Mathematical and Computational Techniques for Multilevel Adaptive Methods

Ulrich Rüde
Technische Universität München

Society for Industrial and Applied Mathematics
siam
Philadelphia 1993

Contents

Chapter 2

Multilevel Splittings

In this chapter, we collect results for the multilevel splitting of finite element spaces. Our presentation is motivated by recent results of Oswald [70], [75], [76] and Dahmen and Kunoth [34]. These papers are in turn related to the quickly developing theory of multilevel preconditioners as studied by Yserentant [111], [112], [113], [114], Xu [108], [109], Bramble, Pasciak, and Xu [25], Dryja and Widlund [37], [38], S. Zhang [115], and X. Zhang [116]. The approach by Oswald, Dahmen, and Kunoth is based on results from the theory of function spaces. The relationship of this abstract theory to multilevel methods is developed in a sequence of papers and reports [34], [69], [68], [70], [71], [72], [74], [73], [75], [76].

As the basic algorithmic structure, we introduce the so-called *multilevel additive Schwarz* method. The idea is to use a hierarchy of levels for a multiscale representation of the problem and to combine the contributions of all levels in a sum. This process implicitly defines an *operator sum* that is well behaved and that has bounded condition number independent of the number of levels. Thus, it is suitable for fast iterative inversion by conjugate gradient-type algorithms.

The recent theoretical approach to these methods by Oswald (see the papers cited above) is based on results in approximation theory, in particular on methods from the theory of Besov spaces. The relevant basic results can be found in Nikol'skiĭ [66] and Triebel [105]. An outline of these results, including a bibliography, is also given in a survey article by Besov, Kudrayavtsev, Lizorkin, and Nikol'skiĭ [15].

From a more general perspective, the multilevel additive Schwarz method is also related to multigrid methods and their theory. Classical multigrid methods can be interpreted as a *multiplicative Schwarz* method where the levels are visited sequentially and the basic structure is a *product of operators.*

Multigrid convergence theory has been studied in so many papers that a complete bibliography cannot be given in the context of this monograph. We refer to Hackbusch [43] and McCormick [56] for the classical theory and to Yserentant [114] for a review of recent developments.

It should be noted that the interpretation of multilevel techniques as

Schwarz methods uses the structure of nested spaces and symmetric operators corresponding to what is known as *variational multigrid.* The classical multigrid theory is more general in this respect, because it assumes relations only between the (possibly nonsymmetric) operators; it assumes no special relations between the grid spaces.

The unified multigrid convergence theory developed from the Schwarz concept seems to need nesting of the spaces and symmetry of the operators (however, see attempts to generalize these limitations by Bramble, Pasciak, and Xu [26] and Xu [110]). At this stage, the new theory also fails to describe some features of the multigrid principle, like the dependency of the performance on the number of smoothing steps per level. Typically, however, the new theory does not need as strong regularity assumptions as the classical multigrid theory. The interested reader is referred to the original papers by Dryja and Widlund [38], Bramble, Pasciak, Wang, and Xu [24], [23], [25], [109], and Zhang [116], where the theory is developed.

Here, we will describe these new techniques, following closely the approach by Oswald, Dahmen, and Kunoth, because they provide an elegant theoretical foundation of the fast adaptive methods that will be discussed in Chapter 3. Our emphasis here is to give a consistent presentation of the abstract foundation in approximation theory and its application to the prototype finite element situations that arise in the solution of the model problems in §1.3. In particular, we will show that the same theoretical background can be used to justify fast iterative solvers, error estimates, and mesh refinement strategies.

2.1. Abstract stable splittings

The basis of multilevel algorithms is a decomposition or splitting of the solution space into subspaces. Multilevel algorithms depend on this structure and its particular features. To classify multilevel splittings, we introduce the notion of a *stable splitting* that we will describe in an abstract setting.

