Remarks on the abstract theory of additive and multiplicative Schwarz algorithms

M. Griebel
Institut für Informatik
Technische Universität München
D-W-8000 München 2 , FRG

P. Oswald
Institut für Angewandte Mathematik
Friedrich-Schiller-Universität Jena
D-O-6900 Jena , FRG

March 1993

Abstract. We propose a modified approach to the abstract convergence theory of the additive and multiplicative Schwarz iterative methods associated with a subspace splitting for solving symmetric positive definit linear problems. In particular, we obtain a satisfactory bound for the convergence rate of the appropriately scaled multiplicative Schwarz method directly in terms of the condition number of the corresponding additive Schwarz operator. The results may be viewed as an appendix to the recent surveys [X, Y].

Key words: Symmetric positive definit linear systems, iterative methods, additive and multiplicative Schwarz methods.

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

In recent years, many efforts have been spent on the design of efficient iterative solvers for large sparse linear systems arising from the discretization of symmetric elliptic boundary value problems. It has turned out that many of the best methods arising in different context (multigrid schemes, multilevel preconditioners, domain decomposition algorithms etc.) can be analyzed from a unified viewpoint if they are interpreted as additive or multiplicative Schwarz methods. We refer to the excellent survey papers by Xu [X] and Yserentant [Y] where also some historical information can be found (cf. in addition [W]).

It is the aim of this short note to give a slightly different derivation of the abstract Schwarz theory, and to analyze the assumptions necessary to get sharp condition number estimates for the additive Schwarz operator on the one hand, and bounds for the convergence rate of the multiplicative Schwarz algorithm on the other. Our technical device is an abstract version of the generating (or indefinite) system used by one of the authors [G] to analyze multilevel Schwarz algorithms.

Let us give the details. We use the following Hilbert space setting [DW, W], the matrix resp. operator notation used in [X, Y] can easily be derived from it. Let V be some fixed Hilbert space, to avoid any convergence discussions we assume that V is finite-dimensional. The scalar product in V is denoted by $(\cdot,\cdot)$. We consider a positive definit, symmetric bilinear form $a(u,v) = (Au,v)\,,\ u,v \in V$, with $A : V \to V$ denoting the corresponding s.p.d. operator acting on V. The spectral condition number of any such operator is defined by

$$\mathrm{cond}\,(A) = \frac{\lambda_{\min}(A)}{\lambda_{\max}(A)} \quad .$$

We look for solution procedures for a variational problem associated with $a(\cdot,\cdot)$:

Find $u \in V$ such that

$$(\mathbf{VP}) \qquad a(u,v) = \Phi(v) \quad , \qquad \forall\, v \in V \ .$$

Here $\Phi(v) = (f,v)$ is an arbitrary linear functional on V ($f \in V$ is given). In operator form, **(VP)** is equivalent to $Au = f$. Due to the Lax-Milgram theory, **(VP)** possesses a unique solution.

The iterative solvers for this problem we are looking for are based on the decomposition of V into a finite linear (not necessarily direct) sum of subspaces $V_j \subset V$:

$$(\mathbf{D}) \qquad V = \sum_{j=0}^{J} V_j \quad .$$

On the subspaces we choose auxiliary s.p.d. forms $b_j(u_j,v_j) = (B_j u_j, v_j)$ given by the s.p.d. operators $B_j : V_j \to V_j$. These forms model the approximative solvers used in

the subspaces, i.e. B_j^{-1} is an approximative inverse for the restriction A_j of A to V_j. Furthermore, let $T_j : V \to V_j$ be given by the variational problems

$$b_j(T_j u, v_j) = a(u, v_j) \quad , \qquad \forall\, v_j \ ,\ u \in V \ ,\ j = 0, \ldots, J$$

(in operator notation $T_j = B_j^{-1} Q_j A$ where $Q_j : V \to V_j$ denotes the orthoprojection onto V_j with respect to the scalar product). Finally, let $\phi_j \in V_j$ be the solution of the problems

$$b_j(\phi_j, v_j) = \Phi(v_j) \quad , \qquad \forall\, v_j \ ,\ j = 0, \ldots, J \ .$$

Now we are able to introduce the *additive Schwarz formulation* **(AS)** of the variational problem **(VP)** associated with the splitting **(D)**, and to state the basic result on the spectrum of the additive Schwarz operator

