Introduction to the Design and Theory of Domain Decomposition Algorithms
A Tutorial Given on February 6, 2011 at UCSD

Olof Widlund
Courant Institute of Mathematical Sciences
New York University
http://www.cs.nyu.edu/cs/faculty/widlund/
widlund@cims.nyu.edu

Scope of This Tutorial

I will primarily look at domain decomposition algorithms for positive definite, symmetric problems arising from low order finite element approximations of elliptic problems. All subproblems solved exactly by Cholesky's algorithm. Some mixed finite element problems also considered.

I will adopt the view that a domain decomposition algorithm provides preconditioners (approximate inverses) M of the large and often very ill-conditioned stiffness matrices A that arise in finite element practice.

They are designed with parallel computing systems in mind and the best of them have proven to scale very well on systems with very many processors. We will focus on scalable algorithms, i.e., those with convergence rates independent of the number of subdomains and on those which also are insensitive to the number of degrees of freedom of the individual subproblems. These days, that number is often in the tens of thousands.

There are many other aspects on domain decomposition, such as the study of multi-physics problem, the search for optimal transmission conditions across the interface between the subdomains, and work on time dependent problems; these topics will not be covered. Nor will work on plate models and problems posed in $H(div)$ and $H(curl)$ and on nonlinear problems.

The presentation is, to a large extent, based on my 2005 monograph, coauthored with Andrea Toselli, and also on more recent work described in Li and W., Inter. J. Numer. Methods Engrg., 2006
Dohrmann and W., SINUM 2009 and Inter. J. Numer. Methods Engrg., 2010.

These and other papers of mine can be downloaded from my web page; see Selected Papers. A good portion will be based on relatively recent work. I will not hesitate proving results.

All this work aims at designing preconditioners M such that $\kappa(M^{-1}A)$, the condition number of the preconditioned operator, is small, while keeping the costs of applying M^{-1} acceptable. A preconditioned Krylov space method is almost always used to accelerate the convergence of the iteration. In practice, the parameters computed in the preconditioned conjugate gradient (PCG) computation can be used to obtain reliable estimates of $\kappa(M^{-1}A)$ by exploiting the connection between the conjugate gradient and Lanczos algorithms.

The development of theory has greatly assisted in the development of improved algorithms. In particular, some of the good choices of primal constraints and scalings for FETI–DP and BDDC methods are unlikely to have been found without theoretical work. The theory can be viewed as a subfield of finite element theory.

Poisson's Equation and a Simple Finite Element Model

By using Green's formula, we can write Poisson's equation as a variational problem: Find $u \in V$ such that $\forall v \in V$,

$$a(u,v) := \int_\Omega \nabla u \cdot \nabla v dx = F(v) := \int_\Omega f v dx + \int_{\partial\Omega_N} g_N v ds.$$

Here f is the load, i.e., the right-hand side and g_N the Neumann data given on $\partial\Omega_N \subset \partial\Omega$. All elements of $V \subset H^1(\Omega)$ vanish on the set $\partial\Omega_D := \partial\Omega \setminus \partial\Omega_N$, first assumed to be non empty. This problem is then uniquely solvable. We will always assume that the domain Ω and all subdomains Ω_i are connected sets. Later, we will also consider, for arbitrary $\rho_i > 0$,

$$a(u,v) := \sum_{i=1}^{N} \rho_i \int_{\Omega_i} \nabla u \cdot \nabla v dx.$$

Introduce a triangulation $\mathcal{T}_h$ of Ω and $V^h \subset V$, the standard piecewise linear finite elements on the triangulation. A linear system $Au = F$ results where u is now the vector of nodal values at all interior nodes and those on $\partial\Omega_N$. The stiffness matrix A is sparse, symmetric, and positive definite and can be very large. The resulting finite element solution $u_h(x)$ is well defined and converges to the solution of the differential equation when the mesh size $h \to 0$.

The smallest eigenvalue $\lambda_1(\Omega)$ of the differential operator, and indirectly that of the stiffness matrix, can be estimated by using Friedrichs' inequality

$$\|u\|^2_{L^2(\Omega)} \leq C_1 a(u,u) + C_2 (\int_{\partial\Omega_D} u ds)^2.$$

For $u \in V$ the second integral vanishes and we get a positive lower bound $1/C_1$ of the Rayleigh quotient $a(u,u)/\|u\|^2_{L^2(\Omega)}$ and of λ_1.

