TENSOR PRODUCT TYPE SUBSPACE SPLITTINGS AND MULTILEVEL ITERATIVE METHODS FOR ANISOTROPIC PROBLEMS

M. GRIEBEL
INSTITUT FÜR INFORMATIK
TU MÜNCHEN
D-80290 MÜNCHEN

P. OSWALD
INSTITUT FÜR ANGEWANDTE MATHEMATIK
FRIEDRICH-SCHILLER-UNIVERSITÄT JENA
D-07740 JENA

Abstract. We describe tensor product type techniques to derive robust solvers for anisotropic elliptic model problems on rectangular domains in $\mathbf{R}^d$. Our analysis is based on the theory of additive subspace correction methods and applies to finite-element- and prewavelet-schemes. We present multilevel- and prewavelet-based methods that are robust for anisotropic diffusion operators with additional Helmholtz term. Furthermore the resulting convergence rates are independent of the discretization level. Beside their theoretical foundation, we also report on the results of various numerical experiments to compare the different methods.

Contents :

Key words: Anisotropic elliptic equations, finite elements, wavelets, preconditioners, multilevel methods.

Subject classification: AMS(MOS) 65F10, 65F35, 65N20, 65N30.

1. Introduction. For the solution of elliptic PDE's, multilevel methods have gained a lot of popularity in the last decade. This is mainly due to their nearly optimal complexity for a number of model problems. In many practical cases, they are based on some kind of multiresolution scale of subspaces

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

of a computational Hilbert space V which is used to produce stable subspace splittings as well as to design fast iterative solvers related to such splittings for variational problems in V. For symmetric elliptic boundary value problems in Sobolev spaces, considerable progress has been achieved in the theoretical understanding of multilevel- and multigrid-methods as well as of other subspace correction methods for finite element discretizations (cf. [Xu, Ys, DSW, GO2, O1]). The underlying theory also applies to various wavelet discretizations, see, e.g., [DmK, Ja, JaL, DmPS] for some papers dealing with wavelet solvers for elliptic problems that are related to our approach.

In this paper we deal with anisotropic elliptic problems. We consider the simple model problem

$$-\sum_{k=1}^{d} c_k u_{x_k x_k} + c_0 u = f\ , \quad c_k > 0\ ,\ c_0 \geq 0\ , k =, \ldots d \tag{1}$$

in its variational form in $H_0^1(\Omega)$ (homogeneous Dirichlet boundary conditions) or $H^1(\Omega)$ (Neumann problem) where Ω is a d-dimensional rectangle. For a given finite-element- (or wavelet-like) discretization, our aim is to construct a fast and robust solver. Robustness means in particular that the convergence rates stay independent of the discretization level and of the coefficients c_k.

We mainly focus on two different approaches that both can be interpreted as subspace splitting methods: First, we consider techniques that are based on the tensor product multilevel nodal basis which is related to the so-called generating system introduced in [G3, G4]. Our methods are variants of the Bramble-Pasciak-Xu (BPX) preconditioner [BPX] and its modifications for the isotropic case, see [Zh, Wa, TCK], which are related to the semi-coarsening of multigrid methods [StT]. Furthermore, we introduce new techniques that stem from a tensor product type prewavelet discretization.

While the prewavelet methods can be applied straightforwardly to our model problem yielding robust solvers with level-independent convergence rates, the situation for the multilevel nodal basis approach is slightly more elaborate. There, robustness is achieved by incorporating further information into the algorithm through a specific choice of the underlying subspace splitting which can be determined easily from the coefficients, and does not increase the complexity of the algorithm. In this respect, our approach is different from Hackbusch's Frequency Decomposition Multigrid Method [Ha2, Ha3] and its recent wavelet variation [RWZ, RZ] where one subspace splitting serves all coefficient choices, at the cost of additional but unnecessary work (cf. [Ha2]).

The theoretical justification of our solvers is given for the additive methods only (see Sections 2 and 3) though the multiplicative variants of the algorithms based on the same subspace splittings work reasonable well, too (c.f. [Xu, Ys, GO2]).

The remainder of this paper is outlined as follows. In the second section, we introduce subspace splittings for higher dimensional problems that are constructed from subspace splittings for one-dimensional problems by means of the tensor product principle and discuss their properties. In the third section, we present the multilevel nodal basis methods and the prewavelet solvers and show that they are robust with respect to varying values c_k, i.e. for anisotropic diffusion as well as for a varying Helmholtz part. There, we consider discretizations on conventional full grids as well as on the recently introduced sparse grids. In the fourth section, we discuss some implementational issues of the different solution methods and present the results of various numerical experiments that support our theory. At the end of this paper, we give some concluding remarks.

2. Subspace splittings via tensor products. Let us briefly recall the definition of the tensor product of two real (separable) Hilbert spaces $H \equiv \{H; a\}$ and $\tilde{H} \equiv \{\tilde{H}; \tilde{a}\}$, see e.g. [We], pp. 51-53 (the notation $\{H; a\}$ indicates that the bilinear form $a(\cdot,\cdot)$ is the scalar product on H). Let

$$F(H,\tilde{H}) = \{\sum_{j=1}^{n} \alpha_j (f_j, \tilde{f}_j) \,|\, \alpha_j \in \mathbf{R}\,,\, (f_j,\tilde{f}_j) \in H \times \tilde{H}\,,\, j = 1,\ldots,n\,,\, n \in \mathbf{N}\}$$

and denote by N the linear subspace of $F(H,\tilde{H})$ spanned by all elements of the form

$$\sum_{j=1}^{n}\sum_{i=1}^{m} \alpha_j \tilde{\alpha}_i (f_j,\tilde{f}_i) - (\sum_{j=1}^{n} \alpha_j f_j, \sum_{i=1}^{m} \tilde{\alpha}_i \tilde{f}_i)\,.$$

It is easy to check that the scalar product

$$a \otimes \tilde{a}(\sum_{j=1}^{n} \alpha_j (f_j,\tilde{f}_j), \sum_{i=1}^{m} \tilde{\beta}_i (g_i,\tilde{g}_i)) = \sum_{j=1}^{n}\sum_{i=1}^{m} \alpha_j \tilde{\beta}_i a(f_j,g_i) \tilde{a}(\tilde{f}_j,\tilde{g}_i) \tag{2}$$

defined on $F(H,\tilde{H})$ turns the factor space $F(H,\tilde{H})/N$ into a pre-Hilbert space. Its completion is called tensor product of H and $\tilde{H}$ and denoted by $H \otimes \tilde{H}$. The notation $f \otimes \tilde{f}$ is preserved for the equivalence class in $F(H,\tilde{H})/N$ containing $(f,\tilde{f})$. Given bounded linear operators T and $\tilde{T}$ acting in H and $\tilde{H}$, respectively, we can introduce their tensor product $T \otimes \tilde{T}$ acting in $H \otimes \tilde{H}$ via the formula

$$T \otimes \tilde{T}(\sum_{j=1}^{n} \alpha_j f_j \otimes \tilde{f}_j) = \sum_{j=1}^{n} \alpha_j T f_j \otimes \tilde{T} \tilde{f}_j$$

and continuity arguments, see [We], p. 262.

If one takes complete orthonormal systems $\{e_j\}$ and $\{\tilde{e}_i\}$ in H and $\tilde{H}$, respectively, then $\{e_j \otimes \tilde{e}_i\}$ is a complete orthonormal system in $H \otimes \tilde{H}$, and the following equivalent definition via Fourier coefficients can be given (cf. [We], Theorem 3.12):

$$H \otimes \tilde{H} = \{\sum_{j,i} \gamma_{ji} e_j \otimes \tilde{e}_i \mid \sum_{j,i} \gamma_{ji}^2 < \infty\} \tag{3}$$

and

$$a \otimes \tilde{a}(\sum_{j,i} \gamma_{ji} e_j \otimes \tilde{e}_i, \sum_{k,l} \gamma'_{kl} e_k \otimes \tilde{e}_l) = \sum_{j,i} \gamma_{ji} \gamma'_{ji} \,. \tag{4}$$

Moreover, if H and $\tilde{H}$ are function spaces over domains Ω, $\tilde{\Omega}$, respectively, it is quite natural to identify $H \otimes \tilde{H}$ as function spaces over the product domain $\Omega \times \tilde{\Omega}$ via the mapping

$$f \otimes \tilde{f} \longmapsto f(x) \cdot \tilde{f}(\tilde{x}) \,. \tag{5}$$

Finally, for reasons of convenience, we introduce one more operation in the set of Hilbert spaces: If the intersection $\hat{H} = H \cap \tilde{H}$ is a non-trivial linear space, and $c, \tilde{c}$ are any two positive real numbers, then $\hat{H}$, equipped with the scalar product

$$\hat{a}(u,v) = ca(u,v) + \tilde{c}\tilde{a}(u,v) \,,$$

is also a Hilbert space. To indicate the dependence on the coefficients $c, \tilde{c}$ we sometimes will write $cH \cap \tilde{c}\tilde{H}$ instead of $\hat{H}$.

These preliminaries are motivated by the fact that the above operations (tensor products and intersection of Hilbert spaces) can be used to create some typical function spaces on product domains (like d-dimensional rectangles in $\mathbf{R}^d$) from their univariate counterparts. Note that the corresponding definitions can be extended from two to $d > 2$ Hilbert spaces without problems. In the following, we give a few examples of such spaces.

Example 1. If $\Omega = [a,b] \times [c,d]$ is a rectangle in $\mathbf{R}^2$ then, for non-negative integers r, s, we have

$$S^{(r,s)}H(\Omega) = H^r(a,b) \otimes H^s(c,d) \tag{6}$$

where $S^{(r,s)}H(\Omega)$ denotes the subspace of all $f \in L_2(\Omega)$ possessing generalized derivatives of order $(r,0), (0,s)$, and (r,s) in $L_2(\Omega)$. The norm in this Hilbert space is defined by

$$\|f\|^2_{S^{(r,s)}H} = \|f\|^2_{L_2} + \|\frac{\partial^r f}{\partial x_1^r}\|^2_{L_2} + \|\frac{\partial^s f}{\partial x_2^s}\|^2_{L_2} + \|\frac{\partial^{r+s} f}{\partial x_1^r \partial x_2^s}\|^2_{L_2}$$

while the norms for the univariate Sobolev spaces $H^r(a,b)$ are introduced by

$$\|f\|^2_{H^r(a,b)} = \|f\|^2_{L_2(a,b)} + \|f^{(r)}\|^2_{L_2(a,b)} \,.$$

These definitions can be extended to noninteger r, s (if proper fractional derivatives are defined or interpolation arguments are used) and to higher dimensions. Such spaces are called Sobolev spaces with dominating mixed derivative (in the above case, this is the derivative $\frac{\partial^{r+s} f}{\partial x_1^r \partial x_2^s}$ of order (r,s)).

Example 2. Another example is the standard energy space of the Poisson equation with homogeneous Dirichlet conditions on the boundary of the rectangle Ω. We have

$$H_0^1(\Omega) = H_0^1(a,b) \otimes L_2(c,d) \cap L_2(a,b) \otimes H_0^1(c,d) \;, \tag{7}$$

with exact coincidence of the scalar products if the equivalent norms

$$\|f\|_{H_0^1(\Omega)} = \|\nabla f\|_{L_2(\Omega)} \;, \quad \|f\|_{H_0^1(a,b)} = \|f'\|_{L_2(a,b)}$$

are used on $H_0^1(\Omega)$ and $H_0^1(a,b)$, respectively.

Example 3. As the central example used in this paper, we consider the energy space associated with the anisotropic Dirichlet problem (1). For $V = H_0^1(\Omega)$, equipped with the norm

$$\|u\|_V^2 = \int_\Omega (\sum_{k=1}^d c_k |\frac{\partial u}{\partial x_k}|^2 + c_0 |u|^2) \, dx \;,$$

on a d-dimensional rectangle $\Omega = I_1 \times \ldots \times I_d$ ($I_k = [a_k, b_k]$), we have the representation

$$V = c_1 S^{(1,0,\ldots,0)} H_0(\Omega) \cap \ldots \cap c_d S^{(0,\ldots,0,1)} H_0(\Omega) \cap c_0 L_2(\Omega) \tag{8}$$

where $S^{(0,\ldots,1,\ldots,0)} H_0(\Omega) = L_2(I_1) \otimes \ldots \otimes H_0^1(I_k) \otimes \ldots \otimes L_2(I_d)$. Analogous statements hold for Neumann boundary conditions as well as for certain cases of mixed boundary conditions.

So far, we could not find an explicit reference for these simple facts. To convince the reader, we give some hints on how to prove the above statements. For Example 2 and 3, the consideration can easily be reduced to the classical periodic case. We may assume that $d = 2$ and $I_1 = I_2 = [0, \pi]$. Using first the odd extension

$$\tilde{u}(\epsilon_1 x_1, \epsilon_2 x_2) = \epsilon_1 \epsilon_2 f(x_1, x_2) \;, \quad (x_1, x_2) \in \Omega \;, \; \epsilon_1, \epsilon_2 = \pm 1$$

onto $[-\pi, \pi]^2$ and then the 2π-periodic continuation, we see that $u \in H_0^1(\Omega)$ if and only if $\tilde{u} \in \tilde{H}^1$ possesses a pure Fourier sine series, i.e. if

$$u(x_1, x_2) = \sum_{n=1}^\infty \sum_{m=1}^\infty a_{nm} \sin n x_1 \, \sin m x_2 \;, \qquad \sum_{m,n=1}^\infty (n^2 + m^2) a_{nm}^2 < \infty$$

on Ω. Now, since the function system $\{\sin nx\}$ defines a complete orthogonal basis in both $L_2(0,\pi)$ and $H_0^1(0,\pi)$, the assertions of examples 2 and 3 are straightforward by the definitions of tensor products (via orthonormal systems) and intersections. Note that in the case of Neumann boundary conditions (when H_0^1 spaces have to be replaced by H^1 spaces), the even extension and the use of $\{\cos nx, n = 0, 1, \ldots\}$ are the appropriate substitutes. For higher order derivatives (e.g., as in Example 1), in order to avoid the problems with periodic extension, one may use other orthogonal systems like suitable orthogonal spline systems, or the extension to $\mathbf{R}$ and Fourier transform arguments.

Since we will not rely on Example 1, we leave details to the interested reader, and refer to [Sm, SmT, Tl, BKLN] for more information on spaces with dominating mixed derivatives.

Note that, for many other elliptic problems with constant coefficients over rectangles, representations analogous to our above specific examples are possible: the corresponding energy spaces can be decomposed via intersection and tensor products into simpler function spaces over one-dimensional intervals.

In the remainder of this section, we will use this observation to obtain stable splittings for such types of energy spaces in a simple way. For the construction of fast elliptic solvers (subspace correction or Schwarz methods), the importance of the stability of splittings is well-recognized by now, see [Xu, Ys, GO2, O1]. For some practical examples that illustrate this application, see also the next section.