Theorem A (Additive Schwarz preconditioner). *The solution of* **(VP)** *coincides with that of the operator equation*

$$\textbf{(AS)} \qquad P_{\sum V_j} u = \phi \ , \qquad \left(P_{\sum V_j} = \sum_{j=0}^{J} T_j \ , \quad \phi = \sum_{j=0}^{J} \phi_j \right) \ .$$

The additive Schwarz operator $P_{\sum V_j}$ is symmetric (with respect to $a(\cdot,\cdot)$) and positive definit. Moreover,

$$\lambda_{\min}(P_{\sum V_j}) = \min_{u \in V} \frac{a(u,u)}{|||u|||^2} \quad , \qquad \lambda_{\max}(P_{\sum V_j}) = \max_{u \in V} \frac{a(u,u)}{|||u|||^2} \quad ,$$

with the new norm $|||u|||$ on V defined by

$$|||u|||^2 = \inf_{u_j \in V_j \,:\, u = \sum_{j=0}^{J} u_j} \sum_{j=0}^{J} b_j(u_j, u_j) \quad .$$

This theorem is very simple but useful : it says under which necessary and sufficient conditions on the splitting **(D)** and the choice for the forms b_j the new problem **(AS)** is well-conditioned. The idea is then to solve **(AS)** by Richardson or cg-iterations which should lead to fast convergence. Since

$$P_{\sum V_j} = \left(\sum_{j=0}^{J} B_j^{-1} Q_j \right) A \equiv C_{\sum V_j} A \quad ,$$

the procedure can be viewed as preconditioning strategy with preconditioner $C_{\sum V_j}$ for the original problem **(VP)**.

Theorem A has many authors, and it is not our intention to find out all contributors (as with any simple thing, people go around for many years, compare [MN, BM, W, Z]). Its proof can be given in a few lines using the explicit formulae

$$a(P^{-1}_{\sum V_j} u, u) = \|\!|u|\!\|^2 \quad , \qquad u \in V \ .$$

(see, e.g. [W, X, GO]). It can also be obtained from the fictitious domain lemma (see [N1, N2]) which has already been known in the Russian literature for some years, and which is useful on its own for different purposes.

Fictitious Domain Lemma. *Let H_0 and H be two Hilbert spaces, with the scalar products denoted by $(\cdot,\cdot)_0$ resp. $(\cdot,\cdot)$, and with bilinear forms $a : H_0 \times H_0 \to \mathbf{R}$ resp. $b : H \times H \to \mathbf{R}$. generated by the s.p.d. operators $A : H_0 \to H_0$ resp. $B : H \to H$. Suppose that there is a surjective linear operator $R : H \to H_0$ such that for all $u_0 \in H_0$ there exists an $v \in H$ with $Rv = u_0$ and*

$$c_T \cdot b(v,v) \le a(u_0,u_0) \quad ,$$

and, on the other hand,

$$a(Ru,Ru) \le c_R \cdot b(u,u) \qquad \forall\, u \in H \quad .$$

Introduce the adjoint operator $R^* : H_0 \to H$ by $(Ru,u_0)_0 = (u,R^*u_0)$ for all $u \in H$ and $u_0 \in H_0$. Then

$$c_T \cdot a(u_0,u_0) \le a(RB^{-1}R^*Au_0,u_0) \le c_R \cdot a(u_0,u_0) \ , \quad \forall\, u_0 \in H_0 \ ,$$

with sharp bounds for the spectrum of $RB^{-1}R^*A$ given by the best possible constants c_T and c_R in the above inequalities.

Proof. (see [N2]). For any $u_0 \in H_0$ we have

$$\begin{aligned} a(RB^{-1}R^*Au_0,u_0) &\le a(RB^{-1}R^*Au_0, RB^{-1}R^*Au_0)^{1/2} \cdot a(u_0,u_0)^{1/2} \\ &\le c_R^{1/2} \cdot b(B^{-1}R^*Au_0, B^{-1}R^*Au_0)^{1/2} \cdot a(u_0,u_0)^{1/2} \\ &= c_R^{1/2} \cdot a(RB^{-1}R^*Au_0,u_0)^{1/2} \cdot a(u_0,u_0)^{1/2} \ , \end{aligned}$$

which gives the upper estimate. On the other hand,

$$\begin{aligned} a(u_0,u_0) &= (Rv,Au_0)_0 = b(v,B^{-1}R^*Au_0) \\ &\le b(v,v)^{1/2} \cdot b(B^{-1}R^*Au_0, B^{-1}R^*Au_0)^{1/2} \\ &\le c_T^{-1/2} \cdot a(u_0,u_0)^{1/2} \cdot a(B^{-1}R^*Au_0,u_0)^{1/2} \end{aligned}$$

yields the lower bound. The sharpness of the bounds is obvious.