We assume that the basic space V is a Hilbert space equipped with a scalar product $\langle \cdot, \cdot \rangle_V$ and the associated norm

$$\|u\|_V = \langle u, u \rangle_V^{1/2}.$$

The elliptic partial differential equation is formulated with a V-elliptic, symmetric, continuous bilinear form $a : V \times V \longrightarrow \mathbb{R}$. Thus there exist constants $0 < c_1 \leq c_2 < \infty$ such that

$$c_1 \langle v, v \rangle_V \leq a(v, v) \leq c_2 \langle v, v \rangle_V \tag{2.1}$$

for all $v \in V$. In view of the model problems (1.1)–(1.2) and their variational form, we study the abstract problem: Find $u \in V$ such that

$$a(u, v) = \Phi(v) \tag{2.2}$$

for all $v \in V$, where the functional $\Phi \in V^*$ is a continuous linear form.

To introduce a multilevel structure we consider a finite or infinite collection $\{V_j\}_{j\in J}$ of subspaces of V, each with its own scalar product $(\cdot,\cdot)_{V_j}$ and the associated norm

$$\|u\|_{V_j} = (u,u)_{V_j}^{1/2}.$$

We further assume that the full space V can be represented as the sum of the subspaces V_j, $j \in J$,

$$V = \sum_{j\in J} V_j. \tag{2.3}$$

REMARK 2.1.1. *Later we will additionally assume that the spaces are nested, that is,* $J \subset \mathbb{N}_0$, $V_i \subset V_j$, *if* $i \le j$. *The theory in this section, however, does not depend on this assumption and without it many more iterative methods can be described in the abstract framework, including classical relaxation methods, block relaxation, and domain decomposition.*

REMARK 2.1.2. *In typical applications,* $\|\cdot\|_V$ *is equivalent to the* H^1-*Sobolev norm. It is our goal to build equivalent norms based on the subspaces* V_j *and their associated norms. The subspace norms* $\|\cdot\|_{V_j}$ *will be properly scaled* L_2-*norms. Based on the associated bilinear forms, we can construct a multilevel operator that uses elementary operations (based on the* L_2-*inner products) on all levels, except possibly the coarsest one; see Definitions* 2.1.3 *and* 2.1.4. *If this operator is spectrally equivalent to the original operator, it can be used to build efficient preconditioners, error estimates, and refinement algorithms.*

The system of subspaces induces a structure in the full space. Any element of V can be represented as a sum of elements in V_j, $j \in J$. Generally, this representation is nonunique. This observation gives rise to the following definition.

DEFINITION 2.1.1. *The* **additive Schwarz norm** $|||\cdot|||$ *in* V *with respect to the collection of subspaces* $\{V_j\}_{j\in J}$ *is defined by*

$$|||v||| \stackrel{\text{def}}{=} \inf\left\{ \left(\sum_{j\in J} \|v_j\|_{V_j}^2\right)^{1/2} \,\middle|\, v_j \in V_j,\ \sum_{j\in J} v_j = v \right\}. \tag{2.4}$$

As we will show, how well a multilevel algorithm converges depends on how well the multilevel structure captures the features of the original problem, that is, how well the additive Schwarz norm approximates the original norm $\langle\cdot,\cdot\rangle_V$ in V. This motivates the following definition.

DEFINITION 2.1.2. *A collection of spaces* $\{V_j\}_{j\in J}$ *is called a* **stable splitting** *of* V *if*

$$\sum_{j\in J} V_j = V$$

and if $\|\cdot\|_V$ *is equivalent to the additive Schwarz norm of* V, *that is, if there exist constants* $0 < c_3 \le c_4 < \infty$ *such that*

$$c_3\|v\|_V^2 \le |||v|||^2 \le c_4\|v\|_V^2 \tag{2.5}$$

for all $v \in V$. *The number*

$$\kappa(V, \{V_j\}_{j\in J}) \stackrel{\text{def}}{=} \inf(c_4/c_3), \tag{2.6}$$

that is, the infimum over all possible constants in (2.5), *is called the* **stability constant** *of the splitting* $\{V_j\}_{j\in J}$.

REMARK 2.1.3. *If V is finite-dimensional, any splitting is stable. The concept of a stable splitting is therefore primarily relevant in an infinite-dimensional setting. In practice, the problem in a finite-dimensional space is not to show the existence of a stable splitting, but to study the size of the stability constant. We will assume that the finite-dimensional discrete space is embedded in an infinite-dimensional space. By showing that the splitting of the infinite-dimensional space is stable, we can derive bounds for the stability constant that are uniform in the number of levels.*

The definition of a stable splitting leaves room for many cases, including pathological ones.

