1

Introduction

1.1 Basic Ideas of Domain Decomposition

The basic ideas of domain decomposition are quite natural and simple. Consider the Poisson equation on a region Ω, in two or three dimensions, with zero Dirichlet data given on $\partial\Omega$, the boundary of Ω. Suppose also that Ω is partitioned into two nonoverlapping subdomains Ω_i :

$$\overline{\Omega} = \overline{\Omega_1 \cup \Omega_2}, \quad \Omega_1 \cap \Omega_2 = \emptyset, \quad \Gamma = \partial\Omega_1 \cap \partial\Omega_2;$$

see Fig. 1.1. We also assume that

$$\text{measure}(\partial\Omega_1 \cap \partial\Omega) > 0, \quad \text{measure}(\partial\Omega_2 \cap \partial\Omega) > 0,$$

and that the boundaries of the subdomains are Lipschitz continuous, and consider the following problem:

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega. \end{aligned} \tag{1.1}$$

Under suitable regularity assumptions on f and the boundaries of the subdomains, typically f square-summable and the boundaries Lipschitz, problem (1.1) is equivalent to the following coupled problem:

$$\begin{aligned} -\Delta u_1 &= f && \text{in } \Omega_1, \\ u_1 &= 0 && \text{on } \partial\Omega_1 \setminus \Gamma, \\ u_1 &= u_2 && \text{on } \Gamma, \\ \frac{\partial u_1}{\partial n_1} &= -\frac{\partial u_2}{\partial n_2} && \text{on } \Gamma, \\ -\Delta u_2 &= f && \text{in } \Omega_2, \\ u_2 &= 0 && \text{on } \partial\Omega_2 \setminus \Gamma. \end{aligned} \tag{1.2}$$

Here u_i is the restriction of u to Ω_i and $\mathbf{n}_i$ the outward normal to Ω_i. This equivalence can be proven by considering the corresponding variational problems; see [392, Sect. 1.2]. The conditions on the interface Γ are called *transmission conditions* and they are also equivalent to the equality of any two independent linear combinations of the traces of the functions and their normal derivatives. In the following, we will also refer to the normal derivative as the *flux*.

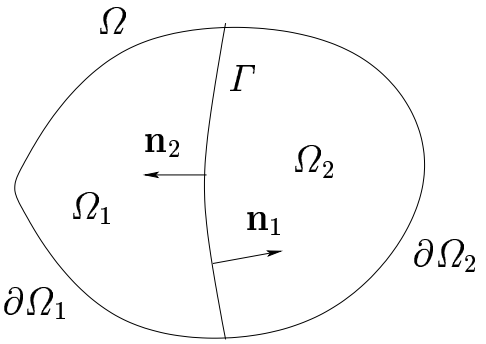

Fig. 1.1. Partition into two nonoverlapping subdomains.

Remark 1.1. The following one-dimensional example shows that some regularity beyond $f \in H^{-1}(\Omega)$ is required. Let u be the weak solution of

$$\begin{aligned} -\frac{d^2u}{dx^2} &= -2\,\delta \qquad \text{in } (-1,1), \\ u(-1) = u(1) &= 0, \end{aligned} \tag{1.3}$$

where $\delta(x)$ is the delta function. The unique weak solution $u \in H_0^1(-1,1)$ is

$$u(x) = \begin{cases} -1 - x & x \le 0, \\ -1 + x & x \ge 0, \end{cases}$$

and its derivative has a jump at $x = 0$.

We note that this particular problem is quite relevant to domain decomposition theory. In many algorithms, we will first eliminate all nonzero components of the right hand side, of a finite element approximation, except those on the interface, in this case $x = 0$. We are then left with an equation for the remaining finite element error, which is a direct analog of equation (1.3).

1.2 Matrix and Vector Representations

In this section, we consider matrix and vector representations of certain operators and linear functionals; we refer to appendix B for additional details.

Starting with any domain decomposition algorithm written in terms of functions and operators, we will be able to rewrite it in matrix form as a preconditioned iterative method for a certain linear system.

We now consider a triangulation of the domain Ω and a finite element approximation of problem (1.1). We always assume that subdomains consist of unions of elements or, equivalently, that subdomain boundaries do not cut through any elements. Such an approximation gives rise to a linear system

$$Au = f \tag{1.4}$$

with a symmetric, positive definite matrix which, for a mesh size of h, typically has a condition number on the order of $1/h^2$. Here,

$$A = \begin{pmatrix} A_{II}^{(1)} & 0 & A_{I\Gamma}^{(1)} \\ 0 & A_{II}^{(2)} & A_{I\Gamma}^{(2)} \\ A_{\Gamma I}^{(1)} & A_{\Gamma I}^{(2)} & A_{\Gamma\Gamma} \end{pmatrix}, \quad u = \begin{pmatrix} u_I^{(1)} \\ u_I^{(2)} \\ u_\Gamma \end{pmatrix}, \quad f = \begin{pmatrix} f_I^{(1)} \\ f_I^{(2)} \\ f_\Gamma \end{pmatrix}, \tag{1.5}$$

where we have partitioned the degrees of freedom into those internal to Ω_1, and to Ω_2, and those of the interior of Γ.

The stiffness matrix A and the load vector f can be obtained by subassembling the corresponding components contributed by the two subdomains. Indeed, if

$$f^{(i)} = \begin{pmatrix} f_I^{(i)} \\ f_\Gamma^{(i)} \end{pmatrix}, \quad A^{(i)} = \begin{pmatrix} A_{II}^{(i)} & A_{I\Gamma}^{(i)} \\ A_{\Gamma I}^{(i)} & A_{\Gamma\Gamma}^{(i)} \end{pmatrix}, \quad i = 1, 2, \tag{1.6}$$

are the right hand sides and the local stiffness matrices for Poisson problems with a Dirichlet condition on $\partial\Omega_i \setminus \Gamma$ and a Neumann condition on Γ, we have

$$A_{\Gamma\Gamma} = A_{\Gamma\Gamma}^{(1)} + A_{\Gamma\Gamma}^{(2)}, \quad f_\Gamma = f_\Gamma^{(1)} + f_\Gamma^{(2)}.$$

In view of the transmission conditions in (1.2), we will look for an *approximation* of the normal derivatives on Γ. Given the local exact solution u_i, its normal derivative can be defined as a linear functional by using Green's formula. Thus, if ϕ_j is a nodal basis function for a node on Γ, we have, using (1.2),

$$\int_\Gamma \frac{\partial u_i}{\partial n_i}\phi_j \, ds = \int_{\Omega_i} (\Delta u_i \phi_j + \nabla u_i \cdot \nabla \phi_j) \, dx = \int_{\Omega_i} (-f\phi_j + \nabla u_i \cdot \nabla \phi_j) \, dx.$$

An approximation, $\lambda^{(i)}$, of the functional representing the normal derivative can be found by replacing the exact solution u_i in the right hand side with its finite element approximation. Letting j run over the nodes on Γ and using the definition of the local stiffness matrix, we introduce the expression

$$\lambda^{(i)} = A_{\Gamma I}^{(i)} u_I^{(i)} + A_{\Gamma\Gamma}^{(i)} u_\Gamma^{(i)} - f_\Gamma^{(i)}. \tag{1.7}$$

This approximation, not surprisingly, coincides with the residual corresponding to the nodes on Γ of a Poisson problem with a Neumann condition on Γ; see (1.6). We also note that $\lambda^{(i)}$ is not obtained from the normal derivative of the finite element solution but as an approximation of the linear functional of the exact solution.

Using these definitions, we can find an approximation of problem (1.2):

$$\begin{aligned}
&A_{II}^{(1)} u_I^{(1)} + A_{I\Gamma}^{(1)} u_\Gamma^{(1)} = f_I^{(1)}, \\
&\\
&u_\Gamma^{(1)} = u_\Gamma^{(2)} = u_\Gamma, \\
&(A_{\Gamma I}^{(1)} u_I^{(1)} + A_{\Gamma\Gamma}^{(1)} u_\Gamma^{(1)} - f_\Gamma^{(1)}) = -(A_{\Gamma I}^{(2)} u_I^{(2)} + A_{\Gamma\Gamma}^{(2)} u_\Gamma^{(2)} - f_\Gamma^{(2)}) = \lambda_\Gamma, \\
&\\
&A_{II}^{(2)} u_I^{(2)} + A_{I\Gamma}^{(2)} u_\Gamma^{(2)} = f_I^{(2)}.
\end{aligned} \tag{1.8}$$

We note that the first and last equations of (1.8) are discretizations of Poisson problems for the interior functions $u_I^{(i)}$ with Dirichlet data which vanishes on $\partial\Omega_i \setminus \Gamma$ and is equal to a common value u_Γ on Γ. Alternatively, the first and third equations provide a discretization of a Poisson problem in Ω_1 for the local function u_1 with Neumann data equal to λ_Γ and vanishing Dirichlet data on $\partial\Omega_1 \setminus \Gamma$. An analogous Neumann problem in Ω_2 is provided by the third and fourth equations. The solution of these local problems with suitable Dirichlet and Neumann data provide the building blocks of the nonoverlapping methods in section 1.3.

We note, finally, that while the equivalence of (1.1) and (1.2) might not be immediate, the equivalence of (1.4) and (1.8) is trivial. If $u_\Gamma^{(1)} = u_\Gamma^{(2)} = u_\Gamma$, then the third equation of (1.8), which ensures the equality of the approximations of the linear functionals representing the normal derivatives, coincides with the third row of the original linear system (1.4) because of (1.5) and (1.6). More importantly, while the continuous problem (1.2) is valid only if the right hand side f is sufficiently regular, its discrete counterpart (1.8) is always valid, since it can be found directly from the finite element problem.

If we, in particular, consider the one-dimensional problem of Remark 1.1 with $\Omega_1 = (-1, 0)$ and $\Omega_2 = (0, 1)$ and a triangulation with a node at $x = 0$, we see that the right hand side f_Γ cannot be built by summing two local components. However, problem (1.8) is still equivalent to the original linear system (1.4) since in (1.8) the local components $f_\Gamma^{(i)}$ are not needed but only their sum f_Γ. A full discussion of the continuous problem in two or more dimensions is relatively complicated. We note that the methods discussed in section 1.3 will have nonzero residuals only at nodal points on the interface after the completion of the first iteration step and this is the analog of a solution of the continuous problem with a nonzero jump in the normal derivative across the interface. However, as we have seen, no genuine technical problems remain once we turn to the finite element case.

1.3 Nonoverlapping Methods

We will first consider some simple *iterative substructuring* methods that rely on a partition into nonoverlapping subdomains. We refer to Chap. 4, 5, and 6 for a systematic presentation and further details and generalizations. As we will show, these methods are indeed preconditioned iterative methods for the boundary value u_Γ or for the normal derivative λ_Γ. We start with these methods since they are derived directly from the coupled problems (1.2) and (1.8). We note that domain decomposition methods based on overlapping partitions were the first to be devised (namely, the alternating Schwarz method on overlapping subdomains); they will be introduced in Sect. 1.4.