In a pure Neumann problem, $\partial\Omega_N = \partial\Omega$, the Laplace operator and the stiffness matrix have a common null space of constants and the problem is uniquely solvable, modulo a constant, iff $F(1) = 0$. The second eigenvalue $\lambda_2(\Omega)$ of the operator is directly related to Poincaré's inequality:

$$\|u\|^2_{L^2(\Omega)} \leq C_1 a(u,u) + C_2 (\int_\Omega u dx)^2.$$

We note that the second term on the right vanishes if u is orthogonal to the null space; then use Courant–Fischer's theorem to obtain $\lambda_2 \geq 1/C_1$.

Poincaré's inequality is much more subtle than Friedrichs' and an estimate of λ_2 enters many domain decomposition estimates.

The largest eigenvalue of the stiffness matrices can be estimated by using Gershgorin's theorem.

It is important to understand what happens to these two inequalities when the diameter of the domain changes under a dilation; a simple change of variables gives the answer. Certain powers of the dilation factor will appear with the constants. Similarly, the full $H^1(\Omega)-$norm should be defined by

$$\|u\|^2_{H^1(\Omega)} := |u|^2_{H^1(\Omega)} + 1/diam(\Omega)^2 \|u\|^2_{L_2(\Omega)} = a(u,u) + 1/diam(\Omega)^2 \|u\|^2_{L_2(\Omega)}.$$

This formula is obtained by using the standard norm for a domain with diameter 1 and a dilation.

Q1: *What happens to Friedrichs' and Poincaré's inequalities under a dilation?*

Using these inequalities and a few additional, elementary arguments, we can show that the condition numbers of the stiffness matrices grow as Ch^{-2} in the case of quasi-uniform meshes. This accounts for the relatively slow convergence of the conjugate gradient method without preconditioning. This is even more true for linear elasticity problems, in particular if the material parameters vary a lot. The ill-conditioning reflects the fact that a second order differential operator maps $H_0^1(\Omega)$ into its dual $H^{-1}(\Omega)$.

Q2: *What happens if we have a Dirichlet condition at one point only?*

We will consider the same type of stiffness matrices for subdomains Ω_i obtained by integrating over $\Omega_i \subset \Omega$. These matrices will be important building blocks for our finite element models and domain decomposition algorithms.

Two Subdomains and Schur Complements

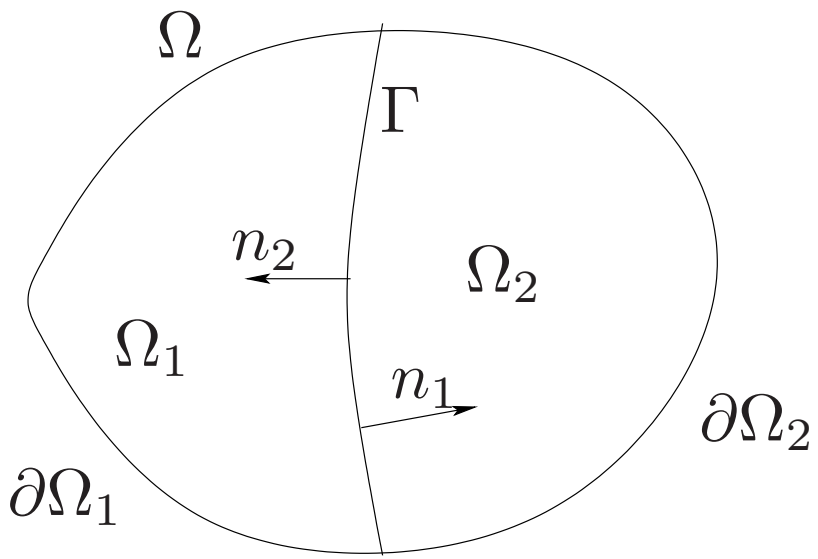


Figure 1: Partition into two non-overlapping subdomains.