Example. Consider a splitting of an arbitrary nontrivial Hilbert space V into two subspaces V_1 and V_2. Let $V_1 = \text{span}\{x\}$ for some $x \in V$, $x \neq 0$, and

$$\|\alpha x\|_{V_1} \stackrel{\text{def}}{=} |\alpha|,$$

where $\alpha \in \mathbb{R}$. Then let $V_2 = V$ and $\|\cdot\|_{V_2} \equiv \|\cdot\|_V$. To show that the splitting $V = V_1 + V_2$ is stable, we must show that (2.5) holds. In this simple case we have

$$|||v||| = \inf_{v_1\in V_1} \sqrt{\|v_1\|_{V_1}^2 + \|v - v_1\|_{V_2}^2} = \inf_{\alpha\in\mathbb{R}} \sqrt{\alpha^2 + \|v - \alpha x\|_V^2}.$$

Therefore, the upper bound holds trivially with $c_4 = 1$ (set $\alpha = 0$). The lower bound can be constructed as follows:

$$\begin{aligned}
|||v|||^2 &= \inf_{\alpha\in\mathbb{R}} \left(\alpha^2 + \|v - \alpha x\|_V^2\right) \\
&\geq \inf_{\alpha\in\mathbb{R}} \left(\alpha^2 + (\|v\|_V - \alpha\|x\|_V)^2\right) \\
&= \inf_{\alpha\in\mathbb{R}} \left((1 + \|x\|_V^2)\alpha^2 - 2\|x\|_V\|v\|_V\alpha + \|v\|^2\right) \\
&= \frac{-(\|x\|_V\|v\|_V)^2 + \|v\|_V^2(1 + \|x\|_V^2)}{1 + \|x\|_V^2} \\
&= c\|v\|_V^2,
\end{aligned}$$

where $c = 1/(1 + \|x\|_V^2)$. Despite the stability, the practical value of the splitting is doubtful. As we will see further below, $V_2 = V$ and $\|\cdot\|_{V_2} = \|\cdot\|_V$ imply that methods based on this splitting involve the solution of a problem that is equivalent to the original one.

The following example is more typical for the situations that are of interest to us.

Example. Let the one-dimensional *Fourier components* ϕ_n be given by

$$\phi_n : \begin{cases} [0,1] & \longrightarrow & \mathbb{R}, \\ \phi_n(x) & = & \sin(n\pi x), \end{cases}$$

$n \in \mathbb{N}$. Consider the space

$$V = \left\{ f = \sum_{n=1}^{\infty} a_n \phi_n \mid \|f\|_V < \infty \right\}, \tag{2.7}$$

where $a_n \in \mathbb{R}$ and

$$\left\| \sum_{n=1}^{\infty} a_n \phi_n \right\|_V \stackrel{\text{def}}{=} \sqrt{\sum_{n=1}^{\infty} n^2 a_n^2}. \tag{2.8}$$

Subspaces of V are now defined by

$$V_j = \left\{ f = \sum_{n=1}^{2^j-1} a_n \phi_n \right\} \subset V, \tag{2.9}$$

where $j \in \mathbb{N}$ and the corresponding norms are defined by

$$\left\| \sum_{n=1}^{2^j-1} a_n \phi_n \right\|_{V_j} \stackrel{\text{def}}{=} 2^j \sqrt{\sum_{n=1}^{2^j-1} a_n^2}. \tag{2.10}$$

Note that $\|\cdot\|_V$ corresponds to an H^1-norm, while $\|\cdot\|_{V_j}$ corresponds to a (scaled) L_2-norm.

Next, we show that the spaces $(V_j, \|\cdot\|_{V_j})$, $j \in \mathbb{N}$, are a stable splitting of $(V, \|\cdot\|_V)$. Clearly, each $f \in V$ can be decomposed in the form

$$f = \sum_{n=1}^{\infty} a_n \phi_n = \sum_{j=1}^{\infty} \underbrace{\sum_{n=1}^{2^j-1} a_{nj} \phi_n}_{= f_j \in V_j}, \tag{2.11}$$

where $\sum_{j=1}^{\infty} a_{nj} = a_n$. Setting $a_{nj} = 0$ if $n \geq 2^j$ we have

