrid V-cycle.

To make this precise, and to move to the multiplicative Schwarz iteration **(MS)**, let us formulate a general *multigrid* $\mathrm{V}^{(1,0)}$*-cycle iteration*, with *coarse grid matrices* $\tilde{A}_j$, *smoothing matrices* $\tilde{S}_j$, $j = 0, \ldots, J$, and *restriction and prolongation matrices* $\tilde{R}_j, \tilde{P}_j$, $j = 1, \ldots, J$, by writing it in the form of a stationary linear iteration (20) and describing the matrix-vector multiplication with the corresponding preconditioner $\tilde{C}_J$.

$\mathbf{V}^{(1,0)}$-cycle preconditioner: Given a vector x_j, the vector $\tilde{x}_j = \tilde{C}_j x_j$, $j = 0, \ldots, J$, is computed by the following recursion:

a) For $j = 0$, set $\tilde{x}_0 = \tilde{S}_0 x_0$ (i.e., $\tilde{C}_0 = \tilde{S}_0$).

b) For $j = 1, \ldots, J$ do

1. $y_j = \tilde{S}_j x_j$ (corresponds to pre-smoothing)
2. $x_{j-1} = \tilde{R}_j (x_j - \tilde{A}_j y_j)$ (residual restriction)
3. $\tilde{x}_{j-1} = \tilde{C}_{j-1} x_{j-1}$ (coarse-grid-correction)
4. $\tilde{y}_j = \tilde{P}_j \tilde{x}_{j-1}$ (prolongation)
5. $\tilde{x}_j = y_j + \tilde{y}_j$ (update)

More general V-cycles are described and investigated in [9] (using the language of space splittings) and [36] (in a more traditional algebraic way). E.g., pre-smoothing can be applied repeatedly, post-smoothing can be implemented after step 4. The number of pre- and post-smoothing steps can vary with j (this is quite natural to add robustness and is called *variable V-cycle iteration*). If the coarse-grid correction is done twice, so-called *W-cycles* result. Standard multigrid algorithms use $\tilde{A}_j = A_j$ and $S_0 = A_0^{-1}$ (exact solution on the coarsest level $j = 0$).

Interesting enough, both **(AS)** and **(MS)** when applied with the splitting (12) are special instances of a $\mathrm{V}^{(1,0)}$-cycle iteration. Let us make precise which **(MS)** is considered: We apply **(MS)** to the second splitting in (12) with the choice

$$b_j(\sum_{i=1}^{m_j} w_{j,i}, \sum_{i=1}^{m_j} w'_{j,i}) = \sum_{i=1}^{m_j} b_{j,i}(w_{j,i}, w'_{j,i}) \ ,$$

where $w_{j,i}, w'_{j,i} \in W_{j,i}$, $i = 1, \ldots, m_j$, $j = 0, \ldots, J$. The order of using the auxiliary subspace problems in the multiplicative algorithms is reverse in $j = J, J-1, \ldots, 0$ (and arbitrary within each level).

Theorem 4 *The algorithms* **(AS)** *and* **(MS)** *when applied to the splitting (12) in the way as indicated above can be interpreted as multigrid* $\mathrm{V}^{(1,0)}$*-cycle iteration if the following choices are made for the components:*

$$\tilde{P}_j = I_j \;, \quad \tilde{R}_j = I_j^T \;, \qquad j = 1, \ldots, J \;,$$

$$\tilde{A}_j = \begin{cases} 0 & \text{for} \quad \textbf{(AS)} \\ I_{j+1}^T \ldots I_J^T A_J I_J \ldots I_{j+1} & \text{for} \quad \textbf{(MS)} \end{cases} \qquad j = 0, \ldots, J \;,$$

and

$$\tilde{S}_j = \begin{cases} \omega \hat{I}_j S_j \hat{I}_j^T & , \qquad j = 1, \ldots, J \\ \omega S_0 & , \qquad j = 0 \end{cases} \qquad .$$

Proof. The case of the additive Schwarz iteration is left upon the reader, see Theorem 3. For the multiplicative case, see [36, 11.6.2] (differences in the notation!). We give the elementary argument. Write the iteration **(MS)** in matrix notation. To make this compact, recall that

$$\omega \hat{T}_j = \omega I_J \ldots I_{j+1} \hat{I}_j S_j \hat{I}_j^T I_{j+1}^T \ldots I_J^T = I_J \ldots I_{j+1} \tilde{S}_j I_{j+1}^T \ldots I_J^T \equiv K_j \;, \quad j = 1, \ldots, J \;,$$

analogously for $j = 0$, and that the order of visiting the subspaces is reverse. Thus, if the vector corresponding to the old iterate $u^{(n)} \in V_J$ is denoted by $x_J = z^{(J+1)}$, the vectors corresponding to the intermediate $v^{(j)}$ in **(MS)** are recursively computed by

$$z^{(j)} = z^{(j+1)} + K_j (f_J - A_J z^{(j+1)}) \;, \quad j = J, \ldots, 0 \;,$$

and $z^{(0)}$ gives the vector corresponding to $u^{(n+1)}$. Introducing the residuals $r^{(j)} = f_J - A_J z^{(j)}$, we obtain

$$r^{(j)} = (\mathrm{Id}_J - A_J K_j) r^{(j+1)} \;, \quad j = J, \ldots, 0 \;,$$

from which the formula

$$r^{(0)} = (\mathrm{Id}_J - A_J K_0) \ldots (\mathrm{Id}_J - A_J K_J) r^{(J+1)}$$

follows which expresses the residual $r^{(0)}$ for $u^{(n+1)}$ by the residual $r^{(J+1)}$ of $u^{(n)}$. Using the same notation, from the description of the $\mathrm{V}^{(1,0)}$-cycle we obtain

$$\tilde{C}_0 = \tilde{S}_0 \;, \quad \tilde{C}_j = \tilde{S}_j + I_j \tilde{C}_{j-1} I_j^T (\mathrm{Id}_j - \tilde{A}_j \tilde{S}_j) \;, \; j = 1, \ldots, J \;.$$

Multiplying the equation for $\tilde{C}_j$ by $I_J \ldots I_{j+1}$ from the left and, accordingly, by $I_{j+1}^T \ldots I_J^T$ from the right, and using the formula for the Galerkin coarse grid matrices $\tilde{A}_j$, we see after trivial transformations that

$$\tilde{K}_j \equiv I_J \ldots I_{j+1} \tilde{C}_j I_{j+1}^T \ldots I_J^T = K_j + \tilde{K}_{j-1} (\mathrm{Id}_J - A_J K_j) \;,$$

$j \geq 1$, $\tilde{K}_0 = K_0$, and

$$\mathrm{Id}_J - A_J \tilde{K}_j = (\mathrm{Id}_J - A_J \tilde{K}_{j-1})(\mathrm{Id}_J - A_J K_j) \;, \quad j \geq 0 \;,$$

if we agree about $\tilde{K}_{-1} = 0$. Since $\tilde{K}_J = \tilde{C}_J$ we see from (20) that the residual propagation of the $\mathrm{V}^{(1,0)}$-cycle iteration (with the specifications of its components given in Theorem 4) is

$$(\mathrm{Id}_J - A_J K_0) \dots (\mathrm{Id}_J - A_J K_J) ,$$

i.e., the same as the one derived above directly from **(MS)**. This proves the Theorem.

We leave it upon the reader to formulate the correct multigrid counterpart of **(SMS)** when applied to the splitting (12) (clearly, a symmetric $\mathrm{V}^{(1,1)}$-cycle should come out).

While the choices made for our multilevel **(MS)** correspond to a $\mathrm{V}^{(1,0)}$-cycle with *Galerkin coarsening* and *Richardson(-Jacobi) smoothing* in the multigrid context, and are therefore relatively restrictive, the convergence theory given in [1, Section 4] is directly applicable (the stability assumptions follow from the Theorems of the previous subsection). Who is interested in the theory of the whole variety of multigrid algorithms should consult [36] or [9] (warning: [9] is quite technical).

Theorem 4 also gives a chance to understand the differences of additive and multiplicative multilevel Schwarz methods, and why the multiplicative version should be preferred in practical computations (this does not mean that we recommend to fully neglect the additive Schwarz methods - as we have seen, their theoretical understanding is an important prerequisite for the multiplicative case, in applications they might be advantegeous if robustness is more important than efficiency). In all other aspects (operation count per iteration, potential for parallelization etc.), there seems to be no serious reason not to use the **(MS)**-based algorithms. The additional computation of residuals by using the coarse-grid matrices $\tilde{A}_j$ in step 2 of the V-cycle implementation usually pays off in improved convergence rates. A theoretical justification of this often numerically observed fact is only available in special cases (it is somehow analogous to the classical comparison between Jacobi and Gauss-Seidel/SOR methods which fit into the general picture according to [1, Example 5]).

We conclude with a short discussion of the *arithmetical complexity* of multilevel iterations if $J \to \infty$. Most of the components of the above algorithms associated with a multilevel system are fixed if a choice for the MRA $\{V_j\}$ is made. This concerns the stiffness matrices A_J and the prolongations/restrictions I_j, I_j^T which are determined by the generalized refinement equations for the bases Φ_j in V_j, $j \geq 0$. The desired sparsity of A_J and I_j and generally time/storage restrictions limit the use of more involved approximation schemes, e.g., high order methods tend to achieve (for smooth solutions) the desired accuracy with smaller J but at the expense of denser A_J and I_j. The appropriate choice of the MRA also depends on the problem at hand.

To simplify the task, let us assume that the MRA is fixed. Then the fixed costs related to precomputing A_J and of its Galerkin projections $\tilde{A}_j$ on coarser levels $j \leq J$ can be estimated (the latter need only be accounted for in the multiplicative methods). The same is true for the costs per iteration, i.e., for residual calculations involving the multiplication by A_J resp. by $\tilde{A}_j$, $j \leq J$, and the application of prolongations and restrictions. The choice of $\{V_j\}$ also fixes the approximation error of the solution.

The choice of Ψ_j has mainly impact on two properties: First it may heavily influence the frame condition $\kappa_J = \kappa(F'_J)$ and, thus, the convergence rate of the multilevel Schwarz methods. Secondly, the *average mask size* of the ψ-masks in (7) determines the costs of the smoothing operations (here it usually does not matter whether $\hat{I}_j, S_j, \hat{I}_j^T$ are applied separately, for each j and each iteration, or their product $\hat{I}_j S_j \hat{I}_j^T$ is precomputed).

What is clear is that the operation count per iteration $\#Ops_J$ of the above multilevel methods for solving a problem on V_J usually satisfies the bound

$$\#Ops_J \leq Cn_J \ , \tag{21}$$

at least, if $m_j \asymp n_j$, if A_J is sparse, with $\leq C_A n_J$ non-zero entries, and if the average size of ϕ- resp. ψ-masks is uniformly bounded ($\leq C_\phi$ resp. $\leq C_\psi$). For **(MS)** one has to require in addition that the precomputed matrices $\tilde{A}_j$ contain $\leq C_{\tilde{A}} n_j$ non-zero entries for all $0 \leq j < J$. Then (21) follows from the recursions for computing the $\mathrm{V}^{(1,0)}$-cycle action and the assumed exponential growth of n_j. The constant C in (21) behaves like a linear combination of the constants $C_A, C_{\tilde{A}}, C_\phi, C_\psi$. In Section 3 we will present some more precise asymptotic formulae which express the costs of multilevel solvers for several proposals of systems F in the case of the linear finite element MRA. Usually, iterative methods for solving (11) that satisfy (21) and that exhibit an iteration count $\#It_J$ which is bounded independently of J and changes only with the desired relative error reduction are called *optimal*. Most of the multilevel and wavelet iterative methods designed for solving variational problems (11) with respect to a given MRA are aiming at this optimality.

The task to find Ψ_j such that they realize an optimal compromise between the iteration count $\#It_J$ (e.g., by minimizing κ_J) and costs per iteration $\#Ops_J$ (e.g., by minimizing C_ψ, i.e., the costs for applying $\hat{I}_j, \hat{I}_j^T$) has been achieved so far only on an experimental level, through numerical tests. A theoretical difficulty is the non-availability of efficient methods for computing frame bounds (rather than solving large scale eigenvalue problems), a practical difficulty is to agree on the benchmarks for such a comparison. Another point to make is that in some cases iterative methods may behave better than predicted by κ_J. Again, to see from the choices of Φ_j, Ψ_j, whether this happens or not is an unsolved problem.

2 Wavelet systems on $\mathbb{R}^1$ - an overview

The case of a *dyadic MRA on* $\mathbb{R}$ allows for the most complete theory, both in the L_2- and Sobolev space setting. It is already well-represented in monographs [30, 14, 43] and in survey papers [38, 31], see also the collections [15, 60, 5, 16]. Most parts of this theory are now well-understood also for $\Omega = \mathbb{R}^d$, $d > 1$ (this will be discussed in Subsection 3.2). Since dyadic MRA on $\mathbb{R}$ are the starting point (and partial case) for most of the multilevel constructions on domains in $\mathbb{R}^d$ (by restriction, boundary modification, and tensor-product techniques), we decided to give a brief list of the most 'famous' one-dimensional examples. The exposition is oriented to our application area, we avoid talking about filters and other signal processing terminology.

A dyadic MRA on $\mathbb{R}$ is a MRA for $L_2(\mathbb{R})$, with two additional properties:

$$u(t) \in V_j \quad \Longleftrightarrow \quad u(2^{-j}t) \in V_0 \qquad \text{(dyadic dilation principle)} \tag{22}$$

and that Φ_0 is generated by the $\mathbb{Z}$-shifts of a finite number of so-called *scaling functions* $\phi^1, \ldots, \phi^M \in L_2(\mathbb{R})$ from which we have (by using the dyadic dilation principle) that

$$\Phi_j = \{\phi_{j,k}^m(t) = \phi^m(2^j t - k) \,:\, k \in \mathbb{Z},\ m = 1, \ldots, M\}\ , \quad j \geq 0\ , \tag{23}$$

can be taken as the designated basis systems in the subspaces V_j of the MRA. In contrast to [1, Example 3], we have dropped the scaling factor (which has to be changed anyhow if we go to Sobolev spaces, see Section 1.1).