Alternatively, if one looks carefully at the conditions and at the proof, one easily obtains the identity

$$a((RB^{-1}R^*A)^{-1}u_0, u_0) = \inf_{v\in H\,:\,Rv=u_0} b(v,v)\,, \quad \forall\, u_0 \in H_0\,.$$

which also implies the above assertions. Indeed, denoting $X = (RB^{-1}R^*A)^{-1}$ and taking an arbitrary $v \in H$ with $Rv = u_0$, we get

$$a(Xu_0, u_0) = (AXu_0, Rv) = b(\underbrace{B^{-1}R^*AXu_0}_{w}, v) \le b(w,w)^{1/2}\cdot b(v,v)^{1/2}$$

with equality for $v = w$ (check that $Rw = u_0$ by definition of X). It remains to take the infimum with respect to all admissible v, and to observe that

$$b(w,w) = (B^{-1}R^*AXu_0, R^*AXu_0) = (X^{-1}Xu_0, AXu_0)_0 = a(Xu_0, u_0)\,.$$

As in Theorem A, one could write down explicit formulae for the minimal and maximal eigenvalues of $RB^{-1}R^*A$.

To see that Theorem A is a consequence of this lemma, put there

$$H_0 = V\,, \quad H = V_0 \times \cdots \times V_J\,, \quad b(\bar u, \bar v) = \sum_{j=0}^{J} b_j(u_j, v_j)\,, \quad R\bar u = \sum_{j=0}^{J} u_j$$

where $\bar u \equiv (u_0, \ldots, u_J)$ denotes an arbitrary element of $V_0 \times \cdots \times V_J$. The scalar product on H_0 is that of V while on H the usual definition for cartesian products of Hilbert spaces is adopted. Now it takes a little bit linear algebra to see that this actually gives Theorem A in all details. E.g., the adjoint $R^* : V \to V_0 \times \ldots \times V_J$ has the components Q_j, and the final preconditioner $RB^{-1}R^*$ coincides with $C_{\sum V_j}$.

The idea behind this proof of Theorem A, the introduction of the auxiliary Hilbert space $H = V_0 \times \cdots \times V_J$, is the theoretical counterpart to the introduction of the generating system in [G] which was successfully used to understand various properties of multilevel Schwarz methods. Though not a great inventure in the present context, it helps to analyze the assumptions for the convergence theory in the multiplicative case, too.

The abstract convergence theory for the multiplicative Schwarz algorithm associated with the splitting **(D)** and the auxiliary forms $b_j(\cdot,\cdot)$ has also a long history (see [X, Y]), a last substantial step has been taken in [BPWX1, BPWX2, BP]. The abstract multiplicative algorithm we have in mind is Algorithm 3.5. (= Algorithm 3.3 if $\omega = 1$) of [X]. There are other versions the theory of which is or similar or can be deduced from that we will deal with. The iteration step consists in a successive subspace correction (instead of the simultaneous one in the additive scheme):

$$u^{k+(j+1)/(J+1)} = u^{k+j/(J+1)} + \omega\cdot(\phi_j - T_j u^{k+j/(J+1)})\,, \quad j = 0, \ldots, J\,;\; k = 0, 1, \ldots\,.$$

The relaxation parameter ω introduces an appropriate scaling of the subproblems. The starting value u^0 is arbitrary. Switching to the operator representations, we get

$$e^{k+1} \equiv u - u^{k+1} = E_{\sum V_j,\omega} e^k \ , \quad k \geq 0$$

where

$$E_{\sum V_j,\omega} = (I - \omega T_J)(I - \omega T_{J-1}) \ldots (I - \omega T_0)$$

for the error reduction in this algorithm.

