1. Adaptive wavelet schemes for elliptic problems

1.1. **The basic wavelet setting.** We briefly state the wavelet setting as far as it is needed for our purposes. In general, a *wavelet basis* $\Psi = \{\psi_\lambda : \lambda \in \mathcal{J}\}$ is a basis for an L_2-space with specific properties outlined below. The indices $\lambda \in \mathcal{J}$ typically encode several types of information, namely the *scale*, often denoted by $|\lambda|$, the *spatial location*, and also the *type* of the wavelet. For instance, on the real line, $|\lambda| = j \in \mathbb{Z}$ denotes the dyadic refinement level and $2^{-j}k$ with $k \in \mathbb{Z}$ stands for the location of the wavelet.

We will ignore any explicit dependence on the type of the wavelet from now on, since this only produces additional constants. Hence, we frequently use $\lambda = (j,k)$ and

$$\mathcal{J} = \{(j,k) : j \geq j_0, k \in \mathcal{J}_j\},$$

where $\mathcal{J}_j$ is some countable index set and $|(j,k)| = j$. Moreover, $\tilde{\Psi} = \{\tilde{\psi}_\lambda : \lambda \in \mathcal{J}\}$ denotes the *dual wavelet basis*, which is biorthogonal to Ψ, i.e.,

$$\langle \psi_\lambda, \tilde{\psi}_{\lambda'} \rangle_{L_2(\Omega)} = \delta_{\lambda,\lambda'}.$$

We assume that the domain $\Omega \subset \mathbb{R}^d$, $d = 1, 2$ under consideration enables us to construct a wavelet basis Ψ with the following properties:

(1) the wavelets form a *Riesz basis* for $L_2(\Omega)$;
(2) the *cardinalities* of the index sets $\mathcal{J}_j$ satisfy

$$\#\mathcal{J}_j \asymp 2^{jd};$$

(3) the wavelets are *local* in the sense that

$$\operatorname{diam}(\operatorname{supp} \psi_\lambda) \asymp 2^{-|\lambda|};$$

(4) the wavelets satisfy the *cancellation property*

$$|\langle v, \psi_\lambda \rangle_{L_2(\Omega)}| \preceq 2^{-|\lambda|(\frac{d}{2}+\widetilde{m})} |v|_{W^{\widetilde{m}}(L_\infty(\operatorname{supp} \psi_\lambda))}$$

for $|\lambda| > j_0$ with some parameter $\widetilde{m} \in \mathbb{N}$;
(5) the wavelet basis induces characterizations of *Besov spaces* $B^t_q(L_p(\Omega))$ of the form

$$\|v\|_{B^t_q(L_p(\Omega))} \asymp \left(\sum_{j=j_0}^{\infty} 2^{j(t+d(\frac{1}{2}-\frac{1}{p}))q} \left(\sum_{k \in \mathcal{J}_j} |\langle v, \tilde{\psi}_{j,k} \rangle_{L_2(\Omega)}|^p \right)^{q/p} \right)^{1/q},$$

for $0 < p, q < \infty$ and all t with $d(\frac{1}{p} - 1)_+ < t < s$ for some parameter $s > 0$.

In (5) the upper bound s depends, in particular, on the smoothness and the approximation properties of the wavelet basis.

Using the fact that $B^t_2(L_2(\Omega)) = H^t(\Omega)$ and by exploiting the norm equivalence (5), a simple rescaling of Ψ immediately yields a Riesz basis

for $H^t(\Omega)$ with $0 < t < s$. If homogeneous Dirichlet boundary conditions are prescribed we will actually assume that the primal wavelet basis Ψ characterizes the Sobolev spaces $H_0^t(\Omega)$ instead of $H^t(\Omega)$. Suitable constructions of wavelets on domains can be found in [1, 5, 8–12], and we also refer to [2] for a detailed discussion.

1.2. **Basis case.** After this short introduction of wavelet bases we briefly show how they can be used to treat elliptic operator equations of the form

$$\mathcal{A}u = f, \tag{1.1}$$

where we will assume $\mathcal{A}$ to be a boundedly invertible operator from some Hilbert space $\mathcal{H}$ into its normed dual $\mathcal{H}'$, i.e.,

$$\|\mathcal{A}u\|_{\mathcal{H}'} \sim \|u\|_{\mathcal{H}}, \quad u \in \mathcal{H}.$$