In technical terms, these assumptions allow for major simplifications through Fourier analysis techniques. This is not the place to go into very details, consult [13, 30, 14, 15, 43]. We just mention a few things closely related to the investigation of a dyadic MRA in connection with Sobolev spaces and multilevel algorithms. For simplicity, let $M = 1$ and write $\phi \equiv \phi^1$ (the case $M > 1$ leads to vector-matrix analogs of the following facts, and is similar in spirit but technically more involved). The refinement equations (4) turn into a single *refinement equation*

$$\phi(t) = \sum_{k \in \mathbf{Z}} a_k \phi(2t - k) \tag{24}$$

which means that we essentially have a single ϕ-mask. In the following, we will assume that ϕ has compact support resp. that the ϕ-mask (a_k) is real and finitely supported. Occasional exceptions will be mentioned. Set $C_\phi = \#\{k \,:\, a_k \neq 0\}$. These assumptions on (24) considerably reduce storage (for I_j) and overhead in multilevel algorithms, and simplify the analysis. E.g., after taking Fourier transforms of both sides, we get the analog of (24) in the frequency domain:

$$\hat{\phi}(\xi) = m(\xi/2)\hat{\phi}(\xi/2)\ , \quad m(\xi) \equiv m_\phi(\xi) = \frac{1}{2}\sum_k a_k e^{-ik\xi}\ . \tag{25}$$

(25) allows us to reconstruct the scaling function $\phi(t)$ from the *symbol* $m(\xi)$ *of the refinement equation*:

$$\hat{\phi}(\xi) = \prod_{j=1}^{\infty} m(2^{-j}\xi)\ . \tag{26}$$

(assuming the normalization $\hat{\phi}(0) = 1$). In our case, $m(\xi)$ is a Laurent polynomial (with real coefficients), and $m(0) = 1$ can be assumed. Necessary and sufficient conditions on the polynomial symbol $m(\xi)$ such that (26) converges in $L_2(\mathbb{R})$ and yields an L_2-solution $\phi(t)$ of the associated refinement equation (24) are due to A. Cohen (see [43, Part II, Chapter 4], the same source deals with H^s-regularity of ϕ which can also be derived from properties of $m(\xi)$). That Φ_j forms a Riesz basis in V_j is essentially equivalent to

$$g_{\phi,0}(\xi), 1/g_{\phi,0}(\xi) \in L_\infty(\mathbb{T}) \ .$$

Necessary and sufficient conditions for the validity of Jackson-Bernstein inequalities as required by Theorem 1 are known, too. E.g., the Jackson inequality

$$\inf_{v_j \in V_j} \|u - v_j\|_{L_2} \leq C 2^{-j\gamma} \|u\|_{H^\gamma} \qquad \forall \ u \in H^\gamma(\mathbb{R}) \ , \ j \geq 0 \ , \tag{27}$$

holds if (and only if)

$$g_{\phi,0}(\xi) - |\hat{\phi}(\xi)|^2 = \mathrm{O}(|\xi|^{2\gamma}) \ , \quad \xi \to 0$$

(see [7]), and the Bernstein inequality

$$\|v_j\|_{H_\gamma} \leq C 2^{j\gamma} \|v_j\|_{L_2} \qquad \forall \ v_j \in V_j \ , \ j \geq 0 \ , \tag{28}$$

holds if (and only if)

$$g_{\phi,\gamma}(\xi) \in L_\infty(\mathbb{T}) \ ,$$

see [50, Lemma 11]. The 2π-periodic function

$$g_{\phi,s}(\xi) = \sum_{k \in \mathbf{Z}} (1 + |\xi + 2\pi k|^2)^s |\hat{\phi}(\xi + 2\pi k)|^2$$

used in all these statements is accessible by knowledge about $m(\xi)$ via (25). Note that the statement $g_{\phi,\gamma}(\xi) \in L_2(\mathbb{T})$ is equivalent to $\phi \in H^\gamma(\Omega)$. What we see is that the regularity of the scaling function ϕ is intimately connected with all other properties of the associated MRA. Without proof, we will give the easy-to-remember formulation of Theorem 1 for the special case of a dyadic MRA.

Corollary 5 *Let ϕ be the compactly supported scaling function associated with a dyadic MRA on $\mathbb{R}$. If $\phi \in H^\gamma(\mathbb{R}^1)$, then for $0 < s < \gamma$ the scaled system*

$$F' = \{2^{(1/2-s)j} \psi_{j,k} \ : \ k \in \mathbb{Z}, \ j \geq 0\}$$

is a frame in $H^s(\mathbb{R})$. The sections F'_J of F' are frames in V_J, considered as subspaces of $H^s(\mathbb{R})$, with frame bounds and condition numbers bounded independently of $J \geq 0$.

Analogous statements can be obtained for dyadic MRA on $\mathbb{R}^d$ [50] and for the case $M > 1$ of MRA generated from several scaling functions.

Needless to say, that Corollary 5 is a less known result. The main-stream research is directed to constructing from a dyadic MRA *orthogonal, semiorthogonal, and biorthogonal wavelet bases* or, generally, *Riesz bases* suitable for $L_2(\mathbb{R})$ and for Sobolev spaces. Below, we will survey some of these constructions. A systematic methodological study of these issues for general MRA was undertaken by Dahmen [23, 24], where further references can be found (see also the recent survey [25]). The crucial moment is the study of pairs of MRA's $\{V_j\}$ and $\{\tilde{V}_j\}$ which are *dual* to each other (for dyadic MRA on $\mathbb{R}$, see [14, Section 5.4]). The construction and investigation of the dual MRA $\{\tilde{V}_j\}$ starting from a given choice of direct sum splittings $V_j = V_{j-1} \dot{+} W_j$, $j \geq 1$, is an interesting issue, even for dyadic MRA. In order to prepare for the examples below, in the dyadic MRA case we create the collections Ψ_j also from a finite number of *wavelets* $\psi^1, \ldots, \psi^L \in V_1$ according to

$$\Psi_j = \{\psi_{j,k}^m(t) = \psi^l(2^{j-1}t - k) \, : \, k \in \mathbb{Z}, \; l = 1, \ldots, L\} \, , \quad j \geq 1 \, . \tag{29}$$

In most cases, it is assumed that these collections form L_2-Riesz systems (i.e., Riesz bases in their span W_j). Note that our notation is consistent with that of the previous section, not with the standard in wavelet papers: W_1 (not W_0) is the span of the integer shifts of the functions ψ^l.

Again, let us restrict to $L = 1$ and denote $\psi = \psi^1$. Since $\psi \in V_1$ we have a unique decomposition

$$\psi(t) = \sum_k b_k \phi(2t - k) \, . \tag{30}$$

In most of the examples we will look for compactly supported ψ resp. for a ψ-mask (b_k) of finite support. In analogy, we set $C_\psi = \#\{k \, : \, b_k \neq 0\}$ and define the symbol of ψ by

$$m_\psi(\xi) = \frac{1}{2} \sum_k b_k e^{-ik\xi} \, .$$

2.1 Orthogonal wavelets

The construction of *orthonormal systems* from a dyadic MRA on $\mathbb{R}$ is a central theme of wavelet theory (and applied harmonic analysis in general). Forerunners of the modern developments in this area are the Haar, Faber-Schauder, Walsh and Franklin systems which have been extended into a whole family of spline bases on $[0, 1]$ and later manifolds in $\mathbb{R}^d$ by Ciesielski and co-workers during the 70-ies (for some mysterical reasons, official wavelet history (see, e.g., [53, 43]) pays tribute to Stroemberg who used these developments in the shift-invariant situation to deal with basis constructions for real Hardy spaces on $\mathbb{R}^d$). Another source from classical analysis are Littlewood-Paley techniques and the modern theory of function spaces based on decomposition techniques. We refer to [68, 53, 43] for more historical information. It should be emphasized that the broad acceptance of the wavelet machinery comes mainly from the successful *synthesis of multiscale modelling, fast algorithms, and harmonic analysis tools* which were evolving in parallel for many years.

A necessary prerequisite is the construction of a scaling function ϕ such that the system of its $\mathbb{Z}$-shifts Φ_0 is an orthonormal basis in V_0. Having such a ϕ then

$$\psi(t) = \sum_k (-1)^k a_{1-k}^* \phi(2t - k)$$

(here, (a_k) is the ϕ-mask and the asterisque denotes complex conjugation) will define a complementary ψ with the same support properties and such that the resulting multilevel system F gives a complete orthogonal system in $L_2(\mathbb{R}^d)$. According to Theorem 1 and its Corollaries 2, 5, the scaled systems F' will then form CONS in $L_2(\mathbb{R})$ (if $s = 0$) and Riesz bases in $H^s(\mathbb{R})$, $-\gamma < s < \gamma$, where γ is the Sobolev regularity of ϕ. Clearly, there are other possible choices for ψ such that the $\mathbb{Z}$-shifts of ψ are a CONS in $W_1 = V_1 \ominus_{L_2} V_0$ but the above one is the default.

A quick calculation shows that, under the above assumptions (i.e., the axioms of a dyadic MRA on $\mathbb{R}$), orthonormality of Φ_0 is equivalent to

$$g_{\phi,0}(\xi) = 1 = |m(\xi)|^2 + |m(\xi + \pi)|^2 , \quad \xi \in \mathbb{T} . \tag{31}$$

The first relation follows by computing the Fourier coefficients of the 2π-periodic function $g_{\phi,0}$, the second is then obvious from

$$g_{\phi,0}(2\xi) = |m(\xi)|^2 g_{\phi,0}(\xi) + |m(\xi + \pi)|^2 g_{\phi,0}(\xi + \pi) ,$$

which can be deduced from (26). Both relations in (31) can be the starting point for constructions. E.g., following Meyer, take any real-valued $C^\infty(\mathbb{R})$ function $g(\xi)$ with support $[-(\pi + \epsilon), \pi + \epsilon]$ ($0 < \epsilon < \pi$ is fixed), and set

$$\hat{\phi}(\xi) = \frac{g(\xi)}{(\sum_k |g(\xi + 2\pi k)|^2)^{1/2}} .$$

This indeed gives a MRA on $\mathbb{R}$, with $\phi \in B_{\pi+\epsilon}$. The symbol of the associated refinement equation is the 2π-periodic continuation of

$$m(\xi) = \begin{cases} \hat{\phi}(2\xi)/\hat{\phi}(\xi) & \text{if} \quad -(\pi + \epsilon)/2 < \xi < (\pi + \epsilon)/2 \\ 0 & \text{if} \quad |\xi - 2\pi| \le (\pi - \epsilon)/2 \end{cases}$$

and in $C^\infty(\mathbb{T})$ by construction (note that for $\epsilon \le \pi/3$ one necessarily has $\hat{\phi(\xi)} = 1$ on the support of $\hat{\phi}(2\xi)$ which simplifies the above formula). Unfortunately, $m(\xi)$ is not a polynomial, and ϕ cannot be compactly supported. See [14, Example 7.3].

In order to produce her famous orthogonal basis of compactly supported wavelets, Daubechies directly constructed polynomial solutions of the form $m(\xi) \equiv m_N(\xi)$ for the second relation in (31), with additional properties and depending on an integer $N \ge 2$. Thus, compact support of the resulting ϕ is obvious. The difficult part is guessing a suitable solution of (31) and establishing the properties of $\phi_N(t)$ via (26). The derivation is well-documented in [30, Chapter 6,7], [43, Part II, Chapter 5], or [14, Section 7.3]. The steps are as follows.

- Check that
$$R(x) = \sum_{k=0}^{N-1} \binom{N+k-1}{k} x^k$$
is a solution of $1 \equiv (1-x)^N R(x) + x^N R(1-x)$.

- Define the non-negative trigonometric polynomial $p(\xi) = R(\sin^2(\xi/2)) \geq 0$, and check that
$$\cos^{2N}(\xi/2)p(\xi) + \sin^{2N}(\xi/2)p(\xi+\pi) = 1 \ .$$

- Use the Riesz theorem to find a Laurent polynomial $q(\xi) = \sum_{k=0}^{N-1} q_k e^{-ik\xi}$ with real coefficients q_k such that $|q(\xi)|^2 = p(\xi)$ (compare [14, Theorem 7.17]).

- Finally, set
$$m_N(\xi) = \left(\frac{1+e^{-i\xi}}{2}\right)^N q(\xi) \ .$$

Since $q(\xi)$ in step 3 is not uniquely defined, there might be several $m_N(\xi)$, see [30, p.159],[43] for tables of the coefficients q_k for several N, and for graphical visualizations. The factor $((1+e^{-i\xi})/2)^N$ guarantees that polynomials of degree $\leq N-1$ can be locally reproduced in V_j which leads to good approximation orders when using the MRA $\{V_j\}$ for discretization purposes. The construction works also for $N=1$ and leads to the refinement equation associated with $\phi_1(t) = \chi_{[0,1]}(t)$. Note that the Sobolev regularity (as well as the C^α-regularity) of ϕ_N asymptotically grows as $\approx N(1-(\log 3)/(\log 4))$ while the support is $[0, 2N-1]$ (clearly, $C_\phi = C_\psi = 2N$, for the latter, see below).

A quite similar construction leads to *coiflets* - see [30, Section 8.2] or [43] - which have larger support but better, explicit polynomial reproduction
$$p(t) = \sum_k p(k)\phi(t-k) \qquad \forall\, p \in \mathbb{P}_{N-1} \ .$$
Finally, we would like to refer to [43, Example 10, p.295ff] for an interesting example of a *non-stationary orthonormal wavelet system* consisting of C^∞ functions for which Corollary 2 is applicable for all $-\infty < s < \infty$. Nonstationary means that the ϕ- and ψ-masks change with j - for the example under consideration their size grows proportional to j. Thus, the price to pay for unlimited regularity of the approximation scheme is the loss of algorithmical optimality (typical A_J will contain $\asymp J^d 2^J$ non-zero elements, and the multilevel V-cycles will take about the same number of arithmetical operations).

It is relatively easy to see that (except for the piecewise constant case) orthonormal systems of compactly supported *spline wavelets* cannot exist. This, and the popularity of spline approximation schemes in applied sciences, has caused a lot of work for weakening the orthogonality concept, see below. More flexibility for constructing orthogonal wavelet systems is available if we go to multiwavelets where typically $M = L > 1$. Recently, such constructions have been proposed by using iterated function systems [32, 51]. See also examples by [17] and [41].

2.2 Interpolatory wavelets

This topic is of interest for two reasons: on the one hand, it is closely related to the orthogonal case, on the other, it is important for the theoretical understanding and practical application of *wavelet collocation methods*.

Suppose that $\{V_j\}$ is a dyadic MRA on $\mathbb{R}$, with scaling function $\phi \in L_2(\mathbb{R}) \cap C(\mathbb{R})$. Then, we call ϕ *interpolatory* if

$$\phi(l) = \delta_l \equiv \begin{cases} 1 & \text{if} \quad l = 0 \\ 0 & \text{if} \quad l \neq 0 \end{cases} . \tag{32}$$

Let I_j denote the interpolation projector into V_j given by

$$(I_j u)(t) = \sum_k u(2^{-j}k)\phi_{j,k}(t)$$

which is well-defined at least on $C_0(\mathbb{R}) \subset L_2(\mathbb{R})$. A natural choice for W_1 resp. ψ is given by

$$W_1 = \operatorname{ran}(I_1 - I_0) , \ \psi(t) = \phi(2t-1) .$$

The potential representation of a continuous function u with respect to the resulting system F will then necessarily have the form

$$u(t) \sim \sum_{k\in\mathbf{Z}} u(k)\phi_{0,k}(t) + \sum_{j=1}^{\infty}\sum_{k\in\mathbf{Z}} (u - I_{j-1}u)(2^{-j}(2k+1))\psi_{j,k}(t) ,$$

and the properties of F with respect to the scale of Sobolev spaces are intimately connected with the properties of the interpolation process $\{I_j\}$. Without proof, let us state that under the assumptions of Corollary 5 and if ϕ is interpolatory, the scaled system F' will be a Riesz bases in $H^s(\mathbb{R})$ if and only if $\frac{1}{2} < s < \gamma$. In higher dimensions, the lower bound is $\frac{d}{2}$.