1.3.1 An Equation for u_Γ: the Schur Complement System

Let us consider the linear system (1.4) with A, u, and f defined in (1.5). In a first step of many iterative domain decomposition methods, the unknowns in the interior of the subdomains ($u_I^{(i)}$) are eliminated. This corresponds to a block factorization of the matrix of (1.5):

$$A = LR = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ A_{\Gamma I}^{(1)} {A_{II}^{(1)}}^{-1} & A_{\Gamma I}^{(2)} {A_{II}^{(2)}}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{II}^{(1)} & 0 & A_{I\Gamma}^{(1)} \\ 0 & A_{II}^{(2)} & A_{I\Gamma}^{(2)} \\ 0 & 0 & S \end{pmatrix} \tag{1.9}$$

and a resulting linear system

$$\begin{pmatrix} A_{II}^{(1)} & 0 & A_{I\Gamma}^{(1)} \\ 0 & A_{II}^{(2)} & A_{I\Gamma}^{(2)} \\ 0 & 0 & S \end{pmatrix} u = \begin{pmatrix} f_I^{(1)} \\ f_I^{(2)} \\ g_\Gamma \end{pmatrix}. \tag{1.10}$$

Here I is the identity matrix and $S = A_{\Gamma\Gamma} - A_{\Gamma I}^{(1)} {A_{II}^{(1)}}^{-1} A_{I\Gamma}^{(1)} - A_{\Gamma I}^{(2)} {A_{II}^{(2)}}^{-1} A_{I\Gamma}^{(2)}$ is the Schur complement relative to the unknowns on Γ. By a direct calculation, we see that S and g_Γ can be found by subassembling local contributions. In particular, recalling the form of the local matrices (1.6) and defining the local Schur complements by

$$S^{(i)} := A_{\Gamma\Gamma}^{(i)} - A_{\Gamma I}^{(i)} {A_{II}^{(i)}}^{-1} A_{I\Gamma}^{(i)}, \quad i = 1, 2, \tag{1.11}$$

we find the Schur complement system for u_Γ to be

$$S u_\Gamma = g_\Gamma, \tag{1.12}$$

with

$$S = S^{(1)} + S^{(2)},$$

$$g_\Gamma = (f_\Gamma^{(1)} - A_{\Gamma I}^{(1)} {A_{II}^{(1)}}^{-1} f_I^{(1)}) + (f_\Gamma^{(2)} - A_{\Gamma I}^{(2)} {A_{II}^{(2)}}^{-1} f_I^{(2)}) =: g_\Gamma^{(1)} + g_\Gamma^{(2)}.$$

We note that once u_Γ is found, by solving (1.12), the internal components can be found by using (1.10):

$$u_I^{(i)} = {A_{II}^{(i)}}^{-1}(f_I^{(i)} - A_{I\Gamma}^{(i)} u_\Gamma); \tag{1.13}$$

these are just solutions of two inhomogeneous Dirichlet problems.

The Schur complement system, which provides an equation for the approximation of the trace of the exact solution on Γ, has been derived purely algebraically by block Gaussian elimination. It is interesting to note that it can also be obtained by using the transmission conditions of the coupled system. Thus, let us write the internal variables $u_I^{(1)}$ and $u_I^{(2)}$ in terms of u_Γ by using the first and last equations of (1.8). Substituting these expressions into the third equation of (1.8), we again arrive at equation (1.12). We note that in this last step, we use the flux condition expressed in the third equation of (1.8). A dual procedure that provides an equation for λ_Γ is given in subsection 1.3.2.

We can also obtain an equation for the trace of the exact solution on Γ working directly with the continuous problem (1.2). The corresponding operator is called a *Steklov-Poincaré* operator. The Schur complement system (1.12) is an approximation of the Steklov-Poincaré equation, *determined* directly by the finite element approximation, particularly, by the approximation of the normal derivatives (1.7). (We refer to Chap. 4 for a systematic presentation of Schur complement systems.)

1.3.2 An Equation for the Flux

We now derive an equation for the normal derivative λ_Γ on Γ by employing a procedure analogous to that of the previous subsection. We use the unknown common boundary value $\lambda_\Gamma = \lambda_\Gamma^{(1)} = -\lambda_\Gamma^{(2)}$ (see the third equation of problem (1.8)) and solve local Neumann problems to find $u^{(1)}$ and $u^{(2)}$:

$$\begin{pmatrix} A_{II}^{(i)} & A_{I\Gamma}^{(i)} \\ A_{\Gamma I}^{(i)} & A_{\Gamma\Gamma}^{(i)} \end{pmatrix} \begin{pmatrix} u_I^{(i)} \\ u_\Gamma^{(i)} \end{pmatrix} = \begin{pmatrix} f_I^{(i)} \\ f_\Gamma^{(i)} + \lambda_\Gamma^{(i)} \end{pmatrix}, \quad i = 1, 2. \tag{1.14}$$

Using a block factorization of the local matrices, we find

$$u_\Gamma^{(i)} = S^{(i)-1}(g_\Gamma^{(i)} + \lambda_\Gamma^{(i)}),$$

with $g_\Gamma^{(i)}$ given as in the previous subsection. Using the second equation of problem (1.8), which makes $u_\Gamma^{(1)}$ and $u_\Gamma^{(2)}$ the same, we find

$$F\lambda_\Gamma = d_\Gamma, \tag{1.15}$$

with

$$F = S^{(1)-1} + S^{(2)-1},$$
$$d_\Gamma = d_\Gamma^{(1)} + d_\Gamma^{(2)} := -S^{(1)-1} g_\Gamma^{(1)} + S^{(2)-1} g_\Gamma^{(2)}. \tag{1.16}$$

We note that this is the same linear system which will appear as (6.29) in section 6.3.1 if we specialize that formula to the case of two subdomains. Once λ_Γ is known, we can find the local solutions $u^{(1)}$ and $u^{(2)}$ by solving the two Neumann problems in (1.14).

In the rest of this section, we will consider some domain decomposition methods of *iterative substructuring* type. We note that many others can be defined by replacing Dirichlet and/or Neumann conditions on Γ by more general ones involving linear combinations of u and the normal derivative; see, e.g, [392, Sect. 1.3] for some of them; our presentation is similar to that of that reference. However, we will only work with the coupled differential and algebraic problems (1.2) and (1.8), respectively. Our purpose is to illustrate that many domain decomposition methods derived from (1.2) are indeed preconditioned Richardson methods for the Schur complement system (1.12) or for the equation (1.15). We note that the methods presented in the following can be derived purely algebraically using suitable splittings of the system (1.8), without any reference to the underlying continuous operator, traces, or normal derivatives. However, employing a functional framework does not only give an interpretation of these iteration procedures, but will also turn out to be crucial as a preparation for the analysis of their convergence rates. For many domain decomposition algorithms this will be carried out using equivalences between certain discrete and Sobolev norms and certain Sobolev type inequalities for finite element functions; the systematic development of this theory will begin in Chap. 4.

All the algorithms to be introduced in the next few subsections involve preconditioners in solving equations (1.12) or (1.15). We note that we could also solve these equations by using a Richardson or conjugate gradient method without preconditioning. The evaluation of S times a vector involves the solution of one Dirichlet problem on each subdomain while that of F times a vector requires the solution of one Neumann problem on each subdomain. The rate of convergence of such algorithms is determined by the condition numbers of S and F. However, even if the Schur complement has a smaller condition number than the original stiffness matrix, the number of iterations will increase in proportion to $1/h$ when the mesh size h decreases; see Sect. 4.3 for details. We also note that the conjugate gradient method will be faster than the Richardson method and that it requires no a priori spectral information while the optimal choice of the parameter θ of the Richardson method involves obtaining estimates of the smallest and largest eigenvalues of the operator; see appendix C.3.

1.3.3 The Dirichlet-Neumann Algorithm

Methods of this type were first considered in [61, 73, 211, 343, 57, 344]. Extensions, also including global coarse solvers, were considered in [176, 173, 465].

The basic Dirichlet-Neumann algorithm consists of two fractional steps corresponding to the two subregions $\Omega_i, i = 1, 2$. Given an initial guess u_Γ^0, we first solve a Dirichlet problem in Ω_1 with Dirichlet data u_Γ^0 on Γ, and then a mixed Neumann-Dirichlet problem on Ω_2 with a Neumann condition on Γ determined by the solution in Ω_1 obtained in the previous step and with Dirichlet conditions on the rest of $\partial\Omega_2$. The new iterate u_Γ^1 is chosen as the trace of the solution in Ω_2, or, more generally, as a linear combination of this trace and u_Γ^0, using a suitably chosen relaxation parameter θ; see appendix C.3. In terms of differential operators (see (1.2)), we can write, for $n \geq 0$:

$$(D) \begin{cases} -\Delta u_1^{n+1/2} = f & \text{in } \Omega_1, \\ u_1^{n+1/2} = 0 & \text{on } \partial\Omega_1 \setminus \Gamma, \\ u_1^{n+1/2} = u_\Gamma^n & \text{on } \Gamma, \end{cases}$$

$$(N) \begin{cases} -\Delta u_2^{n+1} = f & \text{in } \Omega_2, \\ u_2^{n+1} = 0 & \text{on } \partial\Omega_2 \setminus \Gamma, \\ \dfrac{\partial u_2^{n+1}}{\partial n_2} = -\dfrac{\partial u_1^{n+1/2}}{\partial n_1} & \text{on } \Gamma, \end{cases} \tag{1.17}$$

$$u_\Gamma^{n+1} = \theta u_2^{n+1} + (1-\theta) u_\Gamma^n \text{ on } \Gamma,$$

with $\theta \in (0, \theta_{\max})$. Using our approximation for the normal derivatives, i.e. (1.7), we can derive the corresponding iteration for the discrete problem. If we define the vectors of internal degrees of freedom as $v_1 = u_I^{(1)}$ and $w_2 = u_I^{(2)}$, cf. (1.7), we find

$$(D)\ A_{II}^{(1)} v_1^{n+1/2} + A_{I\Gamma}^{(1)} u_\Gamma^n = f_I^{(1)},$$

$$(N)\ \begin{pmatrix} A_{II}^{(2)} & A_{I\Gamma}^{(2)} \\ A_{\Gamma I}^{(2)} & A_{\Gamma\Gamma}^{(2)} \end{pmatrix} \begin{pmatrix} w_2^{n+1} \\ \tilde{u}_\Gamma^{n+1} \end{pmatrix} = \begin{pmatrix} f_I^{(2)} \\ f_\Gamma^{(2)} - \lambda_\Gamma^{n+1/2} \end{pmatrix}, \tag{1.18}$$

$$u_\Gamma^{n+1} = \theta \tilde{u}_\Gamma^{n+1} + (1-\theta) u_\Gamma^n,$$

with

$$\lambda_\Gamma^{n+1/2} = A_{\Gamma I}^{(1)} v_1^{n+1/2} + A_{\Gamma\Gamma}^{(1)} u_\Gamma^n - f_\Gamma^{(1)}.$$

It is clear that (1.18) arises from a splitting of the original system (1.8) and thus provides a consistent iterative method for its solution, i.e., the limit of any convergent sequence will satisfy the correct set of equations.

We next eliminate $v_1^{n+1/2}$ from (1.18) and find

$$\lambda_\Gamma^{n+1/2} = -(g_\Gamma^{(1)} - S^{(1)} u_\Gamma^n).$$

Using then a block factorization of the local matrix $A^{(2)}$ and eliminating w_2^{n+1} from (1.18) yields the following equation:

$$S^{(2)}(u_\Gamma^{n+1} - u_\Gamma^n) = \theta(g_\Gamma - S u_\Gamma^n),$$

which shows that the Dirichlet-Neumann algorithm is a preconditioned Richardson iteration for the Schur complement system (1.12), with the preconditioner ${S^{(2)}}^{-1}$. The preconditioned operator is

$${S^{(2)}}^{-1} S = I + {S^{(2)}}^{-1} S^{(1)},$$

the application of which to a vector involves the solution of a Dirichlet problem, (a multiplication by $S^{(1)}$), and a multiplication by ${S^{(2)}}^{-1}$ which corresponds to solving a problem with Neumann conditions on Γ and Dirichlet conditions of the rest of $\partial\Omega_2$; see section 4.3 for further details.

