Electronic Transactions on Numerical Analysis.
Volume 31, pp. 228-255, 2008.

ISSN 1068-9613.

ETNA
Kent State University
http://etna.math.kent.edu

SCHWARZ METHODS OVER THE COURSE OF TIME*

MARTIN J. GANDER†

To the memory of Gene Golub, our leader and friend.

Abstract. Schwarz domain decomposition methods are the oldest domain decomposition methods. They were invented by Hermann Amandus Schwarz in 1869 as an analytical tool to rigorously prove results obtained by Riemann through a minimization principle. Renewed interest in these methods was sparked by the arrival of parallel computers, and variants of the method have been introduced and analyzed, both at the continuous and discrete level. It can be daunting to understand the similarities and subtle differences between all the variants, even for the specialist.

This paper presents Schwarz methods as they were developed historically. From quotes by major contributors over time, we learn about the reasons for similarities and subtle differences between continuous and discrete variants. We also formally prove at the algebraic level equivalence and/or non-equivalence among the major variants for very general decompositions and many subdomains. We finally trace the motivations that led to the newest class called optimized Schwarz methods, illustrate how they can greatly enhance the performance of the solver, and show why one has to be cautious when testing them numerically.

Key words. Alternating and parallel Schwarz methods, additive, multiplicative and restricted additive Schwarz methods, optimized Schwarz methods.

AMS subject classifications. 65F10, 65N22.

1. The Dirichlet principle and Schwarz's challenge. An important part of the theory of analytic functions was developed by Riemann based on a minimization principle, the Dirichlet principle. This principle states that an harmonic function, which is a function satisfying Laplace's equation $\Delta u = 0$ on a bounded domain Ω with Dirichlet boundary conditions $u = g$ on $\partial\Omega$, is the infimum of the Dirichlet integral $\int_\Omega |\nabla v|^2$ over all functions v satisfying the boundary conditions $v = g$ on $\partial\Omega$. It was taken for granted by Riemann that the infimum is attained, until Weierstrass gave a counterexample of a functional that does not attain its minimum. It was in this context that Schwarz invented the first domain decomposition method [68]:

> Die unter dem Namen Dirichletsches Princip bekannte Schlussweise, welche in gewissem Sinne als das Fundament des von Riemann entwickelten Zweiges der Theorie der analytischen Functionen angesehen werden muss, unterliegt, wie jetzt wohl allgemein zugestanden wird, hinsichtlich der Strenge sehr begründeten Einwendungen, deren vollständige Entfernung meines Wissens den Anstrengungen der Mathematiker bisher nicht gelungen ist.[1]

The Dirichlet principle could be rigorously proved for simple domains, where Fourier analysis was applicable. Therefore Schwarz embarked on the project of finding an analytical tool to extend the Dirichlet principle to more complicated domains.

2. Schwarz methods at the continuous level. There are two main classical Schwarz methods at the continuous level: the alternating Schwarz method invented by Schwarz in [68] as a mathematical tool, and the parallel Schwarz method introduced by Lions in [47] for the purpose of parallel computing.

*Received December 14, 2007. Accepted November 12, 2008. Published online on April 2, 2009. Recommended by Oliver Ernst.

†Université de Genève, Section de Mathématiques, 2-4 rue du Lièvre, CP 64, CH-1211 Genève, Switzerland (Martin.Gander@unige.ch).

[1]The Dirichlet principle, which can be seen as the foundation of the part of functional analysis developed by Riemann, is now widely regarded as not being sufficiently rigorous, and a fully rigorous argument to replace it has so far eluded all efforts of mathematicians.

ETNA
Kent State University
http://etna.math.kent.edu

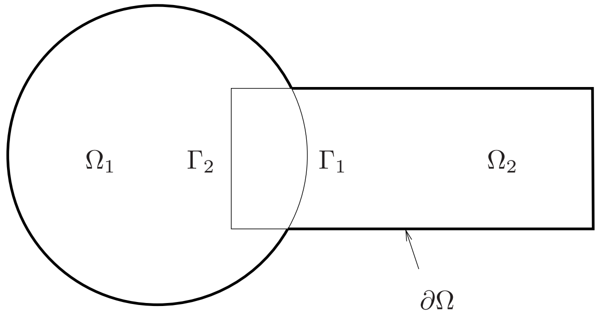

FIGURE 2.1. *The first domain decomposition method was introduced by Schwarz for a complicated domain, composed of two simple ones, namely a disk and a rectangle.*

2.1. The alternating Schwarz method. In order to show that Riemann's results in the theory of analytic functions hold, Schwarz needed to find a rigorous proof for the Dirichlet principle, i.e., he had to show that the infimum of the Dirichlet integral is attained on arbitrary domains. Schwarz presented the fundamental idea of decomposition of the domain into simpler subdomains, for which more information is available:

> Nachdem gezeigt ist, dass für eine Anzahl von einfacheren Bereichen die Differentialgleichung $\Delta u = 0$ beliebigen Grenzbedingungen gemäss integriert werden kann, handelt es sich darum, den Nachweis zu führen, dass auch für einen weniger einfachen Bereich, der aus jenen auf gewisse Weise zusammengesetzt ist, die Integration der Differentialgleichung beliebigen Grenzbedingungen gemäss möglich ist.[2]

In Figure 2.1, we show the original domain used by Schwarz, with the associated domain decomposition into two subdomains, which are geometrically much simpler, namely a disk Ω_1 and a rectangle Ω_2, with interfaces $\Gamma_1 := \partial\Omega_1 \cap \Omega_2$ and $\Gamma_2 := \partial\Omega_2 \cap \Omega_1$. To show that the equation

$$\Delta u = 0 \quad \text{in } \Omega, \qquad u = g \quad \text{on } \partial\Omega, \tag{2.1}$$

can be integrated (note how the particular choice of wording by Schwarz resembles the integration of ordinary differential equations) with arbitrary boundary conditions, Schwarz proposed what is now called the *alternating Schwarz method*, an iterative method which only uses solutions on the disk and the rectangle, where solutions can be obtained using Fourier series. The method starts with an initial guess u_2^0 along Γ_1 (see Figure 2.1), and then computes iteratively for $n = 0, 1, \ldots$ the iterates u_1^{n+1} and u_2^{n+1} according to the algorithm

$$\begin{array}{rcll} \Delta u_1^{n+1} & = & 0 & \text{in } \Omega_1, \\ u_1^{n+1} & = & u_2^n & \text{on } \Gamma_1, \end{array} \qquad \begin{array}{rcll} \Delta u_2^{n+1} & = & 0 & \text{in } \Omega_2, \\ u_2^{n+1} & = & u_1^{n+1} & \text{on } \Gamma_2, \end{array} \tag{2.2}$$

where we omit from now on for simplicity that both u_1^{n+1} and u_2^{n+1} satisfy the given Dirichlet condition in (2.1) on the outer boundaries of the respective subdomains. Schwarz motivated iteration (2.2) using a vacuum pump as an analogy:

> Zum Beweise dieses Satzes kann ein Grenzübergang dienen, welcher mit dem bekannten, zur Herstellung eines luftverdünnten Raumes mittelst einer

[2] After having shown for some simple domains that the partial differential equation $\Delta u = 0$ can be integrated with arbitrary boundary conditions, we have to prove that the same partial differential equation can be solved as well on more complicated domains which are composed of the simpler ones in a certain fashion.

ETNA
Kent State University
http://etna.math.kent.edu

> zweistiefeligen Luftpumpe dienenden Verfahren grosse Analogie hat.[3]

Schwarz proved convergence of his alternating method using the maximum principle, and the proof, also given quite informally by Schwarz with the analogy of the vacuum pump (which seems surprising, considering its purpose of proving rigorously Riemann's deductions), works as follows: denoting by $\underline{u}$ the infimum and by $\overline{u}$ the supremum of the boundary data g given on the outer boundary $\partial\Omega$ in (2.1), Schwarz starts by imposing $\underline{u}$ on Γ_1 which completes the boundary conditions on subdomain Ω_1. One can therefore obtain on the disk u_1^1 (first chamber is pumping). Now he fixes the values of the solution u_1^1 along Γ_2 (first valve closed) and thus on Ω_2 the boundary conditions are complete and one can solve on the rectangular domain to obtain u_2^1 (second chamber is pumping). Schwarz now observes that the difference $u_2^1 - u_1^1$ (or also $u_2^1 - \underline{u}$) is less than $\overline{u} - \underline{u}$ along Γ_1 by the maximum principle. Imposing now the value of u_2^1 along Γ_1 (second valve closed), a new iterate u_1^2 can be obtained on the disk Ω_1. The difference $u_1^2 - u_1^1$ along Γ_2 is now by a factor $q_1 < 1$ smaller by the maximum principle than $G := \overline{u} - \underline{u}$, we thus have $u_1^2 - u_1^1 < Gq_1$ along Γ_2. Proceeding as before on Ω_2, one obtains u_2^2, and by the maximum principle the difference $u_2^2 - u_2^1$ is along Γ_1 smaller by a factor q_2 than the difference $u_1^2 - u_1^1$ along Γ_2, and thus along Γ_1 we have $u_2^2 - u_2^1 < Gq_1q_2$. By induction, and using linearity to see that the quantities q_1 and q_2 are the same for all iterations, one obtains an infinite sequence of functions u_1^n and u_2^n, and it is easy to show ("es ist nun nicht schwer, nachzuweisen") that they converge uniformly to limiting functions defined by

$$\begin{aligned} u_1 &= u_1^1 + (u_1^2 - u_1^1) + (u_1^3 - u_1^2) + \cdots, \\ u_2 &= u_2^1 + (u_2^2 - u_2^1) + (u_2^3 - u_2^2) + \cdots. \end{aligned}$$

Since the series on the right converge for all x and y, because

$$(u_1^{n+1} - u_1^n) < G(q_1q_2)^{n-1}, \qquad (u_2^{n+1} - u_2^n) < G(q_1q_2)^{n-1}q_1.$$

Schwarz now observes that the functions u_1 and u_2 agree on both Γ_1 and Γ_2, and thus must be identical in the overlap. He therefore concludes that u_1 and u_2 must be values of the same function u satisfying Laplace's equation on the entire domain.

The argument of Schwarz still lacked some rigor; in particular at the two corner points where the two subdomains intersect the subdomains do not really overlap (the overlap becomes arbitrarily small), and it is more delicate to use the maximum principle. This problem was studied more carefully by Pierre Louis Lions over a century after Schwarz [48]:

> We study the same question when we relax the condition of overlapping, allowing the "boundaries of the two subdomains" to touch at the boundary of the original domain. As we will see, if the situation is not basically modified for Dirichlet boundary conditions (in this case, our analysis is a minor extension of Schwarz original convergence proof), we will show that drastic changes occur for Neumann boundary conditions.

The Schwarz alternating method can readily be extended to more than two subdomains, only care needs to be taken in the formulation to ensure that the newest available information at the interfaces is always taken, if several choices are possible. We define, for a domain Ω, the J overlapping subdomains Ω_j, $j = 1, 2, \ldots, J$, which also defines the order in which subdomains are updated, and the interfaces Γ_{jk}, $j \neq k$, by

$$\Gamma_{jk} = \partial\Omega_j \cap \left(\Omega_k \setminus \bigcup_{l \in M_{jk}} \Omega_l\right), \tag{2.3}$$

[3] To prove this theorem, one can use an alternating method, which has great analogy with a two-level pump to obtain an air-diluted room.

ETNA
Kent State University
http://etna.math.kent.edu

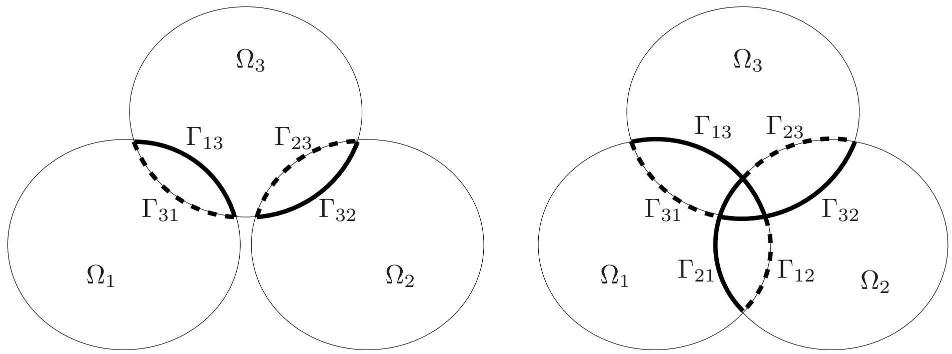

FIGURE 2.2. *Two different three-subdomain configurations.*

with

$$M_{jk} = \begin{cases} \{1, \ldots, j-1, k+1, \ldots, J\}, & \text{if } k > j, \\ \{k+1, \ldots, j-1\}, & \text{if } k < j, \end{cases}$$

as illustrated in Figure 2.2 for two configurations of 3 subdomains. Note that the set M_{jk} can be empty, e.g., if the starting index in its definition is larger than the ending one. The algorithm

$$\begin{aligned} \mathcal{L}u_j^{n+1} &= f && \text{in } \Omega_j, \\ u_j^{n+1} &= u_k^{n+1_{jk}} && \text{on } \Gamma_{jk}, \end{aligned} \tag{2.4}$$

where the symbol 1_{jk} equals one if $j > k$ and zero otherwise, is then a direct generalization of the alternating Schwarz method for the elliptic equation $\mathcal{L}u = f$ and J subdomains. The definition of Γ_{jk} in (2.3) ensures that the interface Γ_{jk} is the part of the interface of Ω_j in Ω_k on which Ω_k provides the newest available update for the algorithm, as the example for three subdomains illustrates in Figure 2.2, i.e., none of the subdomains computed after Ω_k in the cyclic process can provide newer boundary data on Γ_{jk}. Convergence of the method for many subdomains can be proved similarly as in the original argument of Schwarz, provided that the operator $\mathcal{L}$ satisfies a maximum principle. Lions studied the alternating Schwarz method also using a variational approach in [47] and found a very elegant convergence proof using projections. He remarked

> Let us observe, by the way, that the Schwarz alternating method seems to be the only domain decomposition method converging for two entirely different reasons: variational characterization of the Schwarz sequence and maximum principle.