To formulate the convergence result, let us introduce the following notation : Set $\|u\|_E^2 = a(u,u)$ for the energy norm on V associated with $a(\cdot,\cdot)$ (the same notation will be used for the corresponding operator norm), and

$$\tilde{K}_0 = (\lambda_{\min}(P_{\sum V_j}))^{-1} \ , \quad \tilde{K}_1 = \lambda_{\max}(P_{\sum V_j}) \ , \quad \omega_1 = \max_{0 \neq u_j \in V_j \, , \, j=0,\ldots J} \frac{a(u_j,u_j)}{b_j(u_j,u_j)} \ .$$

As we will see below, these constants are close to the corresponding optimal constants K_0, K_1 used in [X] (Fundamental Theorem II). Techniques which lead to estimates for K_0, K_1 (resp.$\tilde{K}_0, \tilde{K}_1$) and ω_1 are discussed in [X, Y]. E.g., many authors [X, Y, W] emphasize the use of so-called strengthened Cauchy-Schwarz inequalities for estimating K_1. However, not the way how to bound K_0, K_1, ω_1 but the above constants themselves are crucial for the behavior of the additive and multiplicative Schwarz methods. The next theorem is a slight modification of results from [X] (in this form, the relationship with the theory of the corresponding additive method is made more explicit).

Theorem M (Convergence of the multiplicative Schwarz method). *For the subspace splitting given above, the multiplicative method converges for the values*

$$0 < \omega < \frac{2}{\omega_1}$$

of the relaxation parameter, with an upper estimate of the convergence rate given by

$$\|E_{\sum V_j,\omega}\|_E^2 \leq 1 - \frac{(2 - \omega_1 \omega)\omega}{K_0 (1 + \omega K_1)^2} \quad .$$

The constants K_0 and K_1 satisfy

$$K_0 = \tilde{K}_0 \quad , \qquad K_1 \leq ([\log_3(2J-3)] + 1) \cdot \tilde{K}_1 \ .$$

In particular, for $\omega = \omega^* \equiv 1/(K_1 + \omega_1)$ one has

$$\|E_{\sum V_j,\omega^*}\|_E^2 \leq 1 - \frac{1}{(2 \cdot K_1 + \omega_1) \cdot K_0} \leq 1 - \frac{1}{(2 \cdot ([\log_3(2J-3)] + 3) \cdot \mathrm{cond}(P_{\sum V_j})} \quad .$$

The latter corollary is of pure theoretical importance: It shows that for *any* variational problem and *any* additive subspace splitting (inclusively the arbitrary choice of auxiliary forms $b_j(\cdot,\cdot)$), there exists a proper scaling of the multiplicative Schwarz algorithm such that the convergence rate depends only on the condition number of the additive Schwarz operator, and logarithmically on the number of subspaces. *Good preconditioning via the additive Schwarz formulation implies relatively fast convergence of the multiplicative Schwarz algorithm* ! Whether the logarithmical factor can be removed is not clear at present, see the proof given below. Note that, as a rule, the convergence rate of the multiplicative algorithm with *optimal* scaling factor ω is better than indicated by the above estimates. This, in particular, makes it necessary to develop much more delicate techniques for the multiplicative case. One approach is the use of strengthened Cauchy-Schwarz inequalities, and the selection of suitable auxiliary subspace decompositions

$$V = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \ldots \oplus \mathcal{V}_J \ , \quad \mathcal{V}_j \subset V_j \ , \ j = 1, \ldots, J$$

which may lead to improved estimates of K_1 (not of K_0). For details, we refer to [Y, X]. In this connection, let us state the estimate

$$\|E_{\sum V_j,\omega}\|_E^2 \leq 1 - \frac{(2-\omega_1\omega)\omega}{\hat{K}_0(1+\omega\sqrt{\omega_1 J \hat{K}_1})^2} \quad .$$

of the convergence rate of the multiplicative method given in [Y], Theorem 5.4., where $1/\hat{K}_0 \leq 1/\tilde{K}_0$ and $\hat{K}_1 \leq \tilde{K}_1$ are the minimal and maximal eigenvalues of the additive Schwarz operator corresponding to the auxiliary splitting $V = \bigoplus \mathcal{V}_j$ (the forms on $\mathcal{V}_j$ are inherited from V_j). Compared with our above formulation, we see that the dependence on J is much worse.