We note that the spectral equivalence between $S^{(1)}$ and $S^{(2)}$, and thus a uniform bound for the condition number of ${S^{(2)}}^{-1} S$, is ensured by the existence of stable, discrete harmonic, finite element extensions $\mathcal{H}_i u_\Gamma$ from the interface Γ into the subdomains Ω_i; these matters are discussed systematically in section 4.6, in particular in Lemma 4.10. Here, and in what follows, the condition number of the preconditioned operator is the ratio of the largest and smallest eigenvalues of the symmetric generalized eigenvalue problem defined by the operator and its preconditioner, in the case at hand by S and $S^{(2)}$; see appendix C.5. We employ the following definition:

Definition 1.2 (Optimality). *An iterative method for the solution of a linear system is said to be optimal, if its rate of convergence to the exact solution is independent of the size of the system.*

We note that for the algebraic systems considered in this monograph optimality is ensured if the rate of convergence is independent of the size of the finite element spaces employed, and therefore of the meshsize h for h approximations or of the polynomial degree for spectral element approximations.

If we denote the appropriate trace seminorm by $|\cdot|_{1/2,\Gamma}$, we will show in Lemma 4.10 that

$$|\mathcal{H}_i u_\Gamma|_{1,\Omega_i}^2 = (\mathcal{H}_i u_\Gamma)^T A^{(i)} (\mathcal{H}_i u_\Gamma) \leq C_i \, |u_\Gamma|_{1/2,\Gamma}^2,$$

and thus

$$u_\Gamma^T S^{(2)} u_\Gamma = |\mathcal{H}_2 u_\Gamma|_{1,\Omega_2}^2 \leq C_2 \, |u_\Gamma|_{1/2,\Gamma}^2 \leq C_2 \tilde{C}_1 \, |\mathcal{H}_1 u_\Gamma|_{1,\Omega_1}^2 = C_2 \tilde{C}_1 \, u_\Gamma^T S^{(1)} u_\Gamma.$$

Here we have used the relation between the Schur complement and the energy of the discrete harmonic extension given in Lemma 4.9 and the trace estimate for Ω_1 given in Lemma 4.10 with a constant $\tilde{C}_1$. Using similar estimates for $u_\Gamma^T S^{(1)} u_\Gamma$, we find

$$\kappa({S^{(2)}}^{-1}S^{(1)}) \leq \frac{C_1\tilde{C}_2}{C_2\tilde{C}_1}.$$

We note that this bound only depends on having shape regular, quasi-uniform local meshes, and, in particular, on the shape and the relative size of the subdomains. In the special case where Ω_1 and Ω_2 have the same shape and size, are symmetric with respect to the interface Γ, and have the same triangulations, then $\kappa({S^{(2)}}^{-1}S^{(1)}) = 1$. Generally, the Dirichlet-Neumann method for two subdomains is optimal since stable extensions can be found: the condition number satisfies

$$\kappa({S^{(2)}}^{-1}S) \leq C,$$

with C a constant independent of the dimension of the finite element problem.

1.3.4 The Neumann-Neumann Algorithm

Methods of this type were first considered in [65, 309, 154]; see [163, 226, 227, 9] for some earlier closely related studies.

The basic Neumann-Neumann algorithm can be described as follows: we start from an initial guess u_Γ^0. A step of the Neumann-Neumann algorithm consists in first solving Dirichlet problems on each Ω_i with data u_Γ^0 on Γ, and then a problem on each subdomain, with Neumann data, on Γ, chosen as the difference of the normal derivatives of the solutions of the two Dirichlet problems. The values on Γ of the solutions of these Neumann problems are then employed to correct the initial u_Γ^0 and find the new iterate u_Γ^1. In terms of differential operators (see (1.2)), we can write, for $n \geq 0$:

$$(D_i)\left\{\begin{array}{ll} -\Delta u_i^{n+1/2} = f & \text{in } \Omega_i, \\ u_i^{n+1/2} = 0 & \text{on } \partial\Omega_i \setminus \Gamma, \\ u_i^{n+1/2} = u_\Gamma^n & \text{on } \Gamma, \end{array}\right\}, \quad i = 1, 2,$$

$$(N_i)\left\{\begin{array}{lll} -\Delta\psi_i^{n+1} = 0 & & \text{in } \Omega_i, \\ \psi_i^{n+1} = 0 & & \text{on } \partial\Omega_i \setminus \Gamma, \\ \dfrac{\partial\psi_i}{\partial n_i}^{n+1} = \dfrac{\partial u_1}{\partial n_1}^{n+1/2} & + \dfrac{\partial u_2}{\partial n_2}^{n+1/2} & \text{on } \Gamma, \end{array}\right\}, \quad i = 1, 2, \tag{1.19}$$

$$u_\Gamma^{n+1} = u_\Gamma^n - \theta(\psi_1^{n+1} + \psi_2^{n+1}) \text{ on } \Gamma,$$

with a suitable $\theta \in (0, \theta_{\max})$. Using our approximation for the normal derivatives, we can derive an iteration for the discrete problem. If we define the vectors of internal degrees of freedom as $v_i = u_I^{(i)}$ and $w_i = \psi_I^{(i)}$, we find

$$
\begin{aligned}
&(D_i)\ A_{II}^{(i)} v_i^{n+1/2} + A_{I\Gamma}^{(i)} u_\Gamma^n = f_I^{(i)}, \quad i = 1, 2, \\
&(N_i)\ \begin{pmatrix} A_{II}^{(i)} & A_{I\Gamma}^{(i)} \\ A_{\Gamma I}^{(i)} & A_{\Gamma\Gamma}^{(i)} \end{pmatrix} \begin{pmatrix} w_i^{n+1} \\ \eta_i^{n+1} \end{pmatrix} = \begin{pmatrix} 0 \\ r_\Gamma \end{pmatrix}, \quad i = 1, 2, \\
&u_\Gamma^{n+1} = u_\Gamma^n - \theta(\eta_1^{n+1} + \eta_2^{n+1}),
\end{aligned}
\tag{1.20}
$$

where the residual r_Γ is defined as

$$
\begin{aligned}
r_\Gamma = &(A_{\Gamma I}^{(1)} v_1^{n+1/2} + A_{\Gamma\Gamma}^{(1)} u_\Gamma^n - f_\Gamma^{(1)}) \\
&+ (A_{\Gamma I}^{(2)} v_2^{n+1/2} + A_{\Gamma\Gamma}^{(2)} u_\Gamma^n - f_\Gamma^{(2)});
\end{aligned}
$$

see the third equation in (1.8).

We next eliminate $v_i^{n+1/2}$ and w_i^{n+1} from (1.20). Problems (D_i) give

$$
r_\Gamma = -(g_\Gamma - S u_\Gamma^n), \tag{1.21}
$$

which shows that the difference r_Γ of the local fluxes is equal to minus the residual of the Schur complement system. Using a block factorization of the local matrices $A^{(i)}$, problems (N_i), then give

$$
\eta_i^{n+1} = S^{(i)-1} r_\Gamma = -S^{(i)-1}(g_\Gamma - S u_\Gamma^n),
$$

Therefore, we find

$$
u_\Gamma^{n+1} - u_\Gamma^n = \theta({S^{(1)}}^{-1} + {S^{(2)}}^{-1})(g_\Gamma - S u_\Gamma^n),
$$

which shows that the Neumann-Neumann algorithm is also a preconditioned Richardson iteration for the Schur complement system, with the preconditioner ${S^{(1)}}^{-1} + {S^{(2)}}^{-1}$. The preconditioned matrix is

$$
FS = ({S^{(1)}}^{-1} + {S^{(2)}}^{-1})S = ({S^{(1)}}^{-1} + {S^{(2)}}^{-1})(S^{(1)} + S^{(2)}), \tag{1.22}
$$

the application of which to a vector involves the solution of two Dirichlet problems and two problems with Neumann data on Γ. We note that the operator F is the same as that in section 1.3.2.

The optimality of this method follows easily from the result on the Dirichlet-Neumann algorithm by using the spectral mapping theorem; the preconditioned matrix in (1.22) can be written as ${S^{(2)}}^{-1} S^{(1)} + 2I + ({S^{(2)}}^{-1} S^{(1)})^{-1}$ where the eigenvalues of ${S^{(2)}}^{-1} S^{(1)}$ are uniformly bounded from above and below.

We note that, given positive weights $\delta_1^\dagger$ and $\delta_2^\dagger$ with

$$
\delta_1^\dagger + \delta_2^\dagger = 1,
$$

the residual r_Γ on the right hand side of the Neumann problems in (1.20) can be replaced by $\delta_1^\dagger r_\Gamma$ and $\delta_2^\dagger r_\Gamma$ for Ω_1 and Ω_2, respectively. Similarly, when finding the new iterate u_Γ^{n+1} the sum of the two corrections η_1^{n+1} and η_2^{n+1} can be replaced by a weighted average:

$$u_\Gamma^{n+1} = u_\Gamma^n - \theta(\delta_1^\dagger \eta_1^{n+1} + \delta_2^\dagger \eta_2^{n+1}).$$

This gives rise to the preconditioner

$$D^{(1)} {S^{(1)}}^{-1} D^{(1)} + D^{(2)} {S^{(2)}}^{-1} D^{(2)},$$

where $D^{(i)}$ is a diagonal matrix, the diagonal entries of which are equal to $\delta_i^\dagger$. Scaling matrices $D^{(i)}$ are commonly employed in practice in order to improve the convergence of Neumann-Neumann algorithms for subdomains with cross points and, in particular, for problems with large changes in the coefficients; see Sect. 6.2.

Generalizations of the Dirichlet-Neumann and Neumann-Neumann algorithms are possible, by using more general Robin conditions. This was already proposed in some early work by Agoshkov and Lebedev [10], Agoshkov [9], and Lions [321]. We note that the use of more general interface conditions is often required for some nonsymmetric or indefinite problems. We refer to Sect. 11.5 for more details and further generalizations of these ideas.

We close this section by introducing an alternative preconditioner for the Schur complement system (1.12). We note that it can be shown that the Poincaré-Steklov operator that corresponds to S is a bijection from $H_{00}^{1/2}(\Gamma)$ to its dual $H_{00}^{-1/2}(\Gamma)$, and that the application of this operator therefore essentially involves the loss of one derivative; see A.2. It is therefore natural to use the inverse of J, the square root of minus an appropriate discrete Laplacian on Γ, as a preconditioner. If the mesh is uniform and has the appropriate number of points, this preconditioner can be implemented using the Fast Fourier transform; see [61]. This idea appears to originate with Dryja [172] and it works quite well. But it also has limitations not shared by the methods just discussed. Thus, in three dimensions, the mesh of the interface Γ must have a quite special tensor structure in order to allow use of a two-dimensional Fast Fourier transform.