While convergence proofs of the Schwarz alternating method strongly depend on the underlying partial differential equation (PDE) to be solved (for an early convergence proof for the case of elasticity, see [70]), similar methods can be defined for any PDE, even time dependent ones, which leads to the class of Schwarz waveform relaxation methods; see for example [38]. The idea of an overlapping subdomain decomposition and an iteration is completely general.

2.2. The parallel Schwarz method. At the time when Lions analyzed the alternating Schwarz method, parallel computers were becoming more and more available, and Lions realized the potential of the Schwarz method on such computers [47]:

> The final extension we wish to consider concerns "parallel" versions of the Schwarz alternating method ..., u_i^{n+1} is solution of $-\Delta u_i^{n+1} = f$ in Ω_i and $u_i^{n+1} = u_j^n$ on $\partial\Omega_i \cap \Omega_j$.

ETNA
Kent State University
http://etna.math.kent.edu

We call this method the *parallel Schwarz method*, in contrast to the *alternating Schwarz method*. For the historical example of Schwarz in Figure 2.1, the method is given by

$$\begin{array}{llll} \Delta u_1^{n+1} = 0 & \text{in } \Omega_1, & \Delta u_2^{n+1} = 0 & \text{in } \Omega_2, \\ u_1^{n+1} = u_2^n & \text{on } \Gamma_1, & u_2^{n+1} = u_1^n & \text{on } \Gamma_2. \end{array} \tag{2.5}$$

The only change is the iteration index in the second transmission condition, which makes this method parallel: given initial guesses u_1^0 and u_2^0, one can now simultaneously compute, for $n = 0, 1, \ldots$, both subdomain solutions in parallel. In this simple two-subdomain case, there is, however, no gain, since the sequence computed on Ω_1 every two steps coincides with the sequence computed on Ω_1 by the alternating Schwarz method. If many subdomains are used, there are no such simple subsequences anymore, and computing in parallel can pay off. There is, however, an important point to address in the multisubdomain case, which Lions discussed carefully in [47]:

> As soon as $J \geq 3$ the situation becomes more interesting. And even if, as we will see in section II, each sequence u_j^n converges in Ω_j to u, this method does not have always a variational interpretation in terms of iterated projections. A related difficulty is that, using the sequences
>
> $$(u_1^n)_n, (u_2^n)_n, \ldots, (u_J^n)_n$$
>
> it is not always possible to define a single-valued function defined on the whole domain Ω in a continuous way. In fact, the necessary and sufficient condition for these two difficulties not to happen is that:
>
> $$\left\{ \begin{array}{l} \text{for all distinct } i, j, k \in \{1, \ldots, J\}, \text{ if } \Omega_i \cap \Omega_j \neq \emptyset \\ \text{and } \Omega_i \cap \Omega_k \neq \emptyset, \text{ then } \Omega_j \cap \Omega_k = \emptyset. \end{array} \right. \tag{2.6}$$

This property ensures that, for each subdomain interface point, there is precisely one neighboring subdomain where the boundary data can be taken from, which is however only rarely satisfied in practice. The simple example in Figure 2.2 on the left satisfies (2.6), whereas the one on the right violates (2.6), and in the latter case, the formulation of the algorithm Lions gives needs to be modified to specify from which neighboring subdomain boundary data is taken. One possibility is to use the same interface definition as for the alternating Schwarz method (2.3), and we obtain the parallel Schwarz method for many subdomains

$$\begin{array}{rcll} \mathcal{L} u_j^{n+1} & = & f & \text{in } \Omega_j, \\ u_j^{n+1} & = & u_k^n & \text{on } \Gamma_{jk}. \end{array}$$

A more general definition from where to obtain neighboring subdomain boundary data, which will be useful later, is to start with a non-overlapping decomposition $\widetilde{\Omega}_j$, construct an associated one that is overlapping by choosing that each Ω_j contains $\widetilde{\Omega}_j$, i.e., $\widetilde{\Omega}_j \subset \Omega_j$, and then to define the interfaces Γ_{jk} by

$$\Gamma_{jk} = \partial\Omega_j \cap \widetilde{\Omega}_k. \tag{2.7}$$

This definition contains the special one from alternating Schwarz, and while one can not prove convergence using variational arguments, the maximum principle technique by Schwarz still applies and convergence follows; see [48].

Note that the alternating Schwarz method with many subdomains can also be used in parallel, one simply needs to assign the same color to subdomains which do not touch, and thus do not need to communicate, and then all those can be updated in parallel.

ETNA
Kent State University
http://etna.math.kent.edu

2.3. Discretization of continuous Schwarz methods. The alternating and parallel Schwarz methods can be discretized to obtain computational tools. In fact, the first motivation in computing was very similar to the motivation of Schwarz in analysis: in the 1970s, fast Poisson solvers were developed, based on the Fast Fourier Transform [2]. These solvers were however restricted to special geometries, important examples being circular or rectangular domains. Golub and Mayers showed in [39] that Schwarz methods presented the ideal computational tool to generalize such fast solvers to more general geometries, using the example of a T shaped domain:

> The two discrete Laplace problems are both Dirichlet problems on a rectangle, and can be solved very efficiently by a fast Poisson solver, using some form of Fast Fourier Transform.

Even for the historical model problem of Schwarz in Figure 2.1, if we discretize the alternating Schwarz method (2.2) using finite differences or finite elements, enforcing the interface conditions directly on the right-hand side, we obtain

$$A_1\boldsymbol{u}_1^{n+1} = \boldsymbol{f}_1 - A_{12}\boldsymbol{u}_2^n, \qquad A_2\boldsymbol{u}_2^{n+1} = \boldsymbol{f}_2 - A_{21}\boldsymbol{u}_1^{n+1}, \tag{2.8}$$

where the matrices A_1 and A_2 are discretizations of the Laplacian, and the matrices A_{12} and A_{21} are zero matrices, except for the unknowns on the interface Γ_1 in A_{12} and for the unknowns on the interface Γ_2 in A_{21}, where the matrix entries are the corresponding entries of the discretization stencil used. Hence the subproblem on domain Ω_1 can be solved by a fast Poisson solver for circular domains, and the subproblem on domain Ω_2 by a fast Poisson solver for rectangular domains. Note that one might need to interpolate in order to transmit data at the interfaces, in which case the interface matrices A_{12} and A_{21} would also include these interpolation matrices. The situation does not change if we have many subdomains, in this case the discrete algorithm is

$$A_j\boldsymbol{u}_j^{n+1} = \boldsymbol{f}_j - \sum_{k=1}^{j-1} A_{jk}\boldsymbol{u}_k^{n+1} - \sum_{k=j+1}^{J} A_{jk}\boldsymbol{u}_k^n,$$

where the A_{jk} correspond to the interface definition (2.3); for a precise algebraic definition in the case of conforming grids; see Assumption 3.1 in the next section.

Similarly, one obtains for the parallel Schwarz method (2.5), in the case of two subdomains, the parallel discrete iteration

$$A_1\boldsymbol{u}_1^{n+1} = \boldsymbol{f}_1 - A_{12}\boldsymbol{u}_2^n, \qquad A_2\boldsymbol{u}_2^{n+1} = \boldsymbol{f}_2 - A_{21}\boldsymbol{u}_1^n, \tag{2.9}$$

and, in the case of many subdomains,

$$A_j\boldsymbol{u}_j^{n+1} = \boldsymbol{f}_j - \sum_{k\neq j} A_{jk}\boldsymbol{u}_k^n, \tag{2.10}$$

where now the A_{jk} correspond to any interface definition of the form (2.7); for the precise algebraic definition, see Remark 3.8 in the next section.

3. Discrete Schwarz methods. If we discretize Laplace's equation (2.1), or a more general elliptic PDE, we obtain a linear system of the form

$$A\boldsymbol{u} = \boldsymbol{f}. \tag{3.1}$$

Schwarz methods have also been introduced directly at the algebraic level for such linear systems, and there are several variants.

3.1. The multiplicative Schwarz method. In order to obtain a domain decomposition-like iteration for the discrete system (3.1), one needs to partition the unknowns in the vector $\boldsymbol{u}$ into subsets, similarly as the continuous domain was partitioned into subdomains. This can be achieved by using simple restriction operators: if we want, for example, to partition the unknowns into a first and a second set, possibly overlapping, we can use the restriction matrices

$$R_1 = \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \end{bmatrix}, \quad R_2 = \begin{bmatrix} & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}, \tag{3.2}$$

which are identically zero, except for the positions indicated by a 1. With these restriction matrices, $R_1\boldsymbol{u}$ gives the first set of unknowns, and $R_2\boldsymbol{u}$ the second one. One can also define a restriction of the matrix A to the first and second set of unknowns using these same restriction matrices,

$$A_j = R_j A R_j^T, \quad j = 1, 2.$$

The *multiplicative Schwarz method* (see for example [5] or [69]), is now defined by

$$\begin{array}{rcl} \boldsymbol{u}^{n+\frac{1}{2}} & = & \boldsymbol{u}^n + R_1^T A_1^{-1} R_1(\boldsymbol{f} - A\boldsymbol{u}^n), \\ \boldsymbol{u}^{n+1} & = & \boldsymbol{u}^{n+\frac{1}{2}} + R_2^T A_2^{-1} R_2(\boldsymbol{f} - A\boldsymbol{u}^{n+\frac{1}{2}}). \end{array} \tag{3.3}$$

This iteration resembles the alternating Schwarz method: one does first a solve with the local matrix A_1 associated with the first set of unknowns, and then a solve with the local matrix A_2 associated with the second set of unknowns. It is however not at all transparent, from formulation (3.3), what information is transmitted in the residual from the first subset of unknowns to the second one and vice versa: nothing like the transmission conditions in the alternating Schwarz method (2.2) is apparent in the multiplicative Schwarz method (3.3). Is it possible that the multiplicative Schwarz method (3.3) is just a discretization of the alternating Schwarz method (2.2)? If yes, how is the algebraic overlap from the restriction matrices (3.2) related to the physical overlap in the alternating Schwarz method ?

Let us take a closer look at the case when the R_j are non-overlapping, i.e., $R_1^T R_1 + R_2^T R_2 = I$, the identity matrix. In this case, we can easily partition the system matrix A, the right hand side $\boldsymbol{f}$ and the vector $\boldsymbol{u}$ accordingly,

$$A = \begin{bmatrix} A_1 & A_{12} \\ A_{21} & A_2 \end{bmatrix}, \qquad \boldsymbol{f} = \begin{bmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \end{bmatrix}, \qquad \boldsymbol{u} = \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \end{bmatrix},$$

and we obtain in the first relation of the multiplicative Schwarz method (3.3) an interesting cancellation at the algebraic level: the restricted residual becomes

$$R_1(\boldsymbol{f} - A\boldsymbol{u}^n) = \boldsymbol{f}_1 - A_1\boldsymbol{u}_1^n - A_{12}\boldsymbol{u}_2^n,$$

and when we apply the local solve A_1^{-1}, a copy of $\boldsymbol{u}_1^n$ is obtained,

$$A_1^{-1}R_1(\boldsymbol{f} - A\boldsymbol{u}^n) = A_1^{-1}(\boldsymbol{f}_1 - A_{12}\boldsymbol{u}_2^n) - \boldsymbol{u}_1^n.$$

Inserting this result back into the first relation of (3.3), we get

$$\begin{array}{rcl} \begin{bmatrix} \boldsymbol{u}_1^{n+\frac{1}{2}} \\ \boldsymbol{u}_2^{n+\frac{1}{2}} \end{bmatrix} & = & \begin{bmatrix} \boldsymbol{u}_1^n \\ \boldsymbol{u}_2^n \end{bmatrix} + \begin{bmatrix} A_1^{-1}(\boldsymbol{f}_1 - A_{12}\boldsymbol{u}_2^n) - \boldsymbol{u}_1^n \\ 0 \end{bmatrix} \\ & = & \begin{bmatrix} A_1^{-1}(\boldsymbol{f}_1 - A_{12}\boldsymbol{u}_2^n) \\ \boldsymbol{u}_2^n \end{bmatrix}, \end{array} \tag{3.4}$$

ETNA
Kent State University
http://etna.math.kent.edu

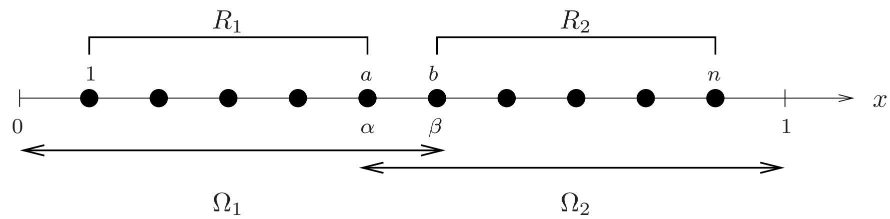

FIGURE 3.1. *A non-overlapping algebraic decomposition is equivalent to an overlapping continuous decomposition for the underlying PDE with minimal overlap of one mesh size.*

where the $\boldsymbol{u}_1^n$ terms have canceled. Similarly, using the second relation of the multiplicative Schwarz method (3.3), we obtain for one full iteration step

$$\begin{bmatrix} \boldsymbol{u}_1^{n+1} \\ \boldsymbol{u}_2^{n+1} \end{bmatrix} = \begin{bmatrix} A_1^{-1}(\boldsymbol{f}_1 - A_{12}\boldsymbol{u}_2^n) \\ A_2^{-1}(\boldsymbol{f}_2 - A_{21}\boldsymbol{u}_1^{n+1}) \end{bmatrix}.$$

This relation can be rewritten in the equivalent form

$$A_1\boldsymbol{u}_1^{n+1} = \boldsymbol{f}_1 - A_{12}\boldsymbol{u}_2^n, \qquad A_2\boldsymbol{u}_2^{n+1} = \boldsymbol{f}_2 - A_{21}\boldsymbol{u}_1^{n+1}, \tag{3.5}$$

which is identical to (2.8), and we have thus proved that the multiplicative Schwarz method (3.3) without overlap is a discretization of the original alternating Schwarz method (2.2) from 1869, albeit one with minimal overlap, as one can best see from the one dimensional sketch in Figure 3.1, even though the R_j were non-overlapping at the algebraic level, a subtle difference between discrete and continuous notations.

