An additive Schwarz analysis for multiplicative Schwarz methods: General case

Susanne C. Brenner

1 Introduction

Multiplicative and additive Schwarz methods are two main classes of iterative methods since the times of Gauss and Jacobi. Traditionally the analyses of these two classes of methods follow different paths. On one hand, the theory for additive Schwarz methods [8, 12, 2, 16, 9, 14, 13, 15, 6, 11], like the theory for the classical Jacobi method, is relatively simple. On the other hand, the theory for multiplicative Schwarz methods [10, 3, 16, 19, 18, 9, 4, 1, 17, 13, 15, 11], like the theory for the classical Gauss-Seidel method, can be quite sophisticated.

An analysis of multiplicative Schwarz methods that is based on the additive theory was carried out in [5]. It is restricted to the case where the subspace corrections are based on symmetric positive definite (SPD) solvers. The goal of this work is to extend the results in [5] to multiplicative Schwarz methods with general subspace corrections. As a by-product we recover the main result in [17], namely a formula for the norm of product operators.

The rest of the paper is organized as follows. First we review the Gauss-Seidel method in Section 2. The analysis of this prototypical multiplicative Schwarz method provides motivations and guidance for the theory in this paper and [5]. We introduce a general framework of multiplicative Schwarz methods in Section 3 and recall the fundamental lemma for additive Schwarz theory in Section 4. The key observation that allows the extension of the formulas in Section 2 to general multiplicative Schwarz methods is presented in Section 5. The main results of the paper are then derived in Section 6. Finally, the connection of our theory to [17] is discussed in Section 7.

Susanne C. Brenner
Department of Mathematics and Center for Computation and Technology, Louisiana State University, Baton Rouge, LA 70803, USA, e-mail: brenner@math.lsu.edu

2 The Gauss-Seidel Method

The additive Schwarz analysis for multiplicative Schwarz methods is motivated and guided by looking at the analysis of the Gauss-Seidel method through the lens of additive Schwarz theory.

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a SPD matrix and $\mathbf{b} \in \mathbb{R}^n$. The (forward) Gauss-Seidel method for the system $\mathbf{Ax} = \mathbf{b}$ is defined by the iteration step

$$\mathbf{x}_{\text{new}} = \mathbf{x}_{\text{old}} + (\mathbf{L}+\mathbf{D})^{-\mathbf{1}}(\mathbf{b} - \mathbf{A}\mathbf{x}_{\text{old}}), \tag{1}$$

where $\mathbf{L}$ and $\mathbf{D}$ are the strictly lower triangular part and the diagonal part of $\mathbf{A}$ respectively. The error propagation for (1) is described by

$$\mathbf{x} - \mathbf{x}_{\text{new}} = \big(\mathbf{I} - (\mathbf{L}+\mathbf{D})^{-\mathbf{1}}\mathbf{A}\big)(\mathbf{x} - \mathbf{x}_{\text{old}}) = (\mathbf{I} - \mathbf{BA})(\mathbf{x} - \mathbf{x}_{\text{old}}), \tag{2}$$

where $\mathbf{I}$ is the $n \times n$ identity matrix and $\mathbf{B} = (\mathbf{L}+\mathbf{D})^{-\mathbf{1}}$.

The norm of the iteration matrix $\mathbf{I} - \mathbf{BA}$ in the matrix norm $\|\cdot\|_{\mathbf{A}}$ induced by the inner product $(\mathbf{v},\mathbf{w})_{\mathbf{A}} = \mathbf{w^t A v}$ is given by the following standard formula:

$$\|\mathbf{I}-\mathbf{BA}\|_{\mathbf{A}}^{\mathbf{2}} = \|(\mathbf{I}-\mathbf{BA})^*(\mathbf{I}-\mathbf{BA})\|_{\mathbf{A}} = \|(\mathbf{I}-\mathbf{B^t A})(\mathbf{I}-\mathbf{BA})\|_{\mathbf{A}}, \tag{3}$$

where $(\mathbf{I}-\mathbf{BA})^*$ denotes the adjoint of $\mathbf{I}-\mathbf{BA}$ with respect to $(\cdot,\cdot)_{\mathbf{A}}$. It follows from (3), the spectral theorem and the Rayleigh quotient formula that

$$\begin{aligned}\|\mathbf{I}-\mathbf{BA}\|_{\mathbf{A}}^{\mathbf{2}} &= \lambda_{\max}\big(\mathbf{I} - (\mathbf{B^t} + \mathbf{B} - \mathbf{B^t AB})\mathbf{A}\big)\\ &= 1 - \lambda_{\min}\big((\mathbf{B^t} + \mathbf{B} - \mathbf{B^t AB})\mathbf{A}\big)\\ &= 1 - \min_{\mathbf{v}\in\mathbb{R}^n} \frac{\mathbf{v^t A v}}{\mathbf{v}^t(\mathbf{B^t} + \mathbf{B} - \mathbf{B^t AB})^{-\mathbf{1}}\mathbf{v}}.\end{aligned} \tag{4}$$