The effect of individual scaling (this is replacing ω by ω_j , $j = 0, 1, \ldots, J$) for the convergence rate in the above multiplicative algorithm resp. for improving the spectral bounds for the additive Schwarz operator (this is replacing $b_j(\cdot,\cdot)$ by $b_j(\cdot,\cdot)/\omega_j$) has not yet been studied in a satisfactory way. Note that constant scaling ($\omega_j = \omega$) does not change the condition of the additive operator though it may substantially influence the behavior of the multiplicative algorithm.

Proof of Theorem M. We reduce it to Theorem 4.4 of [X]. The assumptions of Xu's result are (in our notation)

(i) There exists a constant K_0 such that for any $u \in V$ there is a decomposition $u = \sum_{j=0}^{J} u_j$, $u_j \in V_j$, $j = 0, \ldots, J$ such that

$$\sum_{j=0}^{J} b_j(u_j, u_j) \leq K_0 \cdot a(u,u) \quad .$$

(ii) For any choice $u_j, v_j \in V_j$, $j = 1, \ldots, J$, and any index set $S \subset S_J \equiv \{1, \ldots, J\} \times \{1, \ldots, J\}$ one has

$$\sum_{(i,j)\in S} a(u_i, v_j) \le K_1 \cdot \left(\sum_{j=0}^{J} b_j(u_j, u_j) \right)^{1/2} \cdot \left(\sum_{j=0}^{J} b_j(v_j, v_j) \right)^{1/2}$$

(a close look at the proof in [X] shows that the *only* index set of interest is the 'triangle' $S_{\Delta,J} = \{(i,j) \in S_J \,|\, j < i\}$, the corresponding constant $K_1 \equiv K_{1;J,\Delta}$ will be estimated below by $\tilde{K}_1$).

(iii) There is a $w \in (0,2)$ such that $a(u_j, u_j) \le w \cdot b_j(u_j, u_j)$ for all $u_j \in V_j$, $j = 0, \ldots, J$.

Then, by Theorem 4.4. from [X] the estimate

$$\|E_{\sum V_j,1}\|_E^2 \le 1 - \frac{(2-w)}{K_0(1+K_1)^2}$$

holds true.

We see by Theorem A that (i) is valid with $K_0 = \tilde{K}_0$ while (ii) is true in the case $u_j = v_j$ and for the index set S_J with $K_1 = \tilde{K}_1$. But this particular case implies (ii) for arbitrarily chosen $u_j, v_j \in V_j$ and all 'rectangles' $S = I_1 \times I_2$, $I_1, I_2 \subset \{1, \ldots, J\}$ with the same constant $K_1 = \tilde{K}_1$ in general. To this end, recall the well-known fact that $l(u,u) \le K_1 \cdot \|u\|^2$, $\forall\, u \in H$, implies $l(u,v) \le K_1 \cdot \|u\| \cdot \|v\|$, $\forall\, u, v \in H$ for any symmetric bilinear form $l(\cdot,\cdot)$ on a real Hilbert space H, and apply it to our auxiliary Hilbert space (this time equipped with a 'energy' scalar product)

$$H = V_0 \times \ldots \times V_J \,, \quad \|\bar{u}\| \equiv \|(u_0, \ldots, u_J)\| = \left(\sum_{j=0}^{J} b_j(u_j, u_j) \right)^{1/2}$$

and the bilinear form

$$l(\bar{u}, \bar{v}) = a(\sum_{j=0}^{J} u_j, \sum_{j=0}^{J} v_j) \quad .$$

This gives the above mentioned property for S_J, and hence for all $S = I_1 \times I_2$ (put $u_i = 0$, $i \notin I_1$, $v_j = 0$, $j \notin I_2$).