Without overlap, the multiplicative Schwarz method (3.3) is just a block Gauss-Seidel method, since (3.5) leads in matrix form to the iteration

$$\begin{bmatrix} A_1 & 0 \\ A_{21} & A_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1^{n+1} \\ \boldsymbol{u}_2^{n+1} \end{bmatrix} = \begin{bmatrix} 0 & -A_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1^{n} \\ \boldsymbol{u}_2^{n} \end{bmatrix} + \begin{bmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \end{bmatrix}. \tag{3.6}$$

So one might wonder why the notation with the restriction matrices R_j was introduced. There are two reasons: first, with the R_j and the formulation (3.3) one can easily use overlapping blocks, which is natural for these methods, and difficult to do in the block Gauss-Seidel formulation (3.6); second, with formulation (3.3) there is automatically a global approximate solution $\boldsymbol{u}^n$, a feature which is not available in (2.8) without an additional selection or averaging procedure to define the solution in the overlap.

In the case of more than two subdomains, the multiplicative Schwarz method becomes

$$\boldsymbol{u}^{n+\frac{j}{J}} = \boldsymbol{u}^{n+\frac{j-1}{J}} + R_j^T A_j^{-1} R_j(\boldsymbol{f} - A\boldsymbol{u}^{n+\frac{j-1}{J}}), \quad j = 1, 2, \ldots, J, \tag{3.7}$$

and in order to prove a general equivalence result, we need precise assumptions on the interface operators A_{jk} corresponding to the interface definition (2.3) of the alternating Schwarz method (2.4).

ASSUMPTION 3.1. *We assume that the operators A_{jk} in the alternating Schwarz method (2.4) satisfy*

$$A_j R_j + \sum_{k \neq j} A_{jk} R_k = R_j A, \quad j = 1, \ldots, J, \tag{3.8}$$

which states that all boundary values must be taken from some neighbor, and

$$A_{jk} R_k R_m^T = 0$$

ETNA
Kent State University
http://etna.math.kent.edu

for all $j = 1, \ldots, J$ and $m \in M_{jk}$ defined in (2.3), which is equivalent to stating that always the most recently computed information must be used.

We will need the following technical Lemma.

LEMMA 3.2. *For any given set of subdomain vectors $\boldsymbol{u}_l$, $l = 1, \ldots, J$, and under Assumption 3.1, we have for $k \neq j$*

$$A_{jk}R_k \sum_{l=1}^{j-1} \prod_{i=l+1}^{j-1} (I - R_i^T R_i) R_l^T \boldsymbol{u}_l = \begin{cases} A_{jk}\boldsymbol{u}_k & \text{if } k < j, \\ 0 & \text{otherwise,} \end{cases} \tag{3.9}$$

and

$$A_{jk}R_k \prod_{i=1}^{j-1}(I - R_i^T R_i) \left(\sum_{p=1}^{J} \prod_{q=p+1}^{J} (I - R_q^T R_q) R_p^T \boldsymbol{u}_p \right) = \begin{cases} A_{jk}\boldsymbol{u}_k & \text{if } k > j, \\ 0 & \text{otherwise.} \end{cases} \tag{3.10}$$

Proof. To show (3.9), we start with the case $k < j$: we first split the left-hand side into three parts,

$$\begin{aligned} A_{jk}R_k \sum_{l=1}^{j-1} \prod_{i=l+1}^{j-1} (I - R_i^T R_i) R_l^T \boldsymbol{u}_l &= \sum_{l=1}^{k-1} A_{jk}R_k \prod_{i=l+1}^{j-1} (I - R_i^T R_i) R_l^T \boldsymbol{u}_l \\ +A_{jk}R_k \prod_{i=k+1}^{j-1} (I - R_i^T R_i) R_k^T \boldsymbol{u}_k &+ \sum_{l=k+1}^{j-1} A_{jk}R_k \prod_{i=l+1}^{j-1} (I - R_i^T R_i) R_l^T \boldsymbol{u}_l. \end{aligned} \tag{3.11}$$

The first sum on the right vanishes, because for $k < j$ each product contains the term $I - R_k^T R_k$, and since the terms in the product commute, each term in the sum contains $R_k(I - R_k^T R_k)$ which is zero. In the second term on the right in (3.11), if $k = j - 1$ the product is empty, and since $R_k R_k^T = I$, the term equals $A_{jk}\boldsymbol{u}_k$. If $k < j - 1$, we obtain

$$A_{jk}R_k \prod_{i=k+1}^{j-1} (I - R_i^T R_i) R_k^T \boldsymbol{u}_k = A_{jk}R_k R_k^T \boldsymbol{u}_k = A_{jk}\boldsymbol{u}_k,$$

where we used Assumption 3.1, and thus the second term always equals $A_{jk}\boldsymbol{u}_k$. Now for the last term in (3.11), if $k = j - 1$, the sum is empty and thus the term is zero, and if $k < j - 1$, then

$$\sum_{l=k+1}^{j-1} A_{jk}R_k \prod_{i=l+1}^{j-1} (I - R_i^T R_i) R_l^T \boldsymbol{u}_l = \sum_{l=k+1}^{j-1} A_{jk}R_k R_l^T \boldsymbol{u}_l = 0,$$

where we used Assumption 3.1 twice, and this concludes the proof for (3.9) if $k < j$. If $k > j$, zero is obtained for $j = 1$, because the sum is empty, and for $j > 1$, we get

$$\sum_{l=1}^{j-1} A_{jk}R_k \prod_{i=l+1}^{j-1} (I - R_i^T R_i) R_l^T \boldsymbol{u}_l = \sum_{l=1}^{j-1} A_{jk}R_k R_l^T \boldsymbol{u}_l = 0,$$

where we used Assumption 3.1 again twice. This concludes the proof of (3.9).

To show the second identity (3.10), we first note that, for $k < j$,

$$A_{jk}R_k \prod_{i=1}^{j-1}(I - R_i^T R_i) = 0,$$

ETNA
Kent State University
http://etna.math.kent.edu

since the product includes the term $R_k(I - R_k^T R_k)$, which is the zero matrix. Now if $k > j$, then

$$A_{jk} R_k \prod_{i=1}^{j-1} (I - R_i^T R_i) = A_{jk} R_k,$$

because for $j = 1$ the product is empty, and for $j > 1$, we use Assumption 3.1. We separate the remaining sum into three terms,

$$\begin{aligned} A_{jk} R_k \sum_{p=1}^{J} \prod_{q=p+1}^{J} (I - R_q^T R_q) R_p^T \boldsymbol{u}_p &= \sum_{p=1}^{k-1} A_{jk} R_k \prod_{q=p+1}^{J} (I - R_q^T R_q) R_p^T \boldsymbol{u}_p \\ + A_{jk} R_k \prod_{q=k+1}^{J} (I - R_q^T R_q) R_k^T \boldsymbol{u}_k &+ \sum_{p=k+1}^{J} A_{jk} R_k \prod_{q=p+1}^{J} (I - R_q^T R_q) R_p^T \boldsymbol{u}_p. \end{aligned} \tag{3.12}$$

The first term on the right vanishes as in the proof of the first identity. The second term on the right in (3.12) can be simplified,

$$A_{jk} R_k \prod_{q=k+1}^{J} (I - R_q^T R_q) R_k^T \boldsymbol{u}_k = A_{jk} R_k R_k^T \boldsymbol{u}_k = A_{jk} \boldsymbol{u}_k,$$

using Assumption 3.1. Finally the last term in (3.12) becomes

$$\sum_{p=k+1}^{J} A_{jk} R_k \prod_{q=p+1}^{J} (I - R_q^T R_q) R_p^T \boldsymbol{u}_p = \sum_{p=k+1}^{J} A_{jk} R_k R_p^T \boldsymbol{u}_p = 0,$$

using Assumption 3.1 twice, and this concludes the proof. □

THEOREM 3.3. *If the initial iterates $\boldsymbol{u}_j^0$, $j = 1, \ldots, J$ of the alternating Schwarz method (2.4) and the initial iterate $\mathbf{u}^0$ of the multiplicative Schwarz method (3.7) satisfy*

$$\boldsymbol{u}^0 = \sum_{j=1}^{J} \prod_{i=j+1}^{J} (I - R_i^T R_i) R_j^T \boldsymbol{u}_j^0, \tag{3.13}$$

and Assumption 3.1 holds, then (2.4) and (3.7) generate an equivalent sequence of iterates,

$$\boldsymbol{u}^n = \sum_{j=1}^{J} \prod_{i=j+1}^{J} (I - R_i^T R_i) R_j^T \boldsymbol{u}_j^n. \tag{3.14}$$

REMARK 3.4. Condition (3.13) is not a restriction, it simply relates the initial guess of one algorithm to the initial guess of the other one. If the initial guess is not equivalent for the two methods, they can not produce equivalent iterates.

Proof of Theorem 3.3. The proof uses induction twice: we first assume that for a fixed n, relation (3.14) holds, and show by induction that this implies the relation

$$\boldsymbol{u}^{n+\frac{j}{J}} = \prod_{i=1}^{j} (I - R_i^T R_i) \boldsymbol{u}^n + \sum_{l=1}^{j} \prod_{i=l+1}^{j} (I - R_i^T R_i) R_l^T \boldsymbol{u}_l^{n+1}, \tag{3.15}$$

for $j = 0, 1, \ldots, J$. This relation trivially holds for $j = 0$. Assuming that it holds for $j-1$, we find from (3.7) using the matrix identity (3.8), that

$$\begin{aligned}
\boldsymbol{u}^{n+\frac{j}{J}} &= \boldsymbol{u}^{n+\frac{j-1}{J}} + R_j^T A_j^{-1} R_j(\boldsymbol{f} - A\boldsymbol{u}^{n+\frac{j-1}{J}}) \\
&= \boldsymbol{u}^{n+\frac{j-1}{J}} + R_j^T A_j^{-1}\left(\boldsymbol{f}_j - A_j R_j \boldsymbol{u}^{n+\frac{j-1}{J}} - \sum_{k\neq j} A_{jk} R_k \boldsymbol{u}^{n+\frac{j-1}{J}}\right) \\
&= (I - R_j^T R_j)\boldsymbol{u}^{n+\frac{j-1}{J}} + R_j^T A_j^{-1}\left(\boldsymbol{f}_j - \sum_{k\neq j} A_{jk} R_k \boldsymbol{u}^{n+\frac{j-1}{J}}\right).
\end{aligned}$$

Now, replacing relation (3.15) at $j-1$ in the last sum, and using the induction assumption (3.14) at step n together with Lemma 3.2 leads to

$$\begin{aligned}
\sum_{k\neq j} A_{jk} R_k \boldsymbol{u}^{n+\frac{j-1}{J}} &= \sum_{k\neq j} A_{jk} R_k \left(\prod_{i=1}^{j-1}(I - R_i^T R_i)\boldsymbol{u}^n + \sum_{l=1}^{j-1}\prod_{i=l+1}^{j-1}(I - R_i^T R_i) R_l^T \boldsymbol{u}_l^{n+1}\right) \\
&= \sum_{k=1}^{j-1} A_{jk}\boldsymbol{u}_k^{n+1} + \sum_{k=j+1}^{J} A_{jk}\boldsymbol{u}_k^n.
\end{aligned}$$

Substituting this result back into the expression for $\boldsymbol{u}^{n+\frac{j}{J}}$, we find, using (2.4) and relation (3.15) at $j-1$ again,

$$\begin{aligned}
\boldsymbol{u}^{n+\frac{j}{J}} &= (I - R_j^T R_j)\boldsymbol{u}^{n+\frac{j-1}{J}} + R_j^T A_j^{-1}\left(\boldsymbol{f}_j - \sum_{m=1}^{j-1} A_{jk}\boldsymbol{u}_k^{n+1} - \sum_{m=j+1}^{J} A_{jk}\boldsymbol{u}_k^n\right) \\
&= \prod_{i=1}^{j}(I - R_i^T R_i)\boldsymbol{u}^n + \sum_{l=1}^{j-1}\prod_{i=l+1}^{j}(I - R_i^T R_i) R_l^T \boldsymbol{u}_l^{n+1} + R_j^T \boldsymbol{u}_j^{n+1} \\
&= \prod_{i=1}^{j}(I - R_i^T R_i)\boldsymbol{u}^n + \sum_{l=1}^{j}\prod_{i=l+1}^{j}(I - R_i^T R_i) R_l^T \boldsymbol{u}_l^{n+1},
\end{aligned}$$

which concludes the first proof by induction. The main result (3.14) can now be proved by induction on n. By the assumption on the initial iterate, (3.14) holds for $n = 0$. Thus assuming it holds for n, we obtain from the first part for this n that relation (3.15) holds for $j = 0, 1, \ldots, J$. In particular, for $j = J$, we have

$$\boldsymbol{u}^{n+1} = \prod_{i=1}^{J}(I - R_i^T R_i)\boldsymbol{u}^n + \sum_{l=1}^{J}\prod_{i=l+1}^{J}(I - R_i^T R_i) R_l^T \boldsymbol{u}_l^{n+1},$$

and $\prod_{i=1}^{J}(I - R_i^T R_i) = 0$, which completes the proof. $\square$

3.2. The additive Schwarz method. The multiplicative Schwarz method is sequential in nature, like the alternating Schwarz method, and naturally the question arises if there is a more parallel variant, like the parallel Schwarz method of Lions. Dryja and Widlund studied in [21] a parallel variant, introduced earlier at the continuous level by Matsokin and Nepomnyaschikh [60], which they call the *additive Schwarz method*:

> The basic idea behind the additive form of the algorithm is to work with the simplest possible polynomial in the projections. Therefore the equation $(P_1 + P_2 + \ldots + P_N)u_h = g_h'$ is solved by an iterative method.