A simple calculation yields

$$(\mathbf{B^t} + \mathbf{B} - \mathbf{B^t AB})^{-\mathbf{1}} = (\mathbf{I} + \mathbf{D}^{-\mathbf{1}}\mathbf{U})^{\mathbf{t}}\mathbf{D}(\mathbf{I} + \mathbf{D}^{-\mathbf{1}}\mathbf{U}), \tag{5}$$

where $\mathbf{U} = \mathbf{L^t}$ is the strictly upper triangular part of $\mathbf{A}$. It is easy to see that (5) can be rewritten as

$$(\mathbf{B^t} + \mathbf{B} - \mathbf{B^t AB})^{-\mathbf{1}} = \mathbf{A} + (\mathbf{D}^{-\mathbf{1}}\mathbf{U})^{\mathbf{t}}\mathbf{D}(\mathbf{D}^{-\mathbf{1}}\mathbf{U}). \tag{6}$$

Combining (4) and (5), we have a formula

$$\|\mathbf{I}-\mathbf{BA}\|_{\mathbf{A}}^2 = 1 - \min_{\mathbf{v}\in\mathbb{R}^n} \frac{\mathbf{v}^t\mathbf{Av}}{\mathbf{v}^t(\mathbf{I}+\mathbf{D}^{-1}\mathbf{U})^t\mathbf{D}(\mathbf{I}+\mathbf{D}^{-1}\mathbf{U})\mathbf{v}} \tag{7}$$

for the norm of the iteration matrix $\mathbf{I} - \mathbf{BA}$. Similarly the formula

$$\|\mathbf{I}-\mathbf{BA}\|_{\mathbf{A}}^2 = 1 - \min_{\mathbf{v}\in\mathbb{R}^n} \frac{\mathbf{v^t Av}}{\mathbf{v^t Av} + \mathbf{v^t}(\mathbf{D^{-1}U})^{\mathbf{t}}\mathbf{D}(\mathbf{D^{-1}U})\mathbf{v}}$$
$$= 1 - \frac{1}{1 + \max\limits_{\mathbf{v}\in\mathbb{R}^n, \|\mathbf{v}\|_{\mathbf{A}}=1} \mathbf{v^t}(\mathbf{D^{-1}U})^{\mathbf{t}}\mathbf{D}(\mathbf{D^{-1}U})\mathbf{v}} \tag{8}$$

follows from (4) and (6).

Since $\mathbf{L}+\mathbf{D}$ is the lower triangular part of $\mathbf{A}$, we can apply forward substitutions to obtain

$$(\mathbf{L}+\mathbf{D})^{-1}\mathbf{A} = \sum_{i=1}^{n} \mathbf{X}_i,$$

where $\mathbf{X}_i \in \mathbb{R}^{n\times n}$ is determined recursively by

$$\mathbf{X}_i = \mathbf{T}_i\Big(\mathbf{I} - \sum_{j=1}^{i-1} \mathbf{X}_j\Big),$$

$\mathbf{T}_i = \mathbf{e}_i(\mathbf{e}_i^t\mathbf{A}\mathbf{e}_i)^{-1}\mathbf{e}_i^t\mathbf{A}$, and $\mathbf{e}_1,\ldots,\mathbf{e}_n$ are the canonical basis vectors in $\mathbb{R}^n$.

It follows that

$$\mathbf{I}-\mathbf{BA} = \Big(\mathbf{I} - \sum_{i=1}^{n-1}\mathbf{X}_i\Big) - \mathbf{X}_n = \Big(\mathbf{I} - \sum_{i=1}^{n-1}\mathbf{X}_i\Big) - \mathbf{T}_n\Big(\mathbf{I} - \sum_{j=1}^{n-1}\mathbf{X}_j\Big)$$
$$= (\mathbf{I}-\mathbf{T}_n)\Big(\mathbf{I} - \sum_{i=1}^{n-1}\mathbf{X}_i\Big) = (\mathbf{I}-\mathbf{T}_n)\cdots(\mathbf{I}-\mathbf{T}_1).$$

Hence (7) and (8) are also formulas for the norm of the product $(\mathbf{I}-\mathbf{T}_n)\cdots(\mathbf{I}-\mathbf{T}_1)$.