$$\sum_{j=1}^{\infty} \|f_j\|_{V_j}^2 = \sum_{j=1}^{\infty} 2^{2j} \sum_{n=1}^{2^j-1} a_{nj}^2 = \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} 2^{2j} a_{nj}^2,$$

so that

$$|||f||| = \inf_{f_j \in V_j, \sum f_j = f} \|f_j\|_{V_j}^2 = \inf_{\sum a_{nj} = a_n} \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} 2^{2j} a_{nj}^2 = \sum_{n=1}^{\infty} 2^{2j_n} a_n^2, \tag{2.12}$$

where j_n is the smallest positive integer such that $2^{j_n} > n$. Clearly,

$$n \leq 2^{j_n} \leq 2n,$$

so that the lower bound in (2.5) holds with $c_3 = 1$ and the upper bound holds with $c_4 = 2^2$.

We now continue our discussion of the abstract theory of stable splittings. To describe multilevel algorithms we further introduce V_j-elliptic, symmetric, bilinear forms

$$b_j : V_j \times V_j \longrightarrow \mathbb{R}$$

in the spaces V_j, respectively. These bilinear forms give us the flexibility to describe a wider class of multilevel algorithms within our framework. The particular choice of b_j will lead to different multilevel algorithms. In a first reading one may think of $b_j(\cdot,\cdot) \equiv (\cdot,\cdot)_{V_j}$. Generally, for a properly working multilevel algorithm, we require that the b_j are equivalent to the respective inner product of the subspace, that is, that there exist constants $0 < c_5 \le c_6 < \infty$ such that

$$c_5(v_j, v_j)_{V_j} \le b_j(v_j, v_j) \le c_6(v_j, v_j)_{V_j}, \tag{2.13}$$

for all $v_j \in V_j$, $j \in J$.

Based on these V_j-elliptic bilinear forms, the components of a multilevel algorithm can be defined. In our setup, multilevel algorithms are described in terms of *subspace corrections*, mapping the full space V into each of the subspaces V_j.

DEFINITION 2.1.3. *The mappings $P_{V_j} : V \longrightarrow V_j$ are defined by the variational problem*

$$b_j(P_{V_j}u, v_j) = a(u, v_j), \tag{2.14}$$

for all $v_j \in V_j$, $j \in J$. Analogously, we define $\phi_j \in V_j$ by

$$b_j(\phi_j, v_j) = \Phi(v_j), \tag{2.15}$$

for all $v_j \in V_j$, $j \in J$.

With the subspace corrections we can define the *additive Schwarz operator* as follows.

DEFINITION 2.1.4. *The* **additive Schwarz operator** (*also called BPX operator*) $P_V : V \longrightarrow V$ *with respect to the multilevel structure on V (that is, $a(\cdot,\cdot)$, $\{V_j\}_{j\in J}$, and $b_j(\cdot,\cdot)$) is defined by*

$$P_V = \sum_{j\in J} P_{V_j}. \tag{2.16}$$

Analogously, $\phi \in V$ is defined by

$$\phi = \sum_{j\in J} \phi_j. \tag{2.17}$$

REMARK 2.1.4. *The operator P_V provides the basis for the so-called* additive Schwarz method. *In Theorem* 2.1.1 *below we will also show that P_V can be used to build problems equivalent to the discrete variational problem (see equation* (2.18)*), but with much better conditioning, so that they can be solved efficiently by iterative techniques.*

REMARK 2.1.5. *With a suitably defined bilinear form b_j, it is possible to evaluate P_V efficiently based on its definition as a sum. The explicit construction of P_V, which would be inefficient, is not required.*

REMARK 2.1.6. *Many iterative algorithms, including Jacobi iteration, block-Jacobi, domain decomposition, and variational multigrid methods, can be described in the abstract framework given in Definitions* 2.1.1–2.1.4. *Most of these methods, however, do not generate a stable splitting of the infinite-dimensional function space. The hierarchical structure in the subspace system seems to be essential for obtaining a stable splitting. Otherwise, the complexity of the original problem would have to be captured in the bilinear forms $b_j(\cdot,\cdot)$, and then the evaluation of the P_{V_j} would be as expensive as the solution of the original problem itself (see also the examples above).*

We conclude this abstract discussion by stating and proving two theorems that show the relationship between the concept of a stable splitting and the properties of the additive Schwarz operator. Based on Definitions 2.1.1–2.1.4, the following theorem holds.