Thus, we now consider a domain Ω subdivided into two non-overlapping subdomains Ω_1 and Ω_2. In between the interface Γ.

Consider a finite element approximation of a Poisson problem on Ω (or scalar elliptic, linear elasticity, or even an incompressible Stokes problem.)

Set up a load vector and a stiffness matrix for each subdomain

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

We use a homogeneous Dirichlet condition on $\partial\Omega_i \setminus \Gamma$ but Neumann on Γ.

Subassemble:

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

with $A_{\Gamma\Gamma} = A_{\Gamma\Gamma}^{(1)} + A_{\Gamma\Gamma}^{(2)}$ and $f_\Gamma = f_\Gamma^{(1)} + f_\Gamma^{(2)}$. The degrees of freedom are internal to Ω_1, internal to Ω_2, and those on Γ.

This is a simple example of how stiffness matrices are assembled from those of the subdomains; we add quadratic forms representing the energy contributed by the subdomains.

Eliminate the interior unknowns. This gives two Schur complements:

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

The finite element system of equations can then be reduced to

$$Su_\Gamma = (S^{(1)} + S^{(2)})u_\Gamma = g_\Gamma. \tag{1}$$

If we use exact solvers for the subdomain problems, we can often reduce our discussion to one about Schur complements. We can also take advantage of the reduction in dimension of the Krylov space vectors. Once the interface values are approximated well enough, we can find the values in the interiors by solving a Dirichlet problem for each subdomain. The condition number of a Schur complement of a positive definite symmetric matrix A is always smaller than that of A. Q3: *Why?* In our particular context, the Schur complements will have a condition number on the order of Ch^{-1}. This bound, which is sharp, reflects that we have a mapping from a subspace of $H^{1/2}(\Gamma)$ into $H^{-1/2}(\Gamma)$; $H^{1/2}$ is the trace space of H^1.

It is easy to see that the product of $S^{(i)}$ times a vector can be obtained at essentially the cost of solving a Dirichlet problem; the elements of the Schur complements need not be computed. This is in contrast to when using Cholesky's method for the entire problem. It is known that for any symmetric permutation P, factoring $P^T A P$ will require at least quadratic work in the number of degrees of freedom for any three-dimensional finite element matrix A. Keyword: nested dissection ordering.

The product of S with a vector, as needed when computing a residual for (1), can then be assembled from matrix-vector products with the two subdomain Schur complements.

An important family of domain decomposition methods are the *iterative substructuring* methods – with vocabulary borrowed from structural engineering. They are methods based on non-overlapping subdomains.

By solving a problem with the matrix $A^{(i)}$ with a right-hand side of the form $(0, f_\Gamma^T)^T$, we obtain a solution with the second component equal to ${S^{(i)}}^{-1} f_\Gamma$; this is an easy exercise on block-Gaussian elimination.

Any solution u with such a right-hand side is $discrete\ harmonic$ and is $A^{(i)}-$orthogonal to any v which vanishes on Γ. It therefore provides the $minimal\ energy\ extension$ for given values on Γ. Just note that

$$A_{II}^{(i)} u_I + A_{I\Gamma}^{(i)} u_\Gamma = 0$$

and that v_Γ is assumed to vanish.

Matrix-vector multiplications with $S^{(i)}$ and ${S^{(i)}}^{-1}$ are completely local operations and it does not matter if we have two or many more subdomains; we can use one processor for each subdomain problem and work in parallel.

Coupled system of PDE

Consider Poisson's equation on Ω, in 2D or 3D, with zero Dirichlet data on $\partial\Omega$ and with Ω partitioned into two non-overlapping subdomains Ω_i :

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

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

see Fig. 1. Assume that the boundaries of the subdomains are Lipschitz. Consider

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

Under suitable assumptions on f (square integrable) and on the boundaries of the subdomains (Lipschitz) the Poisson problem is equivalent to a coupled problem:

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

u_i is the restriction of u to Ω_i and n_i outward normal to Ω_i. Conditions on the interface Γ are *transmission conditions*. Equivalently, we could select any two independent linear combinations of the traces of the functions and their normal derivatives. By eliminating the interior variables, the transmission conditions give us Poincaré-Steklov operators, similar to Schur complements.