Below we will derive similar formulas for general multiplicative Schwarz methods. The key observation is that even though the explicit formula (5) does not exist in the general case, we can find an expression for $\mathbf{v^t}(\mathbf{B^t}+\mathbf{B}-\mathbf{B^t AB})^{-1}\mathbf{v}$ through the additive Schwarz theory.

3 Multiplicative Schwarz Methods

Let V be a finite dimensional vector space, $a(\cdot,\cdot)$ be a SPD bilinear form on V, and $\alpha \in V'$, the dual space of V. We consider the following problem:

Find $u \in V$ such that

$$a(u,v) = \langle \alpha, v\rangle \qquad \forall\, v \in V, \tag{9}$$

where $\langle\cdot,\cdot\rangle$ denotes the canonical bilinear form on $V'\times V$.

We can rewrite (9) as

$$Au = \alpha \tag{10}$$

where $A : V \longrightarrow V'$ is defined by

$$\langle Aw, v \rangle = a(w, v) \qquad \forall v, w \in V. \tag{11}$$

The operator A is SPD in the sense that

$$\langle Aw, v \rangle = \langle Av, w \rangle \quad \forall v, w \in V \quad \text{and} \quad \langle Av, v \rangle > 0 \quad \forall v \in V \setminus \{0\}.$$

We will denote by L^* the adjoint of a linear operator $L : V \longrightarrow V$ with respect to $a(\cdot,\cdot)$, i.e,

$$a(Lv, w) = a(v, L^* w) \qquad \forall v, w \in V,$$

and for a linear operator $M : W \longrightarrow V$, the operator $M^t : V' \longrightarrow W'$ is defined by

$$\langle M^t \beta, w \rangle = \langle \beta, Mw \rangle \qquad \forall \beta \in V,' \, w \in W.$$

Let $V_1, V_2, \ldots, V_J$ be subspaces of V such that

$$V = \sum_{j=1}^{J} V_j, \tag{12}$$

and let $a_j(\cdot,\cdot)$ be a nonsingular bilinear form on V_j, i.e.,

$$A_j : V_j \longrightarrow V_j' \quad \text{is invertible} \tag{13}$$

where

$$\langle A_j v_j, w_j \rangle = a_j(v_j, w_j) \qquad \forall v_j, w_j \in V_j.$$

The operator $F_j : V' \longrightarrow V_j$ for $1 \le j \le J$ are defined recursively by

$$a_j(F_j \beta, v_j) = \langle \beta, v_j \rangle - \sum_{k=1}^{j-1} a(F_k \beta, v_j) \qquad \forall v_j \in V_j, \, \beta \in V', \tag{14}$$

and we define $B : V' \longrightarrow V$ by

$$B\beta = \sum_{j=1}^{J} (F_j \beta). \tag{15}$$

The multiplicative Schwarz algorithm is then given by the iteration

$$u_{\text{new}} = u_{\text{old}} + B(\alpha - A u_{\text{old}}), \tag{16}$$

As in the case of the Gauss-Seidel method, we have two expressions for the error propagation operator. The first obvious one is given by

$$u - u_{\text{new}} = (I - BA)(u - u_{\text{old}}), \tag{17}$$

where I is the identity operator on V, and the second one, which is responsible for the name of the algorithm, can be derived as follows.

Let $T_j : V \longrightarrow V_j$ be defined by

$$a_j(T_jv, v_j) = a(v, v_j) \qquad \forall\, v \in V,\, v_j \in V_j. \tag{18}$$

Remark 1. Note that (18) implies that $\operatorname{Ker} T_j$ is the orthogonal complement of V_j with respect to $a(\cdot,\cdot)$. Therefore $\operatorname{Ker} T_j^* = \operatorname{Ker} T_j$ and the restrictions of T_j and T_j^* to V_j are isomorphisms. In particular we have $T_jV = V_j = T_j^*V$. It follows that the pseudo-inverse T_j^{-1} (resp., $(T_j^*)^{-1}$) of T_j (resp., T_j^*) with respect to $a(\cdot,\cdot)$ maps V onto V_j.

It follows from (14) and (18) that

$$F_j\beta = T_j\Big(A^{-1}\beta - \sum_{k=1}^{j-1} F_k\beta\Big) = T_jz_j, \tag{19}$$

where

$$\begin{aligned} z_j = A^{-1}\beta - \sum_{k=1}^{j-1} F_k\beta &= \Big(A^{-1}\beta - \sum_{k=1}^{j-2} F_k\beta\Big) - F_{j-1}\beta \\ &= (I - T_{j-1})z_{j-1} = (I - T_{j-1})\cdots(I - T_1)A^{-1}\beta. \end{aligned} \tag{20}$$

Combining (15), (16), (19) and (20), with $\beta = \alpha - Au_{\mathrm{old}}$, we find

$$\begin{aligned} u - u_{\mathrm{new}} &= u - u_{\mathrm{old}} - \sum_{j=1}^{J} T_jz_j \\ &= \Big(I - \sum_{j=1}^{J} T_j(I - T_{j-1})\cdots(I - T_1)\Big)(u - u_{\mathrm{old}}) \\ &= (I - T_J)\cdots(I - T_1)(u - u_{\mathrm{old}}). \end{aligned} \tag{21}$$

We are interested in formulas for $\|I - BA\|_a = \|(I - T_J)\cdots(I - T_1)\|_a$, where $\|\cdot\|_a$ is the operator norm induced by $a(\cdot,\cdot)$.