In our applications $\mathcal{H}$ is typically a Sobolev space $H^t(\Omega)$ or $H_0^t(\Omega)$ on some domain $\Omega \subset \mathbb{R}^d$, $d = 1, 2$. We shall mainly focus on the special case where

$$a(v, w) := \langle \mathcal{A}v, w\rangle$$

defines a *symmetric* bilinear form on $\mathcal{H}$ which is *elliptic* in the sense that

$$a(v, v) \sim \|v\|_{\mathcal{H}}^2. \tag{1.2}$$

Usually, operator equations of the form (1.1) are solved by a Galerkin scheme, i.e., one defines an increasing sequence of finite-dimensional approximation spaces $S_{\Lambda_l} := \operatorname{span}\{\eta_\mu : \mu \in \Lambda_l\}$, where $S_{\Lambda_l} \subset S_{\Lambda_{l+1}}$, and projects the problem onto these spaces, i.e.,

$$\langle \mathcal{A}u_{\Lambda_l}, v\rangle = \langle f, v\rangle \quad \text{for all} \quad v \in S_{\Lambda_l}.$$

To compute the actual Galerkin approximation, one has to solve a linear system

$$\mathbf{G}_{\Lambda_l}\mathbf{c}_{\Lambda_l} = \mathbf{f}_{\Lambda_l}, \quad \mathbf{G}_{\Lambda_l} = (\langle \mathcal{A}\eta_{\mu'}, \eta_\mu\rangle)_{\mu,\mu'\in\Lambda_l}, \quad \mathbf{f}_{\Lambda_l} = (\langle f, \eta_\mu\rangle)_{\mu\in\Lambda_l}.$$

Then the question arises how to choose the approximation spaces in a suitable way, for doing that in a somewhat clumsy fashion would yield a very inefficient scheme. One natural idea would be to use an *adaptive* scheme, i.e., an updating strategy which essentially consists of the following three steps:

solve	–	estimate	–	refine
$\mathbf{G}_{\Lambda_l}\mathbf{c}_{\Lambda_l} = \mathbf{f}_{\Lambda_l}$		$\|u - u_{\Lambda_l}\| =?$ a posteriori error estimator		add functions if necessary.

Already the second step is highly nontrivial since the exact solution u is unknown, so that clever a posteriori error estimators are needed. Then another challenging task is to show that the refinement strategy leads to a convergent scheme and to estimate its order of convergence,

if possible. In recent years, it has been shown that both tasks can be solved if wavelets are used as basis functions for the Galerkin scheme as we shall now explain.

The first step is to transform (1.1) into a discrete problem. Property (5) implies that a rescaled version of the wavelet basis Ψ gives rise to a Riesz basis for $\mathcal{H}$. By using this fact it is easy to see that (1.1) is equivalent to

$$\mathbf{A}\mathbf{u} = \mathbf{f} \tag{1.3}$$

where

$$\mathbf{A} := \mathbf{D}^{-1}\langle \mathcal{A}\Psi, \Psi\rangle^T\mathbf{D}^{-1},\ \mathbf{u} := \mathbf{D}\mathbf{c},\ u = \mathbf{c}^T\Psi,\ \mathbf{f} := \mathbf{D}^{-1}\langle f, \Psi\rangle^T,$$

and $\mathbf{D} = \left(2^{-t|\lambda|}\,\delta_{\lambda,\lambda'}\right)_{\lambda,\lambda'\in\mathcal{J}}$ is a diagonal scaling matrix.

Now (1.2) implies that

$$\|\mathbf{A}\|_{\mathcal{L}(\ell_{\mathbf{2}}(\mathcal{J}))},\ \|\mathbf{A^{-1}}\|_{\mathcal{L}(\ell_{\mathbf{2}}(\mathcal{J}))} < \infty,$$

and the computation of the Galerkin approximation amounts to solving the system

$$\mathbf{A}_\Lambda\mathbf{u}_\Lambda = \mathbf{f}_\Lambda := \mathbf{f}|_\Lambda, \qquad \mathbf{A}_\Lambda := (\mathbf{D}^{-1}\langle \mathcal{A}\Psi, \Psi\rangle^T\mathbf{D}^{-1})|_\Lambda.$$