Under analogous technical conditions, a necessary and sufficient condition for (32) to hold is

$$m(\xi) + m(\xi+\pi) \equiv 1 , \quad \xi \in \mathbb{T} . \tag{33}$$

Comparing with the second relation in (31), we see how orthogonal and interpolatory MRA are related: If ϕ is associated with an orthogonal MRA then the convolution

$$\phi^{new}(t) = \int_{\mathbb{R}} \phi(s)\overline{\phi(s-t)}\, ds$$

leads to an interpolatory MRA since $\widehat{\phi^{new}}(\xi) = |\hat{\phi}(\xi)|^2$ and, thus, $m_{\phi^{new}}(\xi) = |m(\xi))|^2$. The other direction was already explored in Daubechies' construction. If $m(\xi)$ is the symbol associated with an interpolatory MRA, i.e., satisfies (33), and is a non-negative polynomial then, according to the Riesz theorem, there is a polynomial $m^{new}(\xi)$ for which $m(\xi) = |m^{new}(\xi)|^2$. The simplest example of this connection is that of the Haar- and Faber-Schauder systems (take $N = 1$ in Daubechies' construction). The Faber-Schauder system comes from the interpolatory MRA of linear splines, see [1, Figure 1 b),c)] or Figure 1 below for the associated ψ and ϕ, respectively. For the case $N \geq 2$, see [43, Example 2, p. 282ff].

Interpolatory scaling functions arise naturally in interpolatory subdivision schemes a la Deslauries-Dubuc (see [43, Example 1, p.281-282]). *Smooth interpolatory spline* *wavelets* (with infinite support except for the linear case) have been studied in great detail by Chui and Wang, see [14, section 6.1]. The MRA is given by setting $V_0 = S_m^{m-1}(\mathcal{T}_0) \equiv S_{\mathbf{P}_m}^{m-1}(\mathcal{T}_0)$, where m is here the *degree* of the smooth splines under consideration (not the *order* as in [14] which would be $m+1$). $\mathcal{T}_0$ is given by the $\mathbb{Z}$-shifts of the unit interval. This MRA is also called *B-spline MRA* since the m-th degree B-spline may serve as generating scaling function (to be definite, the corresponding $\phi(t)$ is the m-th degree B-spline with respect to $\mathcal{T}_0$ supported on $[0, m+1]$ and scaled such that $\sum_k \phi(t-k) \equiv 1$, the associated symbol is $m(\xi) = 2^{-m-1}(1 + e^{-i\xi})^{m+1}$). Smooth interpolatory spline wavelets (which are based on taking the cardinal splines instead of B-splines) possess exponential decay at infinity and are still practical through truncation processes. If compact support is desired, relaxation of smoothness requirements (take $S_m^r(\mathcal{T}_0)$ with $r < m-1$ and consider spline multiwavelets) or knot insertion (take suitable subspaces of V_1) are the options. Applications to collocation methods have been tried (see [6]).

2.3 Semiorthogonal wavelets

The semiorthogonal case is of interest for several reasons. As we have seen, it is well-adapted to applications to Sobolev spaces, see Corollary 2. On the other hand, semi-orthogonal spline wavelets of compact support can be constructed.

A system F associated with a dyadic MRA is called *semiorthogonal* if the system Ψ_1 (i.e., the collection of $\mathbb{Z}$-shifts of the ψ's) forms a Riesz basis of the orthogonal complement space $W_1 = V_1 \ominus_{L_2} V_0$. The related function(s) ψ are called *semiorthogonal wavelets* (in other places, these are called *prewavelets*). A detailed study of semiorthogonal wavelets in the case $M = L = 1$ is given in [14, Section 5.6, 6.2] and will not be repeated here. The important message is that, starting from a suitable dyadic MRA with compactly supported scaling function ϕ, we can *always* find a compactly supported semiorthogonal ψ! For the m-th degree B-spline MRA, the symbol m_ψ of the minimally supported ψ is given as follows (compare [14, Section 6.2, (6.2.5)]):

$$m_\psi(\xi) = \left(\frac{1-e^{-i\xi}}{2}\right)^{m+1} \sum_{l=0}^{2m} (-1)^l b_l' e^{-il\xi} \tag{34}$$

where b_l' is the (non-zero) value of the $(2m+1)$-th degree B-spline at $t = l+1$, $l = 0, \ldots, 2m$. The support of this ψ is $[0, 2m-1]$ which is asymptotically twice as large as that of the underlying m-th degree B-spline. Again, for $m = 0$, the Haar basis comes out. The case $m = 1$ is depicted in Figure 1. Generalizations to higher dimensions are discussed in section 3.

2.4 Biorthogonal wavelets

The analysis of multilevel systems in Sobolev spaces makes (implicitly and explicitly) heavy use of duality concepts. According to [1, Section 2], as soon as we have a multilevel frame F in $L_2(\Omega)$ generated from a MRA, we have a dual frame $\tilde{F}$ and we can formally

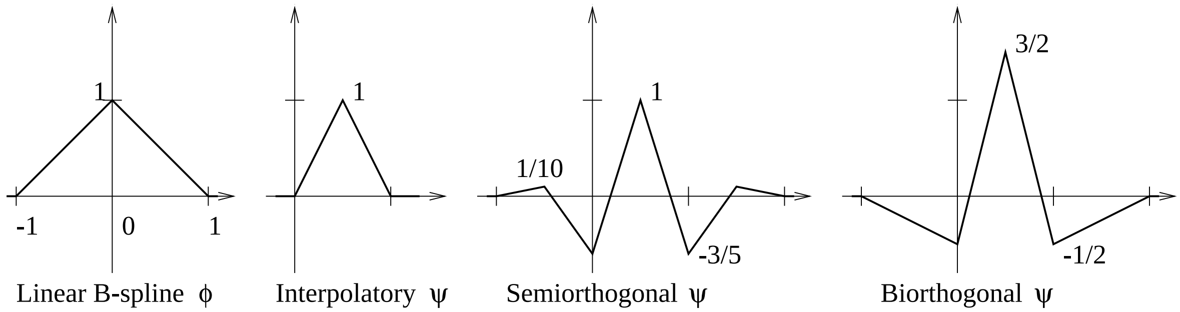

Figure 1: Linear spline MRA: ϕ and various ψ

split it into pieces $\tilde{\Phi}_0 = \{\tilde{\phi}_{0,i}\}$ and $\tilde{\Psi}_j = \{\tilde{\psi}_{j,i}\}$, $j \geq 1$ (as before, we will identify Ψ_0 resp. $\tilde{\Psi}_0$ with Φ_0 resp. $\tilde{\Phi}_0$). If the frame is a Riesz basis then *biorthogonality* of the systems F and $\tilde{F}$ follows from [1, (10)]: $(\psi_{j,i}, \tilde{\psi}_{j',i'})_{L_2} = 0$ except for $(j,i) = (j', i')$ when this scalar product is 1.

Unfortunately, it is not true a priori that the dual system of a dyadic MRA again is related to a dyadic MRA (for a counterexample, see [14, Section 1.4]). In case it is (what we will assume in the following), full information about the quadrupel of subspace sequences $(\{V_j\}, \{W_j\}; \{\tilde{V}_j\}, \{\tilde{W}_j\})$ resp. of scaling functions/wavelets $(\phi, \psi; \tilde{\phi}, \tilde{\psi})$ resp. of symbols $(m, m_\psi; \tilde{m} = m_{\tilde{\phi}}, m_{\tilde{\psi}})$ is very helpful. Beginning with the seminal paper [19], it became convenient to start with a biorthogonal system rather than to derive it as a by-product of some other construction. Two dyadic MRA $\{V_j\}$, $\{\tilde{V}_j\}$ with scaling functions ϕ, $\tilde{\phi}$ are called *biorthogonal* if

$$(\tilde{\phi}, \phi_{0,k})_{L_2} = \delta_k \ , \quad k \in \mathbb{Z} \ . \tag{35}$$

In terms of symbols, up to technical assumptions, (35) is equivalent to

$$m(\xi)\tilde{m}^*(\xi) + m(\xi+\pi)\tilde{m}^*(\xi+\pi) = 1 \ , \quad \xi \in \mathbb{T} \ . \tag{36}$$

Orthonormal MRA are a special case ($\tilde{\phi} = \phi$), and interpolatory MRA, in a certain sense, too ($\tilde{\phi}$ is the delta distribution, with symbol $\tilde{m}(\xi) \equiv 1$).

From a practical point of view, the biorthogonality assumption is sufficiently flexible to construct a variety of examples where both $\phi, \tilde{\phi}$ are compactly supported. It is possible to choose for ϕ the m-th degree B-spline while for the associated $\tilde{\phi}$ the Sobolev smoothness grows linearly with m. In Figure 1, the wavelet ψ of minimal support corresponding to the case $m = 1$ is shown (the Sobolev smoothness of the associated $\tilde{\phi}$ can be calculated as $\tilde{\gamma} \approx 0.440765$).

Having such a pair of biorthogonal dyadic MRA, it is then easy to construct a variety of wavelet systems, in particular, such that ψ and $\tilde{\psi}$ are also compactly supported. By choosing this flexibility in the right way, approximation properties, moment conditions, and smoothness of the resulting systems can be tuned to the application at hand. This is important for applications to variational problems in the Sobolev scale. E.g., the smoothness of ϕ, $\tilde{\phi}$ essentially determines the range of Sobolev spaces where the scaled systems F', $\tilde{F}'$ are Riesz bases.

Theorem 6 *Let be given a biorthogonal pair of dyadic MRA, with compactly supported scaling functions $\phi, \tilde{\phi}$. Let (a_k), $(\tilde{a}_k)$ denote the masks of the corresponding refinement equations. Set*

$$m_\psi(\xi) = e^{-i\xi}\tilde{m}(\xi+\pi)^* , \qquad m_{\tilde{\psi}}(\xi) = e^{-i\xi}m(\xi+\pi)^*$$

or, equivalently,

$$\psi(t) = \sum_{k\in\mathbf{Z}}(-1)^k \tilde{a}^*_{1-k}\phi(2t-k) , \quad \tilde{\psi}(t) = \sum_{k\in\mathbf{Z}}(-1)^k a^*_{1-k}\tilde{\phi}(2t-k) .$$

Then the scaled systems F' resp. $\tilde{F}'$ are biorthogonal Riesz bases in $H^s(\mathbb{R})$ if (and only if) $-\tilde{\gamma} < s < \gamma$ resp. if $-\gamma < s < \tilde{\gamma}$. Here,

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

are called Sobolev regularity exponents of ϕ and $\tilde{\phi}$, respectively.

The above choice of ψ, $\tilde{\psi}$ guarantees compact support, other choices are covered as well (as soon as the biorthogonality $W_j \perp_{L_2} \tilde{W}_{j'}$, $j \neq j'$, and the Riesz basis property of Ψ_j in W_j resp. of $\tilde{\Psi}_{j'}$ in $\tilde{W}_j$ are preserved). The result for orthogonal MRA is completely contained in the above formulation ($\gamma = \tilde{\gamma}$). For a proof of the if-part in Theorem 6, see [24] (this paper uses the language of Jackson-Bernstein inequalities which reduces in the particular case to just knowledge of the Sobolev regularity exponents $\gamma, \tilde{\gamma}$). Consult also the results on characterizing the exact Sobolev regularity of refinable functions, see [19, 33, 39], which show that under the above assumptions one always has $\gamma, \tilde{\gamma} > 0$. In the above mentioned example from [19], the Riesz basis property of the scaled F' consisting of compactly supported linear spline functions is consequently satisfied for Sobolev spaces with $-0.440765 < s < 3/2$. Similar results hold for $d > 1$, see Section 3.2.

In [20, 50], an analogous result has been obtained if we start with a dyadic MRA given by a compactly supported ϕ and a choice of our favorite compactly supported ψ (no compact support is assumed for $\tilde{\phi}$). This setting is quite natural if we are only interested in multilevel solvers based on a ϕ-discretization and multilevel preconditioning. In that case, it is much more important to care about the size of ϕ- and ψ-masks than on properties of the dual system (however, these are implicitly used in the proof of the Riesz basis property of a system). If we investigate ad hoc choices of ϕ-ψ combinations then we have typically no direct access to $\tilde{\phi}$. However, under reasonable assumptions on the two-level decomposition $V_1 = V_0 \dot{+} W_1$ induced by a given choice of ϕ and ψ, one can compute the symbol $\tilde{m}$ of the potential dual MRA from the biorthogonality relations (see, e.g., [14, Section 5.4]), and prove similar results. The difference to the setting of Theorem 6 is that $\tilde{m}$ may not be a polynomial (i.e., the $\tilde{\phi}$-mask may be infinite) but only a rational function in $e^{-i\xi}$. Still, the correct $\tilde{\gamma}$ can be computed from properties of a so-called transfer operator associated with $|\tilde{m}(\xi)|^2$ and may be positive (in which case a dual $\tilde{\phi}$ exists, is in $L_2(\mathbb{R})$ and generates an MRA) as well as non-positive. In the latter case, Theorem 6 is not applicable since a biorthogonal dyadic MRA on $\mathbb{R}$ for which the given ψ is a complimentary function for ϕ does not exist (at least, not in the L_2-setting). Computational methods for obtaining $\tilde{\gamma}$ are discussed in [20, 50]. The following family of examples for the linear B-spline MRA has been considered in [50, Section 3.3]. We give the ψ_l-expressions, $l = 0, 1, \ldots$. The resulting F_l have been called $(l+1)$-point hierarchical bases in [50] since $C_{\psi_l} = l+1$, and $l = 0$ reduces the prototyp of a hierarchical basis, to the Faber-Schauder system):

$$\psi_l(t) = \sum_{k=0}^{l} (-1)^{k-[l/2]} a_k^{(l)} \phi(2t - (k+1-[l/2])) \,, \quad a_k^{(l)} = \left(\begin{array}{c} l \\ k \end{array} \right), \; k = 0, \ldots, l \,.$$

Note that for $l \geq 1$, $(a_k^{(l)})$ (the remaining values are set to 0) is, up to a scaling factor and an index shift, the mask of the B-spline of degree $l-1$. Also, ψ_l satisfies l moment conditions, i.e., $\psi_l \perp_{L_2} \mathbb{P}_{l-1}$. Explicit formulas for the rational dual symbols $\tilde{m}_l(\xi)$ associated with this ϕ-ψ_l choice have been computed in [50, (126)]. These symbols have appeared in the literature in a different context (see [20, 43], i.e., the dual MRA which we are not particularly interested in, has some other interesting applications). In Figure 2, a schematic picture of the graph of ψ_l and the range in which F_l' is a Riesz basis in $H^s(\mathbb{R})$ are shown for $l \leq 3$. Results for larger l and for other B-splines replacing the linear hat function as the choice for ϕ can be deduced from [20, 50].