4 Additive Schwarz Theory

We need the following fundamental result from the additive Schwarz theory.

Lemma 1. *Let $S_j : V_j \longrightarrow V$ and $B_j : V_j \longrightarrow V_j'$ be linear operators for $1 \le j \le J$, and let B_j be SPD. Then the operator $B = \sum_{j=1}^{J} S_jB_j^{-1}S_j^t : V' \longrightarrow V$ is SPD if and only if $V = \sum_{j=1}^{J} S_jV_j$, in which case we have*

$$\langle B^{-1}v, v\rangle = \min_{\substack{v=\sum_{j=1}^{J} S_jv_j \\ v_j \in V_j}} \sum_{j=1}^{J} \langle B_jv_j, v_j\rangle \qquad \forall\, v \in V. \tag{22}$$

Proof. B is clearly symmetric semi-definite, and we have for any $\beta \in V'$,

$$\langle \beta, B\beta \rangle = 0 \Leftrightarrow \sum_{j=1}^{J} \langle S_j^t \beta, B_j^{-1} S_j^t \beta \rangle = 0,$$

which holds if and only if $S_j^t \beta = 0$ for $1 \le j \le J$, since B_j^{-1} is also SPD. We conclude that $\langle \beta, B\beta \rangle = 0$ if and only if

$$\sum_{j=1}^{J} \langle \beta, S_j v_j \rangle = 0 \qquad \forall v_j \in V_j,\ 1 \le j \le J.$$

Therefore $\langle \beta, B\beta \rangle = 0$ implies $\beta = 0$ if and only if $V = \sum_{j=1}^{J} S_j V_j$.

The identity (22) comes from the observations that

$$\langle B^{-1} v, \sum_{j=1}^{J} S_j v_j \rangle = \sum_{j=1}^{J} \langle B_j (B_j^{-1} S_j^t B^{-1} v), v_j \rangle = \sum_{j=1}^{J} \langle B_j w_j, v_j \rangle, \tag{23}$$

where $w_j = B_j^{-1} S_j^t B^{-1} v \in V_j$, and

$$\sum_{j=1}^{J} S_j w_j = \sum_{j=1}^{J} S_j (B_j^{-1} S_j^t B^{-1} v) = \Big(\sum_{j=1}^{J} S_j B_j^{-1} S_j^t \Big) B^{-1} v = B B^{-1} v = v. \tag{24}$$

Indeed it follows from (23) that

$$\langle B^{-1} v, v \rangle = \sum_{j=1}^{J} \langle B_j w_j, v_j \rangle \qquad \text{if} \quad \sum_{j=1}^{J} S_j v_j = v, \tag{25}$$

and in particular, because of (24),

$$\langle B^{-1} v, v \rangle = \sum_{j=1}^{J} \langle B_j w_j, w_j \rangle. \tag{26}$$

Subtracting (26) from (25) we find

$$0 = \sum_{j=1}^{J} \langle B_j w_j, v_j - w_j \rangle \qquad \text{if} \quad \sum_{j=1}^{J} S_j v_j = v. \tag{27}$$

The orthogonality condition (27) implies

$$\sum_{j=1}^{J} \langle B_j v_j, v_j \rangle = \sum_{j=1}^{J} \langle B_j w_j, w_j \rangle + \sum_{j=1}^{J} \langle B_j (v_j - w_j), v_j - w_j \rangle \qquad \text{if} \quad \sum_{j=1}^{J} S_j v_j = v,$$

and hence

$$\sum_{j=1}^{J} \langle B_j w_j, w_j \rangle = \min_{\substack{v=\sum_{j=1} S_j v_j \\ v_j \in V_j}} \sum_{j=1}^{J} \langle B_j v_j, v_j \rangle,$$

which together with (26) implies (22).