A Word of Caution

A simple one-dimensional problem reveals a potential problem. Consider

$$-d^2u/dx^2 = -2\delta, \quad x \in (-1,1), \quad u(-1) = u(1) = 0.$$

Here δ is the Dirac delta function. The solution is

$$u(x) = -1 - x, \ x < 0, \ u(x) = -1 + x, \ x > 0,$$

which has a jump in its first derivative at $x = 0$. Note that the right-hand side of the differential equation is not in L^2.

This type of problem typically arises if we solve the subproblems exactly and the only remaining nonzero residuals are on the interface. We can view it in terms of a single layer potential of classical potential theory. Luckily, this issue disappears in the finite element context.

Refer to the normal derivatives as *fluxes* $\lambda_\Gamma^{(i)}$.

Approximate the flux: with ϕ_j the nodal basis function for a node on Γ

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

In finite element language:

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

This coincides with the residual for the nodes on Γ of a subdomain Poisson problem with a Neumann condition on Γ.

Setting $\lambda_\Gamma^{(1)} + \lambda_\Gamma^{(2)} = 0$ gives us the third equation in the assembled block linear system of equations.

A Dirichlet-Neumann Method

In terms of differential operators, for $n \geq 0$:

$$
\begin{aligned}
(D)\ &\begin{cases} -\Delta u_1^{n+1/2} = f & \text{in } \Omega_1, \\ u_1^{n+1/2} = 0 & \text{on } \partial\Omega_1 \setminus \Gamma, \\ u_1^{n+1/2} = u_\Gamma^n & \text{on } \Gamma, \end{cases} \\
(N)\ &\begin{cases} -\Delta u_2^{n+1} = f & \text{in } \Omega_2, \\ u_2^{n+1} = 0 & \text{on } \partial\Omega_2 \setminus \Gamma, \\ \dfrac{\partial u_2^{n+1}}{\partial n_2} = -\dfrac{\partial u_1^{n+1/2}}{\partial n_1} & \text{on } \Gamma, \end{cases} \\
& u_\Gamma^{n+1} = \theta u_2^{n+1} + (1-\theta) u_\Gamma^n \quad \text{on } \Gamma,
\end{aligned}
$$

Here θ is a relaxation parameter. We can also use conjugate gradients since we can show that the preconditioner is symmetric and positive definite.

Working with matrices, we find that the finite element version gives:

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

Thus, the preconditioned operator is ${S^{(2)}}^{-1} S = I + {S^{(2)}}^{-1} S^{(1)}$ and we need an upper bound for the eigenvalues of ${S^{(2)}}^{-1} S^{(1)}$ to obtain a condition number bound. We can use the right inequality of

$$c u_\Gamma^T S^{(2)} u_\Gamma \leq u_\Gamma^T S^{(1)} u_\Gamma \leq C u_\Gamma^T S^{(2)} u_\Gamma, \ c > 0, \ C < \infty. \tag{2}$$

For a proof of (2), a finite element extension theorem is needed and it can be established by using an extension theorem for H^1, known to hold for a large class of domains including all Lipschitz domains. Then use an interpolant into V^h, due to Scott and Zhang, Math. Comp. 1990, since the H^1- extension is not necessarily a finite element function.

In the Dirichlet-Neumann algorithm, we can view the value of the solution on Γ as the principal unknown. Once it has been determined with sufficient accuracy, we have also found the solution elsewhere in the domain.

We can alternatively use the flux λ_Γ as the principal unknown. We note that the exact solution satisfies, for $i = 1, 2$,

$$A_{II}^{(i)} u_I^{(i)} + A_{I\Gamma}^{(i)} u_\Gamma^{(i)} = f_I^{(i)}$$

and

$$A_{\Gamma I}^{(i)} u_I^{(i)} + A_{\Gamma\Gamma}^{(i)} u_\Gamma^{(i)} = f_\Gamma^{(i)} + (-1)^i \lambda_\Gamma.$$

We note that the FETI algorithms are principal examples of algorithms of this type.