Here is yet another 'toy' example which can be investigated by the above methods. Again, for the linear spline MRA, define the one-parameter family of ψ by the formula

$$\psi(t) = \phi(2t-1) - a(\phi(2t-2) + \phi(2t)) - (1/2 - a)(\phi(2t-3) + \phi(2t+1)) \,. \qquad (37)$$

These ψ are symmetric about $t = 1/2$, and satisfy moment conditions of order 2. The stability assumption on Ψ_1 with respect to W_1 is satisfied if and only if $0 < a < 1$. The family (37) 'interpolates' the following special cases: semiorthogonal linear wavelets (for $a = 3/5$), the functions ψ_2 (after a sign change for $a = 1/2$) and ψ_4 (for $a = 2/3$) from the previous example, and also the biorthogonal wavelet from Figure 1 (for $a = 1/3$, this is the only value for which the dual $\tilde{\phi}$ has compact support). The case $a = 2/3$ is exceptional - this is the only case where ψ satisfies moment conditions of order 4. Figure 3 shows the graph of the lower bound $-\tilde{\gamma} = -\tilde{\gamma}(a)$ for the s-interval $-\tilde{\gamma} < s < 3/2$, such that the Riesz basis property of the scaled systems F' holds in $H^s(\mathbb{R})$. The true value for the exceptional case is $-\tilde{\gamma}(2/3) = -2.539441$, in all other cases we have $-\tilde{\gamma}(a) > -2$ (this example shows that the Sobolev exponents of refinable functions may behave discontinuously, due to changes in the number of moment conditions).

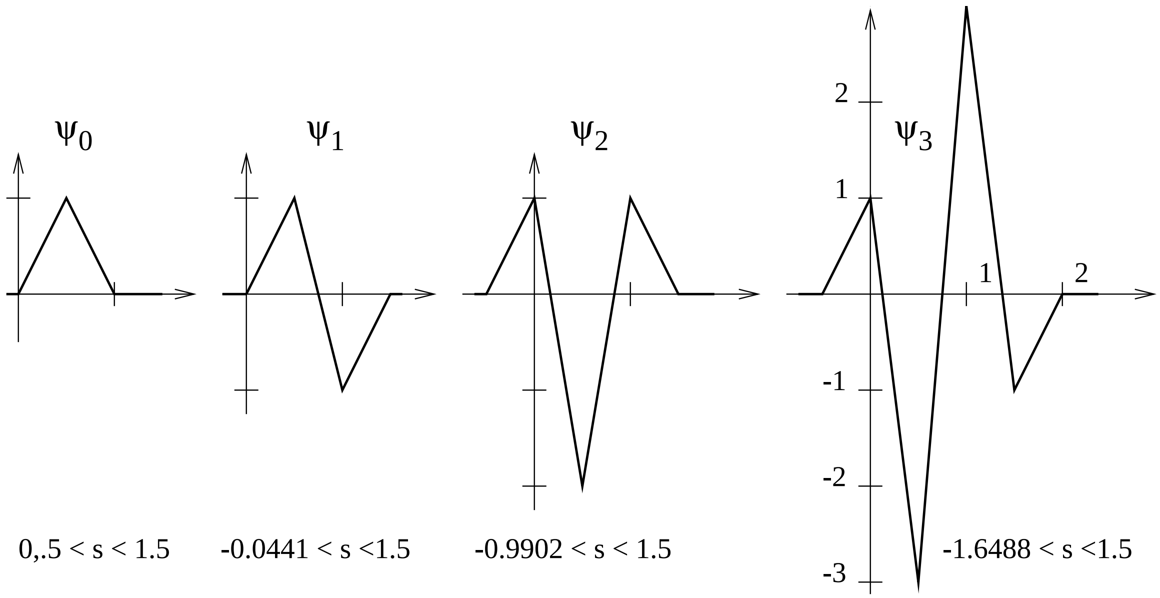

Figure 2: The $(l+1)$-point ψ_l for the linear spline MRA

2.5 Further examples

We wish to mention that there are numerous other examples. Multiwavelets are an area of current research activities. The simplest situation (which is close to our intention to stay in an 'observable' neighborhood of classical spline and finite element schemes) is the dyadic MRA associated with $V_0 = S_2^0(\mathcal{T}_0)$ (quadratic C^0-splines) for which $M = L = 2$ is appropriate. Figure 4 shows our choices for the two scaling functions (ϕ^1 is the linear B-spline while ϕ^2 represents a quadratic 'bubble' function with support on $[0,1]$: $\phi^2(t) = (4t(1-t))_+$), and the two semiorthogonal wavelets

$$\begin{aligned}
\psi^1(t) &= \phi^1(2t) + \frac{3}{4}(\phi^1(2t-1) + \phi^1(2t+1)) - \frac{23}{16}(\phi^2(2t) + \phi^2(2t+1)) \\
&\quad - \frac{7}{16}(\phi^2(2t-1) + \phi^2(2t+2)) \,, \\
\psi^2(t) &= \phi^1(2t-1) - \phi^1(2t+1) - \frac{33}{28}(\phi^2(2t) - \phi^2(2t+1)) \\
&\quad - \frac{17}{28}(\phi^2(2t-1) + \phi^2(2t+2)) \,.
\end{aligned}$$

The remarkable fact is that the supports of ψ^1, ψ^2 are smaller compared to the semiorthogonal linear spline wavelet shown in Figure 1 above. This and the obvious symmetry properties make it easier to introduce *boundary modifications*.

Another option would be to change the dilation factor (the author should admit that he knows only of one paper on multilevel schemes for the numerical solution of PDE

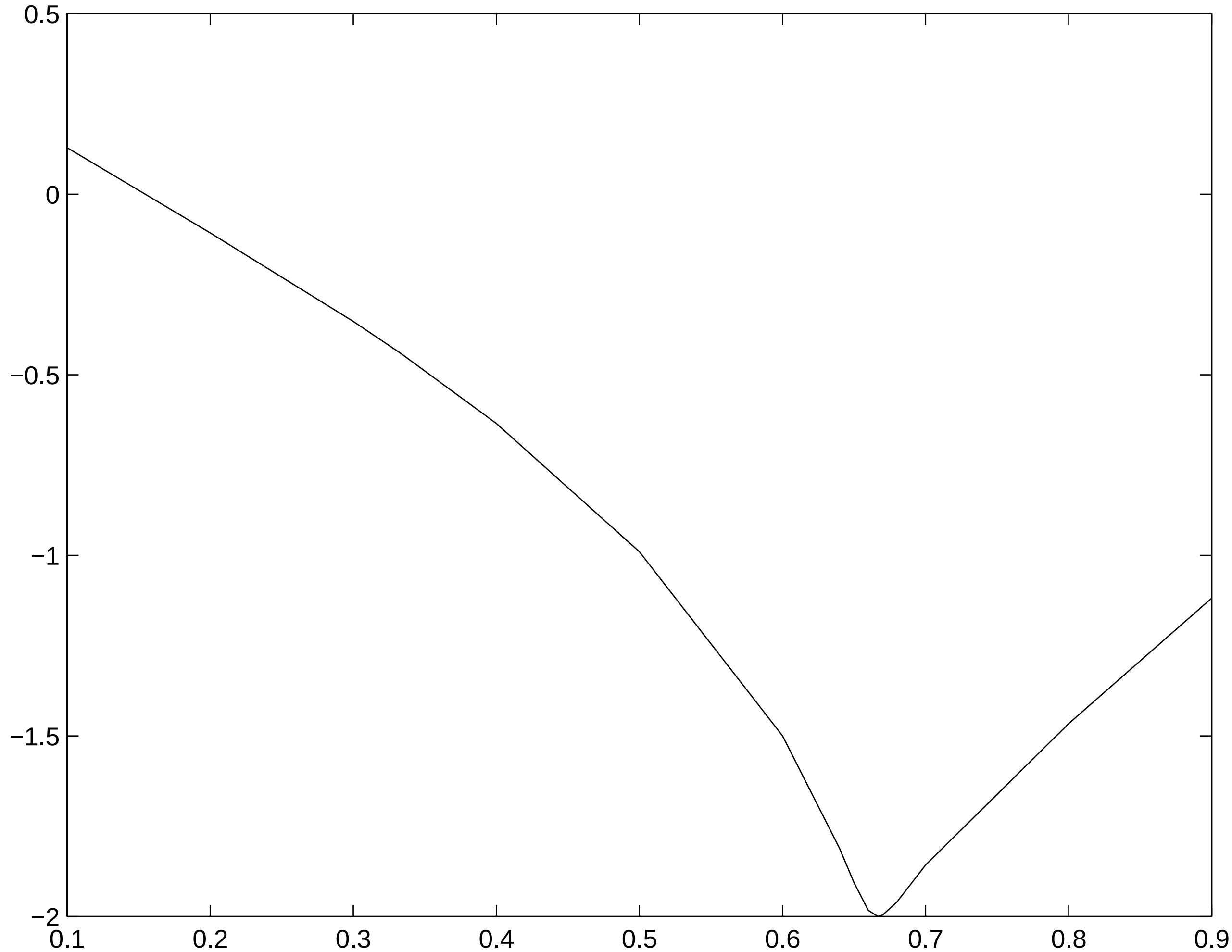

Figure 3: Values of $-\tilde{\gamma}(a)$ for the family (37)

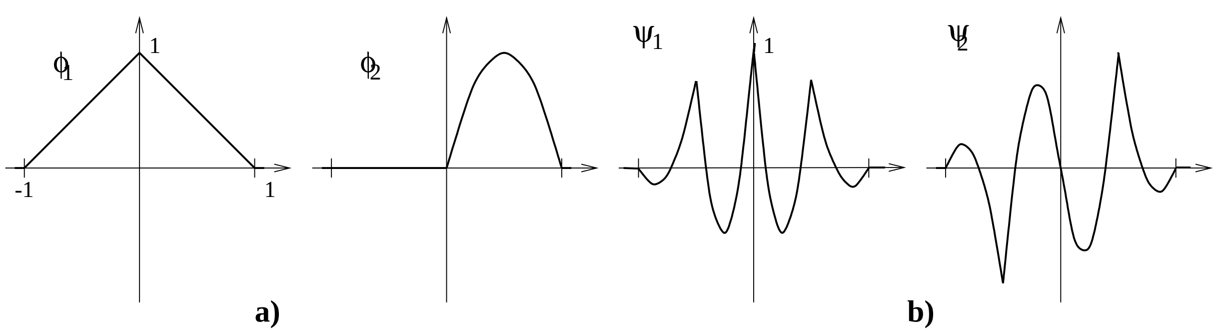

Figure 4: Scaling functions and semiorthogonal wavelets for quadratic C^0 splines

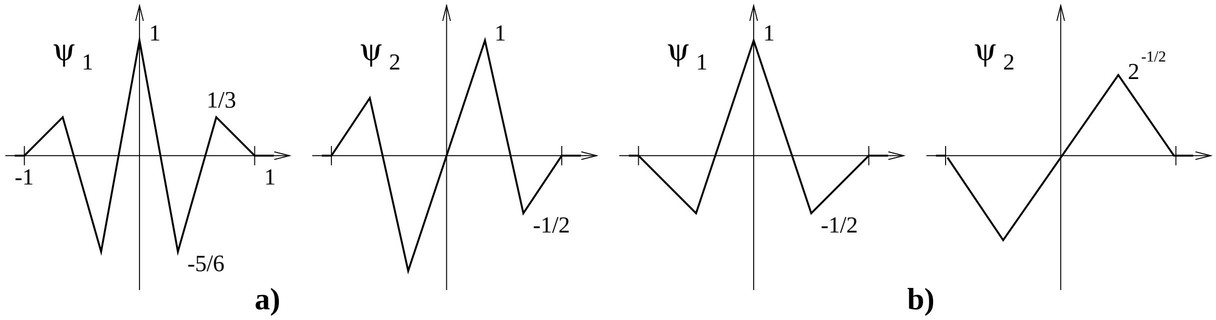

Figure 5: Linear spline ψ^l for a) a triadic MRA and b) a tight frame

where other than dyadic dilation has been tried in a practical application). Figure 5 a) depicts a choice of semiorthogonal spline wavelets ψ^1, ψ^2 related to a linear spline MRA based on *triadic dilation*. I.e., we define now $u \in V_j \Leftrightarrow u(3^{-j}\cdot) \in V_0$ while the other requirements for a MRA remain unchanged. Again, better localization and symmetry properties of the ψ^l are the result, at the expense of having to deal with two instead of one ψ-mask. There are some papers that study MRA on $\mathbb{R}$ associated with *rational dilation parameter* (see, e.g., [4], in higher dimensions there is a much greater variety of dilation procedures.

The next remark concerns the existence of multilevel frames in $H^s(\mathbb{R})$ other than the standard one covered by Corollary 5 which is only good for $s > 0$. In general, there is also some doubt whether frames are a reasonable concept at all for the case $s \le 0$. The question is not yet studied in sufficient detail although it seems to be not a hopeless one. Take again the dyadic linear spline MRA generated from the linear B-spline ϕ shown in Figure 1, and consider the two complementing functions

$$\begin{aligned}
\psi^1(t) &= -\frac{1}{2}\phi(2t-1) + \phi(2t) - \frac{1}{2}\phi(2t+1) \\
\psi^2(t) &= \frac{1}{\sqrt{2}}\phi(2t-1) - \frac{1}{\sqrt{2}}\phi(2t+1) ,
\end{aligned}$$

see Figure 5 b). These complementary functions have been introduced in [59, section 1.3] as partial case of a whole family of *compactly supported tight spline frames* in $L_2(\mathbb{R})$ (actually, [59] considers the bi-infinite case, i.e., the correct statement is that $\{\psi^1(2^j t - k), \psi^2(2^j t - k) \,:\, j,k \in \mathbb{Z}\}$ forms a tight frame in $L_2(\mathbb{R}^d)$). Note that these functions are orthogonal to $\mathbb{P}_0$, and possess symmetry properties.