5 A Fundamental Operator

We begin with the standard formula

$$\|I - BA\|_a^2 = \|(I - BA)^*(I - BA)\|_a, \tag{28}$$

where $(I - BA)^* = I - B^t A$ is the adjoint of $I - BA$ with respect to the bilinear form $a(\cdot,\cdot)$. We can write

$$(I - BA)^*(I - BA) = (I - B^t A)(I - BA) = I - (B^t + B - B^t AB)A. \tag{29}$$

As in the case of the Gauss-Seidel method, the operator $B^t + B - B^t AB$ will play a fundamental role. The key to the additive analysis is to interpret this operator as an additive Schwarz preconditioner. We begin with the following result.

Lemma 2. *We have*

$$\langle \beta, (B^t + B - B^t AB)\beta \rangle = \sum_{j=1}^{J} \Big[2a_j(y_j, y_j) - a(y_j, y_j) \Big] \qquad \forall \beta \in V', \tag{30}$$

where $y_j = F_j \beta$.

Proof. From (14) and (15), we have

$$\begin{aligned} \langle \beta, (B^t + B - B^t AB)\beta \rangle &= 2\langle \beta, \sum_{j=1}^{J} y_j \rangle - a(\sum_{\ell=1}^{J} y_\ell, \sum_{j=1}^{J} y_j) \\ &= 2 \sum_{j=1}^{J} \Big(a_j(y_j, y_j) + \sum_{\ell=1}^{j-1} a(y_\ell, y_j) \Big) - a(\sum_{\ell=1}^{J} y_\ell, \sum_{j=1}^{J} y_j), \end{aligned}$$

which implies (30) by the symmetry of $a(\cdot,\cdot)$.

We assume that

$$\exists\, \omega_j \in (0,2) \text{ such that } \quad a(v_j, v_j) \le \omega_j a_j(v_j, v_j) \qquad \forall v_j \in V_j. \tag{31}$$

Let the operator $B_j : V_j \longrightarrow V_j'$ be defined by

$$\langle B_j v_j, w_j \rangle = a_j(v_j, w_j) + a_j(w_j, v_j) - a(v_j, w_j) \qquad \forall v_j, w_j \in V_j. \tag{32}$$

Clearly B_j is symmetric and it is positive definite because of (31).

Remark 2. Since we are in a finite dimensional setting, condition (31) is equivalent to B_j being SPD. It is also equivalent to

$$\|(I-T_j)v\|_a \le \|v\|_a \quad \forall\, v \in V \quad \text{and} \quad \|(I-T_j)v_j\|_a < \|v_j\|_a \quad \forall\, v_j \in V_j \setminus \{0\}.$$

Note that we can write, by (18),

$$\begin{aligned}\langle B_j v_j, w_j\rangle &= a(T_j^{-1}v_j, w_j) + a(w_j, (T_j^{-1})^* v_j) - a(v_j, w_j)\\ &= a((T_j^*)^{-1}(T_j^* + T_j - T_j^* T_j)T_j^{-1}v_j, w_j) = a(\bar{T}_j T_j^{-1} v_j, T_j^{-1} w_j) \qquad (33)\end{aligned}$$

for all $v_j, w_j \in V_j$, where

$$\bar{T}_j = T_j^* + T_j - T_j^* T_j. \qquad (34)$$

Remark 3. According to Remark 1, we have $\bar{T}_j V \subset V_j$. The relation (33) implies that $a(\bar{T}_j v_j, v_j) = \langle B_j v_j, v_j\rangle > 0$ for $v_j \in V_j \setminus \{0\}$. Therefore the restriction of $\bar{T}_j$ to V_j is an isomorphism and it follows from Remark 1 that $\operatorname{Ker}\bar{T}_j = \operatorname{Ker}T_j = \operatorname{Ker}T_j^*$ is the orthogonal complement of V_j with respect to $a(\cdot,\cdot)$. Consequently the pseudo-inverse $\bar{T}_j^{-1}$ of $\bar{T}_j$ with respect to $a(\cdot,\cdot)$ maps V onto V_j.

From Lemma 2 and (32) we have

$$\langle \beta, (B^t + B - B^t A B)\beta\rangle = \sum_{j=1}^{J} \langle B_j F_j \beta, F_j \beta\rangle = \sum_{j=1}^{J} \langle \beta, F_j^t B F_j \beta\rangle \qquad \forall\, \beta \in V'. \quad (35)$$

It then follows from polarization that

$$B^t + B - B^t A B = \sum_{j=1}^{J} F_j^t B_j F_j = \sum_{j=1}^{J} (F_j^t B_j) B_j^{-1} (B_j F_j) = \sum_{j=1}^{J} S_j B_j^{-1} S_j^t, \qquad (36)$$

where the operator $S_j : V_j \longrightarrow V$ is given by $S_j = F_j^t B_j = (B_j F_j)^t$.