Neumann-Neumann and FETI algorithms can be described using the same framework. The preconditioner for N-N is ${S^{(1)}}^{-1} + {S^{(2)}}^{-1}$. The preconditioned FETI operator is

$$(S^{(1)} + S^{(2)})({S^{(1)}}^{-1} + {S^{(2)}}^{-1}).$$

These two algorithms require the solution of one Neumann and one Dirichlet problem for each subdomain in each iteration; they can also be made quite robust for problems with large variations in the coefficients and for many subdomains once suitable coarse components of the preconditioners are added.

The proofs of the optimality of all these methods reduces to using (2): Given arbitrary values on Γ, estimate the energy contributed by one subdomain in terms of that of the other.

Extension and Trace Theorems

For any Lipschitz domain Ω, there is a bounded extension operator

$$\mathcal{E}_\Omega : W_p^k(\Omega) \to W_p^k(\mathbb{R}^n),$$

such that

$$\|\mathcal{E}_\Omega u\|_{W_p^k(\mathbb{R}^n)} \leq C_\Omega \|u\|_{W_p^k(\Omega)}.$$

Here the spaces W_p^k are Sobolev spaces based on L^p. They are also defined for non-integer k.

In fact, this result holds for a class of much more irregular domains; see Jones, Acta Mathematica, 1981.

For a Lipschitz domain, it is easy to define the trace $\gamma_0 u$ of a smooth function u on the boundary $\partial\Omega$. γ_0 can also be extended to all of $H^1(\Omega)$ and the range of this mapping is $H^{1/2}(\partial\Omega)$. The $H^{1/2}(\partial\Omega)-$seminorm of an element g, e.g., Dirichlet data given on all of $\partial\Omega$, can be defined by $|\mathcal{H}g|_{H^1(\Omega)}$, where $\mathcal{H}$ is the harmonic extension into Ω, or alternatively by

$$|g|_{H^{1/2}(\partial\Omega)} := \left(\int_{\partial\Omega} \int_{\partial\Omega} \frac{|g(x) - g(y)|^2}{|x-y|^n} dS_x dS_y \right)^{1/2}. \tag{3}$$

It is also important to understand when the extension by 0 from Ω to $\mathbb{R}^n$ defines a bounded operator. This is trivially so for $L^2(\Omega)$ but not for $H^s(\Omega)$ except for $s < 1/2$. Similarly, extension of $H^s(\Gamma), \Gamma \subset \partial\Omega$ by zero to $\partial\Omega \setminus \Gamma$ also defines a bounded operator only for $s < 1/2$; we will be interested primarily in cases when Γ is an edge of a 2D domain or a face of a 3D domain.

The Finite Element Case

Something important can be done for elements of finite element spaces V^h, which after all are finite dimensional subspaces of H^1. As a motivation, let us again consider the case of a domain subdivided into two subdomains Ω_1 and Ω_2. Now assume that we have a zero Dirichlet condition on $\partial\Omega_1 \setminus \Gamma$ but a Neumann condition on the rest of $\partial\Omega$. In particular, we can then have arbitrary values on all of $\partial\Omega_2$. We can show that there is a bounded H^1-extension of such finite element functions given on Ω_1 into Ω_2. This is so since we can construct a bounded extension into $H_0^1(\Omega)$.

Can we find a bounded extension from Ω_2 to Ω_1, while enforcing the Dirchlet condition? Only if we allow a logarithmic factor. Note that the minimal norm extension into Ω_1 involves extending the values, given on Γ, by 0 onto the rest of $\partial\Omega_1$ and then computing the discrete harmonic extension. We note that this type of issue will often arise for subdomains with several neighbors; we then need to cut and paste together the traces.

In the continuous case, this will not always work since any element of $H^1(\Omega_1)$ must have a $H^{1/2}(\partial\Omega_1)$ trace and we will soon learn that there are elements of $H^{1/2}(\Gamma)$ for which a bounded extension by zero cannot be found.