Let $\{H; a\}$ be given. In practice, the space H originally carries another basic scalar product, and $a(\cdot,\cdot)$ comes from a symmetric H-elliptic variational problem

$$\textit{Find } u \in H \textit{ such that} \qquad a(u,v) = \Phi(v) \quad \forall\, v \in H \tag{9}$$

to be solved. Now, consider a collection of closed subspaces $H_j \subset H$ such that (topologically) $H = \sum_j H_j$. Furthermore, suppose that $\{H_j; b_j\}$ remain Hilbert spaces when equipped with some scalar products $b_j(\cdot,\cdot)$. Then, the additive subspace splitting

$$\{H; a\} = \sum_j \{H_j; b_j\} \tag{10}$$

is called stable if the norm equivalence

$$a(u,u) \asymp |||u|||^2 \equiv \inf_{u_l \in V_l : u = \sum_l u_l} \Big(\sum_l b_l(u_l, u_l) \Big) \tag{11}$$

holds true, i.e. if the characteristic numbers

$$\kappa = \frac{\lambda_{\max}}{\lambda_{\min}}\ , \quad \lambda_{\min} = \min_{0 \neq u \in V} \frac{a(u,u)}{|||u|||^2}\ , \quad \lambda_{\max} = \max_{0 \neq u \in V} \frac{a(u,u)}{|||u|||^2} \tag{12}$$

are finite and positive. The number κ is called condition number of the splitting (10). It is well-known (see [Xu, GO2]) that the so-called additive Schwarz operator

$$P = \sum_j T_j \ :\ H \rightarrow H\ , \tag{13}$$

with $T_j : H \rightarrow H_j$ given by the variational problems

$$b_j(T_j u, v_j) = a(u, v_j) \quad \forall\, v_j \in H_j\ , \quad \forall\, j\ , \tag{14}$$

associated with the splitting (10) is symmetric positive definite in $\{H; a\}$, and satisfies

$$\lambda_{\max}(P) = \sup_{u \in H\, :\, u,u)=1} a(Pu,u) = \lambda_{\max}\ , \quad \lambda_{\min}(P) = \inf_{u \in H\, :\, u,u)=1} a(Pu,u) = \lambda_{\min}\ . \tag{15}$$

This means that the spectral condition number $\kappa(P)$ of P coincides with the condition number κ of the splitting. Since the operator equation

$$Pu = \phi = \sum_j \phi_j \tag{16}$$

(with ϕ_j defined via variational problems analogous to (14) where the right-hand side $a(u, v)$ is replaced by the linear functional $\Phi(v)$) is equivalent to (9), these constructions lead to fast iterative methods if κ is small, and the subproblems (14) (i.e. the evaluation of $T_j u$) are easy to solve. E.g., one may use Richardson iterations or cg-iterations on (16) in which case the decisive role of κ is well-known, see [Ha1, GL, AB] for the basic results on iterative methods. For a Gauss-Seidel-like iteration associated with (10), the so-called multiplicative Schwarz iteration, we refer to [Xu, Ys, GO2] for estimates of the convergence rates in terms of the characteristic properties of the subspace splitting.

Thus, we are interested in obtaining stable splittings with small values of κ. Moreover, κ should not depend on the number of subspaces and on the values of c_k. Furthermore, the T_j should be 'simple', i.e. its evaluation should be computable in an economic way (the latter may be achieved, for instance, if the dimension of the H_j is kept small, at the cost of a large number of subproblems). Together with other useful techniques like refinement, clustering, and selection described in [GO2, O1], the following technical results will be used for this purpose in the next section.

Proposition 1. If the splittings

$$\{H; a\} = \sum_j \{H_j; b_j\}$$

and

$$\{\tilde{H}; \tilde{a}\} = \sum_i \{\tilde{H}_i; \tilde{b}_i\}$$

are stable and possess the condition numbers κ and $\tilde{\kappa}$, respectively, then the tensor product splitting

$$\{H \otimes \tilde{H}; a \otimes \tilde{a}\} = \sum_j \sum_i \{H_j \otimes \tilde{H}_i; b_j \otimes \tilde{b}_i\} \tag{17}$$

is also stable and possesses the condition number $\kappa\tilde{\kappa}$.

Proof. This follows from certain simple properties of tensor products of operators (cf. [We], Theorem 8.32 (a)) and the observation that the additive Schwarz operator

$$\hat{P} = \sum_{j,i} \hat{T}_{ji} \quad (\hat{T}_{ji} \,:\, H \otimes \tilde{H} \to \tilde{H}_j \otimes \tilde{H}_i)$$

associated with the splitting (17) coincides with $P \otimes \tilde{P}$. Indeed, we have $T_{ji} = T_j \otimes \tilde{T}_i$, since for all $\hat{u} = \sum_s \alpha_s u_s \otimes \tilde{u}_s \in H \otimes \tilde{H}$ and all $\hat{v}_{ji} = \sum_r \beta_r v_{jr} \otimes \tilde{v}_{ir} \in H_j \otimes \tilde{H}_i$

$$\begin{aligned} b_j \otimes \tilde{b}_i(T_j \otimes \tilde{T}_i \hat{u}, \hat{v}_{ji}) &= \sum_{s,r} \alpha_s \beta_r b_j(T_j u_s, v_{jr}) \tilde{b}_i(\tilde{T}_i \tilde{u}_s, \tilde{v}_{ir}) \\ &= \sum_{s,r} \alpha_s \beta_r a(u_s, v_{jr}) \tilde{a}(\tilde{u}_s, \tilde{v}_{ir}) = a \otimes \tilde{a}(\hat{u}, \hat{v}_{ji}) \,. \end{aligned}$$

Summing up and using continuity arguments one has the above-mentioned representation for $\hat{P}$. Now it remains to use

$$\|P \otimes \tilde{P}\|_{H\otimes\tilde{H}} = \|P\|_H \|\tilde{P}\|_{\tilde{H}} = \lambda_{\max}(P)\lambda_{\max}(\tilde{P}) ,$$

(which is valid for all s.p.d. P and $\tilde{P}$), the formula $\hat{P}^{-1} = P^{-1} \otimes \tilde{P}^{-1}$, and the coincidence of the condition numbers of the splittings and the corresponding additive Schwarz operators (cf. (15)). □

Proposition 2. Suppose that the spaces H and $\tilde{H}$ possess similar splittings into a direct sum of the same sequence of subspaces H_j, in the sense that there are a two sequences $\{\alpha_j\}$ and $\{\tilde{\alpha}_j\}$ of positive numbers such that

$$\{H; a\} = \sum_j \{H_j; \alpha_j b(\cdot,\cdot)\}$$

and

$$\{\tilde{H}; \tilde{a}\} = \sum_j \{H_j; \tilde{\alpha}_j b(\cdot,\cdot)\}$$

are both stable (the bilinear form $b(\cdot,\cdot)$ is the same for all j). Then, for all $\alpha, \tilde{\alpha} > 0$, the splitting

$$\alpha H \cap \tilde{\alpha}\tilde{H} = \sum_j \{H_j; (\alpha\alpha_j + \tilde{\alpha}\tilde{\alpha}_j) b(\cdot,\cdot)\} \tag{18}$$

is stable, with condition number $\kappa \le \Lambda_{\max}/\Lambda_{\min}$ where $\Lambda_{\max} = \max\{\lambda_{\max}, \tilde{\lambda}_{\max}\}$ and $\Lambda_{\min} = \min\{\lambda_{\min}, \tilde{\lambda}_{\min}\}$. The result extends to the intersection of $d > 2$ similar Hilbert space splittings.

Proof. Since the splittings are into direct sums of subspaces, there is a unique representation for each $u \in H \cap \tilde{H}$ with respect to $\{H_j\}$ such that

$$\lambda_{\max}^{-1} a(u,u) \le \sum_j \alpha_j b(u_j,u_j) \le \lambda_{\min}^{-1} a(u,u)$$

and

$$\tilde{\lambda}_{\max}^{-1} \tilde{a}(u,u) \le \sum_j \tilde{\alpha}_j b(u_j,u_j) \le \tilde{\lambda}_{\min}^{-1} \tilde{a}(u,u) ,$$

simultaneously. Now, use the definition of $\alpha H \cap \tilde{\alpha}\tilde{H}$ in (18) and add these two-sided inequalities. □

Proposition 2 is often useful when applied together with the following result. In many applications, the Hilbert spaces $\{H; a\}, \{\tilde{H}; \tilde{a}\}, \ldots$ under consideration are subspaces of a basic Hilbert space $\{\mathcal{H}; b(\cdot,\cdot)\}$ which is equipped with some type of multiresolution analysis, i.e. a suitable increasing sequence of subspaces

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

the union of which is dense in $\mathcal{H}$. Together with $\{V_j\}$, we consider the sequence $\{W_j\}$ where $W_0 = V_0$ and W_j is the orthogonal complement of V_{j-1} in V_j (with respect to the scalar product $b(\cdot,\cdot)$ in $\mathcal{H}$). Thus, $\{W_j\}$ forms an orthogonal decomposition of $\mathcal{H}$. □

Proposition 3. Let $\mathcal{H}, \{V_j\}$, and $\{W_j\}$ be given as specified above, and let $\{H;a\}$ be another Hilbert space. Then, for any sequence $\{\alpha_j\}$ of positive numbers satisfying

$$\sum_j \alpha_j^{-1} < \infty \ ,$$

the $|||\cdot|||$-norms of the splitting

$$\{H;a\} = \sum_j \{V_j; \alpha_j b(\cdot,\cdot)\}$$

and the splitting

$$\{H;a\} = \sum_j \{W_j; \beta_j b(\cdot,\cdot)\} \ , \quad \beta_j = (\sum_{l\geq j} \alpha_l^{-1})^{-1}$$

coincide. Thus, the stability of one of these splittings implies the stability of the other, and the condition numbers are the same.

Proof. We will explicitly calculate the square of the $|||\cdot|||$-norm, i.e.

$$|||u|||^2 = \inf_{u_j\in V_j\,:\, u=\sum_j u_j} \sum_j \alpha_j b(u_j,u_j)$$

for the first splitting by defining critical points of the Lagrangian

$$\mathcal{L}(\{u_j\},v) = \sum_j \alpha_j b(u_j,u_j) + 2b(v, u - \sum_j u_j)$$

which is well-defined on the cartesian product of the Hilbert spaces $\{V_j;\alpha_j b(\cdot,\cdot)\}$ (for all j), and $\{\mathcal{H};b\}$. Indeed, critical points $(\{u_j^*\},v^*)$ of $\mathcal{L}$ are characterized by the conditions $u = \sum_j u_j^*$ (in $\mathcal{H}$), and

$$\alpha_j b(u_j^*,v_j) = b(v^*,v_j) \qquad \forall\ v_j \in V_j$$

for each j. Thus, $u_j^* = \alpha_j^{-1} Q_j v^*$ where $Q_j : \mathcal{H} \to V_j$ are the operators of orthogonal projection onto V_j. Denoting by

$$v^* = \sum_j w_j^* \ , \quad u = \sum_j w_j \qquad (w_j^*, w_j \in W_j)$$

the unique orthogonal decompositions of v^* and u, respectively, we obtain $u_j^* = \alpha_j^{-1}\sum_{l\leq j} w_l^*$, and

$$u = \sum_j w_j = \sum_j u_j^* = \sum_j \alpha_j^{-1} \sum_{l\leq j} w_l^* = \sum_j \beta_j^{-1} w_j^* \ .$$

By the uniqueness of the orthogonal decomposition of u, this yields $w_j^* = \beta_j w_j$, $u_j^* = \alpha_j^{-1} \sum_{l \leq j} \beta_l w_l$, and finally

$$|||u|||^2 = \sum_j \alpha_j b(u_j^*, u_j^*) = \sum_j \alpha_j^{-1} \sum_{l \leq j} \beta_l^2 b(w_l, w_l) = \sum_l \beta_l b(w_l, w_l) .$$

But this latter expression is the square of the $|||\cdot|||$-norm of the second splitting. This proves the assertion. □

Note that $\beta_j \leq \alpha_j$ is valid automatically. However, the opposite inequality, i.e. the relation $\alpha_j = O(\beta_j)$ holds if and only if $\sum_{l \geq j} \alpha_l^{-1} = O(\alpha_j^{-1})$ $(j \to \infty)$. In particular, this is true for sequences $\{\alpha_j\}$ that grow at least geometrically, a case which usually appears in practical applications.

Now, the propositions 1-3 can be applied to derive stable subspace splittings for the examples 1-3 mentioned above provided that stable splittings for the Sobolev spaces $H^s(a, b)$ and $H_0^s(a, b)$, respectively, on one-dimensional intervals $[a, b]$ are available. For Sobolev spaces on intervals, the underlying space $\{\mathcal{H}, b\}$ can be identified with $L_2(a, b)$, and the techniques of one-dimensional Fourier analysis and spline approximation theory can be exploited. For Example 3, details are given in the next section.

3. Robust nodal basis and prewavelet schemes. Throughout this section, we concentrate on splittings for the energy space corresponding to the anisotropic model problem (1), see Example 3. Without loss of generality, let $\Omega = [0, 1]^d$ be the unit square, the scaling necessary for an arbitrary rectangle can be achieved by affine-linear mapping and leads to just another set of coefficients $\{c_k\}$ characterizing the anisotropy. Moreover, we can assume that either $c_0 = 1$ or $c_0 = 0$ depending on whether a Helmholtz term is present or not. We mainly deal with homogeneous Dirichlet boundary conditions (i.e. we start with splittings for $H_0^1(0, 1)$ and look for splittings of $H_0^1(\Omega)$). In subsection 3.1, the case of linear finite element discretizations over regularly generated product partitions is discussed whereas both nodal-basis and spline-prewavelet-multilevel-schemes are presented. The employed techniques and results can be generalized to higher degree Lagrange C^0-elements without difficulty. In subsection 3.2, we apply the same techniques to the so-called sparse grid discretizations introduced by Zenger [Ze] and generalize our previous results on preconditioners for the isotropic problem [GO1] to the anisotropic case. In subsection 3.3, the case of Neumann boundary conditions is briefly dealt with.

3.1. Finite element discretizations - the full grid case. We give the details for $d = 2$ and $c_0 = 1$, the other cases are quite similar. Without loss of generality, we may assume that the two quasi-uniform partitions π and $\tilde{\pi}$ of $[0, 1]$, of step-size h and $\tilde{h}$, respectively, belong to increasing sequences $\{\pi_j\}$ and $\{\tilde{\pi}_j\}$, with $h_j, \tilde{h}_j \approx a^{-j}$, $j \geq 0$, for some fixed $a > 1$. More precisely, let $\pi = \pi_J$ and $\tilde{\pi} = \tilde{\pi}_I$. Note that, for the case of standard refinement by uniform bisection, we have $a = 2$.