Now, ellipticity (1.2) and Riesz property yield

$$\|\mathbf{u} - \mathbf{u}_\Lambda\|_{\ell_2(\mathcal{J})} \sim \|\mathbf{A}(\mathbf{u} - \mathbf{u}_\Lambda)\|_{\ell_2(\mathcal{J})} \sim \|\mathbf{f} - \mathbf{A}\mathbf{u}_\Lambda\|_{\ell_2(\mathcal{J})} \sim \|\mathbf{r}_\Lambda\|_{\ell_2(\mathcal{J})},$$

so that the ℓ_2-norm of the *residual* $\mathbf{r}_\Lambda$ serves as an a posteriori error estimator. Each individual coefficient $(\mathbf{r}_\Lambda)_\lambda$ can be viewed as a local error indicator. Therefore a natural adaptive strategy would consist in catching the bulk of the residual, i.e., to choose the new index set $\hat{\Lambda}$ such that

$$\|\mathbf{r}_\Lambda|_{\hat{\Lambda}}\|_{\ell_2(\mathcal{J})} \geq \zeta\|\mathbf{r}_\Lambda\|_{\ell_2(\mathcal{J})}, \quad \text{for some} \quad \zeta \in (0,1).$$

However, such a scheme would not be implementable since the residual involves infinitely many coefficients. To transform this idea into an implementable scheme, the following three subroutines can be utilized:

- **RHS**$[\varepsilon, \mathbf{g}] \to \mathbf{g}_\varepsilon$: determines for $\mathbf{g} \in \ell_2(\mathcal{J})$ a finitely supported $\mathbf{g}_\varepsilon \in \ell_2(\mathcal{J})$ such that
$$\|\mathbf{g} - \mathbf{g}_\varepsilon\|_{\ell_2(\mathcal{J})} \leq \varepsilon.$$
- **APPLY**$[\varepsilon, \mathbf{A}, \mathbf{v}] \to \mathbf{w}_\varepsilon$: determines for a finitely supported $\mathbf{v} \in \ell_2(\mathcal{J})$ a finitely supported $\mathbf{w}_\varepsilon$ such that
$$\|\mathbf{A}\mathbf{v} - \mathbf{w}_\varepsilon\|_{\ell_2(\mathcal{J})} \leq \varepsilon.$$
- **COARSE**$[\varepsilon, \mathbf{v}] \to \mathbf{v}_\varepsilon$: determines for a finitely supported $\mathbf{v} \in \ell_2(\mathcal{J})$ a finitely supported $\mathbf{v}_\varepsilon \in \ell_2(\mathcal{J})$ with at most M significant coefficients, such that
$$\|\mathbf{v} - \mathbf{v}_\varepsilon\|_{\ell_2(\mathcal{J})} \leq \varepsilon. \tag{1.4}$$

Moreover, $M \lesssim M_{\min}$ holds, $M_{\min}$ being the minimal number of entries for which (1.4) is valid.

Then, employing the key idea outlined above, the resulting fundamental algorithm reads as follows:

Algorithm 1.1. SOLVE$[\varepsilon, \mathbf{A}, \mathbf{f}] \to \mathbf{u}_\varepsilon$

$\Lambda_0 := \emptyset$; $\mathbf{r}_{\Lambda_0} := \mathbf{f}$; $\varepsilon_0 := \|\mathbf{f}\|_{\ell_2(\mathcal{J})}$; $j := 0$; $u_0 := 0$;
While $\varepsilon_j > \varepsilon$ do
  $\varepsilon_j := 2^{-(j+1)}\|\mathbf{f}\|_{\ell_2(\mathcal{J})}$; $\Lambda_{j,0} := \Lambda_j$; $\mathbf{u}_{j,0} := \mathbf{u}_j$;
  For $k = 1, ..., K$ do
    Compute Galerkin approximation $\mathbf{u}_{\Lambda_{j,k-1}}$ for $\Lambda_{j,k-1}$;
    Compute $\tilde{\mathbf{r}}_{\Lambda_{j,k-1}} := \mathbf{RHS}[c_1\varepsilon_{j+1}, \mathbf{f}] - \mathbf{APPLY}[c_2\varepsilon_{j+1}, \mathbf{A}, \mathbf{u}_{\Lambda_{j,k-1}}]$;
    Compute smallest set $\Lambda_{j,k}$ s.t. $\|\tilde{\mathbf{r}}_{\Lambda_{j,k-1}}|_{\Lambda_{j,k}}\|_{\ell_2(\mathcal{J})} \geq \frac{1}{2}\|\tilde{\mathbf{r}}_{\Lambda_{j,k-1}}\|_{\ell_2(\mathcal{J})}$;
  od
  $\mathbf{COARSE}[c_3\varepsilon_{j+1}, \mathbf{u}_{\Lambda_{j,k}}] \to (\Lambda_{j+1}, \mathbf{u}_{j+1})$;
  $j := j + 1$;
od