Remark 4. The identity (36) shows that the operator $B + B^t - B^t A B$ is indeed an additive Schwarz preconditioner. Note that (12) and (14) imply $F_1\beta = \cdots = F_J\beta = 0$ if and only if $\beta = 0$, and hence $B^t + B - B^t A B$ is SPD by (35). Therefore the formula (22) in Lemma 1 is valid.

An explicit formula for S_j is provided by the following lemma.

Lemma 3. *We have*

$$S_j = (I - T_1^*) \cdots (I - T_{j-1}^*) \bar{T}_j T_j^{-1}. \qquad (37)$$

Proof. Let $v_j \in V_j$ be arbitrary. It follows from (19), (20) and (33) that

$$\begin{aligned}\langle S_j^t \beta, v_j\rangle = \langle (B_j F_j)\beta, v_j\rangle &= a(\bar{T}_j T_j^{-1} F_j \beta, T_j^{-1} v_j)\\ &= a(z_j, \bar{T}_j T_j^{-1} v_j) = \langle \beta, (I - T_1^*) \cdots (I - T_{j-1}^*) \bar{T}_j T_j^{-1} v_j\rangle,\end{aligned}$$

which implies (37).

6 Formulas for $\|I-BA\|_a$

It follows from (28), (29), the spectral theorem and the Rayleigh quotient formula that

$$\|I-BA\|_a^2 = 1-\min_{v\in V}\frac{\langle Av,v\rangle}{\langle (B^t+B-B^tAB)^{-1}v,v\rangle},$$

which together with (36) and Lemma 1 (cf. Remark 4) implies

$$\|I-BA\|_a^2 = 1-\min_{v\in V}\frac{\langle Av,v\rangle}{\displaystyle\min_{\substack{v=\sum_{j=1}^J S_jw_j\\ w_j\in V_j}}\sum_{j=1}^J\langle B_jw_j,w_j\rangle}. \tag{38}$$

Remark 5. Note that we can rewrite (7) as

$$\|\mathbf{I}-\mathbf{BA}\|_{\mathbf{A}}^2 = 1-\min_{\mathbf{v}\in\mathbb{R}^n}\frac{\mathbf{v^tAv}}{\displaystyle\min_{\mathbf{v}=(\mathbf{I}+\mathbf{D}^{-1}\mathbf{U})^{-1}\mathbf{w}}\mathbf{w^tDw}}, \tag{39}$$

and (38) is precisely the analog of (39).

Next we will replace the implicit decomposition for v that appears in (38) by an explicit decomposition that will lead to an analog of (7). In the case of the Gauss-Seidel method, it is equivalent to inverting the relation $\mathbf{v}=(\mathbf{I}+\mathbf{D}^{-1}\mathbf{U})^{-1}\mathbf{w}$ in (39) to express $\mathbf{w}$ as $(\mathbf{I}+\mathbf{D}^{-1}\mathbf{U})\mathbf{v}$. This motivates the following construction of the explicit decomposition through an "upper triangular" system.

Given $v_j\in V_j$ for $1\le j\le J$, we want to find $w_j\in V_j$ for $1\le j\le J$ such that

$$\sum_{j=1}^J S_jw_j=\sum_{j=1}^J v_j. \tag{40}$$

It is easy to check using (37) that the solution of (40) is given by

$$\bar{T}_jT_j^{-1}w_j = v_j+T_j^*\sum_{k=j+1}^J v_k \qquad \text{for} \quad 1\le j\le J. \tag{41}$$

Combining (33), (38), (40) and (41), we have the following analog of (7):

$$\|I-BA\|_a^2 = 1-\min_{v\in V}\frac{a(v,v)}{\displaystyle\min_{\substack{v=\sum_{j=1}^J v_j\\ v_j\in V_j}}\sum_{j=1}^J a\Big(v_j+T_j^*\sum_{k=j+1}^J v_k,\bar{T}_j^{-1}\Big(v_j+T_j^*\sum_{k=j+1}^J v_k\Big)\Big)}. \tag{42}$$

Remark 6. In the case where $a_j(\cdot,\cdot)$ is SPD for $1\le j\le J$, the formula (42) becomes

$$\|I-BA\|_a^2 = 1-\min_{v\in V}\frac{a(v,v)}{\displaystyle\min_{\substack{v=\sum_{j=1}^J v_j\\ v_j\in V_j}}\sum_{j=1}^J a_j\Big(v_j+T_j\sum_{k=j+1}^J v_k,(2I-T_j)^{-1}\Big(v_j+T_j\sum_{k=j+1}^J v_k\Big)\Big)}\,. \tag{43}$$

The application of the formula (43) to domain decomposition and multigrid can be found in [5].