ETNA
Kent State University
http://etna.math.kent.edu

Using the same notation as before for our two-subdomain model problem, the preconditioned system proposed by Dryja and Widlund in [21] is

$$(R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2) A\boldsymbol{u} = (R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2)\boldsymbol{f}. \tag{3.16}$$

Using this preconditioner for a stationary iterative method yields

$$\boldsymbol{u}^{n+1} = \boldsymbol{u}^n + (R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2)(\boldsymbol{f} - A\boldsymbol{u}^n), \tag{3.17}$$

and we see that now the two subdomain solves can be done in parallel, like in the parallel variant of the Schwarz method proposed by Lions in (2.5). So is the additive Schwarz iteration (3.17) equivalent to a discretization of Lions's parallel Schwarz method? If the R_j are non-overlapping, proceeding like in the multiplicative Schwarz case using the cancellation property observed in (3.4), we obtain

$$\begin{bmatrix} \boldsymbol{u}_1^{n+1} \\ \boldsymbol{u}_2^{n+1} \end{bmatrix} = \begin{bmatrix} A_1^{-1}(\boldsymbol{f}_1 - A_{12}\boldsymbol{u}_2^n) \\ A_2^{-1}(\boldsymbol{f}_2 - A_{21}\boldsymbol{u}_1^n) \end{bmatrix}, \tag{3.18}$$

which can be rewritten in the equivalent form

$$A_1\boldsymbol{u}_1^{n+1} = \boldsymbol{f}_1 - A_{12}\boldsymbol{u}_2^n, \qquad A_2\boldsymbol{u}_2^{n+1} = \boldsymbol{f}_2 - A_{21}\boldsymbol{u}_1^n, \tag{3.19}$$

and is thus identical to the discretization of Lions's parallel Schwarz method (2.9) from 1988.

In the algebraically non-overlapping case, the additive Schwarz method is also equivalent to a block Jacobi method, since one can rewrite (3.19) in the matrix form

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1^{n+1} \\ \boldsymbol{u}_2^{n+1} \end{bmatrix} = \begin{bmatrix} 0 & -A_{12} \\ -A_{21} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1^{n} \\ \boldsymbol{u}_2^{n} \end{bmatrix} + \begin{bmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \end{bmatrix}.$$

The situation changes drastically if the R_j overlap, which is very natural for these methods. The reason is that the cancellation of $\boldsymbol{u}_1^n$ and $\boldsymbol{u}_2^n$ with $\boldsymbol{u}^n$, observed in (3.4) in detail for the multiplicative Schwarz method, and used in (3.18) for the additive Schwarz method with non-overlapping R_j, does not work anymore properly in the overlap, since in the updating formula (3.17),

$$\boldsymbol{u}^{n+1} = \boldsymbol{u}^n + \begin{bmatrix} A_1^{-1}(\boldsymbol{f}_1 - A_{12}\boldsymbol{u}_2^n) - \boldsymbol{u}_1^n \\ \boldsymbol{0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{0} \\ A_2^{-1}(\boldsymbol{f}_2 - A_{21}\boldsymbol{u}_1^n) - \boldsymbol{u}_2^n \end{bmatrix}.$$

Now, the non-zero terms $A_1^{-1}(\boldsymbol{f}_1 - A_{12}\boldsymbol{u}_2^n) - \boldsymbol{u}_1^n$ from the first subdomain solve and the non-zero terms $A_2^{-1}(\boldsymbol{f}_2 - A_{21}\boldsymbol{u}_1^n) - \boldsymbol{u}_2^n$ from the second subdomain solve overlap, and thus the current iterate in $\boldsymbol{u}_1^n$ and $\boldsymbol{u}_2^n$ is subtracted twice in the overlap, and a new approximation from both the first and the second subdomain solve is added. For the model problem of the discretized Poisson equation and two subdomains, it was shown in [23] that the spectral radius of the additive Schwarz iteration operator in (3.17) equals 1, and that the method fails to converge in the overlap, while outside of the overlap it produces the same iterates as Lions's parallel Schwarz method. This was also noticed in [66, page 29]:

> The proof that these variational formulations are equivalent to the original ones is obtained via the verification of the relations
>
> $$U^{k+1} = \begin{cases} \hat{U}^{k+1}_{1|\Omega_1\backslash\Omega_2} & \text{in } \Omega\backslash\Omega_2, \\ \hat{U}^{k+1}_{1|\Omega_{1,2}} + \hat{U}^{k+1}_{2|\Omega_{1,2}} - U^k_{|\Omega_{1,2}} & \text{in } \Omega_{1,2}, \\ \hat{U}^{k+1}_{2|\Omega_2\backslash\Omega_1} & \text{in } \Omega\backslash\Omega_1, \end{cases}$$
>
> (here U^k denotes the additive Schwarz iterate, and $\hat{U}^k_j$ the iterates of Lions's parallel variant).

ETNA
Kent State University
http://etna.math.kent.edu

We now prove a non-convergence result in the general case of more than two subdomains for the additive Schwarz iteration

$$\boldsymbol{u}^{n+1} = \boldsymbol{u}^n + \sum_{j=1}^{J} R_j^T A_j^{-1} R_j(\boldsymbol{f} - A\boldsymbol{u}^n). \tag{3.20}$$

We associate with each unit basis vector $\boldsymbol{e}_l$ the index set I_l containing all indices j such that $R_j\boldsymbol{e}_l \neq 0$, and we denote by M_{AS}^{-1} the additive Schwarz preconditioner in (3.20),

$$M_{AS}^{-1} := \sum_{j=1}^{J} R_j^T A_j^{-1} R_j, \qquad A_j = R_j A R_j^T. \tag{3.21}$$

THEOREM 3.5. *If $R_j A\boldsymbol{e}_l = 0$ for all $j \notin I_l$, then $\boldsymbol{e}_l$ is an eigenvector of the additive Schwarz iteration operator $I - M_{AS}^{-1}A$ with eigenvalue $1 - |I_l|$, where $|I_l|$ is the size of the index set I_l, and hence, if $|I_l| > 1$, the additive Schwarz iteration (3.20) can not converge in general.*

Proof. For $j \in I_l$, we have $R_j^T R_j \boldsymbol{e}_l = \boldsymbol{e}_l$, since $R_j^T R_j$ is the identity on the corresponding subspace, and we obtain

$$\begin{aligned}(I - M_{AS}^{-1}A)\boldsymbol{e}_l &= \boldsymbol{e}_l - \sum_{j=1}^{J} R_j^T A_j^{-1} R_j A\boldsymbol{e}_l \\ &= \boldsymbol{e}_l - \sum_{j\notin I_l} R_j^T A_j^{-1} R_j A\boldsymbol{e}_l - \sum_{j\in I_l} R_j^T A_j^{-1} R_j A R_j^T R_j \boldsymbol{e}_l \\ &= (1 - |I_l|)\boldsymbol{e}_l,\end{aligned}$$

since by definition $A_j = R_j A R_j^T$, and by assumption $R_j A\boldsymbol{e}_l = 0$ for $j \notin I_l$. □

It is interesting to note that in the cases (2.6) where Lions's formulation of the algorithm can be used, e.g., in Figure 2.2 on the left, the additive Schwarz iteration only stagnates in the overlap. As in the two-subdomain case, the only non-converging eigenmodes have eigenvalue minus one, and hence one can still conclude as in [66] that outside the overlap the iterates of the discretized parallel Schwarz method of Lions and the additive Schwarz iteration coincide, and the two methods are equivalent. In the general case however, e.g., in Figure 2.2 on the right, the additive Schwarz method is divergent in the overlap, and whenever a subdomain uses information from within the overlap of other subdomains, i.e., $R_j A\boldsymbol{e}_l \neq 0$ for $j \notin I_l$, there is no longer equivalence between the additive Schwarz iterates and the discretization of Lions's parallel Schwarz method, since then the additive Schwarz method eventually diverges everywhere.

Several examples for a decomposition of the square into four equal overlapping subsquares and a discretized Laplace equation are shown in Figure 3.2. We used the classical five point finite difference stencil on an 11×11 mesh, and a uniform overlap of five mesh widths. In the upper left figure, the initial error for the additive Schwarz iteration (3.20) was chosen equal to 1 in the regions where exactly two subdomains overlap, corresponding to nodes with $|I_l| = 2$ in Theorem 3.5, and zero otherwise. This gives rise to an oscillatory mode $(-1)^n$, shown after two iterations. In the upper right figure, the initial error for the additive Schwarz iteration (3.20) was chosen equal to 1 in the region where four subdomains overlap, corresponding to $|I_l| = 4$ in Theorem 3.5. This gives rise to a growing oscillatory mode $(-3)^n$, shown after two iterations. In the lower left figure, the initial error for the additive Schwarz iteration (3.20) was chosen equal to 1 at an interface lying in the overlap of

ETNA
Kent State University
http://etna.math.kent.edu

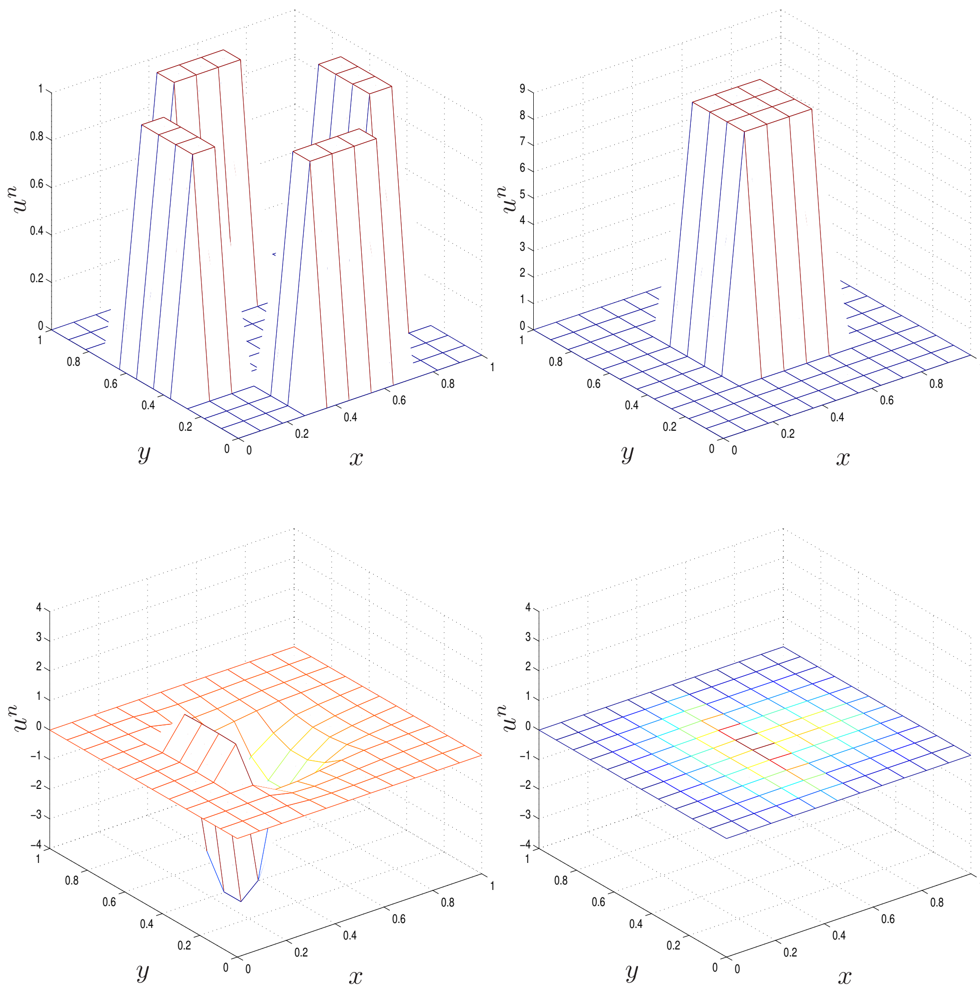

FIGURE 3.2. *Error in the additive Schwarz method after two iterations, for various initial errors, and the result of Lions's parallel Schwarz method on the lower right for the same initial error as additive Schwarz on the lower left.*

two subdomains, i.e., $|I_l| = 2$, but $R_j A\boldsymbol{e}_l \neq 0$ for $j \notin I_l$. This does not lead to an eigenmode of the additive Schwarz iteration operator, and excites other non-converging modes, as shown after two iterations. Even in the interior of the first subdomain the maximum error has already grown from 0 to 0.033 at iteration step 2, and at iteration 18 the maximum error equals 1.05 in the interior of the first subdomain: the iteration diverges everywhere. Finally, for this same initial error and the same decomposition, we show in the lower right figure the result of the discretized parallel Schwarz method (2.10) proposed by Lions, also after two iterations: clearly the method converges very well, the maximum error is already 0.0461 everywhere, and the convergence is geometric, as one can show, as in the continuous case, using a discrete maximum principle. Matsokin and Nepomnyaschikh had introduced a relaxation parameter τ_n in [60] when they studied what was to become the additive Schwarz iterative method, and obtained convergence only "for a suitable choice of τ_n", a fact which seems

ETNA
Kent State University
http://etna.math.kent.edu

to be well known in the numerical linear algebra community, where this factor is called the damping factor; see for example Hackbusch [40], who states "the AS iteration converges for sufficiently small τ", and [28]. The maximum size of the damping factor is precisely related to the problem of the method in the overlap, it has to be smaller than $\max_l \frac{1}{|I_l|}$ for convergence, in order to put the eigenvalues $1-|I_l| \leq -1$ into the unit disk; see Theorem 3.5. Lions's parallel Schwarz method and its discretization however do not need such a damping factor.