1.3.5 A Dirichlet-Dirichlet Algorithm or a FETI Method

We now consider a method dual to the Neumann-Neumann algorithm. We start from an initial guess λ_Γ^0 of the flux on Γ; cf. Sect. 1.3.2. We first solve two Neumann problems on Ω_i with data λ_Γ^0 on Γ, and then a problem on each subdomain, with Dirichlet data, on Γ, chosen as the difference of the trace of the solutions of the two Neumann problems on Γ. The values on Γ of the normal derivatives of the solutions of these Dirichlet problems are then employed to correct the initial λ_Γ^0 and find the new iterate λ_Γ^1. We recall

that λ_Γ is an approximation of the normal derivative in the direction $\mathbf{n}_1$. In order to simplify the notation, we will drop the subscript Γ for the normal derivatives and set

$$\lambda^n = \lambda_\Gamma^n, \quad \lambda_1^n = -\lambda_2^n = \lambda^n.$$

In terms of differential operators (see (1.2)), we can write, for $n \geq 0$:

$$(N_i) \left\{ \begin{array}{lll} -\Delta u_i^{n+1/2} = f & & \text{in } \Omega_i, \\ u_i^{n+1/2} = 0 & & \text{on } \partial\Omega_i \setminus \Gamma, \\ \dfrac{\partial u_i}{\partial n_i}^{n+1/2} = \lambda_i^n & & \text{on } \Gamma, \end{array} \right\}, \quad i = 1, 2,$$

$$(D_i) \left\{ \begin{array}{lll} -\Delta \psi_i^{n+1} = 0 & & \text{in } \Omega_i, \\ \psi_i^{n+1} = 0 & & \text{on } \partial\Omega_i \setminus \Gamma, \\ \psi_i^{n+1} = u_1^{n+1/2} - u_2^{n+1/2} & & \text{on } \Gamma, \end{array} \right\}, \quad i = 1, 2, \tag{1.23}$$

$$\lambda^{n+1} = \lambda^n - \theta\left(\frac{\partial \psi_1}{\partial n_1}^{n+1} + \frac{\partial \psi_2}{\partial n_2}^{n+1}\right) \text{ on } \Gamma,$$

with a suitable $\theta \in (0, \theta_{\max})$. Using our approximation of the normal derivatives, we can derive an iteration for the discrete problem. If we define the vectors of internal degrees of freedom by $v_i = u_I^{(i)}$ and $w_i = \psi_I^{(i)}$, we find

$$\begin{aligned} &(N_i) \begin{pmatrix} A_{II}^{(i)} & A_{I\Gamma}^{(i)} \\ A_{\Gamma I}^{(i)} & A_{\Gamma\Gamma}^{(i)} \end{pmatrix} \begin{pmatrix} v_i^{n+1/2} \\ \gamma_i^{n+1/2} \end{pmatrix} = \begin{pmatrix} f_I^{(i)} \\ f_\Gamma^{(i)} + \lambda_i^n \end{pmatrix}, \quad i = 1, 2, \\ &(D_i)\ A_{II}^{(i)} w_i^{n+1} + A_{I\Gamma}^{(i)} r_\Gamma = 0, \quad i = 1, 2, \\ &\lambda^{n+1} = \lambda^n - \theta(\eta_1^{n+1} + \eta_2^{n+1}), \end{aligned} \tag{1.24}$$

where the residual r_Γ is defined by

$$r_\Gamma = \gamma_1^{n+1/2} - \gamma_2^{n+1/2}$$

and the fluxes η_i^{n+1} by

$$\eta_i^{n+1} = A_{\Gamma I}^{(i)} w_i^{n+1/2} + A_{\Gamma\Gamma}^{(i)} r_\Gamma;$$

cf. Equation (1.7).

We proceed as in the previous section and eliminate $v_i^{n+1/2}$, $\gamma_i^{n+1/2}$, and w_i^{n+1} from (1.24). Using a block factorization of the local matrices $A^{(i)}$, problems (N_i) give

$$r_\Gamma = -(d_\Gamma - F\lambda^n),$$

which shows that the difference r_Γ of the local solutions is equal to the negative of the residual of the system (1.15). Problems (D_i) then provide

$$\eta_i^{n+1} = S^{(i)} r_\Gamma = -S^{(i)}(d_\Gamma - F\lambda^n),$$

Therefore, we find

$$\lambda^{n+1} - \lambda^n = \theta(S^{(1)} + S^{(2)})(d_\Gamma - F\lambda^n),$$

which shows that this Dirichlet-Dirichlet algorithm is also a preconditioned Richardson iteration for the system (1.15), with the preconditioner $S^{(1)} + S^{(2)}$. The preconditioned matrix is

$$SF = S({S^{(1)}}^{-1} + {S^{(2)}}^{-1}) = (S^{(1)} + S^{(2)})({S^{(1)}}^{-1} + {S^{(2)}}^{-1}), \tag{1.25}$$

the application of which to a vector involves the solution of two Neumann problems and two problems with Dirichlet data on Γ. We have called the method presented here the *Dirichlet-Dirichlet* algorithm in analogy to the dual Neumann-Neumann one of the previous section. We remark however that, in the domain decomposition literature, the term Dirichlet-Dirichlet is often used for the Richardson or conjugate gradient method applied to the unpreconditioned Schur complement system (1.12), since the application of S involves the solution of a Dirichlet problem on each subdomain. We also note that the preconditioned operator SF is the same as that of the FETI algorithm with the original Dirichlet preconditioner; cf. Equation (6.36) in Sect. 6.3.1 for the case of two subdomains. In the following, we will refer to the method of this section as the *preconditioned FETI method.* FETI methods will be presented and analyzed systematically in Chap. 6.

We recall that the preconditioned operator for the Neumann-Neumann algorithm is FS. It is a trivial matter to prove that the operators SF and FS have the same eigenvalues and thus the same condition number. The condition numbers of the Neumann-Neumann and Dirichlet-Dirichlet methods, and indeed the entire spectra, are *the same* for the case of two subdomains.

As for Neumann-Neumann algorithms, we can also employ weights. With $\delta_1^\dagger$ and $\delta_2^\dagger$ the weights introduced in Sect. 1.3.4, the residual r_Γ in the Dirichlet problems in (1.24) is often replaced by $\delta_2^\dagger r_\Gamma$ and $\delta_1^\dagger r_\Gamma$ for Ω_1 and Ω_2, respectively. Similarly, when finding the new iterate λ^{n+1} the sum of the two corrections η_1^{n+1} and η_2^{n+1} can be replaced by an average:

$$\lambda^{n+1} = \lambda^n - \theta(\delta_2^\dagger \eta_1^{n+1} + \delta_1^\dagger \eta_2^{n+1}).$$

This gives rise to the preconditioner

$$D^{(2)} S^{(1)} D^{(2)} + D^{(1)} S^{(2)} D^{(1)};$$

see in particular Sect. 6.3.2 and 6.3.3, for more details, and in particular formulas (6.37) and (6.51). Suitable scaling matrices are commonly employed in practice in order to improve the convergence of FETI algorithms for partitions into subdomains with cross points or for problems with large changes in the coefficients; see Sect. 6.3.

We note that, by proceeding as in Sect. 1.3.3, we can also design a Neumann-Dirichlet algorithm by starting the iteration with an initial guess of the Neumann data; see, e.g., [63, 301]. We would then solve a Neumann problem on one subdomain and use the trace of this solution as Dirichlet data for the problem on the other subdomain, etc. The corresponding preconditioned operator is

$$S^{(2)}F = S^{(2)}(S^{(2)-1} + S^{(1)-1}) = I + S^{(2)}S^{(1)-1},$$

the condition number of which can be immediately bounded using a bound for the Dirichlet-Neumann algorithm in Sect. 1.3.3.

Finally, we mention that ideas similar to those in this and the previous subsection were already proposed in some early work by Glowinski and Wheeler [226, 227] for mixed approximations of elliptic problems.

1.3.6 The Case of Many Subdomains

The methods just introduced can be generalized to the case of more than two subdomains. Generalizations of the Neumann-Neumann and FETI methods have become widely used. Here we only give a brief introduction and refer to Chap. 6 for a much more complete description and analysis.

The Dirichlet-Neumann method can also be generalized, under certain restrictions on the partition into subdomains. This method is less widely used than Neumann-Neumann and FETI methods since its performance can deteriorate if the coefficients of the differential equations differ greatly between the subdomains. However, some important ideas underlying much more recent and popular methods, such as the dual-primal FETI algorithms, can be traced back to early work on this family of methods; see Sect. 6.4. In addition, our presentation will show the importance of the cross points of the partition. A cross point is a point common to the boundaries of three or more subregions. If there are no cross points, we essentially have partitioned the region Ω into strips; that case is in many ways similar to the two subdomain case. We note that while we assume that the original region Ω has only one connected component, it is quite natural to consider cases where two subregions Ω_i have several or many components.

Here we will only consider two-dimensional problems. Before proceeding, we need to define some more general operators. We consider problem (1.1) and suppose that Ω is partitioned into a family of nonoverlapping subdomains $\{\Omega_i,\ 1 \leq i \leq N\}$ with

$$\overline{\Omega} = \bigcup_i \overline{\Omega_i}; \quad \Omega_i \cap \Omega_j = \emptyset \quad i \neq j.$$

If $\Gamma_i = \partial\Omega_i \setminus \partial\Omega$, the interface Γ is defined as

$$\Gamma = \bigcup_i \Gamma_i.$$

We note that Γ and the Γ_i are open. A coupled problem as in (1.2) can be found with transmission conditions imposed along each edge $\partial\Omega_i \cap \partial\Omega_j$.

The linear system (1.4) can now be written as

$$\begin{pmatrix} A_{II} & A_{I\Gamma} \\ A_{\Gamma I} & A_{\Gamma\Gamma} \end{pmatrix} \begin{pmatrix} u_I \\ u_\Gamma \end{pmatrix} = \begin{pmatrix} f_I \\ f_\Gamma \end{pmatrix}, \tag{1.26}$$

where we have partitioned the degrees of freedom into those interior to the subdomains and those on Γ. The stiffness matrix and the right hand side are obtained by subassembling the corresponding components relative to the subdomains; see (1.6).

The unknowns in the interior of the subdomains u_I can again be eliminated by block Gaussian elimination and the resulting linear system is

$$\begin{pmatrix} A_{II} & A_{I\Gamma} \\ 0 & S \end{pmatrix} \begin{pmatrix} u_I \\ u_\Gamma \end{pmatrix} = \begin{pmatrix} f_I \\ g_\Gamma \end{pmatrix}. \tag{1.27}$$

As before the Schur complement S and the vector g_Γ can be found by subassembling local contributions; see Chap. 4 for further details. In order to do so, we first define a family of restriction operators. Given a vector of degrees of freedom u_Γ on the interface, we define the restriction $R_i u_\Gamma$ as the vector of degrees of freedom of u_Γ on Γ_i. Here R_i is a rectangular matrix of zeros and ones. If for each subdomain the degrees of freedom are partitioned into those internal to Ω_i and those on Γ_i, as in (1.6), we have

$$\begin{aligned} S &= \sum_{i=1}^{N} R_i^T S^{(i)} R_i, \\ g_\Gamma &= \sum_{i=1}^{N} R_i^T (f_\Gamma^{(i)} - A_{\Gamma I}^{(i)} {A_{II}^{(i)}}^{-1} f_I^{(i)}) \end{aligned} \tag{1.28}$$

where the local Schur complements are defined as in (1.11) and R_i^T is the transpose of R_i.