The subspace of $H^{1/2}(\Gamma)$, which allows a bounded extension by zero is known as $H_{00}^{1/2}(\Gamma)$; it is a true subspace. Formula (3) can be used to develop a formula for the square of the $H_{00}^{1/2}(\Gamma)-$norm: Consider a function which vanishes on $\partial\Omega \setminus \Gamma$. The double integral can then be written as the square of the $H^{1/2}(\Gamma)-$ norm and two additional terms which can be shown to be proportional to

$$\int_\Gamma (|u(x)|^2/dist(x,\Gamma))dS_x \tag{4}$$

by using calculus. Here $dist(x,\Gamma)$ is the distance of $x \in \Gamma$ to $\partial\Gamma$.

While the characteristic function of the set Γ does not belong to $H^{1/2}(\partial\Omega)$, since the weighted L^2-term (4) diverges, useful bounds for similar finite element functions can be found.

In particular, consider the function θ_E where E is an edge of a polygon and $\theta_E(x) = 1$ at all finite element nodes interior to E while it vanishes at all other boundary nodes. Then $\|\theta_E\|^2_{H^{1/2}_{00}(E)}$ can be bounded by $C(1+\log(H/h))$ by estimating the weighted L^2-norm term and by showing that $\|\theta_E\|^2_{H^{1/2}(E)}$ is uniformly bounded. The origin of the logarithm is the second term of

$$\int_0^{H/2} (|\theta_E(s)|^2/s)ds = \int_0^h ((s/h)^2/s)ds + \int_h^{H/2} (1/s)ds.$$

We also note that the difference between $|\theta_E|^2_{H^{1/2}(\Gamma)}$ and $|1|^2_{H^{1/2}(\Gamma)} = 0$ is the sum of two‘ integrals over only the meshes next to the boundary of the edge and that they are easy to estimate by a constant.

There is an alternative approach, which has been developed fully for a face F of a tetrahedron. Construct a function ϑ_F, which equals 1 at all nodes interior to F and vanishes on the rest of the boundary of the tetrahedron. In a neighborhood of any of the edges which is part of ∂F, introduce a cylindrical coordinate system and a function which is linear in the angle and independent of the radius, i.e., the distance to the edge. The finite element interpolant θ_F of this function can be shown to be have an energy bounded by $C(1 + \log(H/h))$; this gives an upper bound for the energy of the discrete harmonic extension of these special boundary values.

To complete the analysis of that relatively complicated two subdomain problem, we also need to estimate $a(\mathcal{H}(\theta_\Gamma u), \mathcal{H}(\theta_\Gamma u))$, the energy of the discrete harmonic extension of the finite element interpolant of the product of the cut-off function and the trace of an arbitrary finite element function. Here let Γ be an edge of the two-dimensional domains or a face in three dimensions.

Consider a two-dimensional case and the use of the first approach. We will estimate the trace norm of $I^h(\theta_\Gamma u)$. An estimate of the weighted L^2-norm term is required and we find, after small modifications of the previous arguments, that

$$a(\mathcal{H}(\theta_\Gamma u), \mathcal{H}(\theta_\Gamma u) \leq C(1 + \log(H/h))\|u_h\|^2_{L^\infty(\Omega)}.$$

We then also need to estimate the maximum of any finite element function in terms of its H^1-norm. In one dimension, we can do so uniformly but we can not do so in two dimensions. Consider the function $\log(\log(1/r))$, which is unbounded but which can shown, by simply using calculus and polar coordinates, to have a bounded H^1-norm.

For V^h and two dimensions, we have the following sharp finite element Sobolev inequality

$$\|u_h\|^2_{L^\infty(\Omega)} \leq C(1+\log(H/h))\|u_h\|^2_{H^1(\Omega)}. \tag{5}$$

We also have, by using Poincaré's inequality, and for the average $\bar{u}_h$ of the values of u_h,

$$\|u_h - \bar{u}_h\|^2_{L^\infty(\Omega)} \leq C(1+\log(H/h))|u_h|^2_{H^1(\Omega)}.$$

Following Brenner and Scott, first assume that the maximum of $|u_h|$ is at the centroid of an element K and that we choose this point as the origin; we also rotate the coordinate system appropriately. Since Ω is Lipschitz, there is then a cone $\mathcal{C} := \{(r, \theta) : 0 < r < H, 0 < \theta < \omega < 2\pi\}$, which belongs to Ω with its vertex at the origin and of height H on the order of the diameter of Ω. Denote the maximum by α. Then,

$$\alpha = u_h(r, \theta) - \int_0^r \frac{\partial u_h}{\partial r}(\rho, \theta) d\rho$$

and

$$\alpha^2 \leq 2u_h(r, \theta)^2 + 2 \left(\int_0^r \frac{\partial u_h}{\partial r} d\rho \right)^2 .$$

Let ηh be the distance of ∂K from the origin and split the integral into two.