Nevertheless, the preconditioned system (3.16), which for the case of many subdomains is of the form

$$\sum_{j=1}^{J} R_j^T A_j^{-1} R_j A\boldsymbol{u} = \sum_{j=1}^{J} R_j^T A_j^{-1} R_j \boldsymbol{f}, \tag{3.22}$$

has a very desirable property for solution with a Krylov method: the preconditioner is symmetric, if A is symmetric. Including a coarse grid correction denoted by $R_0^T A_0^{-1} R_0$ in the additive Schwarz preconditioner, we obtain

$$M_{AS^c}^{-1} := \sum_{j=1}^{J} R_j^T A_j^{-1} R_j + R_0^T A_0^{-1} R_0.$$

Dryja and Widlund showed in [21] a fundamental condition number estimate for this preconditioner applied to the Poisson equation [74], discretized with characteristic coarse mesh size H, fine mesh size h and an overlap δ:

THEOREM 3.6. *The condition number of the additive Schwarz preconditioned system satisfies*

$$\kappa(M_{AS^c}^{-1}A) \leq C\left(1+\frac{H}{\delta}\right), \tag{3.23}$$

where the constant C is independent of h, H and δ.

The additive Schwarz method used as a preconditioner for a Krylov method is therefore optimal in the sense that it converges independently of the mesh size and the number of subdomains, if the ratio of H and δ is held constant. The non-converging modes from Theorem 3.5 in the iterative form of the additive Schwarz operator $I - M_{AS}^{-1}A$ are ≤ -1, and they are bounded from below by the maximum number of subdomains which overlap simultaneously. Thus, in the additive Schwarz preconditioned system (3.22) with operator $M_{AS}^{-1}A$, they are bounded away from zero and not large. The maximum number of subdomains which overlap simultaneously appears also in the constant C in (3.23) [74, page 67], and one could conjecture that this dependence of C is removed with restricted additive Schwarz.

3.3. The restricted additive Schwarz method. In 1998, a new family of discrete Schwarz methods was introduced by Cai and Sarkis [4]:

> While working on an AS/GMRES algorithm in an Euler simulation, we removed part of the communication routine and surprisingly the "then AS" method converged faster in both terms of iteration counts and CPU time.

Using the same notation as before for our two-subdomain model problem, the *restricted additive Schwarz* iteration is

$$\boldsymbol{u}^{n+1} = \boldsymbol{u}^n + (\widetilde{R}_1^T A_1^{-1} R_1 + \widetilde{R}_2^T A_2^{-1} R_2)(\boldsymbol{f} - A\boldsymbol{u}^n),$$

where the new restriction matrices $\widetilde{R}_j$ are like R_j, but with some of the ones in the overlap replaced by zeros, in order to correspond to a non-overlapping decomposition, i.e., $\widetilde{R}_1^T\widetilde{R}_1 +$

ETNA
Kent State University
http://etna.math.kent.edu

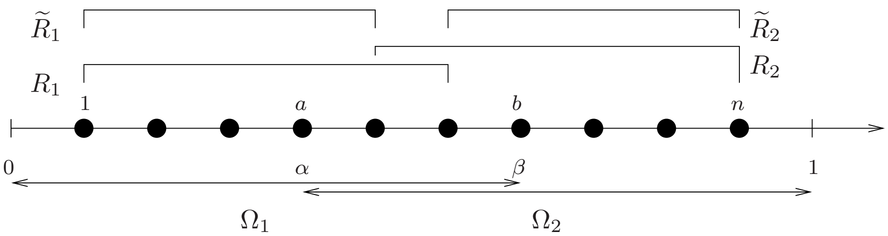

FIGURE 3.3. *Definition of the $\widetilde{R}_j$ for a one dimensional example.*

$\widetilde{R}_2^T\widetilde{R}_2 = I$, the identity; for an illustration in one dimension, see Figure 3.3. As in the case of the additive Schwarz method, this method can readily be generalized to J subdomains,

$$\boldsymbol{u}^{n+1} = \boldsymbol{u}^n + \sum_{j=1}^{J} \widetilde{R}_j^T A_j^{-1} R_j(\boldsymbol{f} - A\boldsymbol{u}^n), \tag{3.24}$$

and the following theorem shows that this method is equivalent to a disretization of Lions's parallel Schwarz method, the non-converging modes in the overlap are eliminated by the $\widetilde{R}_j$.

THEOREM 3.7. *Let A be an invertible matrix, and R_j given restriction matrices, $j = 1, 2, \ldots, J$, such that the $A_j := R_jAR_j^T$ are invertible. Let $\widetilde{R}_j$ be their associated non-overlapping counterparts. If $\boldsymbol{f}_j = R_j\boldsymbol{f}$, $A_{jk} := (R_jA - A_jR_j)\widetilde{R}_k^T$, and $\boldsymbol{u}^0 = \sum_{j=1}^J \widetilde{R}_j^T\boldsymbol{u}_j^0$, then the restricted additive Schwarz method (3.24) and the discretized parallel Schwarz method (2.10) give equivalent iterates, i.e., $\boldsymbol{u}^n = \sum_{j=1}^J \widetilde{R}_j^T\boldsymbol{u}_j^n$.*

Proof. The proof is by induction. For $n = 0$ the result holds by assumption. So, assuming it holds for n, we obtain for $n+1$, using that for any vector $\boldsymbol{u}^n$ the identity $\boldsymbol{u}^n = \sum_{j=1}^J \widetilde{R}_j^T R_j\boldsymbol{u}^n$ holds,

$$\begin{aligned}\boldsymbol{u}^{n+1} &= \boldsymbol{u}^n + \sum_{j=1}^{J} \widetilde{R}_j^T A_j^{-1} R_j(\boldsymbol{f} - A\boldsymbol{u}^n)\\ &= \sum_{j=1}^{J} \widetilde{R}_j^T A_j^{-1}(\boldsymbol{f}_j + A_jR_j\boldsymbol{u}^n - R_jA\boldsymbol{u}^n).\end{aligned}$$

By the induction hypothesis, we have $\boldsymbol{u}^n = \sum_{k=1}^J \widetilde{R}_k^T\boldsymbol{u}_k^n$, which yields

$$\begin{aligned}\boldsymbol{u}^{n+1} &= \sum_{j=1}^{J} \widetilde{R}_j^T A_j^{-1}\left(\boldsymbol{f}_j + (A_jR_j - R_jA)\sum_{k=1}^{J}\widetilde{R}_k^T\boldsymbol{u}_k^n\right)\\ &= \sum_{j=1}^{J} \widetilde{R}_j^T A_j^{-1}\left(\boldsymbol{f}_j - \sum_{k=1}^{J} A_{jk}\boldsymbol{u}_k^n\right),\end{aligned}$$

where we used the definition of A_{jk}. Now, in the last sum the term $k = j$ can be excluded, since

$$A_{jj} = R_jA\widetilde{R}_j^T - A_jR_j\widetilde{R}_j^T = R_jAR_j^TR_j\widetilde{R}_j^T - A_jR_j\widetilde{R}_j^T = A_jR_j\widetilde{R}_j^T - A_jR_j\widetilde{R}_j^T = 0,$$

which concludes the proof by using (2.10). ∎

ETNA
Kent State University
http://etna.math.kent.edu

REMARK 3.8. The choice $A_{jk} = (R_j A - A_j R_j)\widetilde{R}_k^T$ is a very natural one for the discretized parallel Schwarz method of Lions (2.10), since it states precisely that interface values are taken from a non-overlapping decomposition, as in the continuous formulation (2.7).

There is no convergence theorem similar to Theorem 3.6 for restricted additive Schwarz [74]:

> To our knowledge, a comprehensive theory of these algorithms is still missing.

There are only comparison results at the algebraic level between additive and restricted additive Schwarz, which show that restricted additive Schwarz always has better spectral properties [29], and partial results when the shape of the subdomains is modified; see [3]. With the equivalence of the restricted additive Schwarz method (3.24) and the discretized parallel Schwarz method (2.10), we can obtain a first general convergence proof for the restricted additive Schwarz method in the case where a discrete maximum principle holds: the continuous convergence proof of Lions in [48], based on the maximum principle, then also applies to the discretized case, and hence to the restricted additive Schwarz method by Theorem 3.7. A very elegant convergence proof with convergence factor estimates based on the maximum principle and lower and upper solutions for the continuous equivalent of restricted additive Schwarz can be found in [59]; similar techniques were also used for time dependent problems in [38].

Unfortunately, the restricted additive Schwarz preconditioner is non-symmetric, even if the underlying system matrix A is symmetric; therefore, for symmetric problems, there is a trade-off between the additive Schwarz preconditioner having a larger condition number and the restricted additive one being non-symmetric.

4. Problems of classical Schwarz methods.

The classical Schwarz methods we have seen can be applied to many classes of PDEs; their fundamental idea of decomposition and iteration is very general and flexible, both at the continuous and discrete level. With all this flexibility however there are also some problems with Schwarz methods, some of which we discuss in this section, together with ideas of remedies from the literature, which interestingly all point in the same direction.

4.1. Overlap required.

It was Lions who emphasized one of the main drawbacks of the classical Schwarz methods in [49]:

> However, the Schwarz method requires that the subdomains overlap, and this may be a severe restriction - without speaking of the obvious or intuitive waste of efforts in the region shared by the subdomains.

In addition to this evident drawback, Lions might also have thought about problems with discontinuous coefficients, where it would be very natural to use a non-overlapping decomposition with the interface along the discontinuity, or even problems where different models need to be coupled. In such situations, an overlap does not constitute a natural decomposition. Lions therefore proposed a modification of the alternating Schwarz method for a non-overlapping decomposition, as illustrated in Figure 4.1. The only change in the new variant lies in the transmission condition, otherwise it is still an iteration by subdomains,

$$\begin{aligned} \mathcal{L}u_1^{n+1} &= f \quad \text{in } \Omega_1, & \mathcal{L}u_2^{n+1} &= f \quad \text{in } \Omega_2, \\ (\partial_{n_1}+p_1)u_1^{n+1} &= (\partial_{n_1}+p_1)u_2^n \text{ on } \Gamma, & (\partial_{n_2}+p_2)u_2^{n+1} &= (\partial_{n_2}+p_2)u_1^{n+1} \text{ on } \Gamma. \end{aligned} \tag{4.1}$$

With these new Robin transmission conditions, Lions proved in [49] using energy estimates that the new Schwarz method is convergent without overlap for the case of constant parameters p_j and an arbitrary number of subdomains. From this analysis one can not see how the performance depends on the parameters p_j, but Lions makes in the last paragraph the visionary remark:

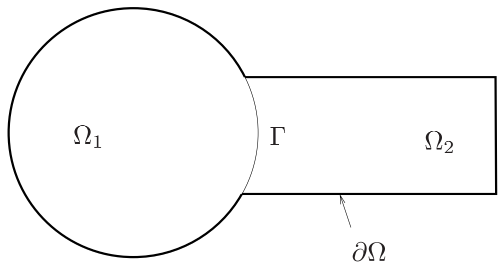

FIGURE 4.1. *A non-overlapping decomposition.*

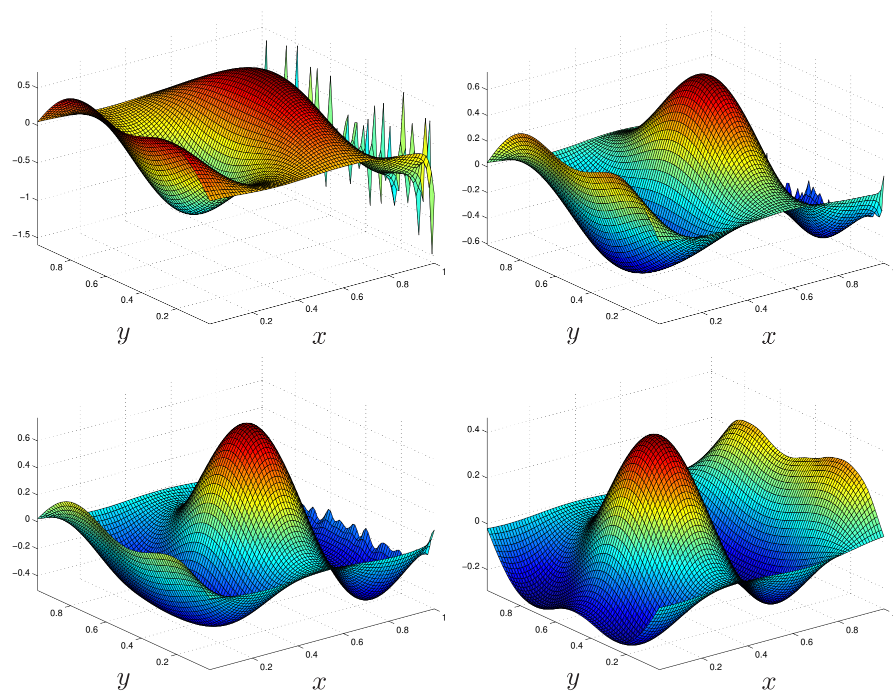

FIGURE 4.2. *Error in the first, second, third and eighth iterate on the left subdomain of the alternating Schwarz algorithm applied to the indefinite Helmholtz equation.*

> First of all, it is possible to replace the constants in the Robin conditions by two proportional functions on the interface, or even by local or nonlocal operators,

and then concludes by showing that for a one dimensional model problem, one can choose the parameters in such a way that the method with two subdomains converges in two iterations, which transforms this iterative method into a direct solver.