Neumann-Neumann Methods

By examining the preconditioned Neumann-Neumann operator for two subdomains (1.22), we can easily find a generalization to the case of many subdomains. We define

$$S_{NN}^{-1} S = \sum_{i=1}^{N} R_i^T {S^{(i)}}^{-1} R_i \; S. \tag{1.29}$$

We note that the application of this operator to a vector involves the solution, on each subdomain Ω_i, of a Dirichlet problem and a problem with Neumann boundary conditions on $\partial\Omega_i \cap \Gamma$. We also note that for subdomains that do not touch $\partial\Omega$, $S^{(i)}$ is singular and ${S^{(i)}}^{-1}$ in (1.29) should be understood as a pseudo-inverse or the inverse of a regularized problem.

We note that the Neumann-Neumann algorithm can also be defined at the continuous level using (1.19), with $i = 1, 2, \ldots, N$, and Neumann conditions for the problem N_i given on each edge $\Gamma_{ij} = \partial\Omega_i \cap \partial\Omega_j$. The new iterate u_Γ^{n+1}, at a node of the interface, is then built from corrections from all subdomains with that node on their boundaries.

The condition number of the Neumann-Neumann method satisfies a logarithmic bound. If h is the diameter of the finite elements and H that of a typical subdomain, we have

$$\kappa(S_{NN}^{-1} S) \leq \frac{C}{H^2} \left(1 + \log \frac{H}{h}\right)^q . \tag{1.30}$$

The method is not therefore scalable according to the following definition:

Definition 1.3 (Scalability). *A domain decomposition iterative method for the solution of a linear system is said to be scalable, if its rate of convergence does not deteriorate when the number of subdomains grows. This typically means that convergence does not deteriorates when H, the typical subdomain size, becomes small.*

The dependence on H^{-2} in (1.30) is typical of domain decomposition preconditioned methods without a coarse solver. Since the Green function for elliptic problems does not in general have a compact support, a residual on the linear system of modestly low frequency will result in an error which cannot be neglected in *any* part of the region. Therefore an iterative method for the solution of the resulting linear system in which information is only exchanged between neighboring subregions must necessarily, for certain initial errors, require a number of steps which is at least equal to the diameter of the dual graph corresponding to the subdomain partition. Here, the dual graph is constructed by introducing a vertex for each subregion and an edge between two subregions that share a part of their boundaries. The diameter of the graph is the maximum distance between pairs of vertices, where the distance is defined as the length of the shortest path between the vertices. In case the diameter of the original domain is one, the diameter of the graph is typically $O(H^{-1})$. Using an argument of contradiction and the upper bound for the error of the conjugate gradient iteration in Lemma C.10, we see that the condition number of a domain decomposition preconditioned operator must grow at least as H^{-2}. We note that for the same reason, the condition number of the stiffness matrix resulting from the finite element approximation of an elliptic problem must grow at least as h^{-2}; see Theorem B.32.

Going back to the bound in (1.30), if the partition does not have cross points, we have $q = 0$ and we have to interpret H as the width of the strips. Otherwise, $q = 3$. A quadratic growth, $q = 2$, is obtained if suitable scaling matrices are incorporated into the preconditioner; see Sect. 6.2 for further details. We note that the Neumann-Neumann method with many subdomains is not in general optimal according to Definition 1.2, since its condition number

may depend on the meshsize h. The condition number bound does not however involve h alone but only the ratio H/h, which gives a measure of the number of unknowns associated with one subdomain. In addition, the condition number increases slowly (logarithmically) with H/h. In such a case, the method is said to be *quasi-optimal*.

We finally note that the preconditioned operator $S_{NN}^{-1}S$ has the typical form of a Schwarz operator; these operators will be discussed in detail in Chap. 2.2. We have

$$S_{NN}^{-1}S = \sum_{i=1}^{N} P_i, \quad P_i = R_i^T {S^{(i)}}^{-1} R_i \, S.$$

Each P_i is associated with a subdomain (or, in other words, to a subspace) and is a projection-like operator; see Sect. 2.2 for further details.

Dirichlet-Neumann Methods

We now assume that there is a red-black partition of the subdomains into two sets $\mathcal{R}$ and $\mathcal{B}$, such that the intersection between the boundaries of two subregions in the same class is either empty or a vertex; see Fig. 1.2.

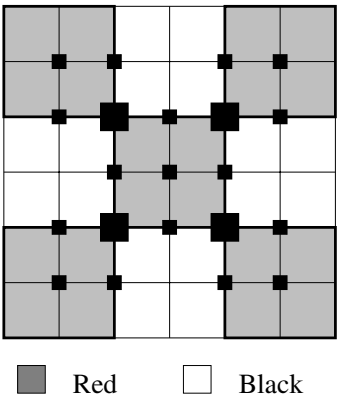

Fig. 1.2. Red-black coloring of the subdomains.

We can then decompose the Schur complement as

$$S = \sum_{i \in \mathcal{R}} R_i^T S^{(i)} R_i + \sum_{i \in \mathcal{B}} R_i^T S^{(i)} R_i,$$

and define the preconditioner essentially in terms of Neumann problems on the subdomains in the set $\mathcal{R}$

$$S_{DN}^{-1} = \left(\sum_{i \in \mathcal{R}} R_i^T S^{(i)} R_i \right)^{-1}.$$

In the particular case where there are no cross points, the matrix S_{DN} is block-diagonal, since the intersection between the boundaries of two red regions is empty and there is therefore no coupling between degrees of freedom on two subdomain boundaries in the same class. In particular, for $i, j \in \mathcal{R}$, we then have

$$R_i R_j^T = \begin{cases} 0 \; i \neq j, \\ I \; i = j, \end{cases}$$

and thus

$$S_{DN}^{-1} = \sum_{i \in \mathcal{R}} R_i^T {S^{(i)}}^{-1} R_i. \tag{1.31}$$

The preconditioned Dirichlet-Neumann operator becomes

$$S_{DN}^{-1} S = I + \left(\sum_{i \in \mathcal{R}} R_i^T {S^{(i)}}^{-1} R_i \right) \left(\sum_{i \in \mathcal{B}} R_i^T S^{(i)} R_i \right), \tag{1.32}$$

and its application involves the solution of a Dirichlet problem on the subdomains in the set $\mathcal{B}$ and a Neumann problem on those in the set $\mathcal{R}$. As for the case of two subdomains, bounds for the condition number of the Dirichlet-Neumann operator rely on finding stable finite element extensions. A uniform bound, with $q = 0$, can be derived for the case of strips while in the general case, discussed below in this section, we have a logarithmic bound (1.30) as for the Neumann-Neumann algorithm. We note however, that in the strip case the algorithm is not scalable and that the bound will contain a factor $1/H^2$ where H is the width of a typical strip of the decomposition; cf. the discussion on the lack of scalability earlier in this subsection.

If there are cross points, then formula (1.31) is no longer valid. Indeed, the inversion of S_{DN} requires the solution of a global problem, which instead of creating a problem turns out to be a blessing since it provides a natural coarse solve, which makes the algorithm scalable, i.e., convergent at a rate independent of the number of subdomains if the subdomains all have good aspect ratios. In particular, we have to solve a Neumann problem on the union of the subdomains of the set $\mathcal{R}$, i.e., on the set glued together at the cross points and given by

$$\bigcup_{i \in \mathcal{R}} \overline{\Omega_i};$$

see the shaded region in Fig. 1.2. This can be done in two stages. All the degrees of freedom of the subdomains of the set $\mathcal{R}$ except for those at cross points (drawn as small squares in Fig. 1.2) are first eliminated in parallel across the subdomains by a step of block Gaussian elimination. All the submatrices involved at this stage are invertible. The resulting Schur complement, which involves only the nodal values at the cross points (drawn as larger squares in Figure 1.2), is sparse since it can be shown that nonzero off-diagonal elements only exist for the pairs of cross points which belong to the boundary of the same subregion. This final step of the elimination typically will involve degrees

of freedom from all parts the region Ω. This mechanism helps explain that bounds can be derived for the condition number of this preconditioned matrix which are independent of the number of subregions. We remark that a similar procedure is employed in FETI-DP methods; see section 6.4. The resulting preconditioned operator is

$$S_{DN}^{-1}S = I + S_{DN}^{-1}\left(\sum_{i\in\mathcal{B}} R_i^T S^{(i)} R_i\right),$$

and its application to a vector requires the solution of Dirichlet problems on the black subdomains, and of the special Neumann problem, just introduced, on the union of the red subdomains, which involves the low-dimensional coarse problem already discussed. The condition number satisfies

$$\kappa(S_{DN}^{-1}S) \leq C\left(1 + \log\frac{H}{h}\right)^2;$$

see [176, 465]. We note that the technical tools needed in deriving such bounds are developed in Chap. 4.

Methods Involving the Normal Derivative

The expression given in (1.7) of the functional representing the normal derivative $\lambda^{(i)}$ on $\partial\Omega_i$ remains valid for the case of many subdomains and in particular when the partition has cross points. We recall that the coupled system (1.8) is obtained from the original system (1.4) by introducing the additional unknowns $\lambda^{(i)}$. A coupled system for the case of many subdomains can easily be found. This involves the equality of the trace functions $u_\Gamma^{(i)}$ at all the nodes on the interface Γ. For the normal derivatives $\lambda^{(i)}$, conditions similar to that of the third equation of (1.8) are a consequence of the fact that the stiffness matrix A is obtained by subassembling contributions from neighboring subdomains. For every node x_h on Γ the corresponding condition becomes

$$\sum_j \lambda^{(j)}(x_h) = 0,$$

where the sum is taken over all the subdomains that share the node x_h. We note that, if x_h is a cross point, the sum is taken over more than two subdomains and it appears to be difficult to give a functional meaning to this condition. In addition, when trying to generalize Equation (1.15) to the case of many subdomains, some of local Neumann problems are not uniquely solvable and therefore the corresponding local Schur complements not invertible, if the subdomain boundaries do not all intersect $\partial\Omega$.

The generalization of what we have called the Dirichlet-Dirichlet algorithm will be carried out systematically in Chap. 6, where we introduce the

important FETI family of methods. These methods provide preconditioned operators which act on suitable approximations of the normal derivative λ_Γ on the interface Γ. There, following the development of FETI methods, we will proceed in a purely algebraic way and approximations of the normal derivative will be given by vectors of Lagrange multipliers of suitable constrained systems. As we will show, in Chap. 6, there are more than one way to define an approximation of the normal derivative λ_Γ when the partition has cross points and the definition will depend on the particular form of the equality constraints chosen for the interface values $u_\Gamma^{(i)}$ across Γ.

We finally recall that the issue of singular, local Neumann problems was already addressed in [226, 227] for mixed approximations of elliptic problems.

1.4 The Schwarz Alternating Method

The earliest domain decomposition method known to the authors is the alternating method of H.A. Schwarz, [416], published in 1870. Schwarz used the algorithm to establish the existence of harmonic functions with prescribed boundary values on regions with nonsmooth boundaries. The regions were constructed recursively by forming unions of pairs of regions starting with regions for which existence could be established by some more elementary means. At the core of Schwarz's work is a proof that this iterative method converges in the maximum norm at a geometric rate.

For more than two subregions, we can in fact define a step of the algorithm by recursion: i) solve on the first subregion; ii) solve on the union of all other subregions, approximately, by recursively invoking a step of the same algorithm.

As pointed out by Pierre-Louis Lions [319], the convergence of this algorithm can be established by two different methods, namely, by a maximum principle and by using Hilbert spaces. The Hilbert space method is the most appropriate here since much of our work relies on the classical calculus of variation and finite elements.