THEOREM 2.1.1. *Assume that the subspaces V_j, $j \in J$, of a Hilbert space V are a stable splitting. Assume further that P_{V_j} and P_V are defined as in Definitions* 2.1.3 *and* 2.1.4 *with bilinear forms b_j satisfying* (2.13). *The variational problem* (2.2) *is equivalent to the operator equation*

$$P_V u = \phi, \tag{2.18}$$

and the spectrum of P_V can be estimated by

$$\frac{c_1}{c_4 c_6} \le \lambda_{\min}(P_V) \le \lambda_{\max}(P_V) \le \frac{c_2}{c_3 c_5}. \tag{2.19}$$

Proof. The unique solvability of (2.18) is a consequence of the positive definiteness asserted in (2.19) that we will prove below. The solution of (2.18) coincides with the solution of (2.2) by the definition of P_V and ϕ; see Definition 2.1.4 and equations (2.16), (2.17), respectively.

We now establish the lower bound of the spectrum asserted in (2.19). Let $u_j \in V_j$, $j \in J$, be an arbitrary decomposition of $u \in V$, that is, $\sum_{j\in J} u_j = u$. Then

$$\begin{aligned} a(u,u) &= \sum_{j\in J} a(u_j,u) = \sum_{j\in J} b_j(P_{V_j}u, u_j) \\ &\le \left(\sum_{j\in J} b_j(P_{V_j}u, P_{V_j}u)\right)^{1/2} \left(\sum_{j\in J} b_j(u_j,u_j)\right)^{1/2} \\ &\le \left(\sum_{j\in J} a(P_{V_j}u, u)\right)^{1/2} \left(\sum_{j\in J} c_6 (u_j,u_j)_{V_j}\right)^{1/2} \end{aligned}$$

Taking the infimum of all decompositions of the form

$$\sum_{j\in J} u_j = u$$

we get

$$\begin{aligned} a(u,u) &\le a(P_V u,u)^{1/2} c_6^{1/2} |||u||| \\ &\le a(P_V u,u)^{1/2} (c_6 c_4)^{1/2} \|u\|_V \\ &\le a(P_V u,u)^{1/2} (c_6 c_4/c_1)^{1/2} a(u,u)^{1/2}. \end{aligned}$$

Therefore,

$$a(P_V u,u) \ge \frac{c_1}{c_4 c_6} a(u,u). \tag{2.20}$$

This establishes the lower bound in (2.19). Thus P_V is invertible so that we can define a uniquely determined $z \in V$ that satisfies $P_V z = u$. Hence,

$$\begin{aligned} a(P_V^{-1}u,u) = a(z,P_V z) &= \sum_{j\in J} a(z, P_{V_j} z) \\ &= \sum_{j\in J} b_j(P_{V_j} z, P_{V_j} z) \\ &\ge c_5 \sum_{j\in J} (P_{V_j} z, P_{V_j} z)_{V_j} \\ &\ge c_5 |||P_V z|||^2 \\ &\ge c_3 c_5 \|P_V z\|_V^2 \\ &\ge \frac{c_3 c_5}{c_2} a(u,u). \end{aligned}$$

We conclude that

$$a(v, P_V v) \le \frac{c_2}{c_3 c_5} a(v,v)$$

for all $v \in V$. This yields the upper bound on the spectrum asserted in (2.19).

REMARK 2.1.7. *Theorem* 2.1.1 *shows that the additive Schwarz method generates a well-conditioned operator if the splitting of the space is stable.*

REMARK 2.1.8. *Results related to Theorem* 2.1.1 *have been shown by other authors, in many cases restricted to special cases like* $V = \mathring{H}^1(\Omega)$. *The interested reader is referred to, e.g., Yserentant* [111], *Bramble, Pasciak, and Xu* [25], *Dryja and Widlund* [38], *and Zhang* [116]. *Our presentation of Theorem* 2.1.1 *has followed Oswald* [75].