4.2. Lack of convergence. Another drawback of the classical overlapping Schwarz method is that there are PDEs for which the method is not convergent. A well known example is the indefinite Helmholtz equation, for which we show, in Figure 4.2, the error on the left subdomain of the alternating Schwarz algorithm with an overlapping two-subdomain decomposition. We started the iteration with a random initial guess, an issue we will come back to in the experiments in Figure 5.2. One can see that while the alternating Schwarz method

ETNA
Kent State University
http://etna.math.kent.edu

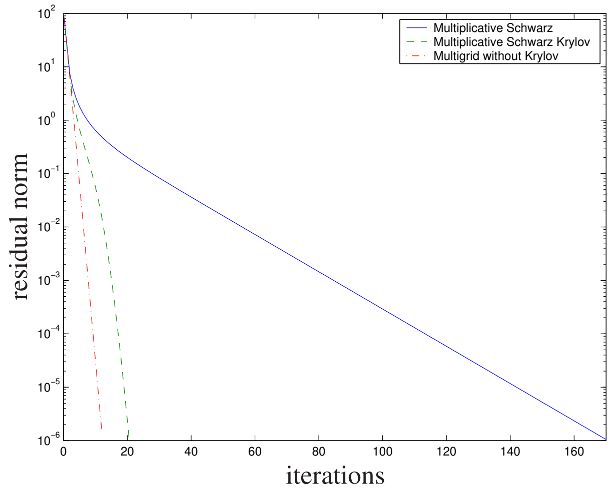

FIGURE 4.3. *Comparison of the multiplicative Schwarz method, used as an iterative solver or as a preconditions, with a multigrid method.*

quickly removes the high frequency components in the error, a particular low frequency component remains. One can show that, for such problems, the overlapping Schwarz method is not effective for low frequencies, the convergence factor being equal to one for these frequencies [30]. Després had already worked on this problem in his PhD thesis [15, 14]:

> L'objectif de ce travail est, après construction d'une méthode de décomposition de domaine adaptée au problème de Helmholtz, d'en démontrer la convergence.[4]

Interestingly, Després also made just one modification to the algorithm, the same that Lions did, except that he fixed the choice of the parameters to be $p_j = i\omega$, where ω is the wave number of the Helmholtz problem, and then used again energy estimates to prove convergence of a non-overlapping variant of the method.

4.3. Convergence speed. The final drawback of the classical Schwarz methods we want to mention is their convergence speed. We show in Figure 4.3 a comparison of the multiplicative Schwarz method with two subdomains, as an iterative solver and as a preconditioner for a Krylov method, with a standard multigrid solver when applied to the discretized positive definite problem $(\eta - \Delta)u = f$, $\eta > 0$. Clearly the multiplicative Schwarz method needs too many iterations to reduce the residual, compared to the multigrid method. Even as a preconditioner for a Krylov method, the method is significantly slower than multigrid. Hagstrom, Tewarson and Jazcilevich proposed for a non-linear problem in [41] an idea to improve the performance of the classical Schwarz method:

> In general, [the coefficients in the Robin transmission conditions] may be operators in an appropriate space of function on the boundary. Indeed, we advocate the use of nonlocal conditions.

Similarly, at the discrete level, Tang proposed generalized Schwarz splittings in [73], with better transmission conditions to improve the performance of the classical Schwarz method:

> In this paper, a new coupling between the overlap[ping] subregions is identified. If a successful coupling is chosen, a fast convergence of the alternat-

[4]The goal of this work is, after construction of a domain decomposition method adapted to the Helmholtz problem, to prove that this new method is convergent.

ETNA
Kent State University
http://etna.math.kent.edu

ing process can be achieved without a large overlap;

see also [67].

Gene Golub also showed me at a recent conference another interesting observation at the discrete level: if one uses for a one dimensional discretized Poisson equation a two-block Jacobi splitting corresponding to a Schwarz method with Dirichlet transmission conditions, the preconditioned problem is of rank two; if one uses however Neumann transmission conditions, it is of rank one, a result which can be generalized to higher dimensions, and cuts in half the maximum number of iterations of a Krylov method with this preconditioner.

In summary, for all the drawbacks we have mentioned from the literature, significant improvements have been achieved by modifying the transmission conditions. This has led to a new class of Schwarz methods we call optimized Schwarz methods, and which we will discuss next, again both at the continuous and discrete level.

5. Optimized Schwarz methods. Optimized Schwarz methods grew out of the comments by Lions and Hagstrom et al. to use more general operators in the Robin transmission conditions. Nataf et al. for example, state in [65]:

> The rate of convergence of Schwarz and Schur-type algorithms is very sensitive to the choice of interface conditions. The original Schwarz method is based on the use of Dirichlet boundary conditions. In order to increase the efficiency of the algorithm, it has been proposed to replace the Dirichlet boundary conditions with more general boundary conditions. (...) It has been remarked that absorbing (or artificial) boundary conditions are a good choice. In this report, we try to clarify the question of the interface conditions.

5.1. Optimized Schwarz methods at the continuous level. We consider the classical alternating Schwarz algorithm (2.2) with the modified transmission conditions

$$\begin{array}{rcll} \mathcal{L}u_1^{n+1} & = & f & \text{in } \Omega_1, \\ \mathcal{B}_1 u_1^{n+1} & = & \mathcal{B}_1 u_2^n & \text{on } \Gamma_1, \end{array} \qquad \begin{array}{rcll} \mathcal{L}u_2^{n+1} & = & f & \text{in } \Omega_2, \\ \mathcal{B}_2 u_2^{n+1} & = & \mathcal{B}_2 u_1^{n+1} & \text{on } \Gamma_2, \end{array} \tag{5.1}$$

where the linear operators $\mathcal{B}_j$ are acting along the interfaces between the subdomains. For a large class of second-order problems, including time dependent ones, one can show for a decomposition into strips that the optimal choice for $\mathcal{B}_j$ is $\partial_{n_j} + DtN_j$, where DtN denotes the non-local Dirichlet to Neumann (or Steklov-Poincaré) operator associated with the second-order elliptic operator $\mathcal{L}$. With this choice and a decomposition into J subdomains, the new Schwarz method converges in J steps, and is thus a direct solver; see [64, 65]. Unfortunately, as the DtN operators are in general non-local in nature, the new algorithm is much more costly to run, and also much more difficult to implement. One is therefore interested in approximating the optimal choice by local operators of the form $\mathcal{B}_j = \partial_{n_j} + p_j + r_j\partial_\tau + q_j\partial_{\tau\tau}$, where ∂_τ denotes the tangential derivative at the interface. One would like to determine the parameters p_j, q_j and r_j such that the method is as effective as possible. This is in general again a difficult problem, but for model situations, such as the plane decomposed into two half planes, or a rectangular domain decomposed into two rectangles, the method can be studied using Fourier analysis. To be more specific, if we consider the positive definite equation $(\eta - \Delta)u = f$, $\eta > 0$, on the plane $\Omega = \mathbb{R}^2$ decomposed into the two half planes $\Omega_1 = (-\infty, L) \times \mathbb{R}$ and $\Omega_2 = (0, \infty) \times \mathbb{R}$, $L \geq 0$, then a Fourier transform in y with Fourier parameter k shows [31] that the contraction factor of (5.1) is of the form

$$\rho = \left(\frac{p + irk + qk^2 - \sqrt{k^2 + \eta}}{p + irk + qk^2 + \sqrt{k^2 + \eta}} \right)^2 e^{-2\sqrt{k^2+\eta}L},$$

ETNA
Kent State University
http://etna.math.kent.edu

where we have assumed for simplicity that $p_j = p$, $q_j = q$ and $r_j = r$. Denoting by $z := ik$, $f_{PDE}(z) := \sqrt{k^2+\eta}$, and letting the polynomial $s(z) := p + irk + qk^2$, we obtain

$$\rho = \rho(z,s) = \left(\frac{s(z) - f_{PDE}(z)}{s(z) + f_{PDE}(z)}\right)^2 e^{-2Lf_{PDE}(z)}, \tag{5.2}$$

and we would get a similar expression for an arbitrary second-order PDE for the same model situation, where f_{PDE} is in general the complex symbol of the associated DtN operator; see [1]. In order to obtain the fastest in a given class of algorithms, determined by the degree n of the polynomial used for the transmission condition (e.g., $n = 0$ for Robin transmission conditions), we need to minimize ρ over all relevant frequencies, i.e., we search for

$$\inf_{s\in\mathbb{P}_n} \sup_{z\in K} |\rho(z,s)|, \tag{5.3}$$

where $\mathbb{P}_n$ is the set of complex polynomials of degree $\leq n$, and K is a bounded or unbounded set in the complex plane. We are thus led to solve a best approximation problem, and the resulting polynomial coefficients give the best possible performance for the associated optimized Schwarz method and the physical problem at hand.

Chebyshev was the first to study best approximation problems, motivated by the mechanics which link the steam engine to the wheel of a locomotive [6], which led him to study the real best approximation problem, i.e., to find the real polynomial p on the interval I which satisfies

$$\min_p \max_{x\in I} |f(x) - p(x)|, \tag{5.4}$$

and he made the Russian style remark (in French),

> … la différence $f(x)-p$ jouira, comme on le sait, de cette propriété : Parmi les valeurs les plus grandes et les plus petites de la différence $f(x)-p$ entre les limites, on trouve au moins $n+2$ fois la même valeur numérique.[5]

without proving it, which is the famous equioscillation property. Only half a century later, De la Vallée Poussin proved in [11] formally existence, uniqueness and equioscillation of the solution for the classical best approximation problem (5.4). Meinardus and Schwedt studied another half a century later in depth linear and non-linear best approximation problems [61], and defined the three fundamental mathematical questions which need to be addressed for such problems:

1. Existiert für jede stetige Funktion $f(x)$ eine Minimallösung?
2. Gibt es zu jedem $f(x)$ genau eine Minimallösung?
3. Wodurch wird die Minimallösung charakterisiert?[6]

The best approximation problem (5.3) from optimized Schwarz methods, which is called a homographic best approximation problem because of the form of the convergence factor in (5.2), was studied only recently; see [1]. For the case without overlap, we have the following result, which answers the three major questions of Meinardus and Schwedt for this case.

[5] … the difference $f(x) - p$ satisfies, *like everybody knows*, the following property: among the largest and smallest values of the difference $f(x) - p$ between their limits, one finds at least $n + 2$ times the same numerical value.

[6] 1. Does a minimal solution exist for any continuous function $f(x)$? 2. Is there precisely one such minimal solution for any $f(x)$? 3. How is this minimal solution characterized?

ETNA
Kent State University
http://etna.math.kent.edu

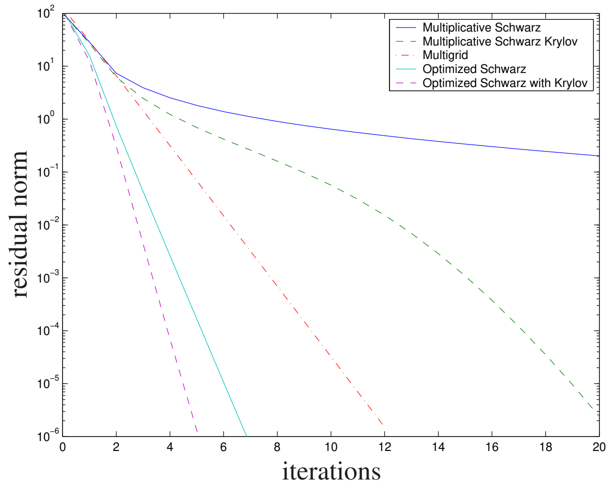

FIGURE 5.1. *Comparison of the classical and optimized multiplicative Schwarz method, used as an iterative solver or as a preconditioner, with a multigrid method.*

THEOREM 5.1. *If $L = 0$ and K is compact, then for every $n \geq 0$ there exists a unique solution s_n^*, and there exist at least $n+2$ points $z_1, \ldots, z_{n+2}$ in K, such that*

$$\left|\frac{s_n^*(z_i) - f(z_i)}{s_n^*(z_i) + f(z_i)}\right| = \left\|\frac{s_n^* - f}{s_n^* + f}\right\|_\infty .$$

With overlap, the situation is more delicate:

THEOREM 5.2. *Let K be a closed set in $\mathbb{C}$, containing at least $n+2$ points. Let f satisfy $\Re f(z) > 0$ and*

$$\Re f(z) \longrightarrow +\infty \quad as \quad z \longrightarrow \infty \quad in\ K.$$

Then, for L small enough, there exists a solution s_n^* and there exist at least $n+2$ points $z_1, \ldots, z_{n+2}$ in K such that

$$\left|\frac{s_n^*(z_i) - f(z_i)}{s_n^*(z_i) + f(z_i)} e^{-Lf(z_i)}\right| = \left\|\frac{s_n^* - f}{s_n^* + f} e^{-Lf}\right\|_\infty = \delta_n(L).$$

In addition, if K is compact, and L satisfies $\delta_n(L)e^{L \sup_{z\in K} \Re f(z)} < 1$, then the solution is unique.