Remark 1.2. (i) We shall not discuss in detail the concrete numerical realization of the three fundamental subroutines. The subroutine **COARSE** consists of a thresholding step, whereas **RHS** essentially requires the computation of a best N-term approximation. The most complicated building block is **APPLY**. The subroutine can be realized and optimality of the resulting algorithm can be proved up to order s^*, if the stiffness matrix $\mathbf{A}$ is s^*-*computable*, i.e., there exists matrices $\mathbf{A}^J$, $J \in \mathbb{N}$, with

$$\|\mathbf{A} - \mathbf{A}^J\|_{\mathcal{L}(\ell_2(\mathcal{J}))} \lesssim M_J^{-s}$$

for all $s < s^*$, where $\mathbf{A}^J$ has $\mathcal{O}(M_J)$ nontrivial entries per column whose combined computation needs at most the same order of operations.

(ii) In **ALGORITHM**1.1, c_1, c_2 and c_3 denote some suitably chosen constants whose concrete values depend on the problem at hand. Also the parameter K has to be chosen in a suitable way. We refer to [3] for details.

It can be shown that **ALGORITHM** 1.1 has the following basic properties:

- **ALGORITHM** 1.1 is guaranteed to converge for a huge class of problems, i.e.,

$$\|\mathbf{u} - \mathbf{u}_\varepsilon\|_{\ell_2(\mathcal{J})} \lesssim \varepsilon.$$

- The order of convergence of **ALGORITHM** 1.1 is *optimal* in the sense that it asymptotically realizes the convergence order

of best N-term wavelet approximation, i.e., if the best N-term approximation satisfies $\mathcal{O}(N^{-s})$, then

$$\|\mathbf{u}-\mathbf{u}_\varepsilon\|_{\ell_2(\mathcal{J})} = \mathcal{O}((\#\mathrm{supp}\mathbf{u}_\varepsilon)^{-s}).$$

- The number of arithmetic operations stays proportional to the number of unknowns, that is, the number of floating point operations needed to compute $\mathbf{u}_\varepsilon$ satisfies $\mathcal{O}(\#\mathrm{supp}\mathbf{u}_\varepsilon)$.

1.3. **Frame case.** The construction of wavelet bases particularly on more involved domains is often very difficult. In this cases the discretiztion of the PDE via a *(Hilbert) frame for* $L_2(\Omega)$ sounds more promising. The latter is a collection of functions

$$\Psi = \{\psi_\lambda : \lambda \in \mathcal{J}\},$$

fulfilling the frame property

(1.5)
$$c||f||_{L_2(\Omega)} \le ||\{\langle f,\psi_\lambda\rangle_{L_2(\Omega)}\}_{\lambda\in\mathcal{J}}||_{\ell_2(\mathcal{J})} \le C||f||_{L_2(\Omega)} \quad \text{for all } f\in L_2(\Omega),$$

with constants $c, C > 0$. In contrast to a basis, a frame allows for redundancies. Quite recently it hast turned out that the construction of a frame of wavelet type in nontrivial domains is much simpler compared to the basis case. On top of that, the redundancy of the frame might be beneficial since it may lead to a sparser representation of functions due to the higher variety of decompositions. It is possible to construct wavelet-type frames, where the primal functions have the same crucial properties as in the basis case like vanishing moments, locality of the support, cancellation properties and characterization of function spaces. Analogue to the basis case one aims at the discrezitation

$$\mathbf{A}\mathbf{u} = \mathbf{f} \tag{1.6}$$

where

$$\mathbf{A} := \mathbf{D}^{-1}\langle \mathcal{A}\Psi,\Psi\rangle^T\mathbf{D}^{-1},\ \mathbf{u} := \mathbf{D}\mathbf{c},\ u = \mathbf{c}^T\Psi,\ \mathbf{f} := \mathbf{D}^{-1}\langle f,\Psi\rangle^T.$$

The main difference to the basis scheme is that due to the redundancy of the frame $\mathbf{\Psi}$, the system matrix $\mathbf{A}$ has a non-trivial kernel, so that (1.6) is not uniquely solvable. Straightforward Galerkin-type approaches might hence run into stability problems.