Computationally, applying the additive Schwarz operator P_V amounts to transferring the residual to all subspaces V_j, applying the inverse of the operator defined by b_j in each subspace, and finally collecting the interpolated results back in V. In the language of multigrid methods, the transfer to V_j is a *restriction.* The b_j implicitly define which kind of smoothing process is used. In the simplest case, the restricted residual is only scaled appropriately, corresponding to a *Richardson smoother.*

From the perspective of classical multigrid, it may be surprising that the *additive* operator has a uniformly bounded condition number, meaning that effective solvers with multigrid efficiency can be obtained by applying steepest

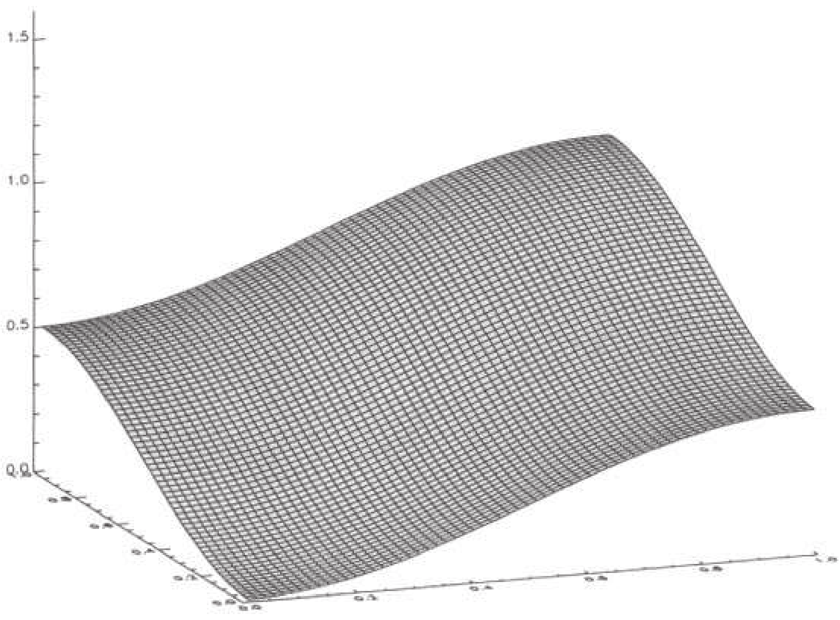

FIG. 2.1. *Idealized spectrum of the discrete Laplace operator.*

descent or conjugate gradient iteration to this system. Classical multigrid is formulated as a *sequential* algorithm where the levels are visited one after the other. The multilevel additive Schwarz method is the corresponding *simultaneous* variant, where the work on the levels is done in parallel (however, see Remark 2.1.9). The sequential treatment of the levels is not necessary for obtaining optimal orders of efficiency.

To illustrate the method, we now visualize the effect of the additive Schwarz process. We assume that V describes the solution space of the discretized two-dimensional Laplacian. In Fig. 2.1 we visualize the spectrum of the discretized Laplace operator. The figure shows the eigenvalues, where the x- and y-axes in the plot represent the frequencies with respect to the two original spatial directions. Each eigenvalue is represented by a point on the graph of the function

$$f(x,y) = \tfrac{1}{2} - (\cos(\pi x) + \cos(\pi y))/4 \tag{2.21}$$

for $h \le x, y \le 1$, where h is the mesh size. The x- and y-coordinates specify the frequency of the Fourier mode relative to the mesh size. Here, we have scaled the discrete Laplacian such that the extreme eigenvalues occur in the northeast corner (near $(1,1)$) and the southwest corner of the frequency domain, where the values are $O(1)$ and $O(h^2)$, respectively. Thus, for this example problem, the condition number grows with $O(1/h^2)$. Consequently, iterative methods, like steepest descent, need a number of cycles to converge, that is, proportional to $O(h^{-2})$.

The additive Schwarz method additionally transfers the residual to coarser levels with mesh sizes $2h, 4h, 8h, \ldots$. This transfer is idealized by restricting the full spectrum represented by $(h,1) \times (h,1)$ to the squares $(h, 0.5) \times (h, 0.5)$,