What we are aiming at is a fast solution method (in form of a good preconditioner) for the discretization of (1) with respect to $\hat{V} = V_{J,I;0}$ $(d = 2)$, the space of bilinear finite elements over $\pi_{J,I} \equiv \pi_J \times \tilde{\pi}_I = \pi \times \tilde{\pi}$ satisfying zero boundary conditions on the boundary of the unit square. To this end, let $V_{j;0} = S_1^0(\pi_j) \cap H_0^1(0,1)$ and $\tilde{V}_{j;0} = S_1^0(\tilde{\pi}_j) \cap H_0^1(0,1)$, respectively, be the corresponding spaces of linear C^0 finite elements over π_j and $\tilde{\pi}_j$, and set $V_{j,i;0} = V_{j;0} \otimes \tilde{V}_{i;0}$. Since both spaces are finite-dimensional, there is no need for specific scalar products at this stage yet. Moreover, set $W_{j;0} = V_{j;0} \ominus_{L_2(0,1)} V_{j-1;0}$ for $j > 0$ and $W_{0;0} = V_{0;0}$. The subspaces $\tilde{W}_{i;0}$ are introduced analogously. Finally, let $W_{j,i;0} = W_{j;0} \otimes \tilde{W}_{i;0}$.

If we denote $\hat{H} \equiv H_0^1(\Omega)$ and

$$\hat{a}(u,v) = \int_\Omega (\nabla u)^T C \nabla v + uv\, dx$$

where C is the 2×2 diagonal matrix with c_1, c_2 on the diagonal, and $\nabla u = (\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2})^T$ is the gradient of u, then $\{\hat{H}; \hat{a}\}$ is the energy space for (1), with $d = 2$ and $c_0 = 1$.

Theorem 1. The energy space $\{\hat{H}; \hat{a}\}$ of (1) admits the following stable splittings, with a bound of the condition numbers that is uniform with respect to the coefficients $c_k > 0$:

$$\{\hat{H}; \hat{a}\} = \sum_{j,i \geq 0} \{W_{j,i;0}; (c_1 a^{2j} + c_2 a^{2i} + 1)(\cdot,\cdot)_{L_2}\} \tag{19}$$

$$\{\hat{H}; \hat{a}\} = \sum_{l \geq 0} \{V_{j_0+l, i_0+l;0}; a^{2l}(\cdot,\cdot)_{L_2}\} \tag{20}$$

$$\{\hat{H}; \hat{a}\} = \sum_{l \geq 0} \{\hat{W}_l; a^{2l}(\cdot,\cdot)_{L_2}\} \tag{21}$$

where $\hat{W}_{l;0} = V_{l+j_0, l+i_0;0} \ominus V_{l-1+j_0, l-1;i_0;0}$ for $l > 0$ and $W_{0;0} = V_{j_0,i_0;0}$. Here, the integers j_0 and i_0 are defined as follows:

$$(j_0, i_0) = \begin{cases} (0, [\log_{a^2}(c_1/c_2)]) & \text{for} \quad c_1 = \max(c_1,c_2) \geq 1\ , \\ ([\log_{a^2}(c_2/c_1)], 0) & \text{for} \quad c_2 = \max(c_1,c_2) \geq 1\ , \\ ([\log_{a^2} c_1^{-1}], [\log_{a^2} c_2^{-1}]) & \phantom{\text{for}} \quad \max(c_1,c_2) < 1\ . \end{cases} \tag{22}$$

The result extends to $d > 2$, to $c_0 = 0$, and to higher degree Lagrange elements, with obvious modifications.

Proof. First of all, in the one-dimensional case, we have the stability of the splitting

$$\{H_0^1(0,1); (\cdot,\cdot)_{H_0^1}\} = \sum_{j \geq 0} \{V_{j;0}; a^{2j}(\cdot,\cdot)_{L_2}\} \tag{23}$$

with $(u,v)_{H_0^1} = (u',v')_{L_2}$ which may be replaced (on $H_0^1(0,1)$) by $(u,v)_{H_1} = (u,v)_{H_0^1} + (u,v)_{L_2}$ as well. The stability of (23) follows from Besov space characterizations via best spline approximations, and is implicitly contained in many papers. For an exposition of spline approximation theory and the connection to Besov spaces on intervals, see [LD]. For the corresponding result on $H^1(0,1)$, see [O2] and the references cited therein. To incorporate zero boundary conditions, the same procedure as described in [O3] for the 2-dimensional case can be applied (alternatively, odd extension can be used to reduce the $H_0^1(0,1)$-case to a periodic setting and arguments from periodic spline approximation apply). The condition number of (23) as well as the corresponding $\lambda_{\max}$ and $\lambda_{\min}$ only depend on the constants characterizing the quasi-uniformity of the underlying partitions $\{\pi_j\}$.

Now, from the stability of (23) the stability of

$$\{H_0^1(0,1);(\cdot,\cdot)_{H_0^1}\} = \sum_{j\geq 0}\{W_{j;0};a^{2j}(\cdot,\cdot)_{L_2}\} \tag{24}$$

follows by Proposition 3. Analogous results hold with respect to $\{\tilde{V}_{i;0}\}$ and $\{\tilde{W}_{i;0}\}$. Finally,

$$\{L_2(0,1);(\cdot,\cdot)_{L_2}\} = \sum_{j\geq 0}\{W_{j;0};(\cdot,\cdot)_{L_2}\}$$

is stable as an orthogonal decomposition (with $\lambda_{\max} = \lambda_{\min} = \kappa = 1$). Thus, by the description of $\{\hat{H},\hat{a}\}$ given in Example 3 of Section 2, and Proposition 2 and 3 it follows that the splitting (19) of Theorem 1 is stable. Then, the stability of the splittings (20) and (21) follows directly by the clustering argument mentioned in [GO2, O1]. Once again, Proposition 3 has to be used to switch between the third and second splitting. The applicability of the clustering argument is based on the fact that, for some constant $\hat{c} > 0$ depending on c_1, c_2 and for the splittings (19), (20) and (21), we have the stability of

$$\{\hat{W}_{l;0};\hat{c}a^{2l}(\cdot,\cdot)_{L_2}\} = \sum_{W_{j,i;0}\subset\hat{W}_{l;0}}\{W_{j,i;0};(c_1a^{2j}+c_2a^{2i}+1)(\cdot,\cdot)_{L_2}\}$$

(compare with the definition of (j_0,i_0) in (22) which yields $\hat{c}a^{2l}\approx c_1a^{2j}+c_2a^{2i}+1$ for the corresponding l and (j,i), and use the L_2-orthogonality of the $W_{j,i;0}$). This proves the Theorem 1. $\square$

Now, to derive practical algorithms, we further split the (still relatively large) subspaces $V_{j,i;0}$, $W_{j,i;0}$ or $\hat{W}_{l;0}$ into smaller subspaces and thus restrict the whole construction to a suitably selected computational subspace. This leads to slightly different implementations of the multilevel principle. As a rule, locally supported basis functions forming Riesz bases in the subspaces and possessing uniform bounds for the Riesz constants are used to obtain splittings into (small or even) one-dimensional subspaces. For $V_{j,i;0}$ the usual nodal basis functions are the most simple candidates with these properties (this leads to nodal basis multilevel schemes or multilevel schemes based on a frame) while for $W_{j,i;0}$ or $\hat{W}_{l;0}$, respectively, various tensor product prewavelets seem to be appropriate for this task (prewavelet schemes).

The choice of the computational subspace may be dictated by different goals, it depends on accuracy or storage requirements, the wish to incorporate adaptivity (nested refinement or compression strategies), or even the 'taste' of the programmer or end-user.

In this subsection, we restrict our attention to any $\hat{V} = V_{J,I;0}$ with arbitrarily fixed J, I. Since this is a usual subspace of bilinear finite elements over a tensor product partition, this case is referred to as full grid case. It is easy to see that stable splittings of $\{\hat{V}; \hat{a}\}$ can be obtained from the splittings of Theorem 1 by taking intersections of $\hat{V}$ with the respective subspaces and afterwards neglecting trivial components. This is obvious for the splittings into W-spaces which are direct sums (cf. the corresponding properties of the selection procedure described in [GO2]) and follows then via Proposition 3 also for the V-splitting. We leave this as a corollary to the reader. Instead, we explicitly describe the resulting multilevel nodal basis methods and prewavelet solvers in the following subsections.

3.1.1. Multilevel nodal basis solvers. To this end, let $\mathcal{V}_j$, $\tilde{\mathcal{V}}_i$ and $\mathcal{V}_{j,i} = \mathcal{V}_j \times \tilde{\mathcal{V}}_i$ the sets of interior nodal points associated with the finite element spaces $V_{j;0}$, $\tilde{V}_{i;0}$, and $V_{j,i;0}$, respectively. To each point P belonging to one of these sets, there corresponds a standard basis function in the respective finite element space which takes value 1 at P and 0 at all remaining nodal points. It will be denoted by $N_{j,P}, \tilde{N}_{i,P}$, and $N_{j,i,P}$, respectively (the latter bilinear basis functions are actually simple tensors of two one-dimensional basis functions), the one-dimensional spaces spanned by such basis functions are $V_{j,P}$, $\tilde{V}_{i,P}$, $V_{j,i,P}$. By using simple quadrature rules, the stability of the splittings

$$\{V_{j,i;0}; (\cdot,\cdot)_{L_2}\} = \sum_{P\in\mathcal{V}_{j,i}} \{V_{j,i,P}; (\cdot,\cdot)_{L_2}\} \tag{25}$$

can be established, with condition numbers independent of (j, i). (25) is known as L_2-stability of the finite element nodal basis, it is equivalent to the above mentioned Riesz basis property. Thus, substituting (25) into the second splitting of Theorem 1 (previously intersected with $\hat{V} = V_{J,I;0}$) we arrive at the following Theorem.

Theorem 2. Under the above assumptions and notation, the computational energy space $\{\hat{V}; \hat{a}\}$ corresponding to the full grid discretization of (1) with respect to $\hat{V} = V_{J,I;0}$, possesses the following stable splittings:

$$\{\hat{V}; \hat{a}\} = \sum_{l=0}^{\max(J-j_0, I-i_0)} \sum_{P\in\mathcal{V}_{\nu_l}} \{V_{\nu_l,P}; a^{2l}(\cdot,\cdot)_{L_2}\}\ , \tag{26}$$

$$\{\hat{V}; \hat{a}\} = \sum_{l=0}^{\max(J-j_0, I-i_0)} \sum_{P\in\mathcal{V}_{\nu_l}} \{V_{\nu_l,P}; \hat{a}\}\ , \tag{27}$$

with $\nu_l = (\min(l+j_0, J), \min(l+i_0, I))$. Once again, the integers j_0 and i_0 are defined as in (22). For an example of the indices of involved subspaces, c.f. Figure 1. The

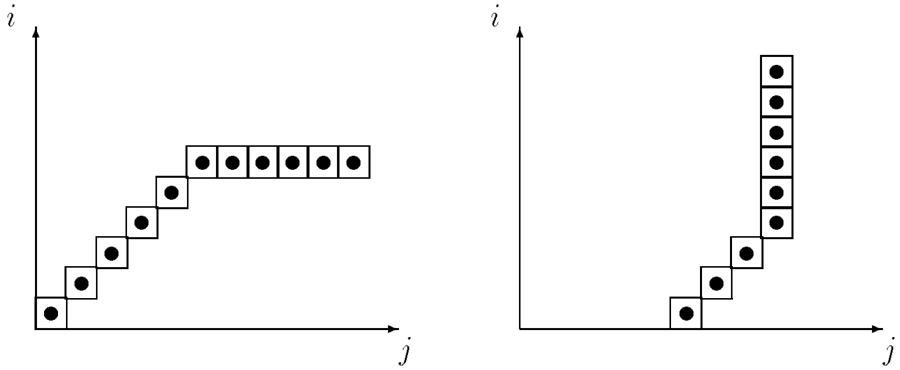


Figure 1. Multilevel nodal basis scheme: necessary subspaces as defined by (22) and ν_l, $J = 10, I = 5, c_1 = c_2$ (left), $I = J = 8, 4^5 \cdot c_1 = c_2$ (right).

condition numbers are independent of c_k, J and I. The result can be generalized to $d > 2$, to $c_0 = 0$ (no Helmholtz term in (1)), and to higher degree Lagrange elements.

Here, several comments are necessary. First, note that the splitting (26) (the stability proof for which was outlined above) is the anisotropic counterpart of the BPX-method [BPX]. A similar construction can be found in [TCK]. The splitting (27) leads to the anisotropic variant of the multilevel diagonal scaling (MDS) method introduced in [Zh] and, independently, in [Wa] where it was called multilevel domain decomposition method. Its stability is a consequence of the stability of the first splitting (26) since one can directly check that $\hat{c}a^{2l}(N_{\nu_l,P}, N_{\nu_l,P})_{L_2} \approx \hat{a}(N_{\nu_l,P}, N_{\nu_l,P})$, i.e. the forms can be interchanged on the one-dimensional subspaces $V_{\nu_j,P}$ without any problems.

Secondly, robustness (with respect to c_k, J, and I) is achieved by incorporating the coefficient information c_1, c_2 and the integers J, I into the algorithm through a specific choice of the subspaces occurring in the splitting – see the definition of the indices ν_l and compare Figure 1. The full splitting $\{\hat{V}; \hat{a}\} = \sum_{j,i=0}^{J,I} \sum_{P \in \mathcal{V}_{j,i}} \{V_{j,i,P}; \hat{a}\}$ is not advisable, since it would result in a condition number of $O(J \cdot I)$ only, compare [GO1]. To obtain better performance, it is sometimes recommended to change (and enlarge) the coarse grid problem by switching to the splitting

$$\{\hat{V}; \hat{a}\} = \{V_{\nu_{l_0};0}; \hat{a}\} + \sum_{l=l_0+1}^{\max(J-j_0, I-i_0)} \sum_{P \in \mathcal{V}_{\nu_l}} \{V_{\nu_l,P}; \hat{a}\} \, . \tag{28}$$

The condition number estimate for these splittings also remains uniform in the new parameter $l_0 = 0, \ldots, \max(J - j_0, I - i_0)$. An analogous modification can be obtained for the BPX-splitting (26).

The underlying multiresolution analysis $\{V_{\nu_l;0}\}$ for all these splittings consists formally of a path in the full, two-dimensional, tensor product family of subspaces $V_{j,i;0}$ generated from $\hat{V} = V_{J,I;0}$ by a certain number of semi-coarsening steps (the next coarser partition is only taken in one, fixed direction while there is no coarsening in the other direction), followed by a certain number of standard-coarsening steps (in both directions), c.f. Figure 1. Thus, our algorithm is basically not more expensive than the traditional BPX- or MDS-preconditioner on a subspace of comparable dimension.