Our statement is that the scaled multilevel system F' constructed from ϕ $(M = 1)$ and the ψ^l $(L = 2)$ is a frame in $H^s(\mathbb{R})$ for all $-1 < s < 3/2$. We will not prove the complete result but show the frame property for part of this interval. The idea is the following: Observe that

$$\psi^1(t) = -\frac{1}{2}(\psi_1(1-t) + \psi_1(t-1)) , \quad \psi^2(t) = \frac{1}{\sqrt{2}}(\psi_1(1-t) - \psi_1(t-1)) ,$$

where $\psi_1(t)$ is the wavelet function introduced in subsection 2.4 (see Figure 2 , $l = 1$). Thus, our F is actually the 'linear combination' of F_1 and another, similar system F_1^- which is obtained by replacing $\psi_1(t)$ by $\psi_1(-t)$. Vice versa, the union of F_1 and F_1^- can be expressed by F in a similar way. From [50] we have that F_1' is a Riesz basis (and therefore a frame!) in $H^s(\mathbb{R})$ for all $-0.044117 < s < 3/2$. The same is true for $(F_1^-)'$ (consider the coordinate transformation $t \to -t$ which is an isometry in $H^s(\mathbb{R})$ and transforms F_1^- into F_1). Looking at the definition of a frame via moments and expressing the moments with respect to one system through moments with respect to the other system, we easily see that F' must be a frame in $H^s(\mathbb{R})$ for at least the same range of s. To cover the full range of $s \in (-1, 3/2)$, one has to relate F with F_3 and suitable perturbations of it and to use a limiting argument. Alternatively, one can view the union $F_1 \cup (-F_1^-)$ (neglect the component for $j = 0$ for a moment) as the result of applying the differentiation operator $\frac{d}{dt}$ to the standard frame associated with the quadratic B-spline MRA. According to Corollary 5, the latter is a frame in $H^s(\mathbb{R})$ for all $0 < s < 5/2$ while the differentiation operator provides a lift from H^s to H^{s-1} (the problems with the Φ_0-part can easily be fixed). This trick gives the same result. Clearly, the same considerations can be applied to the finite sections of these systems. We suggest the problem of creating interesting frames for further investigation.

Finally, for $d = 1$, there are also some promising results concerning the construction of orthogonal, semiorthogonal, and biorthogonal systems with respect to other scalar products. In particular, since we are looking for optimal solvers for elliptic variational problems, (semi-)orthogonality with respect to the energy scalar product $a(\cdot, \cdot)$ would be most desirable. One new direction which we will touch upon in later parts of these lectures is the systematic use of *vaguelettes* initiated by Meyer et al.. Another approach is to consider (semi)orthogonalization with respect to the H^s-seminorm, see [47, 22]. However, for $d > 1$ these attempts to tailor multilevel systems to the operator equation at hand are of limited practical use since they seemingly interfere with the desire to keep ϕ- and ψ-masks local and small [22, 48].

3 Multilevel finite element systems in $H^s(\Omega)$

3.1 Nodal basis frame

We continue with Example 8 of Section 1.1 and give some details on verifying the properties of a finite element (FE) MRA which are necessary for the theory of multilevel systems in Sobolev spaces. For an extensive discussion of the approach, see [56].

As in Example 8 above, let $\Omega \subset \mathbb{R}^d$, $d > 1$, be polyhedral, satisfy the extension property, and be equipped with an increasing sequence of quasi-uniform, regular partitions $\mathcal{T}_j$, $j \geq 0$. Here are examples of finite element spaces V_j for which the requirements of an MRA in $L_2(\Omega)$ hold if we choose the collections of so-called *nodal basis functions* as the default bases Φ_j in V_j. We will not give complete definitions, and refer the interested reader to the literature (generally, [18, 56] could be consulted).

- Lagrange C^0-elements of degree m for $\mathbb{R}^d$-simplices or $\mathbb{R}^d$-rectangles, where $\mathbb{P}_m \subset V_j \subset H^s(\Omega) \cap C^0(\Omega)$ for $s < 3/2$ (see [18]),
- Bogner-Fox-Schmidt rectangles (and their analogs for $d > 2$), where $\mathbb{P}_3 \subset V_j \subset H^s(\Omega) \cap C^1(\Omega)$ for $s < 5/2$,
- triangular Powell-Sabin macro elements (see [34]), $d = 2$, with $\mathbb{P}_2 \subset V_j \subset H^s(\Omega) \cap C^1(\Omega)$ for $s < 5/2$,
- quadrilateral Fraeijs De Veubeke-Sanders macroelements (see [28]), $d = 2$, with $\mathbb{P}_3 \subset V_j \subset H^s(\Omega) \cap C^1(\Omega)$ for $s < 5/2$.

For a schematical display of the local interpolation problems associated with the FE types used below, see Figure 6. The graphical notation for the interpolation conditions reads as follows: Small dots represent interpolation of function values at a point (small circles stand for those points which are omitted if the serendipity element of the same type is considered, see [18] for details), larger circles mean interpolation of the complete set of first resp. second order partial derivatives at a point, one arrow denotes a directional derivative (here, normal derivatives at edge midpoints) while two arrows denote interpolation of the mixed derivative u_{xy}. In all cases the maximal value of m has been shown. The first 4 element types lead to globally C^0 finite element functions and are suitable for conforming Galerkin approximations of second order elliptic boundary value problems. The remaining 4 element types represent C^1 finite elements, and could be applied to fourth order problems such as the biharmonic problem. The last two examples are C^1-macroelements: The local space consists of C^1 continuous piecewise polynomials of a given degree ($m = 2$ resp. $m = 3$ in our examples) with respect to the given secondary partition of the element.

Other FE schemes that fit the rules might be constructed as well but will be too complicated to be included into the list of practical FE constructions (some of the above are already too involved from an practitioner's viewpoint). Compared to what else the interested reader will find in [18], the above list is short, for essentially one reason: very often the monotonicity condition (1) is violated. E.g., Argyris and Bell C^1-elements, Hermite and reduced Hermite C^0-elements, other serendipity elements, Clough-Tocher macroelements, all nonconforming FEs etc. have this defect.

To implement the program of Section 1, we need to verify an appropriate set of Jackson-Bernstein inequalities. We only sketch the steps following [56] where further references can be found. We need a quantity called *k-th order modulus of continuity*:

$$\omega_k(t, f)_{L_2} = \sup_{h \in \mathbb{R}^d \,:\, |h| \le t} \|\Delta_h^k f\|_{L_2(\Omega_h)} \,, \quad f \in L_2(\Omega),\ t > 0 \,. \tag{38}$$

Here, $\Delta_h^k f(x)$ is the k-th order difference operator with stepsize h (i.e., a linear combination of the function values $f(x + lh),\ l = 0, \ldots, k$), and $\Omega_h = \{x \in \Omega \,:\, [x, x + kh] \subset \Omega\}$. The advantage of this function is that its behavior for $t \to 0$ measures L_2-smoothness up to order k on a fine scale, without involving derivatives. It can be used to define *Besov*

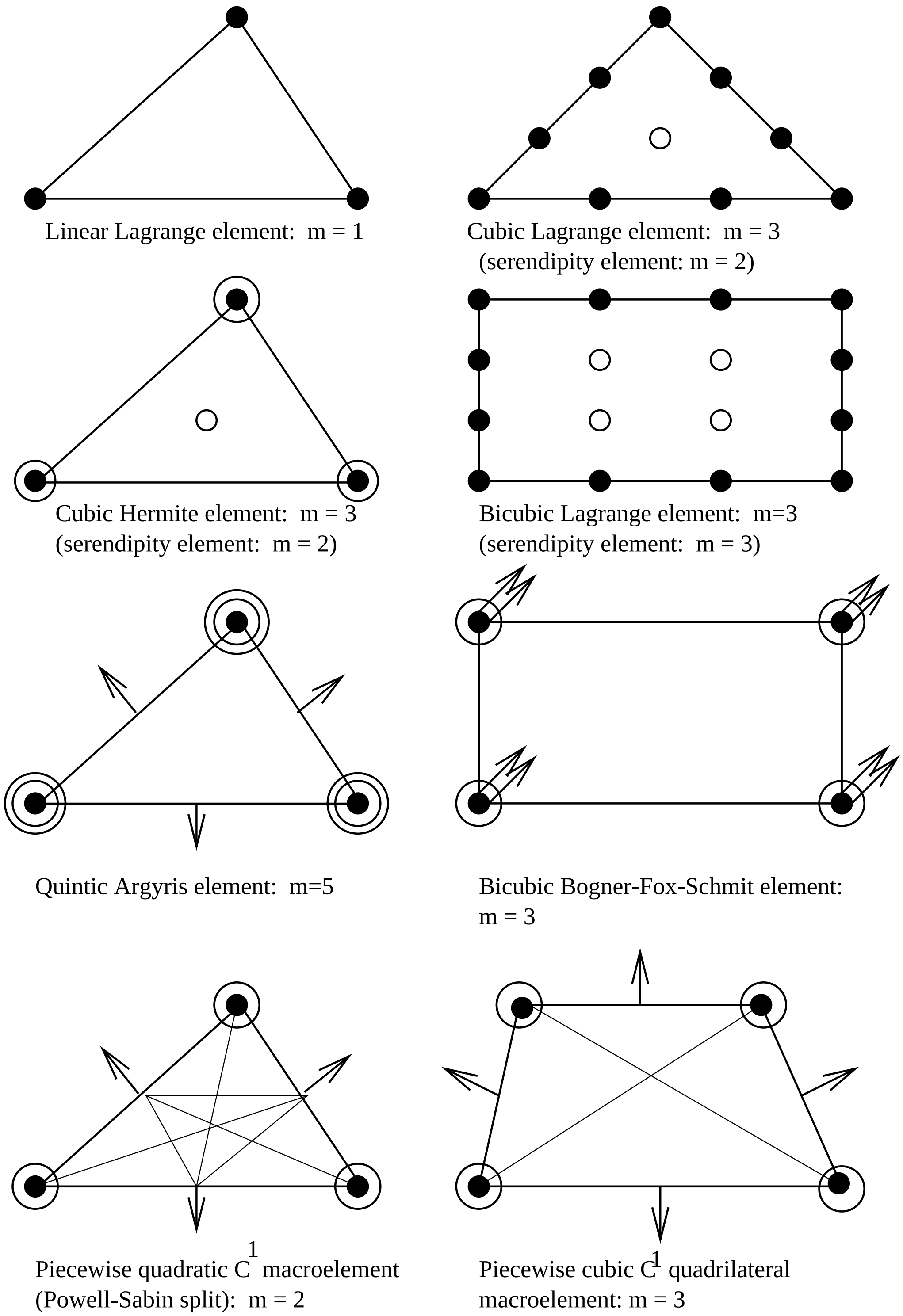

Figure 6: Local interpolation problems for some FE types

spaces $B^s_{2,q}(\Omega)$ in an elementary and explicit way. For the domains under consideration, one has

$$\omega_k(t,f)_{L_2} \asymp K(t,f)_{L_2} \equiv \inf_{g\in H^k(\Omega)} (\|f-g\|_{L_2} + t^k|D^k g|_{L_2})\ , t\to 0\ , \tag{39}$$

which links it to interpolation theory. The notation

$$|D^k g|_{L_2} = (\sum_{\alpha\in\mathbf{Z}^d_+;\,|\alpha|=k} \|\frac{\partial^\alpha g}{\partial x^\alpha}\|^2_{L_2})^{1/2}$$

stands for a seminorm in $H^k(\Omega)$). Without proof, note the equivalence

$$H^s(\Omega) = B^s_{2,2}(\Omega) \equiv \{f\in L_2(\Omega)\ :\ \|f\|^2_{B^s_{2,\infty}} = \|f\|^2_{L_2(\Omega)} + \sum_{j=0}^{\infty} 2^{2sj}\omega_k(2^{-j},f)^2_{L_2} < \infty\} \tag{40}$$

which is valid for any $s>0$ and integer $k>s$ (the norms are equivalent for different such k).

To prove Jackson inequalities for FE spaces, the following approach is standard.

- Construct a sequence of so-called *quasi-interpolant operators* $Q_j\ :\ L_2(\Omega)\to V_j$ by setting

  $$Q_j f = \sum_i \lambda_{j,i}(f)\phi_{j,i}$$

  where a) the coefficient functionals $\lambda_{j,i}(f)$ are linear functionals on $L_2(\Omega)$, and satisfy

  $$|\lambda_{j,i}(f)| \le C\alpha_j^{-1}\|f\|_{L_2(\Omega_{j,i})}\ ,$$

  where $\Omega_{j,i}\subset\Omega$ is called support of $\lambda_{j,i}$, and b) polynomials from some $\mathbb{P}_m$ are reproduced locally:

  $$Q_j p = p \qquad \forall p\in\mathbb{P}_m\ ,$$

  and, in order to define the locality requirement, assume that for each $\Delta\in\mathcal{T}_j$ there exists a suitable set $\Omega_\Delta\subset\Omega$ which contains all those $\Omega_{j,i}$ for which Δ intersects the support of $\phi_{j,i}$. Then, the sets $\Omega_\Delta\ ,\ \Delta\in\mathcal{T}_j$, should have regular boundary, have diameter $\le C2^{-j}$, and form a locally finite family of sets such that

  $$\sum_{\Delta\in\mathcal{T}_j} \|f\|^2_{L_2(\Omega_\Delta)} \le C\|f\|^2_{L_2(\Omega)}\ .$$

  Constants should be independent of j, in practice this follows from the assumptions on $\mathcal{T}_j$ automatically.

  A partial case is the construction of quasi-interpolant *projectors* where additionally $Q_j v_j = v_j$ for all $v_j\in V_j$ (the proper m is defined as the largest integer such that $\mathbb{P}_m\subset V_j$ for all j, see above for these m in particular cases). Necessarily, $\lambda_{j,i}(\phi_{j,i'}) = \delta_{i,i'}$, in particular, $\Omega_{j,i}$ should be a subset of the support $K_{j,i}$ of the nodal basis function $\phi_{j,i}$. In this case, the construction and properties of $\{\Omega_\Delta\}$ are straightforward. Suitable $\lambda_{j,i}$ can be determined from locally (on each $\Delta \in \mathcal{T}_j$) constructing biorthogonal systems to the finite set $\{\phi_{j,i}|_\Delta\}$. For the linear FE case, see [56, p. 15ff]. A flexible use of quasi-interpolant techniques can help a lot!

- Now check that

$$\|Q_j f\|_{L_2} \le C\|f\|_{L_2} \qquad \forall\, f \in L_2(\Omega)\,,\ j \ge 0\,,$$

and

$$\|g - Q_j g\|_{L_2} \le C 2^{-j(m+1)} \|D^{m+1} g\|_{L_2} \qquad \forall\, g \in H^{m+1}(\Omega)\,,\ j \ge 0\,.$$

The latter inequality implies the Jackson estimate with $\gamma = m+1$, and is proved by applying the Bramble-Hilbert lemma locally on Ω_Δ. Indeed, by definition of Ω_Δ and the polynomial reproduction assumption, we have

$$\|g - Q_j g\|_{L_2(\Delta)} \le \|g - p\|_{L_2(\Delta)} + \|Q_j(g-p)\|_{L_2(\Delta)} \le C\|g - p\|_{L_2(\Omega_\Delta)}$$

for all $p \in \mathbb{P}_m$ and $\Delta \in \mathcal{T}_j$ (the last step requires several steps and needs almost all assumptions on $\lambda_{j,i}$ and Φ_j, it is left to the reader as a stimulating exercise). Now, the Bramble-Hilbert lemma applies, and it remains to add all local estimates.