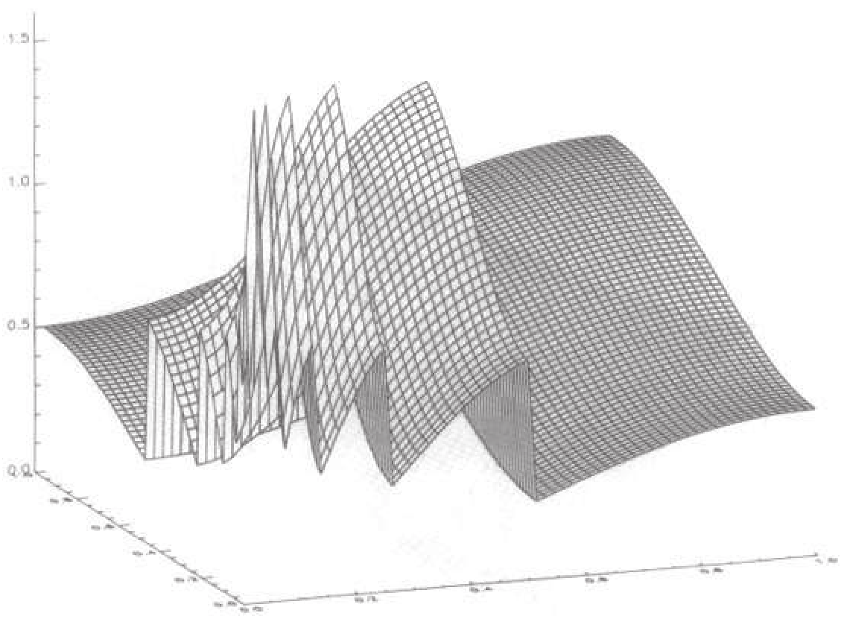

FIG. 2.2. *Idealized spectrum of the additive Schwarz operator associated with the Laplacian.*

$(h, 0.25) \times (h, 0.25), \ldots$. On the coarser levels, the result is rescaled, such that the maximum value in the restricted spectrum is again 1. The results of all levels are finally extended to the full spectrum and are added together. The result of this process is displayed in the function plot of Fig. 2.2.

The multilevel sum of operators, whose eigenvalues are represented in Fig. 2.2, seems to have a minimal eigenvalue bounded away from 0 and a maximal eigenvalue not much larger than the Laplacian itself. The plot suggests that the minimum and maximum value of the combined spectrum are bounded independently of the number of levels.

The method, as it is discussed in this idealized setting, is impractical because the transfer between levels by exact cut-off functions in Fourier space cannot be implemented efficiently. The usual transfer operations are only approximate cut-off functions in Fourier space. This would have to be taken into account in a rigorous analysis.

REMARK 2.1.9. *For computing the additive Schwarz operator P_V applied to a function one may think of evaluating the terms $P_{V_j}u$ in the sum in parallel. This can be exploited most efficiently if the spaces are non-nested, e.g., when $\{V_j\}_{j \in J}$ arise from a domain decomposition. Classical domain decomposition with many subdomains and without coarse mesh spaces, however, does* **not** *cause stable splittings. These usually depend in an essential way on a hierarchical structure, often a nesting of the spaces. Unfortunately, a straightforward parallelization of the sum for nested spaces is often inefficient, since the terms in the sum naturally depend on each other. More precisely, in the case of global meshes, the usual way to compute $P_{V_i}u$ automatically* *generates $P_{V_j}u$ for all $j > i$. Thus, an optimized implementation treating the levels sequentially may be equally fast, but will need only one processor. The multilevel additive Schwarz method must therefore be parallelized using techniques such as those used in the parallelization of multigrid methods (see McBryan et al. [55]). These approaches are usually based on a domain decomposition, and a common problem, then, is that processors tend to go idle on coarse grids, leading to reduced parallel efficiency.*

This argument assumes that we use simple residual corrections *on each level and that the hierarchical structure is induced by* global *mesh refinement. Treating levels in parallel may be attractive, if the process on each level (the smoothing) is computationally significantly more expensive than the restriction operators, or when the mesh structure is highly nonuniform. One such case is the* asynchronous fast adaptive composite grid *(AFAC) method, introduced by McCormick* [57].

To illustrate the relationship between the additive Schwarz norm of Definition 2.1.1 and the additive Schwarz operator of Definition 2.1.4, we now study the special case, when the bilinear forms $a(\cdot,\cdot)$ in V and $b_j(\cdot,\cdot)$ in V_j coincide with the natural bilinear forms on the respective spaces. This is analyzed in the following theorem.