3.1.2. Prewavelet solvers. Now, we present the prewavelet solvers. They are based on suitable bases in $W_{j,i;0}$ or in $\hat{W}_{l;0}$ which should satisfy, as a rule, the Riesz basis property (with well-behaved constants) and possess small masks for the intergrid transfer operations in the resulting multilevel algorithms. Mathematically, the latter means that a prewavelet basis function for $W_{j,i;0}$ should have a nodal basis representation with respect to $\{N_{j,i,P}\}$ with a small number of non-zero coefficients (analogously for $\hat{W}_{l;0}$), i.e. it should possess local support. In our case, we can use tensor product arguments, and reduce the question to the one-dimensional setting. We do not know a general construction for bases in W_j and $W_{j;0}$, respectively, for a more or less arbitrarily given sequence $\{\pi_j\}$, even for linear elements. Thus, we concentrate on the case of sequences of uniform partitions. The corresponding prewavelets are minimally supported one-dimensional spline prewavelets on shift-invariant dyadically refined partitions of $\mathbf{R}$. They have been considered in [CW, Pl], boundary modifications have been given in many papers, see e.g. [Au, CDJW]).

Let us consider the most popular and easy case of uniform dyadic partitions, i.e. let π_j be the uniform partition of $[0,1]$ into 2^{j+1} intervals ($a=2$). Let $\mathcal{V}_j$ denote the set of corresponding interior grid points. Note that we have shifted the index j by 1 to get a nontrivial $V_{0;0}$. To each point $P \in \mathcal{W}_j \equiv \mathcal{V}_j \backslash \mathcal{V}_{j-1}$ ($j>0$) we wish to assign its $\psi_{j,P} \in W_{j;0}$ such that the above requirements are met. Figure 2 shows the nodal values (with respect to $\{N_{j,Q},\ Q \in \mathcal{V}_j\}$) of a suitable choice for these $\psi_{j,P}$, depending on whether P is an interior or boundary (i.e. neighboured to 1 or 0) point of $\mathcal{W}_j$. To the coarsest level, i.e. to the point $P \in \mathcal{V}_0$, we assign the usual nodal basis function.

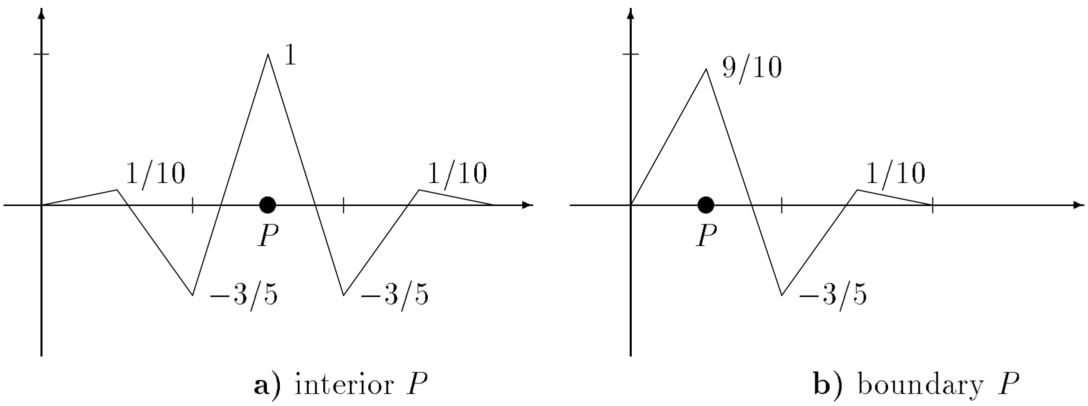


Figure 2. Dyadic prewavelets for $H_0^1(0,1)$.

The Riesz basis property of the splitting

$$\{W_{j;0};(\cdot,\cdot)_{L_2}\} = \sum_{P\in\mathcal{W}_j} \{W_{j,P};(\cdot,\cdot)_{L_2}\} \qquad (\,W_{j,P} = \mathrm{span}\{\psi_{j,P}\}\,) \tag{29}$$

is essentially known, we prove it for the sake of completeness. Consider an arbitrarily fixed j and label the points $P \in \mathcal{W}_{j;0}$ and the prewavelets $\psi_{j,P}$, respectively, from the left to the right by the index $s = 1, \ldots, 2^j$. Analogously, label the points $Q \in \mathcal{V}_j$ and the nodal basis functions $N_{j,Q}$, respectively, by $r = 1, \ldots, 2^{j+1} - 1$. Let $w = \sum_s c_s \psi_s$ (with $\psi_s \equiv \psi_{j,P_s}$) be an arbitrary function from $W_{j;0} \subset V_{j;0}$, and denote its nodal values by $x_r = w(Q_r)$. Set $h = 2^{-j-1}$. Using the Simpson rule, one can easily show that

$$\|w\|_{L_2}^2 \approx h \sum_r x_r^2 \ , \quad \|\psi_s\|_{L_2}^2 \approx h$$

for all s and with reasonable, uniform constants in the two-sided inequalities. Thus, the Riesz basis property (with uniform constants) (29) will follow if we establish

$$\sum_r x_r^2 \approx \sum_s c_s^2 \ ,$$

with constants independent of j. But this is straightforward, look at the nodal values of the prewavelets (see Figure 2) which yield the following expressions for x_r:

$$\begin{aligned} x_1 &= \frac{9}{10} c_1 + \frac{1}{10} c_2 = c_1 + \frac{1}{10}(c_2 - c_1) \\ x_2 &= -\frac{3}{5}(c_1 + c_2) \\ x_3 &= c_2 + \frac{1}{10}(c_1 + c_3) \end{aligned}$$

and so on. Now, use the inequalities $(a + \frac{1}{10} b)^2 \le (1+\epsilon)(a^2 + \frac{1}{100\epsilon} b^2)$, $(a+b)^2 \le 2(a^2 + b^2)$ accordingly to estimate $x_1^2, x_2^2, \ldots$ from above and below, and choose a suitable $\epsilon > 0$ to minimize the bounds. Alternatively (and certainly yielding more precise bounds), we could have calculated the entries of the 5-diagonal symmetric (and almost Toeplitz-type) Gram matrix corresponding to $\{\psi_s\}$.

Now, Proposition 1 yields the L_2-stability of the tensor product prewavelet basis for $\{W_{j,i;0}; (\cdot,\cdot)_{L_2(\Omega)}\}$ when applied to the splittings (29) of $\{W_{j;0}; (\cdot,\cdot)_{L_2(0,1)}\}$ and $\{\tilde{W}_{i;0}; (\cdot,\cdot)_{L_2(0,1)}\}$, with bounds for the Riesz constants which are uniform in (j,i). Substituting this result into the first stable splitting (19) stated in Theorem 1, we have proved the stability of the splitting (30) in the following Theorem.

Theorem 3. Under the above assumptions and notation, the computational energy space $\{\hat{V}; \hat{a}\}$ that corresponds to the full grid discretization of (1) with respect to $\hat{V} = V_{J,I;0}$, possesses the following stable splittings:

$$\{\hat{V}; \hat{a}\} = \sum_{j=0}^{J} \sum_{i=0}^{I} \sum_{P \in \mathcal{W}_j \times \tilde{\mathcal{W}}_i} \{W_{j,i,P}; (c_1 2^{2j} + c_2 2^{2i} + 1)(\cdot,\cdot)_{L_2(\Omega)}\} \ , \tag{30}$$

$$\{\hat{V}; \hat{a}\} = \sum_{j=0}^{J} \sum_{i=0}^{I} \sum_{P \in \mathcal{W}_j \times \tilde{\mathcal{W}}_i} \{W_{j,i,P}; \hat{a}\} \ . \tag{31}$$

Here, $W_{j,i,P}$ is the one-dimensional subspace which is spanned by $\psi_{j,i,P} = \psi_{j,P'} \otimes \tilde{\psi}_{i,\tilde{P}'}$. It corresponds to $P = (P', \tilde{P}') \in \mathcal{W}_j \times \tilde{\mathcal{W}}_i$. The condition numbers are independent of

c_k, J and I. The result can be generalized to $d > 2$, to $c_0 = 0$ (no Helmholtz term in (1)), and to higher degree Lagrange elements.

The stability of the second splitting (31) follows from the stability of (30): Since the splittings represent direct sums of one-dimensional subspaces spanned by the prewavelets $\psi_{j,i,P}$, taking $u = \psi_{j,i,P}$, we have

$$\hat{a}(\psi_{j,i,P}, \psi_{j,i,P}) \approx |||\psi_{j,i,P}|||^2 = (c_1 2^{2j} + c_2 2^{2i} + 1)\|\psi_{j,i,P}\|^2_{L_2(\Omega)}$$

(the $|||\cdot|||$-norm corresponds to the first splitting) and we can change the scalar products on $W_{j,i,P}$ as previously.

The straightforward implementation of these splittings through the additive Schwarz formulation (16), without explicitly calculating the stiffness matrix corresponding to the prewavelet system in $V_{J,I;0}$, is described in Section 4. It turns out that it is (by a small constant factor) more costly than the nodal basis methods of Theorem 2. Note that the splittings (30) and (31) of Theorem 3 are essentially independent of the coefficients c_j. They represent direct sums which might be attractive when the consideration has to be extended to an adaptive refinement application.

We want to give one more example of one-dimensional prewavelets which might lead to attractive robust multilevel schemes. This time, we propose to consider a sequence of triadically refined uniform $\{\pi_j\}$, i.e. let π_j be the uniform partition of $[0,1]$ into 3^j intervals ($a = 3$). Since the number of grid points in $\mathcal{W}_j = \mathcal{V}_j \backslash \mathcal{V}_{j-1}$ per interval of π_{j-1} is 2, we now will associate two prewavelets $\psi^{odd}_{j,P}$ and $\psi^{even}_{j,P}$ (an odd and an even one) with each interior grid point $P \in \mathcal{V}_{j-1}$, and add one more function to each endpoint (the so-called boundary prewavelets). Figure 3 shows the nodal values of the interior prewavelets. For $H^1_0(0,1)$-subspaces, it turns out that the boundary prewavelets can be taken as $\psi^{odd}_{j,0/1}|_{[0,1]}$ while for $H^1(0,1)$-subspaces $\psi^{even}_{j,0/1}|_{[0,1]}$ is appropriate.

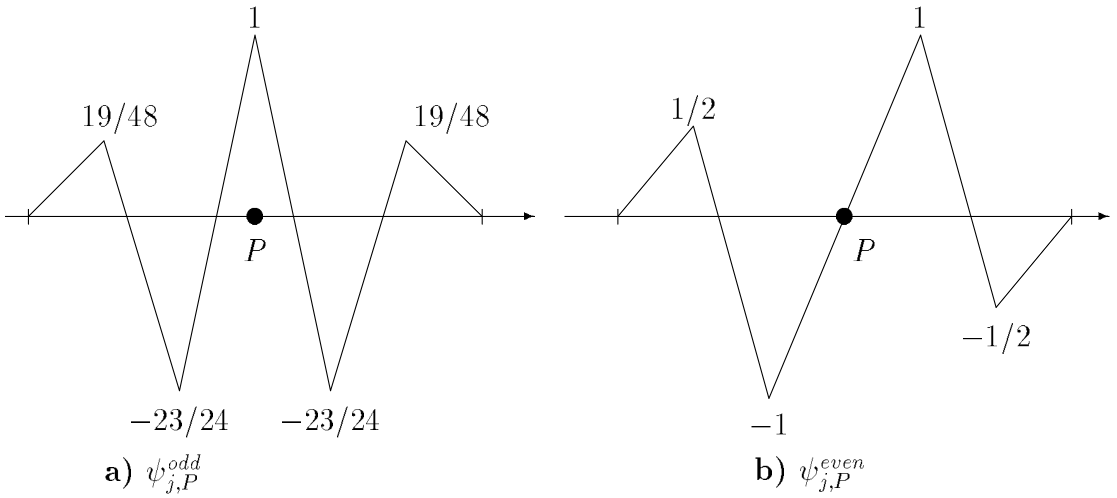


Figure 3. Triadic odd and even prewavelets.

The advantage of this type of prewavelet might be that the dimension of the coarse grid problems is reduced faster than in the dyadic case. However, we did not implement this approach by now.

The use of the tensor product prewavelets is crucial for the robustness of the splittings introduced in Theorem 3. One might think, however, about more economical prewavelet systems resulting from the third stable splitting (21) of Theorem 1. We will be brief since the reasoning repeats the above arguments. Once again, uniform dyadic partitions in both coordinate directions will be used, thus, $a = 2$. Consider the two possible cases for $\hat{W}_{\nu_l;0} = \hat{V}_{\nu_l;0} \ominus \hat{V}_{\nu_{l-1};0}$. For large l, we might have $W_{\nu_l;0} = V_{j,i;0} \ominus_{L_2} V_{j-1,i;0}$ for some index pair (j,i) which corresponds to semi-coarsening in x_1-direction. In this case, the functions $\hat{\psi}_{l,P} = \psi_{j,P'} \otimes \tilde{N}_{i,\tilde{P}'}$ associated with $P = (P', \tilde{P}') \in \mathcal{W}_j \times \tilde{\mathcal{V}}_i$ yield a proper basis of $\hat{W}_{\nu_l;0}$.

For small l, $W_{\nu_l;0} = V_{j,i;0} \ominus_{L_2} V_{j-1,i-1;0}$ (full coarsening) and

$$\hat{\psi}_{l,P} = \begin{cases} \psi_{j,P'} \otimes \tilde{N}_{i,\tilde{P}'} & P = (P', \tilde{P}') \in \mathcal{W}_j \times \tilde{\mathcal{V}}_{i-1} \\ N_{j,P'} \otimes \tilde{\psi}_{i,\tilde{P}'} & P = (P', \tilde{P}') \in \mathcal{V}_{j-1} \times \tilde{\mathcal{W}}_i \\ \psi_{j,P'} \otimes \tilde{\psi}_{i,\tilde{P}'} & P = (P', \tilde{P}') \in \mathcal{W}_j \times \tilde{\mathcal{W}}_i \end{cases}$$

will form a basis in $\hat{W}_{\nu_l;0}$ (for further use, we denote the corresponding sets of nodal points by $\hat{\mathcal{W}}_{\nu_l}$). All these functions have small masks when represented in the associated nodal bases, the shape of their support is not as degenerating as for some of the tensor product prewavelets but still depends on the index pairs (J, I) and (j_0, i_0) and thus on the coefficients c_k. The Riesz basis property follows by tensor product arguments (Proposition 1) from the known one-dimensional results on the basis systems. Thus, we have

Theorem 4. Under the above assumptions and notation, the computational energy space $\{\hat{V}; \hat{a}\}$ that corresponds to the full grid discretization of (1) with respect to $\hat{V} = V_{J,I;0}$, possesses the following stable splittings:

$$\{\hat{V}; \hat{a}\} = \sum_{l=0}^{\max(J-j_0, I-i_0)} \sum_{P \in \hat{\mathcal{W}}_{\nu_l}} \{\hat{W}_{\nu_l,P}; 2^{2l}(\cdot,\cdot)_{L_2}\} \tag{32}$$

$$\{\hat{V}; \hat{a}\} = \sum_{l=0}^{\max(J-j_0, I-i_0)} \sum_{P \in \hat{\mathcal{W}}_{\nu_l}} \{\hat{W}_{\nu_l,P}; \hat{a}\} \tag{33}$$

where $\nu_l = (\min(l + i_0, J), \min(l + i_0, I))$. For the definition of j_0, i_0, see (22). The condition numbers are independent of c_k, J and I. The result can be generalized to $d > 2$, to $c_0 = 0$ (no Helmholtz term in (1)), and to higher degree Lagrange elements.