To derive the analog of (8), we again seek guidance from the analysis in Section 2. The transition from (7) to (8) involves the difference of $\mathbf{I}+\mathbf{D}^{-1}\mathbf{U}$ and $\mathbf{D}^{-1}\mathbf{U}$, which is a diagonal matrix. Therefore we look for operators $Q_j : V_j \longrightarrow V_j$ for $1\le j\le J$ such that

$$\begin{aligned}\sum_{j=1}^J a\Big(v_j+T_j^*\sum_{k=j+1}^J v_k,\bar{T}_j^{-1}\Big(v_j+T_j^*\sum_{k=j+1}^J v_k\Big)\Big)-a(v,v)\\ =\sum_{j=1}^J a\Big(v_j+Q_jv_j+T_j^*\sum_{k=j+1}^J v_k,\bar{T}_j^{-1}\Big(v_j+Q_jv_j+T_j^*\sum_{k=j+1}^J v_k\Big)\Big).\end{aligned} \tag{44}$$

It is straightforward to check that (44) is equivalent to

$$a(Q_jv_j,\bar{T}_j^{-1}Q_jv_j)+2a(v_j+T_j^*\sum_{k=j+1}^J v_k,\bar{T}_j^{-1}Q_jv_j)=-a\Big(v_j+2\sum_{k=j+1}^J v_k,v_j\Big),$$

which would follow from the relations

$$a(Q_jv_j+2v_j,\bar{T}_j^{-1}Q_jv_j)=-a(v_j,v_j), \tag{45}$$

$$a\Big(T_j^*\sum_{k=j+1}^J v_k,\bar{T}_j^{-1}Q_jv_j\Big)=-a\Big(\sum_{k=j+1}^J v_k,v_j\Big). \tag{46}$$

The relation (46) indicates that we should choose $\bar{T}_j^{-1}Q_j=-T_j^{-1}$ and therefore Q_j should be given by

$$Q_j=-\bar{T}_jT_j^{-1}=-(T_j^*+T_j-T_j^*T_j)T_j^{-1}=-(T_j^*T_j^{-1}+I-T_j^*), \tag{47}$$

and then (45) is also satisfied because

$$\begin{aligned}a(Q_jv_j+2v_j,\bar{T}_j^{-1}Q_jv_j)&=-a\big((I-T_j^*T_j^{-1}+T_j^*)v_j,T_j^{-1}v_j\big)\\ &=-a(v_j,T_j^{-1}v_j)+a(T_j^*T_j^{-1}v_j,T_j^{-1}v_j)-a(T_j^*v_j,T_j^{-1}v_j)\\ &=-a(v_j,v_j).\end{aligned}$$

In view of (47), we have

$$v_j + Q_j v_j + T_j^* \sum_{k=j+1}^{J} v_k = -T_j^* T_j^{-1} v_j + T_j^* \sum_{k=j}^{J} v_k = T_j^* \Big(\sum_{k=j}^{J} v_k - T_j^{-1} v_j \Big). \quad (48)$$

Putting (42), (44) and (48) together we arrive at the following analog of (8):

$$\|I-BA\|_a^2 = 1 - \min_{v\in V} \frac{a(v,v)}{a(v,v) + \min\limits_{\substack{v=\sum_{j=1}^J v_j \\ v_j \in V_j}} a\Big(T_j^*\Big(\sum_{k=j}^J v_k - T_j^{-1} v_j\Big), \bar{T}_j^{-1} T_j^*\Big(\sum_{k=j}^J v_k - T_j^{-1} v_j\Big)\Big)}$$

$$= 1 - \frac{1}{1 + \max\limits_{\substack{v\in V \\ \|v\|_a=1}} \min\limits_{\substack{v=\sum_{j=1}^J v_j \\ v_j \in V_j}} a\Big(T_j^*\Big(\sum_{k=j}^J v_k - T_j^{-1} v_j\Big), \bar{T}_j^{-1} T_j^*\Big(\sum_{k=j}^J v_k - T_j^{-1} v_j\Big)\Big)}. \quad (49)$$

7 Connection to the Xu-Zikatanov Theory

The theory in [17] was developed for the product operator $(I-T_J)\cdots(I-T_1)$ on an inner product space $\big(V, a(\cdot,\cdot)\big)$, where $T_j : V \longrightarrow V_j$ and $T_j : V_j \longrightarrow V_j$ is an isomorphism.

A key assumption in [17] is

$$\|T_j v\|_a^2 \le \omega\, a(T_j v, v) \qquad \forall\, v \in V \quad (50)$$

for some $\omega \in (0,2)$.