These key results from best approximation allow us to determine the most effective transmission conditions in each family of transmission conditions (Robin or higher order) for a given PDE, and thus the associated optimized Schwarz method. For the case of the positive definite problem $(\eta - \Delta)u = f$, $\eta > 0$, we obtain with the parameters from [31], for the same comparison with the multigrid method as in Figure 4.3, the results shown in Figure 5.1. Clearly, the performance of the method is greatly enhanced. In addition, now Krylov acceleration does not improve the performance by much, the iterative variant is already close to optimal without Krylov acceleration, like multigrid for this same problem.

Over the last ten years, a lot of research has been devoted to study the optimal choice of parameters in the transmission conditions of optimized Schwarz methods, and there are now results available for many classes of PDEs: for steady symmetric problems [12, 13, 31], and [52, 53] for a first analysis of non-straight interfaces; for advection-reaction-diffusion

ETNA
Kent State University
http://etna.math.kent.edu

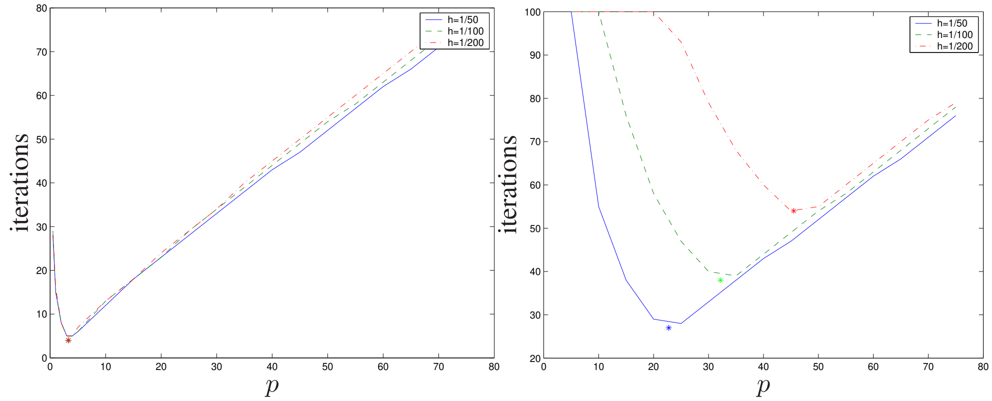

FIGURE 5.2. *Using a zero initial guess and computing a smooth solution, the performance of the optimized Schwarz method does not seem to depend on h on the left. When, however, a random initial guess is used on the right, the dependence appears, as predicted by the theory.*

type problems [22, 42, 43, 44, 45, 51, 62, 63]; for the indefinite Helmholtz case [8, 7, 9, 35, 37, 55, 56]. There are also results for evolution problems, where the algorithms are called optimized Schwarz waveform relaxation: for the heat equation [32], for the unsteady advection reaction diffusion equation [1, 34, 58], for the second-order wave equation [33, 36], and for the shallow water equation [57]. There is also work for problems with discontinuous coefficients [22, 24, 25, 26, 27, 54], and for a more detailed analysis of problems with corners [10]. An interesting relation between optimized Schwarz methods and Schwarz methods for hyperbolic problems using characteristic transmission conditions was found in [16] for the Cauchy-Riemann equation, and then exploited to derive optimized Schwarz methods for Maxwell's equations in [17]. In fluid dynamics, optimized transmission conditions were studied in [19]; in particular, for Euler's equations, see [18, 20]. For an interesting discrete approach, see [67, 73] and the thesis by Tan [72].

Special care must be taken in testing these optimized methods numerically, in order to avoid jumping to false conclusions. In particular, when doing scaling experiments for a diminishing mesh parameter h. Motivated by results in [51], we applied a non-overlapping optimized Schwarz method with Robin transmission conditions and two subdomains to the model problem $(\eta - \Delta)u = f$. It is known in this case that the optimal parameter in the transmission condition is

$$p = \left(\left(k_{\min}^2 + \eta\right) \left(k_{\max}^2 + \eta\right) \right)^{\frac{1}{4}}, \tag{5.5}$$

where the minimal and maximal frequency can be estimated by $k_{\min} \approx \frac{\pi}{l}$ and $k_{\max} \approx \frac{\pi}{h}$, with l denoting the length of the interface, and h the mesh size; see [31]. Hence the optimal parameter depends on the mesh size h. In Figure 5.2, we show on the left how many iterations are needed, as a function of p, when computing a smooth solution starting with a zero initial guess. It seems that the optimal p does not depend on h, and the convergence rate is independent of h, which is in sharp contrast to the analysis in [31]. In Figure 5.2 on the right we show the same set of experiments, but now starting with a random initial guess. Now, clearly, the optimal p depends on h, and the convergence rate deteriorates, as predicted by the analysis in [31]. What has gone wrong with the first experiment? Starting with a zero initial guess and computing a smooth solution, by linearity the error only contains low frequencies, and thus refining h does not add any high frequencies, the behavior of the method remains the same. Starting with a random initial guess ensures that all frequencies

ETNA
Kent State University
http://etna.math.kent.edu

are present in the error, and the method is really tested realistically, since one would only use a mesh fine enough to resolve the features in a real application. The stars in Figure 5.2 denote the optimum according to formula (5.5), where on the left we estimated $k_{\max}$ using the largest frequency in the smooth solution.

5.2. Optimized Schwarz methods at the discrete level. We saw that in the continuous formulation optimized Schwarz methods are obtained from classical ones by changing the transmission conditions. In the discrete Schwarz methods, however, the transmission conditions no longer appear naturally, as we have seen in their formulations in Section 3. One can however show algebraically that it suffices to exchange the subdomain matrices A_j in the multiplicative Schwarz method and the restricted additive Schwarz method, by subdomain matrices representing discretizations of subdomain problems with Robin or more general boundary conditions, in order to obtain the same iterates as discretized optimized Schwarz methods, provided an algebraic condition holds; see [71]. If we take as an example

$$\mathcal{L}u = (\eta - \Delta)u = f, \quad \text{in } (0,1)^2,$$

and use a finite volume discretization, we obtain the discretized system

$$A\boldsymbol{u} = \boldsymbol{f},$$

where the system matrix is of the form

$$A = \frac{1}{h^2}\begin{bmatrix} T_\eta & -I & \\ -I & T_\eta & \ddots \\ & \ddots & \ddots \end{bmatrix}, \quad T_\eta = \begin{bmatrix} \eta h^2 + 4 & -1 & \\ -1 & \eta h^2 + 4 & \ddots \\ & \ddots & \ddots \end{bmatrix}.$$

The classical subdomain matrices used in the discrete Schwarz methods are $A_j = R_j A R_j^T$. In order to obtain optimized subdomain matrices $\tilde{A}_j$, one simply replaces in A_j the interface diagonal blocks T_η by

$$\tilde{T} = \frac{1}{2}T_\eta + phI + \frac{q}{h}(T_0 - 2I), \quad T_0 = T_\eta|_{\eta=0}, \tag{5.6}$$

where p and q are solutions of the associated min-max problem. In numerical linear algebra terms, one modifies slightly the diagonal blocks of an overlapping block Jacobi or block Gauss-Seidel method, where they connect, with neighboring blocks using formula (5.6), and obtains a much more efficient method. The impact of this change is shown in Figure 5.3 for the case of a block Gauss-Seidel or multiplicative Schwarz method with small overlap, where we used for the parameters in (5.6) both low frequency approximations of zeroth and second-order (TO0 and TO2); see [31], and the results of the best approximation problem for this PDE from [31] for a zeroth and second degree polynomial (OO0 and OO2). It is clearly very beneficial to know these parameters.

6. Conclusions. We have shown that discrete Schwarz methods are discretizations of continuous Schwarz methods, with the important exception of the additive Schwarz method with more than minimal overlap, which does not correspond to a continuous iteration per subdomain: in order to remain symmetric for symmetric problems, the method accepts as a compromise non-converging modes in the overlap. These are, however, treated easily when the method is used as a preconditioner for a Krylov method, at the cost of a few more iterations. As an alternative, the restricted additive Schwarz method can be used, which corresponds

ETNA
Kent State University
http://etna.math.kent.edu

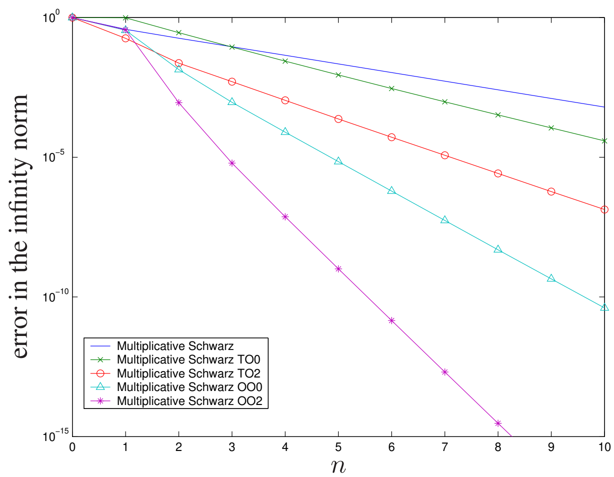

FIGURE 5.3. *Impact of the different diagonal blocks in the optimized multiplicative Schwarz method on the contraction factor of the method.*

to a continuous iteration per subdomain introduced by Lions, namely the parallel Schwarz method, but is non-symmetric, even for symmetric problems.

We have then shown that several drawbacks of the classical Schwarz method, namely the need for overlap, convergence problems for indefinite Helmholtz equations, and slow convergence, have all historically been addressed by introducing one change in the method: different transmission conditions. This motivated the development of optimized Schwarz methods, both at the continuous and discrete level, with significantly enhanced convergence properties. At the discrete level, particular care has to be taken in the case of the additive Schwarz method, because of the non-convergent modes in the overlap, and the need of optimized Schwarz methods to approximate derivatives there.

There are three main open problems in the development of optimized Schwarz methods: first, there is no general convergence proof, neither at the continuous, nor at the discrete level, for overlapping optimized Schwarz methods, although there are interesting results for the special case of two subdomains; see [46] and [50]. Second, the development of coarse grid corrections for optimized Schwarz methods is only at the stage of numerical experiments; see [22]. Finally, it would be very important to develop algebraically optimized Schwarz methods based on the matrix alone, in analogy to the algebraic multigrid methods.

Acknowledgment: The author is very thankful for the detailed comments and precise suggestions of two anonymous referees, which greatly helped to improve this manuscript. He also acknowledges the support of the Swiss National Science Foundation grants 200020-117577/1 and 200021-107988.

REFERENCES

[1] D. BENNEQUIN, M. J. GANDER, AND L. HALPERN, *A homographic best approximation problem with application to optimized Schwarz waveform relaxation*, Math. Comp., 78 (2009), pp. 185–223.
[2] B. L. BUZBEE, G. H. GOLUB, AND C. W. NIELSON, *On direct methods for solving Poisson's equations*, SIAM J. Numer. Anal., 7 (1970), pp. 627–656.
[3] X.-C. CAI, M. DRYA, AND M. SARKIS, *Restricted additive Schwarz preconditioners with harmonic overlap for symmetric positive definite linear systems*, SIAM J. Numer. Anal., 41 (2003), pp. 1209–1231.
[4] X.-C. CAI AND M. SARKIS, *A restricted additive Schwarz preconditioner for general sparse linear systems*, SIAM J. Sci. Comput., 21 (1999), pp. 239–247.