3.2. Sparse grid discretizations. If one has a stable splitting (10) into a *direct* sum of subspaces, then automatically any subsplitting

$$\{H^*; a\} = \sum_j \{H_j^*; b_j\} \qquad (\ H_j^* \subset H_j\ )$$

is stable whereas the condition number $\kappa^* \leq \kappa$ obviously can not increase (note that if $H_j^* = \{0\}$ for some j, the corresponding term will be dropped in the sum). H^* is the closure (in $\{H; a\}$) of the algebraic sum of the H_j^*.

In particular, this observation applies to the first splitting of Theorem 1 as well as to the prewavelet splittings of the previous section that had been obtained by further decomposing $W_{j,i;0}$ into the one-dimensional $W_{j,i,P}$, $P \in \mathcal{W}_j \times \tilde{\mathcal{W}}_i$ (equipped with scaled L_2 scalar products or the original form $\hat{a}$ associated with (1)). E.g., dropping all subspaces where $j > J$ or $i > I$, the result of Theorem 3 directly follows.

Another interesting choice leads to the splitting

$$\{\hat{V}_n; \hat{a}\} = \sum_{i+j \leq n} \{W_{j,i;0}; (c_1 a^{2j} + c_2 a^{2i} + 1)(\cdot,\cdot)_{L_2}\} , \quad \hat{V}_n = \sum_{i+j=n} V_{j,i;0} \, , \; n \geq 0 \, . \quad (34)$$

The use of this so-called sparse grid space $\hat{V}_n$ for discretizing elliptic boundary value problems was proposed in [Ze] and has been elaborated in a series of subsequent papers [Bg, GSZ, G1, G2]. In a recent paper [GO1], we have described a nodal basis scheme for the isotropic case which was optimal (condition numbers independent of n) in the two-dimensional case and suboptimal (logarithmic growth of $\hat{\kappa}_n$) for $d > 2$. The results of this paper extend [GO1] in two directions. It is obvious that using prewavelets we get a robust, optimal method for all $c_k > 0$, $n \geq 0$, and all dimensions $d \geq 2$. On the other hand, the nodal basis method of [GO1] is, after suitable modifications, robust with respect to c_k, and preserves optimality for $d = 2$ and suboptimality for $d > 2$ if $n \to \infty$.

For simplicity, we will consider the case of uniform dyadic partitions as described in 3.1 for two dimensions, see the preparations for Theorem 3. Thus, $a = 2$. Define j_0, i_0 from the coefficients c_1, c_2 as described in (22). As before, let $c_0 = 1$.

Theorem 5. Let $\hat{a}$ be the bilinear form corresponding to (1) as specified in 3.1 ($d = 2$, $c_0 = 1$). Then, the following splittings are stable, uniformly with respect to c_k and n:

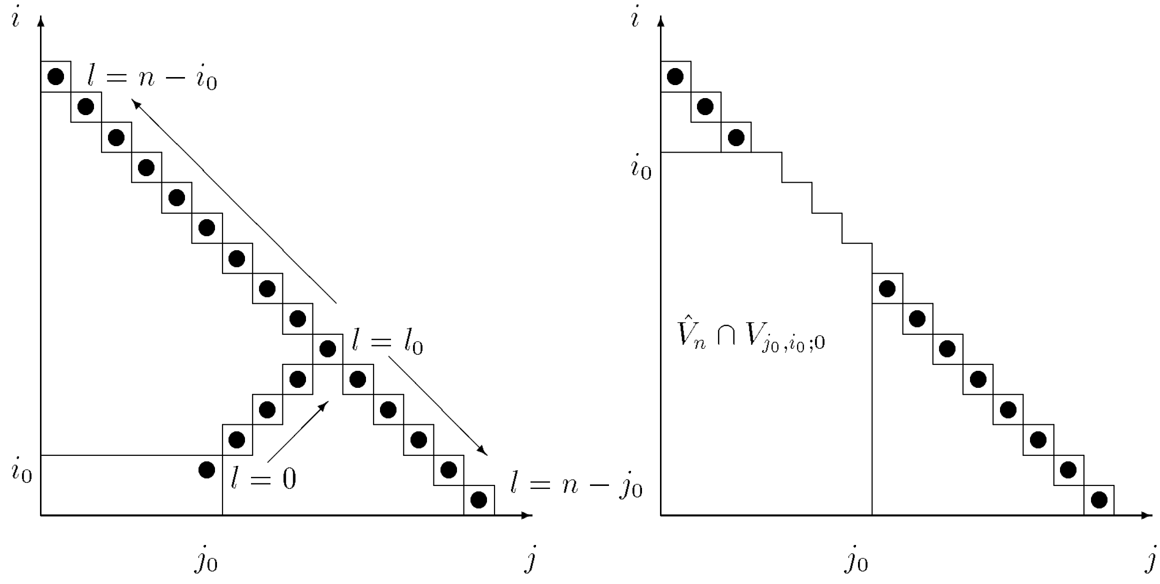
$$\{\hat{V}_n; \hat{a}\} = \sum_{j+i \leq n} \sum_{P \in \mathcal{W}_j \times \tilde{\mathcal{W}}_i} \{W_{j,i,P}; \hat{a}\} \, , \quad (35)$$

$$\begin{aligned} \{\hat{V}_n; \hat{a}\} \;=\; & \{\hat{V}_n \cap V_{j_0,i_0;0}; (\cdot,\cdot)_{L_2}\} + \sum_{l=1}^{n-j_0} \{V_{j_0+l,\min(n-j_0-l,i_0+l);0}; 2^{2l}(\cdot,\cdot)_{L_2}\} \\ & + \sum_{l=1}^{n-i_0} \{V_{\min(j_0+l,n-i_0-l),i_0+l;0}; 2^{2l}(\cdot,\cdot)_{L_2}\} \, . \end{aligned} \quad (36)$$

The results extend to $c_0 = 0$, to $d > 2$ (with the restriction that the analog of the second splitting (36) leads to condition numbers growing at most as $O((\log n)^{d-2})$ with $n \to \infty$).

The first splitting (35) represents the prewavelet method. Its stability proof reduces by the previous reasoning to the application of the last splitting (31) in Theorem 3, with $J = I = n$, for instance.

The splitting (36) leads to a nodal basis multilevel method. Indeed, in the case $j_0 + i_0 \leq n$, the first subspace satisfies $\hat{V}_n \cap V_{j_0,i_0;0} = V_{j_0,i_0;0}$, moreover, for $0 \leq l \leq (n - j_0 - i_0)/2$ the two subproblems corresponding to such l coincide (actually one may be dropped). Figure 4 a) schematically shows the index pairs (j,i) of the subspaces $V_{j,i;0}$ finally involved in this case. We may replace all these $V_{j,i;0}$ by their splittings with respect to the nodal bases (see (25)). The same holds true if $j_0 = 0$ or $i_0 = 0$ in which case $\hat{V}_n \cap V_{j_0,i_0;0}$ coincides with $V_{0,\min(n,i_0);0}$ or $V_{\min(n,j_0),0;0}$. In the remaining cases, the first subproblem is not as trivial as the others (but still harmless since it represents the part where no preconditioning is necessary), see Figure 4 b) for an example. We do not discuss this case further, it does not occur, e.g., for $c_0 = 0$ (in this case either j_0 or i_0 is zero).



a) $j_0 + i_0 \leq n$ **b)** $j_0 + i_0 > n$

Figure 4. Sparse grid splittings: necessary subspaces

The proof of the stability of the second splitting is as follows. We start from (34) and group the components of this splitting according to the size of the factor $c_1 2^{2j} + c_2 2^{2i} + 1$. There may occur different subcases depending on j_0 and i_0, they can be dealt with in the same way. Let us focus, for instance, on $n > j_0 > i_0 = 0$. Define l_0 to be the largest integer such that $j_0 + 2l \leq n$. Then, we obtain the following grouping:

$$
\begin{aligned}
\{\hat{V}_n; \hat{a}\} \quad &= \quad \sum_{j=0}^{j_0} \{W_{j,0;0}; (\cdot,\cdot)_{L_2}\} + \sum_{l=1}^{l_0} \Big(\sum_{\min(l+j_0-j,l-i)=0} \{W_{j,i;0}; 2^{2l}(\cdot,\cdot)_{L_2}\} \Big) \\
&\quad + \sum_{l=l_0+1}^{n-j_0} \Big(\sum_{i=0}^{n-l-j_1} \{W_{l+j_0,i;0}; 2^{2l}(\cdot,\cdot)_{L_2}\} \Big) + \sum_{l=l_0+1}^{n} \Big(\sum_{j=0}^{n-l} \{W_{j,l;0}; 2^{2l}(\cdot,\cdot)_{L_2}\} \Big) \\
&\equiv \quad V_{j_0,0;0} + \sum_{l=1}^{l_0} V_{j_0+l,l;0} \ominus V_{j_0+l-1,l-1;0}
\end{aligned}
$$

$$+\sum_{l=l_0+1}^{n-j_0} V_{j_0+l,n-l-j_0;0} \ominus V_{j_0+l-1,n-l-j_0;0} + \sum_{l=l_0+1}^{n} V_{n-l,l;0} \ominus V_{n-l,l-1;0}$$

(in the latter representation we have used the L_2-orthogonality of the $W_{j,i;0}$, the scalar products are $2^{2l}(\cdot,\cdot)_{L_2}$). The only difference to (34) is the scaling of the L_2-products which is reduced by a factor $\approx c_2^{-1}$.

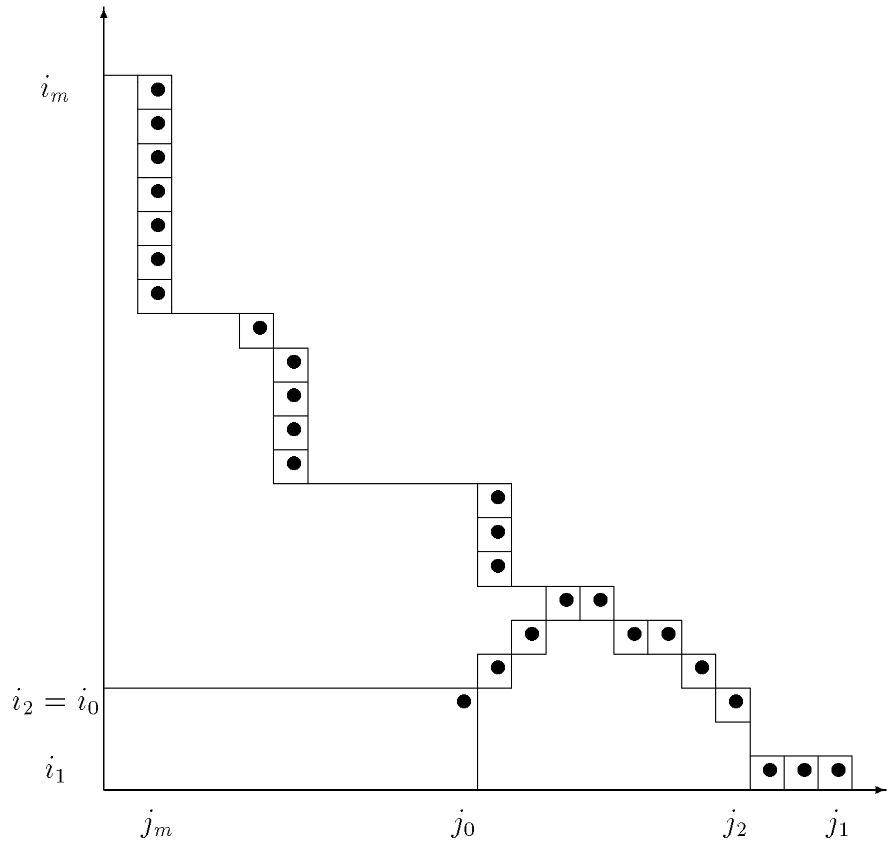


Figure 5. Subspace choice for general $\hat{V}$

It remains to estimate the $|||\cdot|||$-norm of the second splitting (36) of Theorem 5 which can be done in analogy to the proof of Proposition 3. Since the forefactors increase geometrically, the desired norm equivalence results (note that a formal application of Proposition 3 is not directly possible since the spaces involved do not form an increasing sequence as assumed there). The details are left to the reader. Theorem 5 is established. □

We conclude with the remark that an analog of Theorem 5 holds for any $\hat{V}$ which is the finite union of some $V_{j,i;0}$. Indeed, any such $\hat{V}$ can be represented in the form

$$\hat{V} = \sum_{r=1}^{m} V_{j_r,i_r;0}\ , \quad j_1 > j_2 > \ldots > j_m\ ,\ i_1 < i_2 < \ldots < i_m\ .$$

Then, the analog of the first (prewavelet) splitting is simply

$$\{\hat{V};\hat{a}\} = \sum_{(j,i)\,:\,\exists r\,:\, j\le j_r,\, i\le i_r} \; \sum_{P\in\mathcal{W}_j\times\tilde{\mathcal{W}}_i} \{W_{j,i,P};\hat{a}\}$$

while the sequence necessary for the second (nodal basis) splitting will be obtained by the same rules as explained in the proof of Theorem 5. See Figure 5 for an example and compare with Figure 4 a) and b).

Similar arguments can be used in higher dimensions ($d > 2$). Modified subspaces may occur in the attempt to minimize the dimension of the computational subspace under fixed a-priori regularity assumptions and error bounds (see [Bg], the sparse grid space $\hat{V}_n$ itself is optimal in some sense for isotropic problems ($c_1 = c_2$) if the solution u belongs to $S^{(2,2)}H(\Omega)\cap H_0^1(\Omega)$).