By Cauchy–Schwarz's inequality, we have

$$\left| \int_0^r \frac{\partial u_h}{\partial r} d\rho) \right| \leq \eta h |u_h|_{W^1_\infty(K)} + \left(\int_{\eta h}^r |\frac{\partial u_h}{\partial r}|^2 \rho d\rho) \right)^{1/2} \log(H/\eta h)^{1/2}.$$

Squaring, multiplying by r, and integrating over $(H/2, H) \times (0, \omega)$, we find that

$$(\alpha^2 - 4(\eta h)^2 |u_h|^2_{W^1_\infty(K)}) \omega (3/8) H^2 \leq$$
$$2 \int_0^\omega \int_{H/2}^H |u_h(r,\theta)|^2 r dr d\theta \quad + \quad 4 \log(H/\eta h)(3/8) H^2 |u_h|^2_{H^1(\Omega)}.$$

Using an inverse inequality and, if necessary, selecting a smaller constant η, we can conclude that (5) holds for the value at the centroid of elements.

We can then complete the proof by a simple argument, which shows that the difference between the value at the centroid and at an arbitrary point of K can be estimated by $|v|_{H^1(K)}$. We just have to note that the restriction of any finite element function to an individual element belongs to a finite dimensional space where all norms are equivalent.

In three dimensions, the best bound is

$$\|u_h\|^2_{L^\infty(\Omega)} \leq C(1/h)\|u_h\|^2_{H^1(K)}. \tag{6}$$

This bound is sharp; try a standard nodal basis function.

The true analog of (5) in three dimensions is

$$\|u_h\|^2_{L^2(E)} \leq C(1 + \log(H/h))\|u_h\|^2_{H^1(\Omega)}. \tag{7}$$

Here E is an edge of the domain Ω. This is essentially proven by using (5) on slices perpendicular to the edge and through the meshpoints on E and integrating in the direction of E. We can replace the norm on the right-hand side by $|u_h|^2_{H^1(\Omega)}$, if we replace u_h by $u_h - \bar{u}_{h,E}$ in the right-hand side, where $\bar{u}_{h,E}$ is the average of u_h over the edge; note that $u_h - \bar{u}_{h,E}$ does not change if we shift u_h by a constant. Thus, by shifting by the average over Ω and by using Poincaré' inequality, we find that

$$\|u_h - \bar{u}_{h,E}\|^2_{L^2(E)} \leq C(1 + \log(H/h))|u_h|^2_{H^1(\Omega)}.$$

If We Divide into Two, We Might not Conquer

Now let Ω be partitioned into a family of non-overlapping subdomains $\{\Omega_i,\ 1 \leq i \leq N\}$ with

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

With $\Gamma_i = \partial\Omega_i \setminus \partial\Omega$, the interface Γ is defined as $\Gamma := \bigcup_i \Gamma_i$. The linear system, with A_{II} a direct sum of the subdomain matrices $A_{II}^{(i)}$, is now written as

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

We have interior degrees of freedom, collected in u_I, and those on Γ in u_Γ. Block-Gaussian elimination, in parallel across the subdomains, gives

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

The Schur complement S and vector g_Γ are subassembled from subdomain quantities. The restriction operators R_i, of zeros and ones, map values on Γ onto those on $\Gamma_i := \partial\Omega_i \cap \Gamma$. Then, by adding contributions of the subdomains to the system energy, we find that

$$S = \sum_{i=1}^{N} R_i^T S^{(i)} R_i,$$
$$g_\Gamma = \sum_{i=1}^{N} R_i^T (f_\Gamma^{(i)} - A_{\Gamma I}^{(i)} {A_{II}^{(i)}}^{-1} f_I^{(i)}).$$

Neumann-Neumann and Dirichlet-Neumann

How to precondition S? Try to design a N-N algorithm as

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

Not scalable since there is no mechanism for global communication of information across the domain in each iteration step. The number of steps required for good progress with conjugate gradients is at least on the order of $1/H$. Also some $S^{(i)}$ singular. Those subdomains are *floating*.