Now we have to derive the estimate for $S_{\Delta,J}$. The idea is simple: we use induction in J, with the induction step as follows: Decompose the left-hand side of the inequality in (ii) as

$$\sigma_J \equiv \sum_{(i,j)\in S_{J,\Delta}} a(u_i, v_j) = \sum_{i=J-J_1+1}^{J} \sum_{j=1}^{J_1} a(u_i, v_j) + \sum_{(i,j)\in S_{J,\Delta},\, i\le J-J_1} a(u_j, v_j)$$

$$- \sum_{(i,j)\notin S_{J,\Delta}\,,\, i,j\le J_1} a(u_j,v_j) + \sum_{(i,j)\in S_{J,\Delta}\,,\, j>J_1} a(u_j,v_j) \equiv \sigma_{J,R} + \sum_{k=1}^{3} \sigma_{J,k}\ .$$

Here, J_1 is the integer next to $(2J-1)/3$. While $\sigma_{J,R}$ corresponds to a rectangular index set, the other sums correspond to triangular index sets of smaller size. More precisely, one can fix three 'subsquares' $R_k \subset S_J$ of size $\le J_2 = [J/3]+1$ such that the index sets of $\sigma_{J,k}$ are triangular with respect to R_k, cf. Figure 1.

- Figure 1

Let c_j denote the best constant in the inequality in (ii) for a triangular $S_{j,\Delta}$, $2 \le j \le J$. Obviously, the sequence c_j is monotonically increasing, and $c_2 = \tilde{K}_1$. Moreover, by our construction

$$\sum_{k=1}^{3} \sigma_{J,k} \le c_{J_2} \cdot \left\{ \sum_{k=1}^{3} \left(\sum_{i\in I_{1,k}} b_i(u_i,u_i) \right)^{1/2} \cdot \left(\sum_{j\in I_{2,k}} b_j(v_j,v_j) \right)^{1/2} \right\}$$

$$\le c_{J_2} \cdot \left(\sum_{i=2}^{J} b_i(u_i,u_i) \right)^{1/2} \cdot \left(\sum_{j=1}^{J-1} b_j(v_j,v_j) \right)^{1/2}$$

since the index sets $I_{1,k}$ resp. $I_{2,k}$ selected from $\{1,\dots,J\}$ and corresponding to the squares R_k do not intersect. Thus, since $\sigma_{J,R}$ is related to a rectangular set we get

$$c_J \le \tilde{K}_1 + c_{J_2} \quad , \qquad J_2 = [J/3]+1\ .$$

By induction this yields $c_5 = 2\tilde{K}_1$, $c_{14} = 3\tilde{K}_1$, $c_{41} = 4\tilde{K}_1$, and so on. Finally, we get

$$K_1 \equiv K_{1;J,\Delta} = c_J \le ([\log_3(2J-3)]+1) \cdot \tilde{K}_1\ , \qquad J \ge 2\ .$$

Concerning (iii), the only thing we can definitely say is that this condition holds with some $w \le \tilde{K}_1$ which follows from Theorem A if we take $u_j = v_j \in V_j$ to be arbitrary, and set $u_l = v_l = 0$ for the other indices $l \ne j$ $(j = 0,\dots,J)$. Therefore, to fit (iii) we apply scaling, i.e. we replace $b_j(\cdot,\cdot)$ by $b_{j,w}(\cdot,\cdot) = \frac{\omega_1}{w} \cdot b_j(\cdot,\cdot)$ for all j, with $w \in (0,2)$ and ω_1 as defined above. Obviously, (iii) is now fulfilled for the new auxiliary forms while (i) resp. (ii) hold true with new constants $K_{0,w} = \frac{\omega_1}{w} \cdot K_0$ resp. $K_{1,w} = \frac{w}{\omega_1} \cdot K_1$. Since this choice of the forms corresponds to the choice of the relaxation parameter $\omega = \frac{w}{\omega_1} \in (0, \frac{2}{\omega_1})$ we have the first assertion of Theorem M, together with the estimates for K_0 and K_1 already proved. Taking the infimum of the estimate with respect to all indicated ω we get the corollary, too.

Note that the proof of Xu's result which we relied on is also not lengthy. As a whole, we have now a very compact basic convergence theory for the additive and multiplicative

Schwarz methods, with assumptions that are in some sense minimal for the additive variant. There is no additional assumption for the multiplicative case. We hope that the logarithmic factor in Theorem M can be removed to get the optimum in this respect. Our claim is the following assertion to hold true:

There is an absolute constant C^* such that for any Schwarz splitting we have

$$\kappa_{AS} \cdot (1 - \rho_{MS}^2) \geq C^* \quad ,$$

with $\kappa_{AS} = \mathrm{cond}(P_{\sum V_j})$ *and* $\rho_{MS} = \inf_\omega \|E_{\sum V_j, \omega^*}\|_E$.