The L_2-boundedness is much simpler to prove. From the Riesz basis assumption of Φ_j and the bound for $\lambda_{j,i}$:

$$\|Q_j f\|_{L_2}^2 \le C \sum_i \alpha_j^2 |\lambda_{j,i}(f)|^2 \le C \sum_i \|f\|_{L_2(\Omega_{j,i})}^2\,.$$

We conclude by using the fact that $\{\Omega_{j,i}\}$ is a locally finite family of sets, too (why?).

- The Jackson type estimate for $H^\gamma(\Omega)$, $0 < \gamma < m+1$, as well as the one in terms of moduli of smoothness

$$e_j(f)_{L_2} \le \|f - Q_j f\|_{L_2} \le C\omega_{m+1}(2^{-j}, f)_{L_2}\,, \quad f \in L_2(\Omega),\ j \ge 0\,,$$

follow by interpolation (for the latter, use (39) with $k = m+1$ after applying the above inequalities to the decomposition

$$f - Q_j f = (f-g) + (g - Q_j g) + Q_j(g - f) \qquad \forall\, g \in H^{m+1}(\Omega)\ ).$$

The Bernstein inequality (in terms of moduli of smoothness) for spaces of piecewise polynomials has been established in [56, Section 2.4]. For our examples (and all other reasonable FE constructions), it reads

$$\omega_k(t, v_j)_{L_2} \le C \min(1, (2^j t)^{\min(r+3/2,k)}) \|v_j\|_{L_2}\ , \quad t > 0\,,$$

for all $v_j \in V_j$, where r denotes the global smoothness (i.e., derivatives of order r are continuous while for $r+1$ they may be discontinuous) of the FE functions in V_j. Thus, we have $r=0$ for Lagrange C^0 elements, and $r=1$ for the remaining examples in the above list. The reader is recommended to prove this inequality for linear FE, where $k=2$ and $r=0$ are appropriate. The case $t \geq C2^{-j}$ is trivial since

$$\|\Delta_h^k f\|_{L_2(\Omega_h)} \leq 2^k \|f\|_{L_2}$$

by definition of the k-th order difference. By choosing k sufficiently large and substituting into (40) it follows that

$$\|v_j\|_{H^s} \leq C 2^{js} \|v_j\|_{L_2} \qquad \forall\, v_j \in V_j,\ j \geq 0\,,\ 0 < s < r + 3/2\,.$$

Thus, we are ready to apply Theorem 1. We will do it in the language of space splittings. To this end, let again $V_{j,i}$ be the one-dimensional subspaces associated with the nodal basis functions $\phi_{j,i}$.

Theorem 7 *Let $\{V_j\}$ be a FE MRA as specified above. Let $r = -1, 0, \ldots$ denote the global smoothness order of the FE scheme, and m the largest integer for which the above quasi-interpolant construction works. Set*

$$\gamma = \min(r + 3/2, m + 1)$$

(for all examples in the above list, $\gamma = r + 3/2$, i.e., $\gamma = 3/2$ for C^0-elements, and $\gamma = 5/2$ for C^1-elements), and assume $0 < s < \gamma$. Let $a(\cdot,\cdot)$ be any symmetric $H^s(\Omega)$-elliptic bilinear form. Then the splitting

$$\{H^s(\Omega); a\} = \sum_{j=0}^{\infty} \{V_j; 2^{2sj}(\cdot,\cdot)_{L_2}\} = \sum_{j=0}^{\infty} \sum_i \{V_{j,i}; 2^{2sj}(\cdot,\cdot)_{L_2}\}$$

resp. its finite sections

$$\{V_J; a\} = \sum_{j=0}^{J} \{V_j; 2^{2sj}(\cdot,\cdot)_{L_2}\} = \sum_{j=0}^{J} \sum_i \{V_{j,i}; 2^{2sj}(\cdot,\cdot)_{L_2}\}\,, \quad J \geq 0\,,$$

are stable, with upper and lower stability constants and condition numbers that are uniformly bounded with respect to J. The condition numbers depend on Ω, on the parameters characterizing the regularity and quasi-uniformity of the sequence $\{T_j\}$, on the FE type, on s, and on the ellipticity constants of the form $a(\cdot,\cdot)$.

This result applies to both bounded and unbounded polyhedral domains satisfying the formulated conditions. Slit domains for which the extension property does not hold, are formally excluded (see [56] for arguments that show that the results remain true). The associated system $F = \{\phi_{j,i}\}$ will be called *nodal basis FE frame* (the precise meaning of this catchy phrase is that the scaled systems F' will be frames in $H^s(\Omega)$

for the corresponding parameter range). The above splittings will be called *basic FE multilevel splittings*, to underline that they are central for the understanding of whole classes of optimal preconditioning methods (this will be demonstrated in the third part of these lectures).

The main practical application of this type of result is restricted to integer s, thus $s = 1$ (second–order elliptic boundary value problems) and $s = 2$ (fourth–order problems, if $r = 1$). Then the stiffness matrices A_J are sparse, and the first of the two approaches to multilevel solvers mentioned in subsection 1.2) makes sense. In connection with this problem class, *essential boundary conditions* are an issue, and the space splittings need to be tuned to them. It turns out (see [56]) that this can be done on the basis of Theorem 7, without attracting essentially new ideas. The restriction we have to pose is that the part $\Gamma \subset \partial\Omega$ of the boundary where essential boundary conditions are included into the test space should be resolved by the partition $\mathcal{T}_0$, i.e., Γ should be the union of $(d-1)$-dimensional faces of simplices from $\mathcal{T}_0$. For simplicity, consider homogeneous Dirichlet boundary conditions $u = 0$ on $\Gamma = \partial\Omega$. Then

$$\hat{V}_j = \{v_j \in V_j \;:\; v_j|_\Gamma = 0\}\;, \quad j \geq 0\;,$$

is the promising FE MRA, and the subspace

$$H^s_\Gamma(\Omega) = \{u \in H^s(\Omega) \;:\; u|_\Gamma = 0\}\;, \quad s > \frac{1}{2}\;,$$

of $H^s(\Omega)$ the appropriate 'energy space'. The restriction on s is natural, and is required for a meaningful trace definition.

We may assume that the Riesz basis $\hat{\Phi}_j$ is a subset of Φ_j, for all $j \geq 0$ (in all of the above examples, this can be achieved by modifying the definition of $\phi_{j,i}$ for boundary nodes, for Lagrange C^0-elements, no changes are required at all). What we claim is that the stability of the splitting

$$\{H^s_\Gamma(\Omega); a_\Gamma\} = \sum_{j=0}^{\infty}\{\hat{V}_j; 2^{2sj}(\cdot,\cdot)_{L_2}\} = \sum_{j=0}^{\infty}\sum_{i\,:\,\phi_{j,i}|_\Gamma=0}\{V_{j,i}; 2^{2sj}(\cdot,\cdot)_{L_2}\} \tag{41}$$

is valid for any symmetric $H^s_\Gamma(\Omega)$-elliptic form follows from the result of Theorem 7 where additionally $s > 1/2$ has to be required. The corresponding statements on the sections of the splitting (41) are also valid.

As can easily be seen, the new splitting is a subsplitting (obtained by selection, see the terminology of [1, Section 5]) of the old one, thus the lower stability bound is trivial. In order to prove the upper bound, use the inclusion $H^s_\Gamma(\Omega) \subset H^s(\Gamma)$ and the upper bound result from Theorem 7: For any $u \in H^s_\Gamma(\Omega)$, there exist $u_j \in V_j$, $j \geq 0$, such that

$$u = \sum_{j=0}^{\infty} u_j\;, \quad \sum_j 2^{2sj}\|u_j\|^2_{L_2} \leq \tilde{B}a(u,u) \leq C\tilde{B}\|u\|^2_{H^s} \leq C\tilde{B}a_\Gamma(u,u)\;.$$

Now, we will 'correct' the representation to a representation with respect to the new MRA $\{\hat{V}_j\}$ as follows. Consider the partial sums

$$v_j = \sum_{k=0}^{j} u_k = \sum_i c_{j,i}\phi_{j,i}$$

and project them into $\hat{V}_j$ by setting

$$\hat{v}_j = \sum_{i\,:\,\phi_{j,i}|_\Gamma = 0} c_{j,i}\phi_{j,i} \ .$$

Finally, set $\hat{u}_j = \hat{v}_j - \hat{v}_{j-1}$, $j \geq 1$ and $\hat{u}_0 = \hat{v}_0$. Then, at least in $L_2(\Omega)$, we have again

$$u = \sum_{j=0}^{\infty} \hat{u}_j \ , \qquad \hat{u}_j \in \hat{V}_j \ , \ j \geq 0 \ .$$

Now, there comes the technically tricky point. Since

$$\hat{u}_j = (\hat{v}_j - v_j) + u_j - (\hat{v}_{j-1} - v_{j-1}) \ , \quad j \geq 1 \ ,$$

one needs estimates for the L_2-norms of the differences of the functions

$$w_j = v_j - \hat{v}_j = \sum_{i\,:\,\phi_{j,i}|_\Gamma \not\equiv 0} c_{j,i}\phi_{j,i}$$

which are essentially defined by there values on Γ and do not vanish only in a small boundary corridor of size $\leq C2^{-j}$. Due to the local, element-wise definition of FE-functions, one easily sees that in all of the cases covered by Theorem 7, we have $\|w_j\|_{L_2}^2 \asymp 2^{-j}\|w_j|_\Gamma\|_{L_2(\Gamma)}^2$ and

$$w_j|_\Gamma = v_j|_\Gamma = -\sum_{k>j}^{\infty} u_k|_\Gamma \ .$$

Thus, the L_2-norm of the traces of w_j can be estimated from the L_2-norms of traces of u_k, $k > j$, and it remains to use the inverse direction $\|u_k|_\Gamma\|_{L_2(\Gamma)} \leq C2^k\|u_k\|_{L_2}$ to finally come back to the information we started with. The intermediate result is

$$\|w_j\|_{L_2}^2 \leq C\sum_{k>j} 2^{2tk}\|u_k\|_{L_2}^2 \ , \quad t > \frac{1}{2} \ ,$$

the final (after choosing an appropriate $1/2 < t < s$, substitution, and a change of the order of summation)

$$\sum_{j=0}^{\infty} 2^{2sj}\|\hat{u}_j\|_{L_2}^2 \leq C\sum_{j=0}^{\infty} 2^{2sj}(\|w_j\|_{L_2}^2 + \|u_j\|_{L_2}^2) \leq C\sum_{k=0}^{\infty} 2^{2sk}\|\hat{u}_k\|_{L_2}^2 \leq Ca_\Gamma(u,u) \ .$$

This proves the upper stability bound for the splitting in (41).

Implicitly, this proof contains the core idea of constructing trace spaces, restriction and extension operators from multilevel FE decompositions and their finite sections which are of importance in many problems, both on an algorithmical and theoretical level.

After having established the basic FE multilevel splittings (including boundary conditions if required), a lot of further desires come to mind: hierarchical bases (the FE counterpart of interpolatory MRA), semi-orthogonal FE schemes (suitable for H^s-problems with $s \leq 0$), biorthogonal Riesz systems (direct sum splittings) and so on. Some recent research for the linear FE MRA on polygonal domains in $\mathbb{R}^2$ is given in the next subsection. This material partly generalizes to other FE schemes and to $d = 3$. The last subsection briefly deals with more general domains in $\mathbb{R}^d$.

3.2 Linear FE multilevel systems

In this subsection we go to more details in a very specific case. We consider the FE MRA for C^0 Lagrange elements of degree $m = 1$, i.e. for linear finite elements. Everything simplifies, Theorem 7 is applicable for $0 < s < 3/2$, and the case $s = 1$ covers applications to standard second-order elliptic boundary value problems. Since this particular case is a central problem in the FE-community, studied in great detail and re-used in other applications, it deserves special attention.

We give some list of linear FE multilevel systems (other than the nodal basis frame), with references to algorithms which have been suggested before (and after) the terminology of multilevel systems was created, and some rough impression on the performance one can expect from one or the other multilevel scheme. On the other hand, we also borrow ideas from the wavelet community to further enlarge the number of examples. The reader can find a partly more exhaustive discussion in [49, 50], special examples appear also in greater detail in work by Stevenson [61, 64], Vassilevski [69], Dahmen et al. [12, 29] or were already covered by the material of [56, Section 4.2].

Unless stated otherwise, we consider the two-dimensional case $d = 2$, Ω is a bounded polygonal domain (or $\mathbb{R}^2$ if we discuss the wavelet counterparts of some examples), the refinement for producing the triangulations $\mathcal{T}_j$ is uniform and dyadic (as shown in Figure 13). The nodal basis functions $\phi_{j,i} \equiv \phi_{j,P}$ are the standard hat functions (value 1 at the associated vertex $P = P_{j,i}$ of $\mathcal{T}_j$). These sets of basis functions fit the concept of interpolatory scaling functions, as for spline spaces in one dimension, orthogonal bases in V_j would have global support. It is easy to geometrically count the dimension of V_j: $n_j \equiv \dim V_j = \#\Phi_j = \#\mathcal{V}_j$, where $\mathcal{V}_j = \{P_{j,i}\}$ denotes the vertex set of $\mathcal{T}_j$. All our further examples will aim at providing Riesz bases and are based on a direct-sum splitting

$$V_j = V_{j-1} \dot{+} W_j \ , \quad W_j = [\Psi_j] \ , \ j \geq 1 \ ,$$

and $m_j = \dim W_j = \#\Psi_j = n_j - n_{j-1}$ equals the number of edges in $\mathcal{T}_{j-1}$. Thus, it is quite natural to talk about functions $\psi_{j,i} \equiv \psi_{j,e}$ associated with the edges e from $\mathcal{T}_{j-1}$ or, equivalently, with the midpoints M_e of these edges (this set will be denoted by

$\mathcal{W}_j = \mathcal{V}_j \backslash \mathcal{V}_{j-1}$, $j \geq 1$). Thus, all our examples will be, in a certain sense, *edge-oriented* (just to mention, the 'geometric language' of vertices, edges, faces, etc., is also important for a purely applied reason: FE programs are often part of larger CAD systems, and geometrically oriented data structures are often natural for implementations on unstructured, adaptive grids). The numerical computations of condition numbers and iteration counts for the **SCG** method documented in the tables below have been done on a square domain (and with the spaces $\{\hat{V}_j\}$ rather than with $\{V_j\}$). The bilinear forms considered correspond to the Poisson problem with homogeneous Dirichlet boundary conditions ($s = 1$) and to the L_2-scalar product ($s = 0$). The stopping criterion for the pcg-iterations was given by a relative error reduction by $\leq 10^{-6}$.