Lemma 4. *Under assumption* (50)*, we have*

$$T_j v = 0 \Leftrightarrow a(v, v_j) = 0 \quad \forall\, v_j \in V_j.$$

Proof. If $a(v_j, v) = 0$ for all $v_j \in V_j$, then $T_j v = 0$ by (50). Therefore, by a dimension argument, the kernel of T_j is the orthogonal complement of V_j with respect to $a(\cdot,\cdot)$.

In view of Lemma 4, we can define $a_j(\cdot,\cdot)$ by

$$a_j(T_j v, v_j) = a(v, v_j) \qquad \forall\, v \in V,\ v_j \in V_j.$$

Then $a_j(\cdot,\cdot)$ is nonsingular since

$$a_j(w_j, v_j) = 0 \quad \forall\, w_j \in V_j \quad \Rightarrow \quad a(v, v_j) = 0 \quad \forall\, v \in V \quad \Rightarrow v_j = 0.$$

On one hand we have $a(T_j v, T_j v) = \|T_j v\|_a^2$, and on the other hand we have $a(T_j v, v) = a(v, T_j v) = a_j(T_j v, T_j v)$. Hence (50) is equivalent to (31) since $V_j = T_j V$.

We conclude that the framework in [17] is identical to the framework in Section 3 and Section 5, and $\|(I-T_J)\cdots(I-T_1)\|_a$ is given by the formulas (42) and (49). In particular, the formula (49) is identical to the identity (1.1) in [17]. We note that another derivation of this identity can be found in [7].

Acknowledgements

This work is supported in part by the National Science Foundation under Grant No. DMS-16-20273. The author would also like to acknowledge the support provided by the Hausdorff Research Institute of Mathematics at Univeristät Bonn during her visit in Spring 2017.

References

1. M. Benzi, A. Frommer, R. Nabben, and D.B. Szyld. Algebraic theory of multiplicative Schwarz methods. *Numer. Math.*, 89:605–639, 2001.
2. P. Bjørstad and J. Mandel. On the spectra of sums of orthogonal projections with applications to parallel computing. *BIT*, 31:76–88, 1991.
3. J.H. Bramble, J.E. Pasciak, J. Wang, and J. Xu. Convergence estimates for product iterative methods with applications to domain decomposition. *Math. Comp.*, 57:1–21, 1991.
4. J.H. Bramble and X. Zhang. The Analysis of Multigrid Methods. In P.G. Ciarlet and J.L. Lions, editors, *Handbook of Numerical Analysis, VII*, pages 173–415. North-Holland, Amsterdam, 2000.
5. S.C. Brenner. An additive analysis of multiplicative Schwarz methods. *Numer. Math.*, 123:1–19, 2013.
6. S.C. Brenner and L.R. Scott. *The Mathematical Theory of Finite Element Methods* (*Third Edition*). Springer-Verlag, New York, 2008.
7. L. Chen. Deriving the X-Z identity from auxiliary space method. In *Domain decomposition methods in science and engineering XIX*, volume 78 of *Lect. Notes Comput. Sci. Eng.*, pages 309–316. Springer, Heidelberg, 2011.
8. M. Dryja and O.B. Widlund. An additive variant of the Schwarz alternating method in the case of many subregions. Technical Report 339, Department of Computer Science, Courant Institute, 1987.
9. M. Griebel and P. Oswald. On the abstract theory of additive and multiplicative Schwarz algorithms. *Numer. Math.*, 70:163–180, 1995.
10. P.-L. Lions. On the Schwarz alternating method. I. In *First International Symposium on Domain Decomposition Methods for Partial Differential Equations* (*Paris, 1987*), pages 1–42. SIAM, Philadelphia, PA, 1988.
11. T.P.A. Mathew. *Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations*. Springer-Verlag, Berlin, 2008.
12. S. Nepomnyaschikh. On the application of the bordering method to the mixed boundary value problem for elliptic equations and on mesh norms in $W_2^{1/2}(S)$. *Sov. J. Numer. Anal. Math. Modelling*, 4:493–506, 1989.
13. Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, Philadelphia, 2003.
14. B. Smith, P. Bjørstad, and W. Gropp. *Domain Decomposition*. Cambridge University Press, Cambridge, 1996.
15. A. Toselli and O.B. Widlund. *Domain Decomposition Methods - Algorithms and Theory*. Springer, New York, 2005.
16. J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Review*, 34:581–613, 1992.
17. J. Xu and L. Zikatanov. The method of alternating projections and the method of subspace corrections in Hilbert space. *J. Amer. Math. Soc.*, 15:573–597 (electronic), 2002.
18. H. Yserentant. Old and new convergence proofs for multigrid methods. *Acta Numerica*, 2:285–326, 1993.
19. X. Zhang. Multilevel Schwarz methods. *Numer. Math.*, 63:521–539, 1992.