We finally note that between the work of Schwarz, [416], and Lions, [319], there was some quite significant work, in particular by Sobolev [425] and Babuška [29]; see also [319] for additional references.

1.4.1 Description of the Method

The classical Schwarz method, for two subregions, can be described as follows: Consider the Poisson problem (1.1) on a bounded Lipschitz region Ω with zero Dirichlet boundary conditions. There are two fractional steps corresponding to two overlapping subregions, Ω_1' and Ω_2' of the original region $\Omega = \Omega_1' \bigcup \Omega_2'$; see Fig. 1.3. Given an initial guess u^0, which vanishes on $\partial\Omega$, the iterate u^{n+1} is determined from the previous iterate u^n in two sequential steps in which the approximate solution on the two subregions is updated:

$$\begin{cases} -\Delta u^{n+1/2} = f & \text{in } \Omega_1', \\ u^{n+1/2} = u^n & \text{on } \partial\Omega_1', \\ u^{n+1/2} = u^n & \text{in } \Omega_2 = \Omega_2' \setminus \overline{\Omega_1'}, \end{cases}$$
$$\begin{cases} -\Delta u^{n+1} = f & \text{in } \Omega_2', \\ u^{n+1} = u^{n+1/2} & \text{on } \partial\Omega_2', \\ u^{n+1} = u^{n+1/2} & \text{in } \Omega_1 = \Omega_1' \setminus \overline{\Omega_2'}. \end{cases} \tag{1.33}$$

Thus, the Dirichlet data for these problems are obtained from the original data given on $\partial\Omega \cap \partial\Omega_i'$, and the values from the previous fractional step on the remaining part $\Gamma_i = \partial\Omega_i' \setminus \partial\Omega$ of the subdomain boundaries. We note that this algorithm also can be viewed as a mapping of values on Γ_1 (or Γ_2) onto values on the same set; see Sect. 1.6.1.

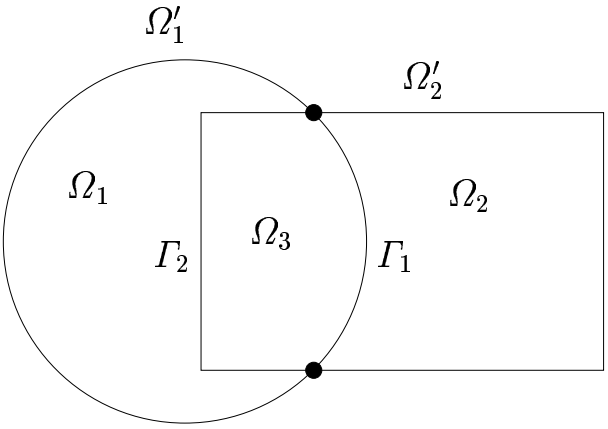

Fig. 1.3. Overlapping partition for the Schwarz alternating method with two subdomains

1.4.2 The Schwarz Alternating Method as a Richardson Method

In order to show that the algorithm is indeed a Richardson iteration, we rewrite it in variational form. For the original problem, we use the space $H_0^1(\Omega)$ and the bilinear form

$$a(u,v) = \int_\Omega \nabla u \cdot \nabla v \; dx, \tag{1.34}$$

which defines the $H^1(\Omega)-$seminorm; see appendix A.1. We also introduce a finite element triangulation $\mathcal{T}$, of maximum diameter h, such that the local

boundaries $\partial\Omega_i'$ do not cut through any element in $\mathcal{T}$. The mesh $\mathcal{T}$ thus defines two local meshes $\mathcal{T}_i$ on Ω_i', $i = 1, 2$. We next define the spaces of continuous, piecewise linear finite elements on $\mathcal{T}$, $\mathcal{T}_1$, and $\mathcal{T}_2$, which vanish on $\partial\Omega$, $\partial\Omega_1'$, and $\partial\Omega_2'$, respectively, and denote them by V, V_1, and V_2. We note that there are two natural extension operators

$$R_i^T : V_i \longrightarrow V, \quad i = 1, 2,$$

which take local functions on the Ω_i' with zero boundary values and extend them by zero to the rest of Ω to give global functions in V. We note that the restriction operators $R_i : V \longrightarrow V_i$ take global functions on Ω and give local functions in V_i that vanish on $\partial\Omega_i'$ and are equal to the original ones at the nodes *inside* Ω_i'. We also need the two local bilinear forms

$$a_i(u, v) = \int_{\Omega_i'} \nabla u \cdot \nabla v \, dx, \quad u, v \in V_i.$$

We can now write the Schwarz method in terms of two orthogonal projections P_i, $i = 1, 2$. For $i = 1, 2$, the projections are

$$P_i = R_i^T \tilde{P}_i,$$

where $\tilde{P}_i : V \longrightarrow V_i$, are defined by

$$a_i(\tilde{P}_i u, v_i) = a(u, R_i^T v_i), \quad v_i \in V_i \, .$$

Indeed, if we consider a finite element discretization of the first problem in (1.33), we see that $u^{n+1/2} - u^n$ vanishes in $\Omega_2 = \Omega_2' \setminus \overline{\Omega_1'}$ and, when restricted to Ω_1', belongs to V_1. With u the exact finite element solution, this function satisfies, for $v_1 \in V_1$,

$$a_1(R_1(u^{n+1/2} - u^n), v_1) = \int_{\Omega_1'} f v_1 \, dx - a_1(u^n, R_1^T v_1) = a(u - u^n, R_1^T v_1),$$

and thus

$$R_1(u^{n+1/2} - u^n) = \tilde{P}_1(u - u^n).$$

Since $(u^{n+1/2} - u^n)$ vanishes in $\Omega \setminus \Omega_1'$, we obtain

$$u^{n+1/2} - u^n = R_1^T R_1(u^{n+1/2} - u^n) = P_1(u - u^n).$$

The error $u^{n+1/2} - u$ is then given by

$$u^{n+1/2} - u = (I - P_1)(u^n - u).$$

Proceeding in a similar way for the finite element discretization of the second problem in (1.33), we find

$$u^{n+1} - u = (I - P_2)(u^{n+1/2} - u) = (I - P_2)(I - P_1)(u^n - u),$$

which shows that the error propagation operator of this *multiplicative* Schwarz method is

$$(I - P_2)(I - P_1) \;=\; I - (P_1 + P_2 - P_2 P_1).$$

Therefore, the algorithm can be viewed as a simple iterative method for solving

$$P_{mu} u := (P_1 + P_2 - P_2 P_1) u = g.$$

With an appropriate right-hand side g, $u \in V$ will be the solution of the original finite element problem. We note that the error propagation operator is not symmetric. However, after the first step, the error will belong to $range(I - P_2)$ and the operator P_{mu} is symmetric on that subspace. The algorithm immediately extends to three and more subspaces; in these more general cases symmetry cannot be recovered in this way.

1.5 Block Jacobi Preconditioners

An important variant of the Schwarz methods is the *additive Schwarz method.* We first consider a very simple example, namely a two-block Jacobi method, and try to understand how well it works and how it can be improved. This will give a first example of an additive method.

We will work with the matrix form (1.4) of our finite element problem. The stiffness matrix A is positive definite, symmetric, which are properties inherited by any conforming finite element method, from the bilinear form $a(\cdot, \cdot)$ in (1.34); see appendix B for additional details. Here, and in the following, we exploit the one-to-one correspondence between the finite element space V and the corresponding space of degrees of freedom consisting of the nodal values of a function and use the same notation for finite element spaces and spaces of degrees of freedom, and functions and corresponding vectors of degrees of freedom.

We consider the block-Jacobi/conjugate gradient method: the stiffness matrix A is preconditioned by a matrix A_J^{-1}, where A_J is the a direct sum of two diagonal blocks of A. Each block corresponds to a set of degrees of freedom, which define a subspace V_i. We write the space V as a direct sum of the subspaces $V_i, i = 1, 2,$

$$V = R_1^T V_1 \oplus R_2^T V_2,$$

where R_i^T are the natural extension operators

$$R_i^T : V_i \longrightarrow V, \quad i = 1, 2,$$

which take the sets of local degrees of freedom of V_i and extend them by zero to the remaining degrees of freedom in Ω to give global vectors of nodal values

in V. If the block of A related to V_i is denoted by A_i, the preconditioner A_J^{-1} can be written as

$$A_J^{-1} = \begin{pmatrix} A_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & A_2^{-1} \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}^{-1}. \tag{1.35}$$

The matrix A_J is obtained from A by removing the coupling between the two subspaces V_1 and V_2. We expect that the weaker the coupling, the better the preconditioner.

We now write A_J in a more compact form. Let the restriction operators

$$R_i : V \longrightarrow V_i, \quad i = 1, 2,$$

be the adjoints of R_i^T with respect to the Euclidean scalar product; R_i takes a vector of global degrees of freedom and extracts the degrees of freedom associated with V_i. We immediately see that

$$A_i = R_i A R_i^T, \quad i = 1, 2,$$

and that

$$A_J^{-1} = R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2. \tag{1.36}$$

The choice of the subspaces is a key issue and so is the choice of basis of V, in particular in the spectral element case. In order to improve the convergence, we can use subspaces that do not form a direct sum of V, e.g., by choosing the subspaces V_1 and V_2 defined in the previous section, consisting of degrees of freedom associated with the interior of two overlapping subdomains Ω_1' and Ω_2'. The preconditioner A_J^{-1} can still be written as in (1.36).

In order to establish a connection with the Schwarz alternating method of Sect. 1.4, we will show that the preconditioned additive operator

$$P_{ad} = A_J^{-1} A$$

can also be written using two projections,

$$P_i := R_i^T A_i^{-1} R_i A \ : V \longrightarrow V, \quad i = 1, 2.$$

In Lemma 2.1, we will show that the P_i are the same as those of Sect. 1.4. We immediately see that

$$P_{ad} = P_1 + P_2.$$

Since this operator is the sum of two projections, which are orthogonal in the inner product (1.34), we obtain an upper bound of 2 for the largest eigenvalue. If there are N subspaces, we similarly obtain a bound of N. However, this bound can be often be improved by using a coloring technique; see section 2.5.1.

Before concluding this section, we return to the case of two subspaces that form a direct sum decomposition of V, and show how in this simple case the

smallest eigenvalue of the block-Jacobi preconditioned system P_{ad} is related to the representation properties of the subspace decomposition. This will be generalized to more general decompositions in section 2.3; see Assumption 2.2 and Lemma 2.5.

Since we now consider a direct sum of two subspaces, there is a unique decomposition $u = R_1^T u_1 + R_2^T u_2$, for every $u \in V$, with

$$u_1 = R_1 u, \quad u_2 = R_2 u.$$

The corresponding block-Jacobi preconditioner is given by (1.35). Let C_0 be a positive constant such that, for every $u \in V$,

$$u_1^T A_1 u_1 + u_2^T A_2 u_2 \leq C_0^2 \, u^T A u.$$

We note that C_0^2 measures the stability of the splitting and that we use the scalar product induced by the original matrix A. Since we work with a direct sum, we can write

$$\begin{aligned} u^T A(P_{ad}^{-1})u &= u^T (R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2)^{-1} u \\ &= u^T (R_1^T A_1 R_1 + R_2^T A_2 R_2) u \\ &= u_1^T A_1 u_1 + u_2^T A_2 u_2, \end{aligned}$$

which, combined with the stability property of the splitting, gives

$$\sup_{u \in V} \frac{u^T A(P_{ad}^{-1})u}{u^T A u} \leq C_0^2,$$

and thus a lower bound for the smallest eigenvalue of P_{ad}. Thus the condition number is bounded from above by $2/C_0^2$.