3.3. Neumann conditions. Until now, we only considered splittings for $H_0^1(\Omega)$, i.e. for the case of pure Dirichlet boundary conditions. If Neumann boundary conditions are imposed on part or all of $\partial\Omega$, a singular perturbation problem may result. This is the case, for instance, if we have a pure Neumann boundary value problem, with $H^1(\Omega)$ as energy space, $c_0 = 1$, and c_1 or c_2 tending to zero. Note that for $c_0 = 0$ the Neumann problem has a one-dimensional null-space spanned by constants.

These problems can be easily dealt with if the prewavelet constructions are properly used. To concentrate on the main issues, let us consider the situation of Neumann conditions on the whole of $\partial\Omega$, set $c_0 = 1$ (Helmholtz term involved), and consider the linear spline spaces V_j, $\tilde{V}_i$, $V_{j,i} = V_j\otimes\tilde{V}_i$ with respect to the uniform partitions π_j, $\tilde{\pi}_i$, and their tensor products as introduced in 3.1, respectively. Note that V_0 is now 3-dimensional, for convenience we split it into the orthogonal sum of $W_{-1} = \{const.\}$ and $W_0 = V_0\ominus W_{-1}$. As before, we set $W_j = V_j\ominus V_{j-1}$, $j > 0$. The sets $\mathcal{V}_j$ of nodal points corresponding to V_j now additionally contain the end-points of the interval $[0,1]$, $\mathcal{W}_{-1}$ contains the only point 0, $\mathcal{W}_0$ the points 1 and 1/2, finally $\mathcal{W}_j = \mathcal{V}_j\backslash\mathcal{V}_{j-1}$ for $j > 0$. The prewavelets associated with these subspaces will be fixed as follows: $\psi_{-1,0}(t) = 1$ ($P = 0\in\mathcal{W}_{-1}$), $\psi_{0,P}(t) = 2t-1$ ($P = 1\in\mathcal{W}_0$ for $t\in[0,1]$ while $\psi_{0,P}(t)$ for $P = 1/2\in\mathcal{W}_0$ is shown in Figure 6 a). If $P\in\mathcal{W}_j$ for some $j > 0$ is interior (not a neighbour of 0 or 1), then it coincides with the function shown in Figure 2 a), the boundary wavelet differs from that in Figure 2 b) and is shown in Figure 6 b) for P next to 0.

The Riesz basis property remains true for the modified prewavelet bases in $\{W_j;(\cdot,\cdot)_{L_2}\}$. This leads to the following Theorem.

Theorem 6. Under the above assumptions and notation, the computational energy space $\{\hat{V};\hat{a}\}$ which corresponds to the full grid discretization of (1) with respect to $\hat{V} = V_{J,I}$, possesses the stable splitting

$$\{\hat{V};\hat{a}\} = \sum_{j=-1}^{J}\sum_{i=-1}^{I}\sum_{P\in\mathcal{W}_j\times\tilde{\mathcal{W}}_i}\{W_{j,i,P};\hat{a}\}\ . \tag{37}$$

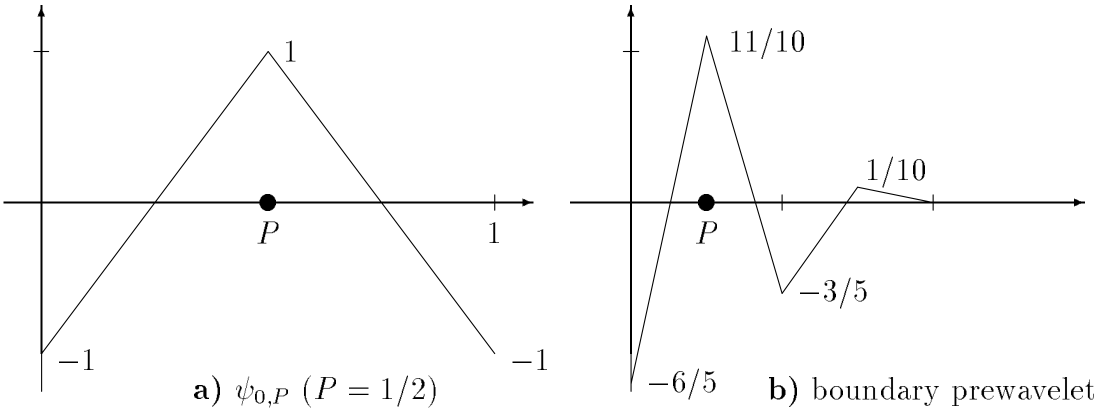


Figure 6. Modified prewavelets for H^1-problems.

Here, $W_{j,i,P}$ is the one-dimensional subspace spanned by $\psi_{j,i,P} = \psi_{j,P'} \otimes \tilde{\psi}_{i,\tilde{P}'}$ corresponding to $P = (P', \tilde{P}') \in \mathcal{W}_j \times \tilde{\mathcal{W}}_i$. The condition numbers are independent of c_k, J and I. The result can be generalized to $d > 2$ and to higher degree Lagrange elements. For $c_0 = 0$ the result remains true if $\hat{V} = V_{J,I}/\{const.\}$ is considered, and the term corresponding to $j = i = -1$ is dropped in the splitting.

Proof. We look careful at the analog of Example 3 in Section 2 for Neumann boundary conditions. The same arguments as used there lead to

$$\{H^1(\Omega); \hat{a}\} \equiv c_1 H^1(0,1)^* \otimes L_2(0,1) \cap c_2 L_2(0,1) \otimes H^1(0,1)^* \cap L_2(\Omega)$$

where $H^1(0,1)^*$ stands for $H^1(0,1)$ equipped with the semi-definite scalar product $(u,v)_{H^1(0,1)^*} = (u',v')_{L_2(0,1)}$. Since the splitting

$$\{H^1(0,1); (\cdot,\cdot)_{H^1(0,1)}\} = \sum_{j=-1}^{\infty} \sum_{P \in \mathcal{W}_j} \{W_{j,P}; 2^{2j}(\cdot,\cdot)_{L_2(0,1)}\}$$

(the analog of (23) for $H^1(0,1)$) is stable, and

$$\|u\|^2_{H^1(0,1)} \approx |u|^2_{H^1(0,1)^*} + (\int_0^1 u(t)\,dt)^2$$

is an equivalent norm on $H^1(0,1)$, we see that

$$\{H^1(0,1)^*; (\cdot,\cdot)_{H^1(0,1)^*}\} = \{W_{-1,0}; 0\} + \sum_{j=0}^{\infty} \sum_{P \in \mathcal{W}_j} \{W_{j,P}; 2^{2j}(\cdot,\cdot)_{L_2(0,1)}\}$$

is formally a stable splitting for $H^1(0,1)^*$. Using Proposition 1 and 2, we obtain the stability of

$$\{H^1(\Omega; \hat{a}\} = \{W_{-1,-1;(0,0)}; (\cdot,\cdot)_{L_2}\} + \sum_{j=0}^{\infty} \sum_{P \in \mathcal{W}_j \times \tilde{\mathcal{W}}_{-1}} \{W_{j,-1,P}; (c_1 2^{2j} + 1)(\cdot,\cdot)_{L_2}\} \quad (38)$$

$$+ \sum_{i=1}^{\infty} \sum_{P \in \mathcal{W}_{-1} \times \tilde{\mathcal{W}}_i} \{W_{j,i,P}; (c_2 2^{2i} + 1)(\cdot,\cdot)_{L_2}\}$$

$$+ \sum_{j=0}^{J} \sum_{i=0}^{I} \sum_{P \in \mathcal{W}_j \times \tilde{\mathcal{W}}_i} \{W_{j,i,P}; (c_1 2^{2j} + c_2 2^{2i} + 1)(\cdot,\cdot)_{L_2}\}$$

from which the assertion of Theorem 6 follows by selecting the subsplitting corresponding to $V_{J,I}$ and switching back to the energy scalar product $\hat{a}(\cdot,\cdot)$ on the $W_{j,i,P}$. □

At the moment, it is not clear whether the analog of the nodal basis splittings (26) and (27), respectively, is robust for all choices of coefficients c_k (or will become robust after some cheap modification). The extension of Theorem 6 to $\hat{V}_n$ is straightforward.

4. Numerical experiments. We now consider implementations of the additive Schwarz preconditioned conjugate gradient method that corresponds to the multilevel nodal basis splittings and to the prewavelet splittings and report the results of our numerical experiments regarding the convergence behaviour of the resulting algorithms.

4.1. Remarks on implementational issues. Of course, any efficient implementation of the additive Schwarz operator that results from the various stable splittings introduced in section 3, either the multilevel nodal basis schemes or the prewavelet schemes, must follow some sort of multigrid- or pyramid-type implementation. Thus, it can be avoided to construct the operator P explicitly, i.e. to assemble a corresponding matrix, which would be more densely populated than the usual finite element stiffness matrix, anyway. Instead, the action of P onto a function or its corresponding vector with respect to some basis representation can be described by the application of certain simple operators that essentially map a representation with respect to the multilevel nodal basis or the prewavelet basis to the standard nodal basis representation, and vice versa. Additionally some simple scaling operators are necessary.

To be more precise and for reasons of simplicity, let us consider the case of the full grid discretization. Then, following the explanations in [G3, G4], the additive Schwarz operator P and the operator equation (16), respectively, can be implemented by some sort of multilevel preconditioner applied to the linear system $Lu = f$ of discrete equations that arises from the finite element nodal basis on the fine grid space $V_{J,I;0}$. For the multilevel nodal basis schemes, we obtain the representation

$$Pu = \phi \qquad \Longleftrightarrow \qquad S\hat{D}^{-1}S^T \cdot Lu = S\hat{D}^{-1}S^T f \tag{39}$$

and for the prewavelet schemes, we get

$$Pu = \phi \qquad \Longleftrightarrow \qquad TD^{-1}T^T \cdot Lu = TD^{-1}T^T f\ . \tag{40}$$

Now, u denotes the vector of nodal values of the unknown function, f denotes the associated vector that arises from the discretization of the right hand side by finite element nodal basis functions in $V_{J,I;0}$ and L denotes the usual sparse finite element stiffness matrix resulting from the discretization with finite element nodal basis functions in $V_{J,I;0}$.

The operator S is a mapping from the vector space that is associated with the (in general non-unique) multilevel nodal basis representation to the vector space that is associated with the unique nodal basis representation on the fine grid space. Moreover, it has the vector as input that contains all the coefficients of the functions of a splitting with respect to the nodal bases in the respective subspaces (c.f. (11)) and the vector of nodal values with respect to the fine grid space $V_{J,I;0}$ as output. The involved sequence of subspaces is determined by (22) and the definition of ν_l.

This summation process can be implemented by a recursion from $l = 0, .., \max(J - j_0, I - i_0)$ involving only interpolation operators, i.e. multigrid prolongation operators between successive subspaces V_{ν_l} and $V_{\nu_{l+1}}$ with the local masks

$$s_{j,i} = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}, \quad s_{j,\cdot} = \begin{bmatrix} \frac{1}{2} & 1 & \frac{1}{2} \end{bmatrix} \text{ and } s_{\cdot,i} = \begin{bmatrix} \frac{1}{2} \\ 1 \\ \frac{1}{2} \end{bmatrix} \tag{41}$$

(with $\nu_l = (j, i)$) for the standard refinement case, the x-semi- and the y-semi-refinement case, respectively. The operator S^T is implemented as the transposed of S.

The operator $\hat{D}$ is a mapping from the vector space which is associated with the multilevel nodal basis representation to itself and implements a scaling. It just contains the weights of the respective splittings as diagonal entries. They can be computed in a setup step (involving $O(2^{I+J})$ operations) and stored for further use. Its action onto a multilevel nodal basis vector can be computed in $\alpha_1(2^J - 1)(2^I - 1)$ operations. The constant α_1 ranges from $4/3$ to 2 depending on the sequence of subspaces that has to be taken into account, c.f. (22) and the definition of ν_l. Note that S and $\hat{D}$ should be indexed by $\cdot_{J,I,c_1,c_2}$ to indicate their dependence on the suitable chosen subspaces but we omit this for reasons of simplicity.

Thus, all we do algorithmically, is that we set up the data structures for the respective sequence of spaces according to (22) and the definition of ν_l, precompute and store the values of $\hat{D}$, and just use the preconditioner $S\hat{D}^{-1}S^T$ within the conjugate gradient method for the solution of $Lu = f$. Thus, we apply S^T, then $\hat{D}^{-1}$ and at last S to the corresponding multilevel nodal basis vectors and standard nodal basis vectors.

When it comes to the prewavelet solvers, things look slightly different. First, we do not need to set up a sequence of spaces that has to be chosen properly as given by (22) and the definition of ν_l. Instead, we will use *all* subspaces $V_{j,i;0}, j = 0, .., J, i = 0, .., I$ with its corresponding multilevel nodal basis. Then, we can express the operator T simply by

$$T = \tilde{S}Q$$

where $\tilde{S}$ is defined analogously to S but with all possible subspaces $V_{j,i;0}, j, i = 0, .., J, I$ and its corresponding multilevel nodal basis taken into account. The operator Q is a mapping from the vector space which is associated with the prewavelet basis representation to the vector space which is associated with the multilevel nodal basis representation with respect to all $V_{j,i;0}, j, i = 0, .., J, I$. For our tensor product dyadic prewavelets, it involves the application of the mask

$$q_{j,i} = \begin{bmatrix} \frac{1}{100} & -\frac{6}{100} & \frac{1}{10} & -\frac{6}{100} & \frac{1}{100} \\ -\frac{6}{100} & \frac{36}{100} & -\frac{6}{10} & \frac{36}{100} & -\frac{6}{100} \\ \frac{1}{10} & -\frac{6}{10} & 1 & -\frac{6}{10} & \frac{1}{10} \\ -\frac{6}{100} & \frac{36}{100} & -\frac{6}{10} & \frac{36}{100} & -\frac{6}{100} \\ \frac{1}{100} & -\frac{6}{100} & \frac{1}{10} & -\frac{6}{100} & \frac{1}{100} \end{bmatrix} = \begin{bmatrix} \frac{1}{10} & -\frac{6}{10} & 1 & -\frac{6}{10} & \frac{1}{10} \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{10} \\ -\frac{6}{10} \\ 1 \\ -\frac{6}{10} \\ \frac{1}{10} \end{bmatrix} \quad (42)$$

for each 'interior' prewavelet function $\phi_{j,i,P}$ assigned to the level (j, i). The modifications of the mask for the prewavelets that are situated near the boundary are obvious, c.f. Figure 2. It casts the prewavelet functions or, more exactly, its associated coefficient on level (j, i), to its coefficients with respect to the standard nodal basis in $V_{j,i;0}$. Note that this operation can be implemented (parallel on all levels (j, i)) by a successive application of the mask to all points $P \in \mathcal{W}_j \times \tilde{\mathcal{W}}_i$ with a successive summation of the resulting nodal basis values. This involves only one additional floating point variable to store intermediate results.