- **Hierarchical basis** (Yserentant [72]). Consider the system

$$F_{HB} = \bigcup_{j=0}^{\infty} \{\phi_{j,i} \,:\, P_{j,i} \in \mathcal{W}_j\}$$

the finite sections $F_{J,HB} = F_{HB} \cap V_J$ of which form algebraic bases in V_J (and, therefore, Riesz bases in $\{V_J; (\cdot,\cdot)_{H^s}\}$ for any J and $s < 3/2$). Yserentant [72] has essentially proved that for $s = 1$ and $d = 2$ the Riesz bounds $\tilde{A}_J, \tilde{B}_J$ and the condition numbers $\kappa_J = \kappa(\mathcal{P}_{J,HB})$ of the scaled sections $F'_{J,HB}$ satisfy the asymptotic relations

$$\tilde{A}_J \asymp 1 \,, \quad \tilde{B}_J \asymp J^2 \quad \Longrightarrow \quad \kappa_J \asymp J^2 \,.$$

For the case $d = 3$, the growth is even exponential: $\kappa_J \asymp 2^J$.

Yserentant's preconditioner is identical, up to minor details such as the specific choices for $S_0 = A_0^{-1}$ and the diagonal matrices S_j, with the **AS** multilevel preconditioner for the splitting associated with $F_{J,HB}$. It is suboptimal for two-dimensional H^1-problems (i.e., the iteration count grows mildly rather than exponentially with J), and a huge improvement over solving the discretized system $A_J x_J = f_J$ without preconditioning. In contrast, solving generic L_2-elliptic problems based on $F_{J,HB}$ is not a good idea, compare the the numerical evidence provided at the end of this section. This observation is in full coincidence with our theory [56, Section 3.4.3] since for $d = 2$ (as expected from the material of Section 2.2) F'_{HB} is a frame in $H^s(\Omega)$ if and only if $1 < s < 3/2$. Thus the interval of optimality is far away from the L_2-case ($s = 0$).

Let us note some good things. The resulting algorithm is cheap with respect to the operation count $\#Ops_J$ (e.g., the operations $\hat{I}_j^T$, $\hat{I}_j$ do not require any serious action, just selecting a subvector resp. copying a vector into a longer zero-vector), is relatively robust, and simple to implement which explains its popularity (see the description of Bank's PLTMG). However, almost the same advantages are characteristic for the standard nodal basis frame which we will denote by F_{BPX} since the corresponding **AS** method, the so-called **BPX preconditioner**, was introduced by Bramble, Pasciak, Xu [10]. For H^1-problems, the practical condition numbers are not only bounded with J but also surprisingly small, see the corresponding table below for the values in the case of a two-dimensional square. Our experiments show that the detoriation of condition numbers for the L_2-case is also moderate (the computations reproduce exactly the theoretical behavior $\kappa_J \asymp J$ which follows from the considerations in [1]). Implementations are available for $d = 2$ and $d = 3$, e.g., the adaptive code Kaskade 3.x can be obtained via elib.

- **Three-point hierarchical basis** (Stevenson [61, 62]). In an attempt to improve the convergence behavior of the hierarchical basis method as well as to add robustness and to preserve the basis property, Stevenson proposes the following modification: For any e from $\mathcal{T}_{j-1}$, $j \geq 1$, with midpoint $M_e \in \mathcal{V}_j$ and endpoints $P, P' \in \mathcal{V}_{j-1}$, set
$$\psi_{j,e} = \phi_{j,e} + c_{j,e}\phi_{j,P} + c'_{j,e}\phi_{j,P'}$$
with constants $c_{j,e}, c'_{j,e}$ such that $\psi_{j,e}$ is discretely L_2-orthogonal to V_{j-1}. For a uniform grid, the constants are $c_{j,e} = c'_{j,e} = -1/2$. The resulting multilevel FE system will be denoted by F_{3HB} (the number 3 is chosen because $C_\psi = 3$). In [61], for $d = 2$ and uniform grids, the Riesz basis property in $H^s(\Omega)$ is essentially established for the range $-s_0 < s < 3/2$, with an unspecified $s_0 > 0$. Experiments for $d = 3$ [62] and recent theoretical investigations [63] yield an analogous result for the nested refinement case and $d \leq 3$. Thus, the use of $F_{J,3HB}$ improves the BPX-algorithm for the L_2 case while being only slightly more complicated and a bit slower in the H^1 case (the condition numbers reported in [61, 62] for the above mentioned generic H^1-problem are about 17 for $d = 2$ and 50 for $d = 3$ and $J \leq 7$). Due to the validity of the Riesz basis property also for negative Sobolev exponents, the systems $F_{J,3HB}$ might be attractive for integral and boundary integral equations, particularly for $d = 2$ (actually, in [26], the authors have used a numerical scheme for solving an L^2-elliptic integral equation based on this system before its detailed investigation by Stevenson). On uniform partitions and away from the boundary, the $\psi_{j,e}$ are orthogonal to $\mathbb{P}_1$, in the general case at least to constants. In [50], we have investigated some properties of the wavelet counterpart of F_{3HB} associated with the corresponding dyadic linear FE MRA on $\mathbb{R}^d$, $d \leq 3$. Using the machinery of multivariate wavelet theory, we could compute the exact range $-0.992036 < s < 3/2$ for the Sobolev exponent s such that F'_{3HB} is a Riesz basis in $H^s(\mathbb{R}^d)$. The result does holds for $d \leq 3$, and shows the flexibility and potential of this modification for various standard applications.

- **L_2-semiorthogonal constructions** ([42], [46], [64], [35]) According to the general results of section 1.1, it would be desirable to construct a L_2-stable basis $\{\psi_{j,e}\}$ of the L_2-orthogonal complement space $W_j = V_j \ominus_{L_2} V_{j-1}$, $j \geq 1$, since in that case the resulting (scaled) system F'_{L_2} will be a Riesz basis in $H^s(\Omega)$ for all $-3/2 < s < 3/2$. The proof of the existence of such basis functions $\psi_{j,e}$, with support in a finite union of simplices near the associated edge e (uniformly in j) for all sequences of quasi-uniform triangulations $\{\mathcal{T}_j\}$ is still an open question. However, the following local procedure worked extremely well for $d = 2$. Let $\Omega_{j,e}$ be the union of all triangles in $\mathcal{T}_{j-1}$ the closure of which has at least one intersection point with the closed edge e. Let P, P' denote the endpoints of e, and $\mathcal{E}_{j-1,e}$ the set of all edges e' in $\mathcal{T}_{j-1}$ emanating from either P or P', with the exception of e. Then, we determine

$$\psi_{j,e} = \phi_{j,M_e} + a_P \phi_{j,P} + a_{P'} \phi_{j,P'} + \sum_{e' \in \mathcal{E}_{j-1,e}} b_{e'} \phi_{j,M_{e'}}$$

such that it is L_2-orthogonal to V_{j-1}, and

$$\sum_{e' \in \mathcal{E}_{j-1,e}} b_{e'}^2 \to \min \ .$$

If this construction is applied in the setting of the dyadic linear FE MRA on $\mathbb{R}^2$, the corresponding systems Ψ_j have indeed the desired Riesz basis properties as shown in [46]. Below, the support of the $\psi_{j,e}$ associated with a vertical edge and the 13 non-zero coefficients in the ψ-mask are shown. The masks for the other edges can be obtained by rotation. For comparison, the masks are also depicted for some of the following multilevel suystems.

No rigorous proof of the Riesz basis property of such F'_{L_2} is available for more general polygonal domains resp. for $d > 2$. A similar construction has been considered in a thesis by Junkherr [42], in connection with the solution of integral equations. Very recently, Floater and Quak [35] proposed another local scheme of constructing piecewise linear semi-orthogonal prewavelets, and proved the Riesz basis property of the resulting F'_{L_2} under weak conditions on the triangulations. Before that, for shift-invariant simplicial partitions of $\mathbb{R}^d$, whole families of semiorthogonal box spline wavelets have been introduced by several authors, however, all these examples exhibit larger supports and masks.

Stevenson [64] has proposed another explicit construction which works for more general situations and is in addition more economical to implement then the ones previously discussed. To fix the idea, here is the construction. In a first step, let us construct a biorthogonal system $\tilde{\Phi}_{j-1} \equiv \{\tilde{\phi}_{j-1,P} \, : \, P \in \mathcal{V}_{j-1}\} \subset V_j$ for Φ_j, i.e.,

$$(\phi_{j-1,P}, \tilde{\phi}_{j-1,Q})_{L_2} = \delta_{PQ} \quad \forall \ P, Q \in \mathcal{V}_{j-1} \ .$$

It turns out that under the above assumptions on the refinement process, one can find $\tilde{\phi}_{j-1,P}$ such that its support is in the support of $\phi_{j-1,P}$, $P \in \mathcal{V}_{j-1}$. Clearly, the biorthogonal system $\tilde{\Phi}_{j-1}$ could be used to define a quasi-interpolant projection onto V_{j-1}:

$$Q_{j-1} u = \sum_{P \in \mathcal{V}_{j-1}} (u, \tilde{\phi}_{j-1,P})_{L_2} \phi_{j-1,P} \ ,$$

see below. The remarkable fact is that the range of the operator $\mathrm{Id}_j - \tilde{Q}_{j-1} : V_j \to V_j$, where

$$\tilde{Q}_{j-1}u = \sum_{P\in\mathcal{V}_{j-1}} (u, \phi_{j-1,P})_{L_2}\tilde{\phi}_{j-1,P} ,$$

is orthogonal to V_{j-1}. Hence, setting

$$\tilde{\psi}_{j,e} = \phi_{j,M_e} - \tilde{Q}_{j-1}\phi_{j,M_e}$$

where e is any edge in $\mathcal{T}_{j-1}$, one can prove under reasonable assumptions on the refinement process that the resulting systems Ψ_j form Riesz bases in the orthogonal complement spaces $V_j \ominus_{L_2} V_{j-1}$, uniformly in $j \geq 1$. For details, we refer to the original paper [64] by Stevenson (compare also [29] for generalizations).

- **Biorthogonal Riesz bases** obtained by 'repairing' F_{HB} (Vassilevski, Wang [70, 71, 69], Carnicer, Dahmen, Peña [12], Sweldens [66], Dahmen, Stevenson [29]). In these papers, the above semiorthogonal complement space W_j is replaced by

$$W_j^{approx} = [\{\psi_{j,e}\}] , \quad \psi_{j,e} = (\mathrm{Id} - Q_{j-1}^{approx})\phi_{j,M_e} ,$$

where $Q_{j-1}^{approx} : V_j \to V_{j-1}$ are suitable substitutes for the L_2-orthogonal projection Q_{j-1} onto V_{j-1} (if the latter are used, the W_j of the previous example would have resulted). This can be viewed as a *coarse-grid correction* substracted from the functions in the hierarchical basis F_{HB}. If worked out in the setting of a dyadic MRA on $\mathbb{R}^d$, such corrections always lead to finitely supported masks for the potential dual scaling function $\tilde{\phi}$ which does not yet guarantees the existence of $\tilde{\phi}$ in the L_2-sense but greatly simplifies the analysis, and is desirable for a number of applications. General contributions in this direction are the papers [66], [12].

The choice of Q_{j-1}^{approx} in [70, 71] is described in the form

$$Q_{j-1;m}^{approx} = p_{m-1}(G_{j-1})I_j^* G_j , \qquad p_{m-1}(t) = \frac{1}{\beta}\sum_{l=0}^{m-1}(1 - \frac{t}{\beta})^l ,$$

where β is an upper bound for the spectrum of the Gram matrix G_{j-1} of the nodal basis Φ_{j-1} of V_{j-1}, and I_j^* is the adjoint of the natural embedding of $V_{j-1} \to V_j$. The main result of [70, part I] is, reformulated in our language, that the scaled version of the resulting system F^{approx} is a Riesz basis in $H^s(\Omega)$ for a given $0 < s \leq 1$ if m is chosen sufficiently large. The proof assumes the validity of the result for the nodal basis FE frame but otherwise it is general. The numerical experiments [70, part II] show that already small m lead to significant improvements (compared to the use of F_{HB}), however, the BPX-method is not outperformed. Below, we will look at the cheapest case $m = 1$ of this construction, with the exact value $\beta = h_{j-1}^2$, suitable for uniform triangulations in $\mathbb{R}^2$. [12, Section 5.2] discusses (as a specific example within a more general approach) the use of a standard quasi-interpolant operator for Q_{j-1}^{approx} .

In [50], a 1-parameter family of wavelet functions ψ_e was introduced which generalizes both examples (again in the setting of a dyadic linear FE MRA on $\mathbb{R}^2$). Some details are given These ψ_e possess vanishing moments of order 2 (orthogonality of the ψ to $\mathbb{P}_1$). Numerical computations of the range of the Sobolev exponent s for the Riesz basis property to hold in $H^s(\mathbb{R}^d)$ are reported on in [50], and provide interesting insight into the potential of these quite similar proposals. According to our computations, $a = -3/16$ is the optimal choice and gives the largest s-interval. Surprisingly enough, this special case was already considered in [21] (it appears to be the only choice of the parameter a where the dual $\tilde{\phi}$ is compactly supported). However, we do not know of a natural generalization of this example to more general domains and triangulations. Numerical tests reported on below show the quite good performance of the corresponding multilevel preconditioner. One should add that for coarse-grid corrected multilevel schemes, the complexity of $\hat{I}_j$ is not determined by the average size C_ψ of the ψ-masks as defined above but, provided a slightly modified implementation, by a smaller number. In the particular case, it is equivalent to setting $C_\psi = 5$ which has motivated us to call this type of linear FE multilevel system a 5-point hierarchical basis, too.

At the end of this section, we also show some results from [50] for a **2-point hierarchical basis**. It is similar to the Stevenson system, but uses a correction only at one endpoint of the edge e, say, at P. Orthogonality to $\mathbb{P}_0$ can be achieved on general grids, however, there are no good rules for the choice of P so far. Finally, we give some more precise information about $\#Ops_J$ for a generic implementation of the cg-method with any of the multilevel preconditioners surveyed in this section. The 'rule of thumb' for calculating the factor N_ψ associated with the choice of the ψ is $N_\psi = 2C_\psi + 1$. Due to a certain large overhead (scalar products etc.) of the cg-method, most of the above examples of multilevel systems would lead to comparable overall complexity estimates per iteration. This may change for the multiplicative versions of the algorithms or if parallel computing is a must. At least for general refinement schemes, there is also a considerable setup time and an increased storage for all components of the multilevel preconditioner C_J.