Simple calculations (take a linear system in $\mathbf{R}^2$ and examine Jacobi versus SOR-Gauss-Seidel methods) show that $C^* \leq 1/2$, and that no upper bound is possible (compare, however, [BPWX1], Theorem 3.1). There are certainly shortcomings that one can not expect to be excluded in a theory of this generality: the estimates in the multiplicative case are still very rough, they do not depend on the order of carrying out the correction steps, and so on.

We conclude with a few remarks.

- According to Theorems A and M, the analysis of Schwarz algorithms for a *given* splitting may be reduced to the proof, with best possible constants, of the norm equivalence

  (NE) $$a(u,u) \approx \|\|u\|\|^2 = \inf_{u_j \in V_j \,:\, u = \sum_{j=0}^J u_j} \sum_{j=0}^J b_j(u_j, u_j) \; .$$

  Another basic question is to choose the splitting and the approximate solvers which correspond to the auxiliary bilinear forms $b_j(\cdot,\cdot)$ on the subspaces such that, on the one hand, a good estimate **(NE)** can be derived and, on the other hand, an overall algorithm with cheap components will be obtained. The requirements of what is cheap strongly depend on the computer environment available and the problem classes. There are the questions of data structures, serial versus vectoriell resp. parallel algorithms, and of adaptivity. Nevertheless, in view of the above theory, it is good to have many potential candidates of splittings for a given problem class, with a precise knowledge on **(NE)** before one starts to design efficient practical solvers in a given environment.

- If one has already found a 'good' splitting one may obtain others by simple procedures like *refinement*, *clustering*, and, under certain restrictions, *selection*. Refinement means that some or all V_j are replaced in turn by their splittings $V_j = \sum_i V_{j,i}$ with possibly other auxiliary forms $b_{j,i}(\cdot,\cdot)$ on these new subspaces. Clustering is the inverse process - some of the subspaces in the splitting are grouped together.

The whole theory of such techniques can be described by the obvious inequalities (cf. **(NE)** and Theorem A)

$$\mathrm{cond}\,(P_{\sum_{j,i} V_{j,i}}) \le \mathrm{cond}\,(P_{\sum_j V_j}) \cdot \max_j\ \mathrm{cond}\,(P_{\sum_i V_{j,i}})$$

$$\mathrm{cond}\,(P_{\sum_j V_j}) \le \mathrm{cond}\,(P_{\sum_{j,i} V_{j,i}}) \cdot \max_j\ \mathrm{cond}\,(P_{\sum_i V_{j,i}})\ .$$

Selection can be used, e.g., for dealing with certain adaptive algorithms. Let $\mathcal{J} \subset \{0,\ldots,J\}$. Starting from a given decomposition **(D)** one may ask for properties of the splitting

$$V_{\mathcal{J}} = \sum_{j\in\mathcal{J}} V_j \subset V$$

for the new (selected) subspace $V_{\mathcal{J}} \subset V$ (the auxiliary bilinear forms will be held fixed). It is obvious by Theorem A that

$$\lambda_{\max}(P_{\sum_{\mathcal{J}} V_j}) \le \lambda_{\max}(P_{\sum V_j})$$

for any index selection $\mathcal{J}$. If **(D)** is a direct sum of subspaces than an analogous but opposite estimate holds for the minimal eigenvalues, too. In general, due to the overlap in the subspaces V_j which may be greatly reduced by the selection but was necessary to keep the condition of $P_{\sum V_j}$ small, this is not true, and one has to carefully run the estimation from below in **(NE)** for each selection set. Though difficulties remain, the selection procedure is better than nothing.

We will not give examples of what is obvious: the above tricks have implicitely be used all the time.