The operator D is a mapping from the vector space which is assigned to the prewavelet basis representation onto itself. It implements a scaling of the prewavelet basis. Thus, it just contains the weights of the respective splittings as diagonal entries.

Now, this implementation of the prewavelet preconditioner roughly costs

$$W_{PWL} \equiv 2|\tilde{S}| + 2|Q| + |D|$$

operations, where

$$\begin{aligned} |\tilde{S}| &= \sum_{j=0}^{J-1} \sum_{i=0}^{I-1} (2^j - 1)(2^i - 1) \cdot |s_{j,i}| \\ &\approx 4 \cdot 2^{J-1} 2^{I-1} \cdot |s_{J,I}| = 2^{J+I} \cdot |s_{J,I}|, \end{aligned}$$

$$\begin{aligned} |Q| &= \sum_{j=0}^{J-1} \sum_{i=0}^{I-1} (2^{j-1} - 1)(2^{i-1} - 1) \cdot |q_{j,i}| \\ &\approx 4 \cdot 2^{J-2} 2^{I-2} |q_{J,I}| = \frac{1}{4} 2^{J+I} \cdot |q_{J,I}|, \end{aligned}$$

$$|D| = (2^J - 1)(2^I - 1) \cdot 1 \approx 2^{J+I},$$

and $|s_{j,i}|$ and $|q_{j,i}|$ denote the number of operations necessary for the application of the masks $s_{j,i}$ and $q_{j,i}$, respectively.

The implementation of the multilevel nodal basis preconditioner involves

$$W_{MLN} \equiv 2|S| + |\hat{D}|$$

operations where

$$
\begin{aligned}
|S| &= \sum_{l=0}^{\max(J-j_0,I-i_0)-1} \sum_{P\in\mathcal{V}_{\nu_l}} |s_{\nu_l}| \\
&= \begin{cases} \sum_{l=0}^{J-1} 2^{2l}|s_{l,l}| \approx \frac{4}{3}\cdot 2^{2(J-1)}\cdot |s_{J,J}| = \frac{1}{3}\cdot 2^{2J}\cdot |s_{J,J}| & \text{for } I=J, c_1=c_2, \\ \sum_{l=0}^{J-1} 2^I\cdot 2^l\cdot |s_{l,\cdot}| \approx 2\cdot 2^I\cdot 2^{J-1}\cdot |s_{J,\cdot}| = 2^{2J}\cdot |s_{J,\cdot}| & \text{for } I=J, 4^J c_1 = c_2, \end{cases}
\end{aligned}
$$

$$
|\hat{D}| = \alpha\cdot(2^J-1)(2^I-1) \approx \alpha\cdot 2^{J+I} \qquad \alpha\in[4/3,2].
$$

Here, without loss of generality, we only consider two extreme cases of subspace splittings for the multilevel nodal basis preconditioner, involving either pure standard-coarsening steps (i.e. $I=J, c_1=c_2$) or pure semi-coarsening steps (i.e. $I=J, 4^J c_1 = c_2$). All other cases just result in a certain mixture of standard- and semi-coarsening steps, anyway.

This allows for a rough analysis of the number of operations involved in the different methods. Since $|q_{j,i}|\approx 25$, $|s_{j,i}|\approx 9$ and $|s_{j,\cdot}|\approx 3$ (multiplications and additions) we obtain

$$
W_{PWL} \approx 2^{J+I}(2\cdot 9 + 2\cdot\frac{25}{4} + 1)
$$

and

$$
W_{MLN} \approx 2^{2J}(2\cdot 3+\alpha) \quad \begin{cases} \text{for } I=J, c_1=c_2, \\ \text{as well as for } I=J, 4^J c_1=c_2. \end{cases}
$$

Altogether, we can conclude that this implementation of the prewavelet preconditioner is at least four times more expensive than that of the multilevel nodal basis scheme. Note that it is also possible to exploit the fact that the masks $s_{j,i}$ and $q_{j,i}$ are tensor products of one-dimensional masks, c.f. (41) and (42). This approach only results in slightly different operation count estimates (i.e. $|\tilde{S}|\approx 6\cdot 2^{J+I}$, $|Q|\approx 5\cdot 2^{J+I}$, thus $W_{PWL}\approx 2^{J+I}(2\cdot 6+2\cdot 5+1)$, and the prewavelet preconditioner is at least three times more expensive than the multilevel nodal basis preconditioner).

The conjugate gradient solution process, however, involves, beside the application of a preconditioner, additionally the evaluation of the matrix-vector multiplication and a few scalar products. Thus, with respect to one iteration step of the overall conjugate gradient algorithm, the prewavelet solver is just slightly more costly than the multilevel nodal basis solver (roughly a factor 1.7 to 2 depending on the particular implementation).

Of course, following the ideas of wavelet pyramid schemes, the action of T can be implemented somewhat differently as well. To some extent, a certain sort of refinement relation defining the prewavelet on a coarse level as a linear combination of prewavelets on the next finer level (plus at least one nodal basis function, since our dyadic prewavelets form a basis) might help to reduce the work involved in the prewavelet transformation T slightly. But we expect no seriously better implementation with respect to the overall operation count. Moreover, when it comes to algorithms that have to deal with non-constant coefficients, we expect modifications to be implemented more straightforwardly within our framework.

Note at last, that the representation of the preconditioners in terms of the operators S, $\tilde{S}$ D, $\hat{D}$, T and Q gives further insight in the relation between the multilevel nodal basis schemes, i.e. BPX-type methods, and prewavelet schemes: With help of the relation $T = \tilde{S}Q$, the prewavelet solver can be interpreteted in terms of the BPX-preconditioner as follows:

$$TD^{-1}T^TL = \tilde{S}QD^{-1}Q^T\tilde{S}^TL = \tilde{S}\hat{C}\tilde{S}^TL \ , \qquad \text{where } \hat{C} \equiv QD^{-1}Q^T.$$

The choice of the BPX-method is, to use the simple and cheap $\hat{D}^{-1}$ instead of the more complicated and slightly more expensive $\hat{C} = Q^TD^{-1}Q$, compare (39). Additionally, by (22) and the definition of ν_l, only a properly chosen sequence of subspaces $\{V_{l;0}\}$ is used in S, whereas the prewavelet method uses all possible subspaces $V_{j,i;0}, j, i = 0, .., J, I$. Furthermore, as we will see in the following subsection, the resulting condition numbers and iteration counts are worse for the prewavelet schemes than for the multilevel nodal basis methods. Thus, prewavelet algorithms are less efficient than the multilevel nodal basis methods. However, they are more convenient to deal with when it comes to more general problems than the simple Laplacian. To gain robustness for anisotropic problems, there is no need to carefully choose a suitable sequence of spaces in the splitting like we did by (22) and the definition of ν_l. Thus, we believe that they can be modified much easier than the multilevel nodal basis schemes when it comes to problems with non-constant coefficients or locally varying anisotropy directions and to applications that involve adaptive refinement.

4.2. Results. We now consider the model problem

$$-c_1u_{xx} - c_2u_{yy} + c_0u = f, \text{ in } \Omega = (0,1)^2, \text{ where } c_1, c_2 > 0, c_0 \geq 0$$

with homogeneous Dirichlet boundary conditions. In the following, we compute the characteristic numbers $\lambda_{\min}, \lambda_{\max}, \kappa$, (c.f. (12)) for various splittings introduced in the previous sections. These numbers give an upper bound for the convergence rate of the method of conjugate gradients preconditioned by the corresponding additive Schwarz preconditioner (13), i.e. of the CG-method on the corresponding equation (16). Furthermore, we also give the number *it* of iterations necessary to reduce the norm of the initial residual by a factor 10^{-6}.

4.2.1. Multilevel nodal basis scheme, full grid discretization. First, we consider the multilevel nodal basis schemes for the full grid case $\hat{V} = V_{J,I;0}, J = I$ (with uniform mesh width $h = 1/2^{J+1}$). The corresponding splittings are given in Theorem 2, formulas (26) and (27).

For the Poisson-problem, i.e. $c_1 = c_2 = 1$, $c_0 = 0$, we obtain just the the well-known BPX-preconditioner from the splitting (26) and its MDS-variant [Zh] from the splitting (27). For splitting (27), Table 1 shows the resulting characteristic numbers and iteration counts for comparison only. Here, for I=J, the splitting involves $J+1$ different levels. Note that the same condition numbers and iteration counts result for the splitting (26).

TABLE 1
Poisson-problem, multilevel nodal basis scheme, full grid discretization.

$I = J =$	1	2	3	4	5	6	7	8
λ_{min}	0.823	0.769	0.755	0.751	0.751	0.751	0.75	0.75
λ_{max}	1.70	2.28	2.71	3.06	3.34	3.58	3.78	3.95
κ	2.06	2.96	3.59	4.07	4.45	4.76	5.0	5.3
it	7	10	12	13	14	15	15	16

The corresponding eigenvalues are just multiplied by a constant factor.

In Table 2 we give the values of the characteristic numbers and the iteration count for $c_1 = 1, c_0 = 0$ and varying values for c_2, i.e the anisotropic diffusion problem on the full grid space $V_{I,J;0}$ with fixed level numbers $I = J = 5$. Here, we show the results for splitting (27). Again, note that the same condition numbers and iteration counts result for the splitting (26). The corresponding eigenvalues are just multiplied by a constant factor.

TABLE 2
Anisotropic diffusion problem, multilevel nodal basis scheme, full grid discretization, $I = J = 5$.

$c_2 =$	4^0	4^1	4^2	4^3	4^4	4^5	4^6	4^7	4^8	4^9	4^{10}
λ_{min}	0.751	0.800	0.589	0.524	0.507	0.502	0.501	0.501	0.501	0.501	0.501
λ_{max}	3.34	3.56	4.19	4.91	5.54	6.08	6.42	6.53	6.57	6.58	6.58
κ	4.45	4.45	7.11	9.39	10.9	12.1	12.8	13.0	13.1	13.1	13.1
it	14	14	18	21	23	24	25	25	26	26	26

We see that the characteristic numbers and thus the convergence rates are independent of c_2. Note that for higher values of c_2, we computed the same results as for $c_2 = 4^{10}$. Thus, we obtain a robust solution method for the anisotropic diffusion problem.

4.2.2. Prewavelet scheme, full grid discretization. Now, we consider the prewavelet basis that results from the dyadic prewavelets for $H_0^1(0,1)$ by means of the tensor product approach. The corresponding splittings for the full grid space $\hat{V} = V_{J,I;0}$ are given in Theorem 3, formulas (30) and (31). For the case $c_1 = c_2 = 1, c_0 = 0$, i.e. the Poisson-problem, the results are shown for varying values $I = J$ (i.e. involving $J+1$ levels) in Table 3.

In comparison to the results for the BPX-preconditioner of Table 1, we see that this method results in values that are worse by roughly a factor of 5-6. Furthermore, its implementation is (by a small constant factor of about 2) more costly. Thus, its efficiency is worse as well. However, this method can be implemented more straightforwardly when it comes to less simple problems than the Poisson-problem. In contrary

Table 3
Poisson-problem, prewavelet scheme, full grid discretization.

$I = J =$	1	2	3	4	5	6	7	8
	splitting (30) with forefactors $2^{2j} + 2^{2i}$							
λ_{min}	1.59	1.26	1.07	0.971	0.927	0.888	0.871	0.862
λ_{max}	12.97	18.46	20.44	21.08	22.95	24.46	25.55	26.53
κ	8.17	14.6	19.0	21.7	24.8	27.5	29.3	30.8
it	7	22	28	31	33	35	35	36
	splitting (31)							
λ_{min}	0.261	0.166	0.139	0.118	0.110	0.105	0.103	0.102
λ_{max}	2.17	2.43	2.60	2.79	2.93	3.05	3.15	3.24
κ	8.31	14.6	18.7	23.6	26.7	28.9	30.6	31.9
it	7	22	28	31	34	35	36	37

to the multilevel nodal basis schemes, the coefficient information is now not involved in the choice of the subspaces arising in the respective splittings. Furthermore, this approach is more directly applicable when it comes to adaptive refinement methods.

In Table 4, we give the values of the characteristic numbers and the iteration count for $c_1 = c_2 = 1$ and varying values for c_0, i.e. the Helmholtz-problem, on the full grid space $V_{I,J;0}$ with fixed level numbers $I = J = 5$.

Table 4
Helmholtz problem, prewavelet scheme, full grid discretization, $I = J = 5$.

c_0	0	10	100	1000	10^4	10^5	10^6	10^7	10^8
	splitting (30) with forefactors $2^{2j} + 2^{2i} + c_0$								
λ_{min}	0.927	0.920	0.717	0.491	0.406	0.373	0.359	0.354	0.354
λ_{max}	22.9	22.4	21.0	19.2	10.6	3.25	1.93	1.79	1.78
κ	24.8	24.3	29.3	39.0	26.0	8.7	5.4	5.1	5.0
it	33	33	35	41	34	20	15	14	14
	splitting (31)								
λ_{min}	0.110	0.110	0.111	0.124	0.154	0.260	0.339	0.352	0.353
λ_{max}	2.93	2.91	2.82	2.57	2.41	2.07	1.82	1.78	1.77
κ	26.7	26.4	25.3	20.8	15.6	8.0	5.4	5.1	5.0
it	34	33	30	26	18	15	14	14	14

We see that the characteristic numbers and thus the convergence rates are independent of c_0. They even tend to become better as c_0 tends to infinity. This is not surprising since our operator then gets similar to unity. As the prewavelet basis functions are L_2-orthogonal *between* different levels, the characteristic numbers in the limit case $c_0 \to \infty$ then accounts for the coupling of the prewavelet basis functions *on* each level. Especially for the splitting (31), a monotonic decrease of κ can be seen, whereas, for the splitting (30), first a slight increase of κ takes place before its values start dropping. Anyway, we obtain a robust solution method for the Helmholtz-problem.

In Table 5, we give the values of the characteristic numbers and the iteration count for $c_1 = 1, c_0 = 0$ and varying values for c_2, i.e the anisotropic diffusion problem, on the full grid space $V_{I,J;0}$ with fixed level numbers $I = J = 5$.

We see that the characteristic numbers and thus the convergence rates are inde-

TABLE 5
Anisotropic diffusion problem, prewavelet scheme, full grid discretization, $I = J = 5$.