Nonetheless, classical iterative schemes like the damped Richardson iteration

(1.7)
$$\mathbf{u}^{(j+1)} := \mathbf{u}^{(j)} + \omega(\mathbf{f}-\mathbf{A}\mathbf{u}^{(j)}), \quad 0<\omega<\frac{2}{\|\mathbf{A}\|_{\mathcal{L}(\ell_2(\mathcal{J}))}}, \quad j=0,1,\dots$$

or variations thereof, like steepest descent or conjugate gradient iterations, can still be applied in a numerically stable way, and the associated expansions $u^{(j)} = \mathbf{c}^{(j)T}\Psi$ will converge to the solution u under quite general assumptions. See [4, 6, 7, 13] for further information. By judiciously choosing the respective tolerances, convergence

can even be preserved under perturbation of the exact iterations when, e.g., each evaluation of the infinite-dimensional right-hand side $\mathbf{f}$ and each matrix-vector product $\mathbf{Av}$ are replaced by the numerical approximations **COARSE** and **APPLY**. The resulting algorithm reads as follows:

Algorithm 1.3. R_SOLVE$[\varepsilon, \mathbf{A}, \mathbf{f}] \to \mathbf{u}_\varepsilon$

% Let $\theta < 1/3$ and $K \in \mathbb{N}$ be fixed such that $3\rho^K < \theta$.
% $i := 0,\ \mathbf{v}^{(0)} := 0,\ \varepsilon_0 := \|(\mathbf{A}|_{\mathrm{ran}(\mathbf{A})})^{-1}\|\|\mathbf{f}\|_{\ell_2(\mathcal{J})}$
While $\varepsilon_i > \varepsilon$ do
  $i := i + 1$
  $\varepsilon_i := 3\rho^K \varepsilon_{i-1}/\theta$
  $\mathbf{f}^{(i)} := \mathbf{RHS}[\mathbf{f}, \frac{\theta\varepsilon_i}{6\alpha K}]$
  $\mathbf{v}^{(i,0)} := \mathbf{v}^{(i-1)}$
  For $j = 1, \ldots, K$ do
    $\mathbf{v}^{(i,j)} := \mathbf{v}^{(i,j-1)} - \alpha(\mathbf{APPLY}[\mathbf{A}, \mathbf{v}^{(i,j-1)}, \frac{\theta\varepsilon_i}{6\alpha K}] - \mathbf{f}^{(i)})$
  od
  $\mathbf{v}^{(i)} := \mathbf{COARSE}[\mathbf{v}^{(i,K)}, (1-\theta)\varepsilon_i]$
od
$\mathbf{u}_\varepsilon := \mathbf{v}^{(i)}$

It can be shown that **ALGORITHM** 1.3 has the same basic properties as **ALGORITHM** 1.1.

References

1. C. Canuto, A. Tabacco, and K. Urban, *The wavelet element method. I: Construction and analysis*, Appl. Comput. Harmon. Anal. **6** (1999), 1–52.
2. A. Cohen, *Numerical Analysis of Wavelet Methods*, Studies in Mathematics and its Applications, vol. 32, North-Holland, Amsterdam, 2003.
3. A. Cohen, W. Dahmen, and R. DeVore, *Adaptive wavelet methods for elliptic operator equations – Convergence rates*, Math. Comput. **70** (2001), no. 233, 27–75.
4. ———, *Adaptive wavelet methods II: Beyond the elliptic case*, Found. Comput. Math. **2** (2002), no. 3, 203–245.
5. A. Cohen, I. Daubechies, and J.-C. Feauveau, *Biorthogonal bases of compactly supported wavelets*, Commun. Pure Appl. Math. **45** (1992), 485–560.
6. S. Dahlke, M. Fornasier, and T. Raasch, *Adaptive frame methods for elliptic operator equations*, Adv. Comput. Math. **27** (2007), no. 1, 27–63.

7. S. Dahlke, M. Fornasier, T. Raasch, R. Stevenson, and M. Werner, *Adaptive frame methods for elliptic operator equations: The steepest descent approach*, IMA J. Numer. Anal. **27** (2007), no. 4, 717–740.
8. W. Dahmen and R. Schneider, *Wavelets with complementary boundary conditions – function spaces on the cube*, Result. Math. **34** (1998), 255–293.
9. _____, *Composite wavelet bases for operator equations*, Math. Comput. **68** (1999), 1533–1567.
10. _____, *Wavelets on manifolds I: Construction and domain decomposition*, SIAM J. Math. Anal. **31** (1999), 184–230.
11. Y. Meyer, *Wavelets and Operators*, Cambridge Studies in Advanced Mathematics, vol. 37, Cambridge University Press, 1992.
12. M. Primbs, *Stabile biorthogonale Spline-Waveletbasen auf dem Intervall*, Dissertation, Universität Duisburg-Essen, 2006.
13. R. Stevenson, *Adaptive solution of operator equations using wavelet frames*, SIAM J. Numer. Anal. **41** (2003), no. 3, 1074–1100.