We conclude this section with some remarks on the two methods presented so far.

Remark 1.4.

1. The additive and multiplicative algorithms can be generalized to the case of more than two subspaces in a straightforward way. An additional component, associated with a coarse space V_0, can also be added to any of the algorithms. The latter is often necessary for good convergence, in particular, to make the convergence rate independent of the number of subproblems; cf. discussion in subsection 1.3.6 and appendix C.5.
2. Additive and multiplicative algorithms can be determined by the same decomposition into subspaces but they are different. They correspond to different preconditioned operators, P_{ad} and P_{mu}, determined by different *polynomial* functions of the projections $\{P_i\}$.
3. More generally, instead of the local solvers $\{A_i^{-1}\}$ in the definition of the preconditioners, we can employ modified operators $\{\tilde{A}_i^{-1}\}$, corresponding, for instance, to *inexact solvers*. We also note that we might consider replacing the sum of the values of $P_1 u$ and $P_2 u$ at the nodes of an overlapping subdomain by a convex combination of the two values.

4. For two subspaces and exact solvers, Bjørstad and Mandel [58] have shown that the eigenvalues of the additive operator $P_1 + P_2$ can be expressed by those of the multiplicative operator $P_1 + P_2 - P_2P_1$, and vice versa.

1.6 Some Results on Schwarz Alternating Methods

In this final section, we give two results on the convergence of the Schwarz alternating method. They represent early efforts to prove the convergence of Schwarz overlapping methods before the Schwarz theory of the next chapter was fully developed. While we believe that these results are of interest, this section can also be bypassed since it is not necessary for the understanding of the main part of this monograph.

1.6.1 Analysis for the Case of Two Subdomains

In Chap. 2, we will develop a systematic framework to estimate the rate of convergence of the Schwarz alternating method and many other algorithms, and we will also devote Chap. 3 to the description and analysis of a number of algorithms based on overlapping subdomains. Here, we will show, following Bjørstad and Widlund [62], that a, perhaps surprising, connection can be made to the algorithms discussed and analyzed in Sect. 1.3. In this subsection, we will consider the case of two subregions; see Fig. 1.3.

Examining Schwarz's method (1.33), we see that we can view it as a mapping from $\Gamma_1 = \partial\Omega_1' \setminus \partial\Omega$ onto itself; once the correct value of u_{Γ_1} has been found, the iteration has converged. This observation is equally valid for the finite element approximation. We note that the values of the iterates on Γ_1 change only in the second fractional step of the algorithm.

We also see that as soon as the first full step of the algorithm has been completed, the error in (1.33) will be harmonic in the three regions Ω_1, Ω_2, and $\Omega_3 = \Omega_1' \cap \Omega_2'$. Similarly, in the finite element case, we see that the residuals vanish at all nodes inside these three regions after the first step of the algorithm has been completed. We can therefore confine our discussion to the case for which the right hand sides differ from zero only on the interfaces Γ_1 and Γ_2. In fact, after the first fractional step the residual will also vanish at all mesh points on Γ_2, and we see that the error $e^{n+1/2}$ will satisfy the linear system of equations

$$Ae^{n+1/2} = \begin{pmatrix} 0 \\ 0 \\ r_{\Gamma_1}^{n+1/2} \end{pmatrix},$$

where A is the global stiffness matrix on Ω and we have partitioned the degrees of freedom into subvectors representing nodal values in the interior of Ω_1', the

interior to Ω_2, and on the interface Γ_1. Using a block factorization of A (cf. equation (1.12)), we find

$$(S_{\Gamma_1}^{(1)} + S_{\Gamma_1}^{(2)})e_{\Gamma_1}^{n+1/2} = r_{\Gamma_1}^{n+1/2}. \tag{1.37}$$

Here, $S_{\Gamma_1}^{(1)} + S_{\Gamma_1}^{(2)}$ is the Schur complement of A with respect to Γ_1, obtained by subassembling the local Schur complements for Ω_1' and Ω_2. We can express the same observation by saying that the restriction of the stiffness matrix for the entire region to the space with vanishing residuals on these two subregions can be expressed in terms of the sum of these two Schur complements.

The error on Γ_1 will only be partially eliminated in the second fractional step, since we are solving a problem on the subregion Ω_2' rather than on all of Ω. With vanishing residuals in Ω_1' and Ω_2, we can write the correction in the second fractional step as

$$A_2(u_2^{n+1} - u_2^{n+1/2}) = \begin{pmatrix} 0 \\ 0 \\ r_{\Gamma_1}^{n+1/2} \end{pmatrix},$$

where A_2 is the stiffness matrix of Ω_2' and we have partitioned the degrees of freedom of Ω_2' into those interior to Ω_2, those interior to Ω_3, and those of the interface Γ_1. Using a block factorization of A_2, we find

$$u_{\Gamma_1}^{n+1} - u_{\Gamma_1}^{n+1/2} = (S_{\Gamma_1}^{(2)} + S_{\Gamma_1}^{(3)})^{-1} r_{\Gamma_1}^{n+1/2}; \tag{1.38}$$

here $S_{\Gamma_1}^{(2)} + S_{\Gamma_1}^{(3)}$ is the Schur complement of A_2 with respect to Γ_1, obtained by subassembling the local Schur complements for Ω_2 and Ω_3. We note that $S_{\Gamma_1}^{(2)}$ in (1.38) is the same as in (1.37).

Since $e_{\Gamma_1}^{n+1/2} = e_{\Gamma_1}^{n}$, a simple computation allows us to find the error propagation operator, regarded as a mapping from Γ_1 to itself:

$$e_{\Gamma_1}^{n+1} = \left(I - (S_{\Gamma_1}^{(2)} + S_{\Gamma_1}^{(3)})^{-1}(S_{\Gamma_1}^{(1)} + S_{\Gamma_1}^{(2)})\right) e_{\Gamma_1}^{n}. \tag{1.39}$$

It is easy to show that $S_{\Gamma_1}^{(3)} \geq S_{\Gamma_1}^{(1)}$, (an inequality in terms of quadratic forms), since the minimal extension from Γ_1 to the interior of Ω_3 necessarily has a larger energy than its counterpart for the larger region Ω_1'. Thus, the space of possible extensions is larger in the latter case; see Lemma 4.9. Therefore, the error propagation operator has only positive eigenvalues. The convergence is uniform since we can also show, by arguments quite similar to those for the Neumann-Dirichlet and Neumann-Neumann methods, that the eigenvalues of

$$(S_{\Gamma_1}^{(2)} + S_{\Gamma_1}^{(3)})^{-1}(S_{\Gamma_1}^{(1)} + S_{\Gamma_1}^{(2)}) \tag{1.40}$$

are bounded uniformly away from zero.

1.6.2 The Case of More than Two Subdomains

We will now show that the convergence rate of the Schwarz alternating method for three and more subdomains can be estimated in terms of the convergence rates of certain problems on two subdomains; this result was first given in [464]. We note that results of a somewhat different nature for the same problem can be based on the abstract theory of Chap. 2 and the results of Chap. 3. In particular, several methods with coarse spaces will be introduced and analyzed in the latter chapter.

We will consider the case of three and more subregions and a symmetric variant of the algorithm. Exact solvers are employed and the operators P_i are thus projections. For the case of three subdomains, as in Fig. 1.4, the error propagation operator is given by

$$(I - P_1)(I - P_2)(I - P_3)(I - P_2)(I - P_1); \tag{1.41}$$

see Sect. 1.4.2.

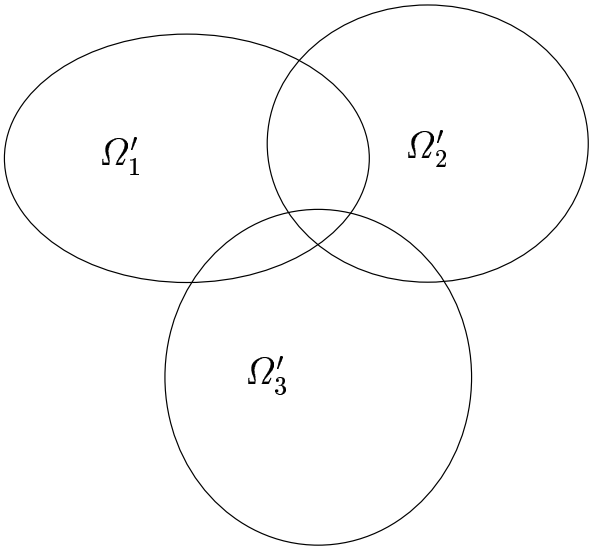

Fig. 1.4. Three overlapping subdomains.

This represents an algorithm where we solve Dirichlet problems in Ω_1', Ω_2', and Ω_3', in that order, prior to returning to solve on the second and first subdomains. Since P_3 is a projection, we can also write this operator as $E_3^* E_3$, where the transpose $*$ is understood in the sense of the bilinear form $a(\cdot, \cdot)$, and $E_3 = (I - P_3)(I - P_2)(I - P_1)$. This algorithm can clearly be generalized to the case of $k > 3$ subdomains and we will derive a bound for the condition number of the symmetric multiplicative Schwarz operator

$$I - E_k^* E_k \quad \text{where} \quad E_k = (I - P_k) \ldots (I - P_1) \tag{1.42}$$

in terms of the condition numbers of problems on pairs of subdomains for which our analysis of subsection 1.6.1 can be used. Since the P_i are orthogonal projections, the maximum eigenvalue of $I - E_k^* E_k$ is 1, and the smallest eigenvalue of this operator equals $1/\kappa$, where κ is its condition number. It is now easy to show that $\|E_k\|_a \leq (1 - 1/\kappa)^{1/2}$, with $\| \cdot \|_a$ the norm induced by the scalar product $a(\cdot, \cdot)$. Thus, our bound for the condition number of $I - E_k^* E_k$, given in Theorem 1.6, will immediately translate into an estimate of the rate of convergence of the multiplicative Schwarz method.

Before we formulate and prove the main result of this subsection, we will examine the recursive structure of the algorithm in some detail. We first solve a problem in Ω_1' resulting in a zero residual in that subregion. For the final, $(2k-1)$th, fractional step, we return to the linear system for the same subregion Ω_1'. The second through $(2k-2)$th fractional steps provide an updated solution in $\Omega \setminus \Omega_1'$. In the final fractional step, we only need the original right hand side and the new values on $\partial\Omega_1'$ obtained in the intermediate steps. All the values in Ω_1' are overwritten and do not enter the computation in the final fractional step in any other way. Similarly, the third through $(2k-3)$th step can be viewed as solely providing updated values in the set $\Omega \setminus (\Omega_1' \cup \Omega_2')$ including new Dirichlet data on part of $\partial\Omega_2'$. We can therefore view the Schwarz algorithm in terms of two regions, Ω_1' and $\Omega_2' \cup \ldots \cup \Omega_k'$ where we obtain approximative values in $\Omega \setminus \Omega_1'$ by a recursive call of one step of the symmetric, multiplicative Schwarz algorithm using the subregions Ω_2' and $\Omega_3' \cup \ldots \cup \Omega_k'$, etc.