- The use of the fictitious domain lemma opens the way for dealing with outer approximation schemes in the same conceptional way. Thus, it is not really necessary to explicitely split V into *subspaces*. Instead, given any collection of Hilbert spaces V_j equipped with their own scalar products $(\cdot,\cdot)_j$ and auxiliary s.p.d. forms $b_j(u_j,v_j) = (B_j u_j, v_j)_j$ $(u_j, v_j \in V_j)$, it suffices to define the linear mapping

$$R\ :\ V_0 \times \ldots \times V_J \to V \qquad (R\bar{u} = \sum_{j=0}^{J} R_j u_j)$$

such that it is onto V, and to check the constants in the two-sided norm comparison

$$a(u,u) \approx \inf_{v_j\in V_j\,:\,u=\sum_{j=0}^{J} R_j v_j} \sum_{j=0}^{J} b_j(v_j,v_j) \qquad \forall\ u \in V\ .$$

The preconditioner for the additive Schwarz algorithm takes now the form $C_{\sum V_j} = \sum_{j=0}^{J} R_j B_j^{-1} R_j^*$, with $R_j^*\ :\ V \to V_j$ defined by $(R_j u_j, v) = (u_j, R_j^* v)_j \quad \forall u_j \in V_j\ ,\ v \in V,\ j = 0,\ldots,J$.

Formally, the mapping R transfers the external construction $\{V_j, b_j\}$ into an internal *subspace* splitting

$$V = \sum_{j=0}^{J} R_j V_j \equiv \sum_{j=0}^{J} \tilde{V}_j \ , \qquad \tilde{b}_j(\tilde{u}_j, \tilde{v}_j) = b_j(u_j, v_j)$$

where, for any $\tilde{u}_j \in R_j V_j$, we denote by $u_j \in V_j$ the unique element satisfying $R_j u_j = \tilde{u}_j$ and $b_j(u_j, \hat{u}_j)$ for all $\hat{u}_j \in \operatorname{Ker} R_j$. This shows the role of the R_j, and it is clear that their proper choice is crucial for the success. One has to compare also [BPX, D]. We are tempted to look for a systematic approach to Schwarz iterative methods for nonconforming methods via this idea but will do this elsewhere.

References

[BM] P.E.Bjorstad, J.Mandel, On the spectra of orthogonal projections with applications to parallel computing. BIT 31 (1991) 76-88.

[BP] J.H.Bramble, J.E.Pasciak, New estimates for multilevel methods including the V-cycle. Manuscript, 1991.

[BPWX1] J.H.Bramble, J.E.Pasciak, J.Wang, J.Xu, Convergence estimates for product iterative methods with applications to domain decomposition. Math. Comp. 57 (1991), 1-21.

[BPWX2] J.H.Bramble, J.E.Pasciak, J.Wang, J.Xu, Convergence estimates for multigrid algorithms without regularity assumptions. Math. Comp. 57 (1991) 23-45.

[BPX] J.H.Bramble, J.E.Pasciak, J.Xu, The analysis of multigrid algorithms with nonested spaces or noninherited quadratic forms. Math. Comp. 56 (1991), 1-34.

[D] W.Dörfler, Hierarchical bases for elliptic problems. Math.Comp. 58 (1992) 513-529.

[DW] M.Dryja, O.Widlund, Additive Schwarz methods for elliptic finite element problems in three dimensions. Tech.Rep. 570, Courant Inst., New York Univ., June 1991.

[G] M.Griebel, Multilevel algorithms considered as iterative methods on indefinite systems. Report SFB 342/29/91A TU München, October 1991.

[GO] M.Griebel, P.Oswald, On additive Schwarz preconditioners for sparse grid discretizations. Numer.Math. (to appear) - Forsch.-Erg. FSU Jena, Math/92/7.

[MN] A.M.Matsokin, S.V.Nepomnyaschikh, Schwarz alternating method in a subspace. Sov. Math. (Izv. vuz.) 29 (1985) 78-84.

[N1] S.V.Nepomnyaschikh, Decomposition and fictitious domain methods for elliptic boundary value problems. In: Domain Decomposition V SIAM, Philadelphia, 1991, 62-72.

[N2] S.V.Nepomnyaschikh, Fictitious components and subdomain alternating methods. Sov. J. Numer. Anal. Math. Modelling 5, 1 (1990), 53-68.

[W] O.Widlund, Some Schwarz methods for symmetric and nonsymmetric elliptic problems. Tech. Rep. 581, Courant Inst., New York Univ., September 1991.

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

[Y] H.Yserentant, Old an d new convergence proofs for multigrid methods. Acta Numerica 1993 (to appear).

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