[5] T. F. CHAN AND T. P. MATHEW, *Domain decomposition algorithms*, Acta Numer., 3 (1994), pp. 61–143.
[6] P. L. CHEBYSHEV, *Théorie des mécanismes connus sous le nom de parallélogrammes*, Mem. Acad. Sci. St. Petersb., 7 (1854), pp. 539–568.
[7] P. CHEVALIER, *Méthodes Numériques pour les Tubes Hyperfréquences. Résolution par Décomposition de Domaine*, Ph.D. thesis, Université Paris VI, 1998.
[8] P. CHEVALIER AND F. NATAF, *Symmetrized method with optimized second-order conditions for the Helmholtz equation*, in Domain Decomposition Methods 10, J. Mandel, C. Farhat, X.-C. Cai, eds., (Boulder, 1997), Amer. Math. Soc., Providence, 1998, pp. 400–407.
[9] P. CHEVALIER AND F. NATAF, *Symmetrized method with optimized second-order conditions for the Helmholtz equation*, Contemp. Math., (1998), pp. 400–407.
[10] C. CHNITI, F. NATAF, AND F. NIER, *Improved interface conditions for a non-overlapping domain decomposition of a non-convex polygonal domain*, C. R. Math. Acad. Sci. Paris, 342 (2006), pp. 883–886.
[11] C. J. DE LA VALLÉE-POUSSIN, *Sur les polynômes d'approximation et la représentation approchée d'un angle*, Bull. Acad. Belg., 12 (1910), pp. 808–845.
[12] Q. DENG, *An analysis for a nonoverlapping domain decomposition iterative procedure*, SIAM J. Sci. Comput., 18 (1997), pp. 1517–1525.
[13] Q. DENG, *An optimal parallel nonoverlapping domain decomposition iterative procedure*, SIAM J. Numer. Anal., 41 (2003), pp. 964–982.
[14] B. DESPRÉS, *Décomposition de domaine et problème de Helmholtz*, C. R. Acad. Sci. Paris, 1 (1990), pp. 313–316.
[15] B. DESPRÉS, *Méthodes de Décomposition de Domaine pour les Problèmes de Propagation d'Ondes en Régimes Harmoniques*, Ph.D. thesis, Université Paris IX, 1991.
[16] V. DOLEAN AND M. J. GANDER, *Why classical Schwarz methods applied to hyperbolic systems can converge even without overlap*, in Domain Decomposition Methods in Science and Engineering XVII, U. Langer, M. Discacciati, D. Keyes, O. Widlund, W. Zulehner, eds., Springer Lecture Notes in Computational Science and Engineering 60 (2007), pp. 467–475.
[17] V. DOLEAN, M. J. GANDER, AND L. GERARDO-GIORDA, *Optimized Schwarz methods for Maxwell's equations*, SIAM J. Sci. Comput., to appear (2009).
[18] V. DOLEAN, S. LANTERI, AND F. NATAF, *Construction of interface conditions for solving compressible Euler equations by non-overlapping domain decomposition methods*, Int. J. Numer. Meth. Fluids, 40 (2002), pp. 1485–1492.
[19] ———, *Optimized interface conditions for domain decomposition methods in fluid dynamics*, Int. J. Numer. Meth. Fluids, 40 (2002), pp. 1539–1550.
[20] ———, *Convergence analysis of a Schwarz type domain decomposition method for the solution of the Euler equations*, Appl. Num. Math., 49 (2004), pp. 153–186.
[21] M. DRYJA AND O. B. WIDLUND, *Some domain decomposition algorithms for elliptic problems*, in Iterative Methods for Large Linear Systems, L. Hayes and D. Kincaid, eds., Academic Press, San Diego, 1989, pp. 273–291.
[22] O. DUBOIS, *Optimized Schwarz Methods for the Advection-Diffusion Equation and for Problems with Discontinuous Coefficients*, Ph.D. thesis, McGill University, June 2007.
[23] E. EFSTATHIOU AND M. J. GANDER, *Why restricted additive Schwarz converges faster than additive Schwarz*, BIT, 43 (2003), pp. 945–959.
[24] I. FAILLE, E. FLAURAUD, F. NATAF, F. SCHNEIDER, AND F. WILLIEN, *Optimized interface conditions for sedimentary basin modeling*, in Domain Decomposition Methods in Science and Engineering, N. Debit, M. Garbay, R. Hoppe, D. Keyes, Y. Kuznetsov, J. Périaux, eds., International Center for Numerical Methods in Engineering (CIMNE), Barcelona, Spain, 2002, pp. 461–468.
[25] E. FLAURAUD AND F. NATAF, *Optimized interface conditions in domain decomposition methods. application at the semi-discrete and at the algebraic level to problems with extreme contrasts in the coefficients*, Tech. Report 524, CMAP École Polytechnique, Paris, March 2004.
[26] E. FLAURAUD, F. NATAF, I. FAILLE, AND R. MASSON, *Domain decomposition for an asymptotic geological fault modeling*, C. R. Mech. Acad. Sci., 331 (2003), pp. 849–855.
[27] E. FLAURAUD, F. NATAF, AND F. WILLIEN, *Optimized interface conditions in domain decomposition methods for problems with extreme contrasts in the coefficients*, J. Comput. Appl. Math., 189 (2006), pp. 539–554.
[28] A. FROMMER AND D. B. SZYLD, *Weighted max norms, splittings, and overlapping additive Schwarz iterations*, Numer. Math., 83 (1999), pp. 259–278.
[29] A. FROMMER AND D. B. SZYLD, *An algebraic convergence theory for restricted additive Schwarz methods using weighted max norms*, SIAM J. Numer. Anal., 39 (2001), pp. 463–479.
[30] M. J. GANDER, *Optimized Schwarz methods for Helmholtz problems*, in Domain Decomposition Methods in Science and Engineering, N. Debit, M. Garbay, R. Hoppe, D. Keyes, Y. Kuznetsov, J. Périaux, eds., International Center for Numerical Methods in Engineering (CIMNE), Barcelona, Spain, 2002, pp. 245–252.

[31] ———, *Optimized Schwarz methods*, SIAM J. Numer. Anal., 44 (2006), pp. 699–731.
[32] M. J. GANDER AND L. HALPERN, *Méthodes de relaxation d'ondes pour l'équation de la chaleur en dimension 1*, C. R. Acad. Sci. Paris Sér. I Math., 336 (2003), pp. 519–524.
[33] ———, *Absorbing boundary conditions for the wave equation and parallel computing*, Math. Comp., 74 (2004), pp. 153–176.
[34] ———, *Optimized Schwarz waveform relaxation methods for advection reaction diffusion problems*, SIAM J. Numer. Anal., 45 (2007), pp. 666–697.
[35] M. J. GANDER, L. HALPERN, AND F. MAGOULÈS, *An optimized Schwarz method with two-sided Robin transmission conditions for the Helmholtz equation*, Int. J. for Num. Meth. in Fluids, 55 (2007), pp. 163–175.
[36] M. J. GANDER, L. HALPERN, AND F. NATAF, *Optimal Schwarz waveform relaxation for the one dimensional wave equation*, SIAM J. Numer. Anal., 41 (2003), pp. 1643–1681.
[37] M. J. GANDER, F. MAGOULÈS, AND F. NATAF, *Optimized Schwarz methods without overlap for the Helmholtz equation*, SIAM J. Sci. Comput., 24 (2002), pp. 38–60.
[38] M. J. GANDER AND H. ZHAO, *Overlapping Schwarz waveform relaxation for the heat equation in n-dimensions*, BIT, 42 (2002), pp. 779–795.
[39] G. H. GOLUB AND D. MAYERS, *The use of preconditioning over irregular regions*, in Computing Methods in Applied Sciences and Engineering, VI, R. Glowinski and J. L. Lions, eds., North-Holland, Amsterdam, New York, Oxford, 1984, pp. 3–14.
[40] W. HACKBUSCH, *Iterative Solution of Large Sparse Linear Systems of Equations*, Springer, Berlin, 1994.
[41] T. HAGSTROM, R. P. TEWARSON, AND A. JAZCILEVICH, *Numerical experiments on a domain decomposition algorithm for nonlinear elliptic boundary value problems*, Appl. Math. Lett., 1 (1988).
[42] C. JAPHET, *Conditions aux limites artificielles et décomposition de domaine: Méthode OO2 (optimisé d'ordre 2). Application à la résolution de problèmes en mécanique des fluides*, Tech. Report 373, CMAP École Polytechnique, Paris, 1997.
[43] C. JAPHET AND F. NATAF, *The best interface conditions for domain decomposition methods: Absorbing boundary conditions*. in 'Absorbing Boundaries and Layers, Domain Decomposition Methods. Applications to Large Scale Copmutations', L. Tourrette and L. Halpern, eds., Nova Science Publishers, Inc., New York (2001), pp. 348–373.
[44] C. JAPHET, F. NATAF, AND F. ROGIER, *The optimized order 2 method. application to convection-diffusion problems*, Future Gen. Comput. Sys., 18 (2001), pp. 17–30.
[45] C. JAPHET, F. NATAF, AND F.-X. ROUX, *The Optimized Order 2 Method with a coarse grid preconditioner. Application to convection-diffusion problems*, in Ninth International Conference on Domain Decompositon Methods in Science and Engineering, P. Bjorstad, M. Espedal, and D. Keyes, eds., John Wiley & Sons, 1998, pp. 382–389.
[46] J.-H. KIMN, *A convergence theory for an overlapping Schwarz algorithm using discontinuous iterates*, Numer. Math., 100 (2005), pp. 117–139.
[47] P.-L. LIONS, *On the Schwarz alternating method. I.*, in First International Symposium on Domain Decomposition Methods for Partial Differential Equations, R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, eds., SIAM, Philadelphia, 1988, pp. 1–42.
[48] ———, *On the Schwarz alternating method. II.*, in Domain Decomposition Methods, T. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., SIAM, Philadelphia, 1989, pp. 47–70.
[49] ———, *On the Schwarz alternating method. III: a variant for nonoverlapping subdomains*, in Third International Symposium on Domain Decomposition Methods for Partial Differential Equations , held in Houston, Texas, March 20-22, 1989, T. F. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., SIAM, Philadelphia, 1990.
[50] S. LOISEL AND D. B. SZYLD, *On the convergence of algebraic optimizable Schwarz methods with applications to elliptic problems*, Research Report 07-11-16, Department of Mathematics, Temple University, November 2007.
[51] G. LUBE, L. MUELLER, AND F.-C. OTTO, *A non-overlapping domain decomposition method for the advection-diffusion problem*, Computing, 64 (2000), pp. 49–68.
[52] S. H. LUI, *A Lions non-overlapping domain decomposition method for domains with an arbitrary interface*, IMA J. Numer. Anal. 29 (2009), pp. 332–349.
[53] ———, *On the condition number of an optimized Schwarz method*, preprint, Department of Mathematics, University of Manitoba, 2007.
[54] Y. MADAY AND F. MAGOULÈS, *Non-overlapping additive Schwarz methods tuned to highly heterogeneous media*, C. R. Acad. Sci. Paris Sér. I Math., 341 (2005), pp. 701–705.
[55] F. MAGOULÈS, P. IVÁNYI, AND B. TOPPING, *Convergence analysis of Schwarz methods without overlap for the Helmholtz equation*, Comput. & Structures, 82 (2004), pp. 1835–1847.
[56] ———, *Non-overlapping Schwarz methods with optimized transmission conditions for the Helmholtz equation*, Comput. Methods Appl. Mech. Engrg., 193 (2004), pp. 4797–4818.
[57] V. MARTIN, *Méthodes de Décomposition de Domaines de Type Relaxation d'Ondes pour des Équations de*

l'Océnographie, Ph.D. thesis, Université Paris XIII, December 2003.
[58] V. MARTIN, *An optimized Schwarz waveform relaxation method for unsteady convection diffusion equation*, Appl. Numer. Math., 52 (2005), pp. 401–428.
[59] T. MATHEW, *Uniform convergence of the Schwarz alternating method for solving singularly perturbed advection-diffusion equations*, SIAM J. Numer. Anal., 35 (1998), pp. 1663–1683.
[60] A. M. MATSOKIN AND S. V. NEPOMNYASCHIKH, *A Schwarz alternating method in a subspace*, Soviet Math. (Iz. VUZ), 29 (1985), pp. 78–84.
[61] G. MEINARDUS, *Approximation von Funktionen und ihre Numerische Behandlung*, Springer, Berlin, Göttingen, Heidelberg, New York, 1964.
[62] F. NATAF, *Absorbing boundary conditions in block Gauss-Seidel methods for convection problems*, Math. Models Methods Appl. Sci., 6 (1996), pp. 481–502.
[63] F. NATAF AND F. NIER, *Convergence rate of some domain decomposition methods for overlapping and nonoverlapping subdomains*, Numer. Math., 75 (1997), pp. 357–77.
[64] F. NATAF AND F. ROGIER, *Factorization of the convection-diffusion operator and the Schwarz algorithm*, Math. Models Methods Appl. Sci., 5 (1995), pp. 67–93.
[65] F. NATAF, F. ROGIER, AND E. DE STURLER, *Optimal interface conditions for domain decomposition methods*, Tech. Report 301, CMAP École Polytechnique, Paris, 1994.
[66] A. QUARTERONI AND A. VALLI, *Domain Decomposition Methods for Partial Differential Equations*, Oxford Science Publications, 1999.
[67] H. SAN AND W. P. TANG, *An overdetermined Schwarz alternating method*, SIAM J. Sci. Comput., 7 (1996), pp. 884–905.
[68] H. A. SCHWARZ, *Über einen Grenzübergang durch alternierendes Verfahren*, Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich, 15 (1870), pp. 272–286.
[69] B. F. SMITH, P. E. BJØRSTAD, AND W. GROPP, *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*, Cambridge University Press, 1996.
[70] S. L. SOBOLEV, *L'algorithme de Schwarz dans la théorie de l'elasticité*, Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS, IV (1936), pp. 243–246.
[71] A. ST-CYR, M. J. GANDER, AND S. THOMAS, *Optimized multiplicative, additive and restricted additive Schwarz preconditioning*, SIAM J. Sci. Comput., 29 (2007), pp. 2402–2425.
[72] K. H. TAN, *Local Coupling in Domain Decomposition*, Ph.D. thesis, Utrecht University, April 1996.
[73] W. P. TANG, *Generalized Schwarz splittings*, SIAM J. Sci. Stat. Comput., 13 (1992), pp. 573–595.
[74] A. TOSELLI AND O. WIDLUND, *Domain Decomposition Methods - Algorithms and Theory*, vol. 34 of Springer Series in Computational Mathematics, Springer, Berlin, Heidelberg, 2004.