We will now give a matrix representation of the process. Given a region $\Psi = \Psi_1' \cup \Psi_2'$, and associating the subvectors x_1, x_2, and x_3 with the degrees of freedom of Ψ_1', $\Psi_2 := \Psi \setminus \overline{\Psi_1'} = \Psi_2' \setminus \overline{\Psi_1'}$, and $\partial\Psi_1' \setminus \partial\Psi$, respectively, we can write the coefficient matrix for the region Ψ as

$$A = \begin{pmatrix} A_{11} & 0 & A_{13} \\ 0 & A_{22} & A_{23} \\ A_{13}^T & A_{23}^T & A_{33} \end{pmatrix} ; \tag{1.43}$$

cf. (1.5). We can make the substitutions $\Psi_1' = \Omega_i'$, $\Psi_2' = \Omega_{i+1}' \cup \ldots \cup \Omega_k'$, and thus $\Psi_2 = (\Omega_{i+1}' \cup \ldots \cup \Omega_k') \setminus \overline{\Omega_i'}$, for $1 \leq i \leq k-1$, as suggested by our discussion of the recursive formulation of the algorithm.

We now consider the first step of an exact block Cholesky factorization of the matrix (1.43) and obtain

$$A = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ A_{13}^T A_{11}^{-1} & 0 & I \end{pmatrix} \begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & A_{23} \\ 0 & A_{23}^T & A_{33}^{(2)} + S^{(1)} \end{pmatrix} \begin{pmatrix} I & 0 & A_{11}^{-1} A_{13} \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix} . \tag{1.44}$$

As in Sect. 1.2, we write the matrix A_{33} as a sum of $A_{33}^{(1)}$ and $A_{33}^{(2)}$, which represent the contributions from integrals over Ψ_1' and $\Psi_2 = \Psi \setminus \overline{\Psi_1'}$, respectively.

As before, the Schur complement $S^{(1)} = A_{33}^{(1)} - A_{13}^T A_{11}^{-1} A_{13}$ is an intermediate matrix in a block Gaussian elimination of the matrix

$$\begin{pmatrix} A_{11} & A_{13} \\ A_{13}^T & A_{33}^{(1)} \end{pmatrix}, \tag{1.45}$$

which represents the contribution from Ψ_1' to the stiffness matrix of the entire problem. We note that solving a linear system of equations with the Schur complement $S^{(1)}$ as the coefficient matrix gives the same result as solving a system with the matrix (1.45), after extending the right hand side by zero, and then discarding the components of the solution corresponding to the subvector x_1. This can easily be seen by considering a block factorization of the matrix (1.45).

As already shown in subsection 1.6.1, the multiplicative Schwarz algorithm for two subregions can be interpreted in terms of the solution of a linear system with a modified factored matrix obtained by replacing the Schur complement $S^{(1)}$, which corresponds to the region Ψ_1', by $S^{(3)}$, which is the Schur complement corresponding to $\Psi_3 = \Psi_1' \cap \Psi_2'$. The matrix, which represents the preconditioner for the symmetric two-region Schwarz algorithm, is then of the form,

$$\hat{A} = \begin{pmatrix} I & 0\,0 \\ 0 & I\,0 \\ A_{13}^T A_{11}^{-1} & 0\,I \end{pmatrix} \begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & A_{23} \\ 0 & A_{23}^T & A_{33}^{(2)} + S^{(3)} \end{pmatrix} \begin{pmatrix} I\,0 & A_{11}^{-1} A_{13} \\ 0\,I & 0 \\ 0\,0 & I \end{pmatrix}. \tag{1.46}$$

We remark that the second and third factors of the matrix $\hat{A}$ in (1.46) can be modified to obtain a nonsymmetric factorization

$$\begin{pmatrix} I & 0\,0 \\ 0 & I\,0 \\ A_{13}^T A_{11}^{-1} & 0\,I \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ 0 & A_{22} & A_{23} \\ 0 & A_{23}^T & A_{33}^{(2)} + S^{(3)} \end{pmatrix} \begin{pmatrix} A_{11} & 0 & A_{13} \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix}. \tag{1.47}$$

The three fractional steps, which correspond to the symmetric variant of the Schwarz method for the case of two subregions, are given directly by the three factors of the matrix (1.47).

We have now set the stage for using standard techniques to estimate the condition number of the two subregion Schwarz method in terms of a generalized Rayleigh quotient involving the two matrices (1.44) and (1.46). Since these matrices have the same first and third factors, we might as well consider the generalized Rayleigh quotient

$$\frac{\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & A_{23} \\ 0 & A_{23}^T & A_{33}^{(2)} + S^{(1)} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}}{\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & A_{23} \\ 0 & A_{23}^T & A_{33}^{(2)} + S^{(3)} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}}. \tag{1.48}$$

As noted in Sect. 1.6.1, it is easy to show, by elementary variational arguments, that $S^{(3)} \geq S^{(1)}$. Therefore the upper bound of the Rayleigh quotient is 1 and this bound is attained for $y_2 = 0,\ y_3 = 0$. The lower right two-by-two principal minors determine the lower bound on the spectrum of the relevant generalized eigenvalue problem. The relevant bound can then be obtained from the reduced generalized Rayleigh quotient

$$\frac{\begin{pmatrix} y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} A_{22} & A_{23} \\ A_{23}^T & A_{33}^{(2)} + S^{(1)} \end{pmatrix} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix}}{\begin{pmatrix} y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} A_{22} & A_{23} \\ A_{23}^T & A_{33}^{(2)} + S^{(3)} \end{pmatrix} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix}} . \tag{1.49}$$

It is important to note that the matrix of the denominator of formula (1.49) is the Schur complement obtained by eliminating the variables of $\Psi_3 = \Psi_1' \cap \Psi_2'$ from the stiffness matrix corresponding to all of Ψ_2'. It is equally important to note that the changes in the matrix (1.46) when we go from a symmetric multiplicative Schwarz method on the two subregions Ω_1' and $\Omega_2' \cup \Omega_3'$ to one on three subregions Ω_1', Ω_2', and Ω_3' are confined to the same lower right two-by-two principal minor since from our discussion of the recursive nature of the algorithm, we have learned that the exact solution on the subregion $\Omega_2' \cup \Omega_3'$ should be replaced by a symmetric multiplicative Schwarz step, with three fractional steps. In both cases, we will solve a problem with zero residuals in Ψ_1' and discard the part of the solution which corresponds to the nodes in that region; the resulting matrices are thus Schur complements obtained by eliminating the variables in the subregion of $\Omega_2' \cup \Omega_3'$ that overlaps Ω_1'. In the case of three subdomains, the relevant generalized Rayleigh quotient is of the form

$$\frac{\begin{pmatrix} y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} A_{22} & A_{23} \\ A_{23}^T & A_{33}^{(2)} + S^{(1)} \end{pmatrix} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix}}{\begin{pmatrix} y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} \tilde{A}_{22} & \tilde{A}_{23} \\ \tilde{A}_{23}^T & \tilde{A}_{33}^{(2)} \end{pmatrix} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix}} , \tag{1.50}$$

where the elements of the matrix of the denominator is the Schur complement for the three-subdomain case discussed above; we will not need any detailed knowledge on the elements of this matrix.

This generalized Rayleigh quotient can be written as the product of the Rayleigh quotient in formula (1.49) and

$$\frac{\begin{pmatrix} y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} A_{22} & A_{23} \\ A_{23}^T & A_{33}^{(2)} + S^{(3)} \end{pmatrix} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix}}{\begin{pmatrix} y_2 \\ y_3 \end{pmatrix}^T \begin{pmatrix} \tilde{A}_{22} & \tilde{A}_{23} \\ \tilde{A}_{23}^T & \tilde{A}_{33}^{(2)} \end{pmatrix} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix}} . \tag{1.51}$$

The minimal and maximal values of the generalized Rayleigh quotient (1.49) can be estimated as in subsection 1.6.1. We also note that the minimal and

maximal values of the generalized Rayleigh quotient (1.50) can be bounded from below by the product of the minimal values of those of (1.49) and (1.51) and that an upper bound can be obtained similarly. Using the following lemma, we can estimate the extremal values of (1.51) by those for the two-subregion symmetric multiplicative Schwarz method on the subregions Ω_2' and Ω_3', i.e., again by using the result of subsection 1.6.1.

Lemma 1.5 *Let two symmetric, positive definite matrices A and $\tilde{A}$ be given with the same block structure,*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{12}^T & \tilde{A}_{22} \end{pmatrix},$$

and assume that

$$cA \leq \tilde{A} \leq CA.$$

Then, their Schur complements, defined by

$$S = A_{22} - {A_{12}}^T {A_{11}}^{-1} A_{12}$$

and

$$\tilde{S} = \tilde{A}_{22} - \tilde{A}_{12}^T \tilde{A}_{11}^{-1} \tilde{A}_{12},$$

satisfy

$$cS \leq \tilde{S} \leq CS.$$

Proof. We write the matrix A as a sum of two positive semi-definite matrices,

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^T & {A_{12}}^T {A_{11}}^{-1} A_{12} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & S \end{pmatrix}$$

and decompose $\tilde{A}$ in the same way. We note that the first of these matrices is semi-definite: this can be seen easily by carrying out a block Cholesky elimination, which results in a zero Schur complement. Rewriting the assumption of the lemma, using this decomposition, it follows that

$$\begin{aligned} c{x_2}^T S x_2 &\leq c \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^T & {A_{12}}^T {A_{11}}^{-1} A_{12} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + {x_2}^T S x_2 \right) \\ &\leq \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{12}^T & \tilde{A}_{12}^T \tilde{A}_{11}^{-1} \tilde{A}_{12} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + {x_2}^T \tilde{S} x_2. \end{aligned}$$

If we now select

$$x_1 = -\tilde{A}_{11}^{-1} \tilde{A}_{12} x_2,$$

then the first of the quadratic forms of the right hand side vanishes and we obtain one of the inequalities. The other follows in exactly the same way. □

We have now completed all that is required for a proof of the following theorem for the case of three subdomains.

Theorem 1.6 *The condition number of the symmetric Schwarz operator satisfies,*

$$\kappa(\Omega_1', \Omega_2', \ldots, \Omega_k') \leq \kappa(\Omega_1', \Omega_2' \cup \ldots \cup \Omega_k')\, \kappa(\Omega_2', \Omega_3' \cup \ldots \cup \Omega_k') \ldots \kappa(\Omega_{k-1}', \Omega_k'). \tag{1.52}$$

Here the expression on the left is the condition number of the operator $I - E_k^* E_k$. *The first factor on the right hand side is the condition number of the symmetric multiplicative Schwarz operator for the pair of subregions* Ω_1' *and* $\Omega_2' \cup \Omega_3' \cup \ldots \cup \Omega_k'$, *etc.*

No new ideas are required for the proof of the general case. Thus, we find for the case of four subdomains, by a minor modification of the arguments above and by using the notation of the theorem, that

$$\kappa(\Omega_1', \Omega_2', \Omega_3', \Omega_4') \leq \kappa(\Omega_1', \Omega_2' \cup \Omega_3' \cup \Omega_4')\kappa(\Omega_2', \Omega_3', \Omega_4').$$

We can then use the result for three subdomains to replace the final factor of this estimate by a product of two factors that each only involve two subdomains.

Springer

http://www.springer.com/978-3-540-20696-5

Domain Decomposition Methods - Algorithms and Theory

Toselli, A.; Widlund, O.

2005, XV, 450 p., Hardcover
ISBN: 978-3-540-20696-5