3.3 More general domains

We will not give a detailed treatment of this topic. We will go back to the wavelet setting of a dyadic MRA on $\mathbb{R}^d$. Examples can be obtained from dyadic MRA on $\mathbb{R}$. First variant: *Tensor product wavelet systems* such as

$$F^{(1)} \otimes F^{(2)} = \{\psi_{j,i;j',i'}(x_1, x_2) = \psi^{(1)}_{j,i}(x_1)\psi^{(2)}_{j',i'}(x_2)\}$$

for $d = 2$ (with applications to *anisotropic problems* and *large-d-problems* using sparse grid resp. hyperbolic cross approximations). Second variant: *Tensor product dyadic MRA* can be generated from $V_0 = V_0^{(1)} \otimes V_0^{(2)}$ with scaling function

$$\phi(x_1, x_2) = \phi^{(1)}(x_1)\phi^{(2)}(x_2)$$

and detail space

$$W_1 = W_1^{(1)} \otimes V_0^{(2)} + V_0^{(1)} \otimes W_1^{(2)} + W_1^{(1)} \otimes W_1^{(2)} \; ,$$

(to each term in the sum of spaces, a natural generating ψ is associated). This is the standard for isotropic behavior. The supports of scaling and wavelet functions of level j are d-dimensional rectangles of size $\asymp 2^{-j}$ in each coordinate direction. Alternatively, *true multivariate MRAs*, with less directional preference or more general dilation matrices, are a theoretical option. We do not know attractive examples for possible PDE applications, except for FE MRA as discussed before. *Box spline MRA* such as the linear FE MRA are intermediate, they are, in some sense, tensor products with more than d directions (look at the definition of box splines in the Fourier transform domain, [8]).

How to adapt a multilevel system F coming from a dyadic MRA on $\mathbb{R}^d$ to a domain? The case $[0,1]^d$ (or of domains that are logical cubes) is treatable, look at the literature on wavelets on $[0,1]$ and use tensor products. For some special types of wavelet functions (such as discussed in subsection 2.5) it is almost obvious what to do. Exercise: Construct spaces suitable for an L-shaped domain such as $[-1,1]^2 \backslash [0,1]^2$! To just give the reader an orientation, here are, in our opinion, the main-stream options.

- **Interior constructions including a boundary modification.** Here, one starts with dividing the shifts of scaling functions/wavelets into two groups: interior and boundary group. The rule of thumb is to declare a wavelet/scaling function as interior if the distance of its support to the boundary is larger than a fixed constant times the diameter of the support. Then think of a reasonable transformation of the boundary functions. See the Cohen/Dahmen/DeVore construction of a biorthogonal Riesz basis for rather general domains. For extensions of the frame concept in connection with solvers for elliptic problems of $2m$-th order, see [57].

- **Glueing patches together.** For this CAGD-motivated approach, see the survey [25], natural for manifolds but also for other patch-decomposable domains.

- **Domain embedding techniques** We mean mainly *fictitious domain and fictitious space methods*: The problem is mapped into a *larger* problem (say, the wavelet-discretized analog on a rectangular domain containing Ω), and the problem are adequate *transfer operations (extension and restriction)*, see [58] for an example where wavelets are used for the extended problem, compare also [57]. Appending boundary conditions (see, e.g., [45]) may also be formulated as a variant of the fictitious space method involving a larger, unconstrained but mixed problem.

- **Do not touch general domains by wavelet methods!** This recommendation should not be taken too seriously but in the long run this might happen, for several reasons, among them practical performance. There are enough problems involving smoothness which live on well-structured domains. From time to time, an analogous discussion (structured versus unstructured grids) can be heard in conferences devoted to classical finite element/difference discretization techniques. At least for very complicated domains (and problems which are far away from the generic H^s-elliptic case), the use of prefixed, problem-independent multilevel systems is not justifiable anymore.

References

[1] P. Oswald, Frames and space splittings in Hilbert spaces, Lectures, Part 1.

[2] R. Bank, PLTMG: A software package for solving elliptic partial differenetial equations. User's guide 7.0, Frontiers in Appl. Math. 15, SIAM, Philadelphia, 1994.

[3] R. A. Adams, Sobolev spaces, Academic Press, New York, 1975.

[4] P. Auscher, Wavelet bases for $L_2(\mathbb{R})$ with rational dilation factor, in [60], p. 439–451.

[5] J. J. Benedetto, M. W. Frazier (eds.), Wavelets: Mathematics and Applications, CRC Press, Boca Raton, 1994.

[6] S. Bertoluzza, G. Naldi, A wavelet collocation method for the numerical solution of partial differential equations, ACHA 3 (1996), 1–9.

[7] C. de Boor, R. DeVore, A. Ron, Approximation from shift-invariant subspaces of $L_2(\mathbb{R}^d)$, TAMS 341 (1994), 787–806.

[8] C. de Boor, K. Höllig, S.Riemenschneider, Box splines, Springer, New York, 1993.

[9] J. Bramble, Multigrid methods, Pitman, Longman Sci.&Tech., Harlow, 1993.

[10] J. Bramble, J. Pasciak, J. Xu, Parallel multilevel preconditioners, Math. Comp. 55 (1990), 1–22.

[11] A. Brandt, W. Lubrecht, Multilevel matrix multiplication and fast solution of integral equations, J. Comput. Phys. 90 (1990), 348–370.

[12] J. M. Carnicer, W. Dahmen, J. M. Peña, Locally finite decompositions of nested spaces, Appl. Comput. Harmon. Anal. 3, 1996, 127–153.

[13] A. S. Cavaretta, W. Dahmen, C. A. Micchelli, Stationary subdivision, Memoirs AMS, vol. 93, n. 453, AMS, Providence, 1991.

[14] C. K. Chui, An introduction to wavelets, Academic Press, Boston, 1992.

[15] C. K. Chui (ed.), Wavelets: A tutorial in theory and applications. Academic Press, Boston, 1992.

[16] C. K. Chui, L. L. Schumaker (eds.), Approximation Theory VIII, vol. 2: Wavelets and multilevel approximation. World Scientific, Singapore, 1995.

[17] C. K. Chui, J.-A. Lian, A study of orthonormal multi-wavelets, J. Appl. Numer. Math. (to appear).

[18] P. Ciarlet, The finite element method for elliptic problems, North-Holland, Amsterdam, 1978.

[19] A. Cohen, I. Daubechies, J.-C. Feauveau, Biorthogonal bases of compactly supported wavelets, Comm. Pure Appl. Math. 45, 1992, 485–560.

[20] A. Cohen, I. Daubechies, A new technique to estimate the regularity of refinable functions, Rev. Iberoamericana 12, 1996, 527–591.

[21] A. Cohen, J. M. Schlenker, Compactly supported wavelets with hexagonal symmetry, Constr. Approx. 9, 1993, 209–236.

[22] S. Dahlke, I. Weinreich, Wavelet bases adapted to pseudodifferential operators, ACHA 1 (1994), 267–283.

[23] W. Dahmen, Stability of multiscale transformations, J. Fourier Anal. Appl. 4, 1996, 341–361.

[24] W. Dahmen, Multiscale analysis, approximation, and interpolation spaces, in [16], p. 47–88.

[25] W. Dahmen, Wavelet and multiscale methods for operator equations, Acta Numerica, Cambridge Univ. Press, 6, 1997, pp. 55–228.

[26] W. Dahmen, B. Kleemann, S. Prössdorf, R. Schneider, A multiscale method for the double layer potential equation on a polyhedron, in Advances in Computational Mathematics (H. P. Dikshit, C. A. Micchelli, eds.), World Scientific, Singapore, 1994, pp. 1–40.

[27] W. Dahmen, A. Kunoth, Multilevel Preconditioning, Numer. Math. 63 (1992), 315-344.

[28] W. Dahmen, P. Oswald, X.-Q. Shi, C^1-hierarchical bases, J. Comput. Appl. Math. 51 (1994), 37–56.

[29] W. Dahmen, R. Stevenson, Element-by-element construction of wavelets satisfying stability and moment conditions, IGPM-Report 145, RWTH Aachen, November 1997.

[30] I. Daubechies, Ten Lectures on Wavelets, CBMS-NSF Regional Conf. Ser. in Applied Math., vol. 61, SIAM, Philadelphia, 1992.

[31] R. DeVore, B. Lucier, Wavelets, Acta Numerica 91, Cambr. Univ. Press, Cambridge, 1991, 1–56.

[32] G. C. Donovan, J. S. Geronimo, D. P. Hardin, P. R. Massopust, Construction of orthogonal wavelets using fractal interpolation functions, SIAM J. Math. Anal. 27 (1996), 1158–1192.

[33] T. Eirola, Sobolev characterization of solutions of dilation equations, SIAM J. Math. Anal. 23 (1992), 1015–1030.

[34] G. Farin, Triangular Bernstein-Bezier patches, CAGD 3 (1986), 83–127.

[35] M. S. Floater, E. Quak, Piecewise linear prewavelets on arbitrary triangulations, Numer. Math. (submitted).

[36] W. Hackbusch, Iterative solution of large sparse systems of equations. Appl. Math. Sci. vol. 95, Springer, New York, 1994.

[37] W. Hackbusch, Z. P. Nowak, On the fast matrix multiplication in the boundary element method by panel clustering, Numer. Math. 54, 1989, 463–491.

[38] B. Jawerth, W. Sweldens, An overview of wavelet based multiresolution analyses, SIAM Review, 36 (1994), 377–412.

[39] R.-Q. Jia, Refinable shift-invariant spaces: from splines to wavelets, in [16], p. 179–208.

[40] R. Q. Jia, C. A. Micchelli, Using the refinement equation for the construction of pre-wavelets II: Powers of two, In: Curves and Surfaces (P. J. Laurent, A. LeMehaute, L.L.Schumaker, eds.), Academic Press, Boston 1992, 209–246.

[41] R.-Q. Jia, S. D. Riemenschneider, D.-X. Zhou, Smoothness of multiple refinable functions and multiple wavelets, Preprint, Univ. of Alberta, Edmonton, 1997.

[42] J. Junkherr, Effiziente Lösung von Gleichungssystemen, die aus der Diskretisierung von schwach singulären Integralgleichungen 1. Art herrühren (in German), Dissertation, Univ. Kiel, 1994.

[43] J.-P. Kahane, P.-G. Lemarie-Rieusset, Fourier Series and Wavelets, Gordon & Breach, Luxembourg, 1995.

[44] J. Ko, A. Kurdila, P. Oswald, Scaling function and wavelet preconditioners for second order elliptic problems, In: Multiscale Wavelet Methods for PDE (W. Dahmen, A. Kurdila, P. Oswald, eds.), Academic Press, San Diego, 1997, 413–438.

[45] A. Kunoth, Multilevel preconditioning - Appending boundary conditions by Lagrange multipliers, Advances in Comp. Math. 4 (1995), 145–170.

[46] U. Kotyczka, P. Oswald, Piecewise linear prewavelets of small support, in [16], p. 235–242.

[47] R. Lorentz, W. Madych, Spline wavelets for ordinary differential equations, Arbeitspapiere GMD No. 562, GMD, St. Augustin, 1991.

[48] R. Lorentz, P. Oswald, Nonexistence of compactly supported box spline prewavelets in Sobolev spaces, in: Curves and Surfaces (A. Le Mehaute, C. Rabut, L. Schumaker, eds.), Vanderbilt Univ. Press, Nashvillew, 1997, pp. 235–244.

[49] R. Lorentz, P. Oswald, Multilevel finite element Riesz bases in Sobolev spaces, In: DD9 Proceedings (P. Bjorstad, M. Espedal, D. Keyes, eds.), Wiley, New York, 1997 (to appear).

[50] R. Lorentz, P. Oswald, Criteria for hierarchical bases in Sobolev spaces. GMD-Arbeitsbericht 1059, GMD, Sankt Augustin, March 1997; ACHA (submitted).

[51] P. Massopust, Fractal functions, fractal surfaces, and Wavelets, Acad. Press, San Diego, 1994.

[52] Y. Meyer, Ondelettes, vol. 1,2, Hermann, Paris, 1990.

[53] Y. Meyer, Wavelets: Algorithms & Applications, SIAM Philadelphia, 1993.

[54] P. Oswald, On discrete norm estimates related to multilevel preconditioners in the finite element method, In: Constr. Theory of Functions, Proc. Int. Conf. Varna 1991 (K. G. Ivanov, P. Petrushev, B. Sendov, eds.), Bulg. Acad. Sci., Sofia 1992, pp. 203–214.

[55] P. Oswald, Stable subspace splittings for Sobolev spaces and some domain decomposition algorithms, In: DD7 Proceedings (D. Keyes, J. Xu, eds.), Cont. Math., vol. 180, AMS, Providence 1994, pp. 87–98.

[56] P. Oswald, Multilevel finite element approximation: theory & applications, Teubner Skripte zur Numerik, Teubner, Stuttgart 1994.

[57] P. Oswald, Multilevel solvers for elliptic problems on domains, in: Multiscale Wavelet Methods for PDE (W. Dahmen, A. Kurdila, P. Oswald, eds.), Academic Press, San Diego, 1997, 3–58.

[58] A. Rieder, A domain embedding method for Dirichlet problems in arbitrary space dimensions, Preprint, Univ. Saarbrücken, July 1996.

[59] A. Ron, Z. Shen, Affine systems in $L_2(\mathbb{R}^d)$: the analysis of the analysis operator, J. Funct. Anal. (to appear), Preprint December 1995, Univ. of Wisconsin, Madison.

[60] M. R. Ruskai, G. Beylkin, R. Coifman, I. Daubechies, S. Mallat, Y. Meyer, L. Raphael (eds.), Wavelets and Their Applications, Jones & Bartlett, Boston, 1992.

[61] R. Stevenson, A robust hierarchical basis preconditioner on general meshes, Report 9533, Dep. Math., Univ. Nijmegen, September 1995, Numer. Math.).

[62] R. Stevenson, Experiments in 3D with a three-point hierarchical basis preconditioner, Appl. Numer. Math. 23, 1997, 159–176.

[63] R. Stevenson, Stable three-point wavelet bases on general meshes, Report 9627, Dep. Math., Univ. Nijmegen, October 1996.

[64] R. Stevenson, Piecewise linear (pre-)wavelets on general meshes, Report 9701, Dep. Math., Univ. Nijmegen, January 1997.

[65] J. Stöckler, Multivariate Wavelets. in: Wavelets - A Tutorial in Theory and Applications (C. K. Chui, ed.), Academic Press, Boston 1992, pp. 325–355.

[66] W. Sweldens, The lifting scheme: A custom-design construction of biorthogonal wavelets, Appl. Comput. Harmon. Anal. 3, 1996, 186–200.

[67] H. Triebel, Interpolation Theory, Function Spaces, Differential operators, 2nd edition, Barth, Heidelberg, 1995.

[68] H. Triebel, Theory of Function Spaces II, Birkhäuser, Basel, 1992.

[69] P. S. Vassilevski, On two ways of stabilizing the hierarchical basis multilevel methods, SIAM Review 39 (1997), 18–53.

[70] P. S. Vassilevski, J. Wang, Stabilizing the hierarchical basis by approximate wavelets, I: Theory. Numer. Lin. Alg. Appl. 4, II: Implementation and numerical results, SIAM J. Sci. Comput. (submitted).

[71] P. S. Vassilevski, J. Wang, Wavelet-like methods in the design of efficient multilevel preconditioners for elliptic PDEs, In: Multiscale Wavelet Methods for PDE (W. Dahmen, A. Kurdila, P. Oswald, eds.), Academic Press, San Diego, 1997, 59–105.

[72] H. Yserentant, On the multilevel-splitting of finite element spaces, Numer. Math. 49 (1986), 379–412.