THEOREM 2.1.2. *If*

$$b_j(\cdot,\cdot) \equiv (\cdot,\cdot)_{V_j} \tag{2.22}$$

for all $j \in J$ and

$$a(\cdot,\cdot) \equiv \langle\cdot,\cdot\rangle_V, \tag{2.23}$$

then

$$|||u|||^2 = \langle P_V^{-1}u, u\rangle_V \tag{2.24}$$

for all $u \in V$.

Proof. With (2.22) and (2.23) the definition of P_{V_j} reads

$$(P_{V_j}u, v_j)_{V_j} = \langle u, v_j\rangle_V \tag{2.25}$$

for all $u \in V$, $v_j \in V_j$, and $j \in J$. Let z be defined by $P_V z = u$. This is possible because P_V is positive definite according to Theorem 2.1.1. We have

$$\begin{aligned}\langle P_V^{-1}u, u\rangle_V &= \langle z, P_V z\rangle_V \\ &= \sum_{j\in J}\langle z, P_{V_j}z\rangle_V \\ &= \sum_{j\in J}(P_{V_j}z, P_{V_j}z)_{V_j} \\ &\geq |||u|||^2 = \inf\left\{\sum_{j\in J}\|u_j\|_{V_j}^2 \;\middle|\; u_j \in V_j,\ \sum_{j\in J}u_j = u\right\}\end{aligned}$$

because

$$u = P_V z = \sum_{j\in J}P_{V_j}z.$$

Finally, we show that this particular splitting attains the infimum. We do this by choosing an arbitrary splitting $v_j \in V_j$, $j \in J$, such that $\sum_{j\in J} v_j = u$, and showing that it yields a larger or equal sum of norms.

$$\begin{aligned}
\sum_{j\in J}(v_j, v_j)_{V_j} &= \sum_{j\in J}\Big((P_{V_j}z, P_{V_j}z)_{V_j} + 2(v_j - P_{V_j}z, P_{V_j}z)_{V_j} \\
&\qquad\qquad + (v_j - P_{V_j}z, v_j - P_{V_j}z)_{V_j}\Big) \\
&= \sum_{j\in J}\Big((P_{V_j}z, P_{V_j}z)_{V_j} + 2\langle v_j - P_{V_j}z, z\rangle_V \\
&\qquad\qquad + (v_j - P_{V_j}z, v_j - P_{V_j}z)_{V_j}\Big) \\
&= 2\langle\underbrace{(u - P_V z)}_{=0}, z\rangle_V \\
&\qquad + \sum_{j\in J}\left((P_{V_j}z, P_{V_j}z)_{V_j} + \underbrace{(v_j - P_{V_j}z, v_j - P_{V_j}z)_{V_j}}_{\geq 0}\right) \\
&\geq \sum_{j\in J}(P_{V_j}z, P_{V_j}z)_{V_j}.
\end{aligned}$$

This concludes the proof.

REMARK 2.1.10. *If the bilinear forms that define P_V coincide with the natural ones in V_j and V, respectively, then Theorem 2.1.2 shows that P_V^{-1} defines the bilinear form associated with the norm $|||\cdot|||$.*

2.2. Finite element approximation spaces in two dimensions

In this section we will apply the concept of stable splittings in the context of finite element spaces. Typically, the full space $V = H^1(\Omega)$ will correspond to the function space associated with the partial differential equation, and the $\{V_j\}_{j\in\mathbb{N}_0}$ will be an infinite collection of subspaces generated by successively refining a finite element approximation of the differential equation. To apply the results of §2.1 we will consider splittings of V in a *nested* sequence of spaces

$$V_0 \subset V_1 \subset V_2 \subset \cdots \subset V = H_0^1(\Omega) \tag{2.26}$$

generated by a regular family of nested triangulations

$$\mathcal{T}_0 \subset \mathcal{T}_1 \subset \mathcal{T}_2 \subset \cdots$$

of a bounded domain $\Omega \subset \mathbb{R}^2$. The proper structure for such finite element partitions is discussed in Chapter 4. For a detailed presentation of the properties required for finite element partitions, see also Ciarlet [33]. We assume that continuous, piecewise linear elements are used, corresponding to the second order model problems of §1.3. For generalizations to more complicated situations, see the references listed at the beginning of this chapter.