$c_2 =$	4^0	4^1	4^2	4^3	4^4	4^5	4^6	4^7	4^8	4^9	4^{10}
	splitting (30) with forefactors $2^{2j} + c_2 2^{2i}$										
λ_{min}	0.927	0.925	0.915	0.906	0.901	0.895	0.891	0.892	0.892	0.892	0.892
λ_{max}	22.95	22.98	23.09	23.21	23.26	23.33	23.35	23.38	23.38	23.39	23.39
κ	24.8	24.8	25.2	25.6	25.8	26.1	26.2	26.2	26.2	26.2	26.2
it	33	33	34	34	34	34	34	34	34	34	34
	splitting (31)										
λ_{min}	0.110	0.110	0.109	0.109	0.108	0.108	0.108	0.108	0.108	0.108	0.108
λ_{max}	2.93	2.95	3.01	3.08	3.119	3.14	3.16	3.17	3.18	3.18	3.18
κ	26.7	26.9	27.5	28.2	28.8	29.0	29.2	29.3	29.4	29.4	29.4
it	34	34	35	35	36	36	36	36	36	36	36

pendent of c_2. Note that, for higher values of c_2, we obtained the same results as for $c_2 = 4^{10}$. Thus, we obtain a perfectly robust solution method for the anisotropic diffusion problem.

4.2.3. Sparse grid discretization. Now, we consider the multilevel nodal basis scheme and the prewavelet scheme for the case of the sparse grid space $\hat{V}_n$, c.f. (34). The corresponding splittings are given in Theorem 5, formulas (35), and (36). For the implementation of the operators that correspond to S and $\tilde{S}$, now the hierarchical basis representation of [Ze] must be used instead of the standard nodal basis representation.

For the case $c_1 = c_2 = 1, c_0 = 0$, i.e. the Poisson-problem, the results are shown for the sparse grid discretization on $\hat{V}_n$ with different values n in Table 6. Note that $n + 1$ levels are involved.

TABLE 6
Poisson-problem, multilevel nodal basis scheme and prewavelet scheme, sparse grid discretization.

$n =$	1	2	3	4	5	6	7	8
	Multilevel nodal basis scheme, splitting (35)							
λ_{min}	1	0.960	0.861	0.712	0.665	0.620	0.560	0.544
λ_{max}	2.31	2.58	3.71	3.89	4.72	4.92	5.59	5.75
κ	2.31	2.68	4.30	5.47	7.09	7.94	10.0	10.6
it	4	8	12	15	17	19	21	22
	Prewavelet scheme, splitting (36)							
λ_{min}	0.4	0.262	0.192	0.161	0.140	0.126	0.118	0.113
λ_{max}	1.82	2.19	2.40	2.59	2.78	2.91	3.02	3.12
κ	4.55	8.39	12.5	16.2	19.8	23.1	25.5	27.6
it	4	13	21	26	29	31	33	35

In Table 7 we give the values of the characteristic numbers and the iteration count for $c_1 = 1, c_0 = 0$ and varying values for c_2, i.e the anisotropic diffusion problem, for the sparse grid discretization on $\hat{V}_n$ with fixed $n = 7$, i.e. 8 involved level.

We clearly see that the characteristic numbers and thus the convergence rates are independent of the values of c_2. Thus, we obtained a robust method for anisotropic diffusion problems on sparse grid spaces as well.

TABLE 7

Anisotropic diffusion problem, multilevel nodal basis scheme and prewavelet scheme, sparse grid discretization, $n = 7$.

$c_2 =$	4^0	4^1	4^2	4^3	4^4	4^5	4^6	4^7	4^8	4^9	4^{10}
	Multilevel nodal basis scheme, splitting (35)										
λ_{min}	0.560	0.544	0.532	0.515	0.511	0.508	0.504	0.503	0.502	0.501	0.501
λ_{max}	5.60	5.67	6.14	6.26	6.62	6.73	6.93	7.06	7.16	7.21	7.22
κ	10.0	10.4	11.5	12.1	13.0	13.2	13.8	14.0	14.3	14.3	14.4
it	21	22	23	24	25	25	26	27	27	27	27
	Prewavelet scheme, splitting (36)										
λ_{min}	0.118	0.118	0.117	0.117	0.117	0.117	0.117	0.117	0.117	0.117	0.117
λ_{max}	3.02	3.04	3.09	3.12	3.13	3.13	3.14	3.14	3.14	3.14	3.14
κ	25.6	25.8	26.4	26.7	26.8	26.8	26.8	26.8	26.8	26.8	26.8
it	33	33	34	34	34	34	34	35	34	34	34

5. Concluding remarks. At last, we give a few remarks on possible extensions and open problems.

In this paper, we restricted ourselves to second order problems on rectangles. The extension to certain fourth order problems (anisotropic plate bending or Stokes discretizations via the stream function approach for $d = 2$) is straightforward as long as the tensor product and intersection tricks apply. However, we do not have a theory for general domains, and it is doubtful whether our approach can easily be adapted to this case.

Another practically interesting and at the moment open question is the investigation of robust multilevel methods working also for non-constant anisotropies c_k. The above theorems might give some indication of how to guess reasonable modifications for this difficult problem via localization, however, we do not know about definitive theoretical results. Note that the concept of operator adapted wavelet constructions might be helpful in this respect, compare [DlW1, DlW2, DlK1, DlK2, JwS] for some one-dimensional schemes.

We mentioned that our theory works also for other wavelet-like decompositions (indeed, what we need to start with are L_2-based stable splittings for Sobolev spaces on intervals which are, as a rule, the by-product of any wavelet scheme generated by a sufficiently smooth refinable function ϕ). However, we are not too optimistic about the practical efficiency of such schemes for solving elliptic problems on domains in comparison to the nodal basis (or prewavelet) finite element methods proposed in this paper. There are, at least, two problems not yet satisfactorily resolved for the application of the general wavelet concept in this field: first, the incorporation of boundary conditions, i.e. the adoption from the shift-invariant setting on $\mathbf{R}$ to its application on an interval (see, e.g., [Au, CDJW]) or on more general domains, and, second, the higher arithmetic work per iteration (including an efficient implementation) which usually does not compare favourably to the simple finite element V-cycle-like algorithms our multilevel nodal basis schemes are based on. The majority of recent papers on wavelet solvers does not deal too seriously with these important issues, and it is our belief that much work has still to be done to fully explore the potential of the multilevel wavelet concepts. We refer also to Section 4 where the differences in the implementations of nodal basis versus prewavelet constructions are briefly discussed.

At last, a few words about adaptivity. First of all, we are extremely flexible in selecting computational subspaces (along with an optimal subspace splitting yielding a robust iterative method) since we can rely on stable splittings of the whole, infinite-dimensional energy spaces (Theorem 1). This was demonstrated in 3.2 for the sparse grid spaces. On the other hand, nested refinement (for the nodal basis splittings) and compression (in the (pre-)wavelet cases) producing problem-dependent sequences of computational subspaces and splittings do not influence the convergence rates of the associated iterative solvers (use the selection argument, see [GO2]). A problem which might be more difficult in connection with anisotropy and tensor product techniques as used in this paper is to find and justify appropriate error indicators which are a substantial part of any adaptive code.

Acknowledgements: We are indepted to S. Zimmer for his programming assistance and many fruitful discussions.

REFERENCES

[Au] P. Auscher, Wavelets with boundary conditions on the interval. In: [Ch], 217-236.

[AB] O. Axelsson, V. A. Barker, Finite element solution of boundary value problems. Theory and computation. Acad. Press, New York 1984.

[BKLN] O. V. Besov, L. D. Kudryavzev, P. I. Lizorkin, S. M. Nikol'skij, Investigations in the theory of spaces of differentiable functions of several variables. Proc. Steklov Ins. Math. 1 (1990) 73-139.

[BPX] J. H. Bramble, J. E. Pasciak, J.Xu, Parallel multilevel preconditioners. Math. Comp. 55 (1990), 1-22.

[Bg] H.-J. Bungartz, Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung. Dissertation, TU München 1992.

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

[CW] C. K. Chui, J. Z. Wang, On compactly supported spline wavelets and a duality principle. Trans. Amer. Math. Soc. 330 (1992), 903-916.

[CDJW] A. Cohen, I. Daubechies, B. Jawerth, P. Vial, Multiresolution analysis, wavelets and fast algorithms on the interval, C. R. Acad. Sci. Paris 316 (1993), 417-421.

[DlK1] S. Dahlke, A. Kunoth, A biorthogonal wavelet approach for solving boundary value problems, Preprint Inst. f. Geom. Prakt. Math. RWTH Aachen 1993.

[DlK2] S. Dahlke, A. Kunoth, Biorthogonal wavelets and multigrid, Preprint No. A 21-93, FB Math. FU Berlin 1993.

[DlW1] S. Dahlke, I. Weinreich, Wavelet-Galerkin-methods: An adapted biorthogonal wavelet basis. Constructive approximation 9 (1993), 237-262.

[DlW2] S. Dahlke, I. Weinreich, Wavelet bases adapted to pseudo-differential operators. Preprint FU Berlin, 1992.

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

[DmPS] W. Dahmen, S. Prössdorf, R. Schneider, Multiscale methods for pseudodifferential equations, In: L. L. Schumaker, G. Webb (eds.): Recent advances in wavelet analysis, Acad. Press, New York 1994, 191-235.

[DSW] M. Dryja, B. Smith, O. Widlund, Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions. Preprint, Courant Inst., New York Univ. 1993.

[GL] G. H. Golub, C. F. Van Loan, Matrix Computations. The Johns-Hopkins Univ. Press, Baltimore 1983.

[G1] M. Griebel, A parallelizable and vectorizable multi–level algorithm on sparse grids. In: W. Hackbusch (ed.). Parallel Algorithms for PDE, Proc. 6th GAMM Seminar Kiel, Notes on Numerical Fluid Mechanics 31, Vieweg, Braunschweig 1991, 94-100.

[G2] M. Griebel, Parallel multigrid methods on sparse grids. In: Multigrid Methods III, International Series of Numerical Mathematics 98, Birkhäuser Verlag, Basel, 1991, 211-221.

[G3] M. Griebel, Multilevel algorithms considered as iterative methods on semidefinite systems. SIAM Int. J. Sci. Stat. Comput. 15(3), (1994), 547-565.

[G4] M. Griebel, Multilevelmethoden als Iterationsverfahren über Erzeugendensystemen, Teubner Skripten zur Numerik, Teubner Verlag, Stuttgart, 1994.

[GO1] M. Griebel, P. Oswald, On additive Schwarz preconditioners for sparse grid discretizations. Numer. Math. 66 (1994), 449-464.

[GO2] M. Griebel, P. Oswald, On the abstract theory of additive and multiplicative Schwarz algorithms. Numer. Math. (to appear).

[GSZ] M. Griebel, M. Schneider, C. Zenger, A combination technique for the solution of sparse grid problems. In: P. de Groen and R. Beauwens (eds.), Iterative Methods in Linear Algebra, IMACS, Elsevier, North Holland, 1992, 263-281.

[Ha1] W. Hackbusch, Iterative solution of large sparse systems of equations. Springer, New York 1994.

[Ha2] W. Hackbusch, The frequency decomposition multigrid method. Part I: application to anisotropic equations. Numer. Mathematik 56 (1989), 229-245.

[Ha3] W. Hackbusch, The frequency decomposition multigrid methods. Part II: Convergence analysis based on the additive Schwarz method. Numer. Mathematik 63 (1992), 433-453.

[Ja] S. Jaffard, Wavelet methods for fast resolution of elliptic equations. SIAM J. Numerical Analysis 29 (1992), 965-986.

[JaL] S. Jaffard, P. Laurencot. Orthogonal wavelets, analysis of operators, and applications to numerical analysis. In: [Ch], 543-601.

[JwS] B. Jawerth, W. Sweldens, Wavelet multiresolution analyses adapted for the fast solution of boundary value ordinary differential equations. Proc. Sixth Copper Mountain Multigrid Conf. (April 1993).

[LD] G. G. Lorentz, R. A. DeVore, Constructive approximation. Grundlehren V. 303, Springer, Berlin 1993.

[O1] P. Oswald, Multilevel Finite Element Approximation. Theory & Applications. Teubner Skripten zur Numerik, Teubner, Stuttgart 1994 (in print).

[O2] P. Oswald, On estimates for one-dimensional spline approximation. In: J. W. Schmidt, H. Späth (eds.): Splines in Numerical Analysis. Math. Research, v. 52, Akad.-Verlag, Berlin 1989, 111-124.

[O3] P. Oswald, On function spaces related to finite element approximation theory. Z. Anal. Anwendungen 9 (1990), 43-64.

[Pl] G. Plonka, Generalized spline wavelets. Preprint 94/8, FB Mathematik, Univ. Rostock, 1994.

[RWZ] A. Rieder, R. O. Wells, X. Zhou, A wavelet approach to robust multilevel solvers for anisotropic elliptic problems. Preprint 1993, Rice University Houston 1993.

[RZ] A. Rieder, X. Zhou, On the robustness of the damped V-cycle of the wavelet frequency decomposition multigrid method. TR CML TRS 3-10, Rice University Houston 1993.

[Sm] H.-J. Schmeisser, Vector-valued Sobolev and Besov Spaces, In: Seminar Analysis of the Karl-Weierstraß- Institute 1985/86 (B.-W.-Schulze, H. Triebel, eds.) Teubner Texte Math. Bd. 96, Teubner Leipzig 1987, 4-44

[SmT] H.-J. Schmeisser, H. Triebel, Topics in Fourier Analysis and Function Spaces, Akad. Verlagsges. Geest & Portig K.-G., Leipzig 1987. Ch. 2,4.

[StT] K. Stüben and U. Trottenberg, Multigrid Methods: Fundamental algorithms, model problem analysis and applications, In: Multigrid methods, Lecture Notes in Mathematics 960, Springer-Verlag, 1982.

[Tl] V. N. Temlyakov, Approximation of periodic functions. Nova Science Publishers, New York 1993.

[TCK] C. Tong, T. Chan and C. Kuo, Multilevel filtering elliptic preconditioners, SIAM J. Sci. Stat. Comput. 13 (1992), 227-245.

[Wa] J. Wang, Convergence analysis of Schwarz algorithm and multilevel decomposition iterative methods II: non-selfadjoint and indefinite problems. SIAM J. Numer. Anal. 30 (1993), 953-970.

[We] J. Weidmann, Linear operators in Hilbert spaces. Springer, New York 1980.

[Xu] J. Xu, Iterative methods by space decomposition and subspace correction. SIAM Review 34 (1992), 581-613.

[Ys] H. Yserentant, Old and new convergence proofs for multigrid methods. Acta Numerica. Cambridge University Press, New York 1993, 285-326.

[Ze] C. Zenger, Sparse grids. In: W. Hackbusch (ed.). Parallel Algorithms for PDE, Proc. 6th GAMM Seminar Kiel, Notes on Numerical Fluid Mechanics 31, Vieweg, Braunschweig 1991, 241-251.

[Zh] X. Zhang, Multilevel additive Schwarz methods. Numer. Math. 63 (1992), 521-539.