Color subdomains red and black. Use Dirichlet conditions on black and Neumann on red and glue together the red subdomains at the cross points. Gives scalable algorithm in 2D. Condition number bound: $C(1+\log(H/h))^2$.

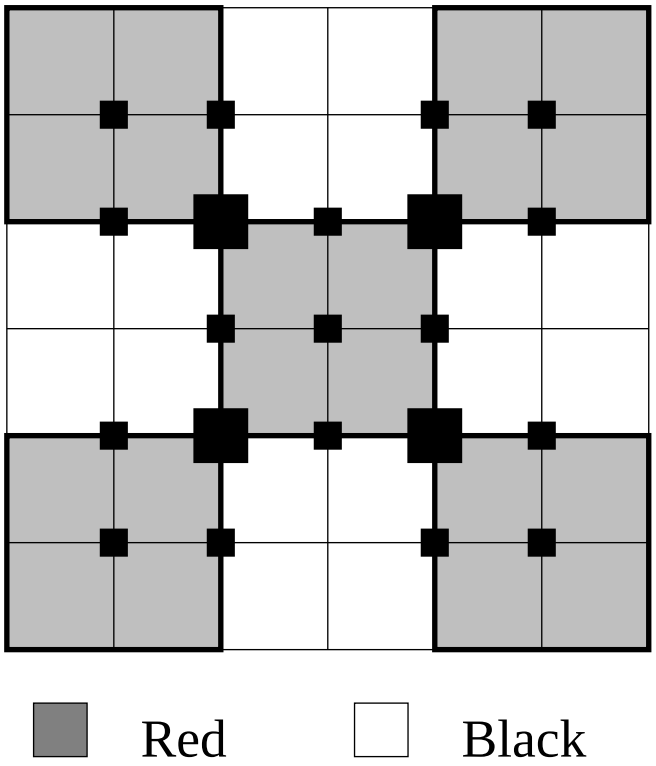


Figure 2: Red-black coloring of the subdomains.

A Word About Block–Cholesky

Consider the block–Cholesky factorization of a symmetric, positive definite block matrix:

$$\begin{bmatrix} A & B^T \\ B & C \end{bmatrix}.$$

We find

$$\begin{bmatrix} A & B^T \\ B & C \end{bmatrix} = \begin{bmatrix} I_A & \\ BA^{-1} & I_C \end{bmatrix} \begin{bmatrix} A & \\ & C - BA^{-1}B^T \end{bmatrix} \begin{bmatrix} I_A & A^{-1}B^T \\ & I_C \end{bmatrix},$$

where I_A and I_C are appropriate identity matrices. The matrix $S := C - BA^{-1}B^T$ is a Schur complement.

It can also be useful to have a formula for the inverse:

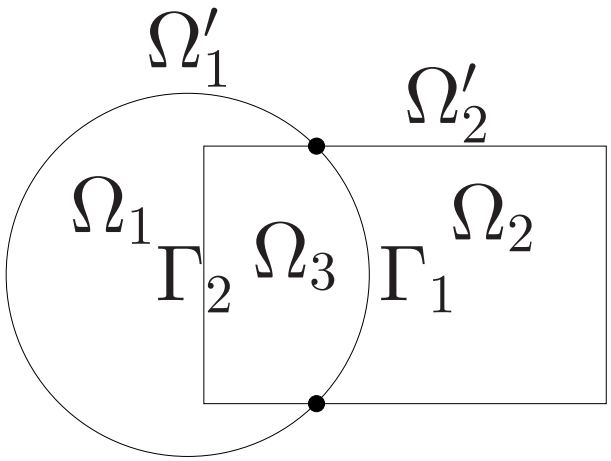
$$\begin{aligned}\begin{bmatrix} A & B^T \\ B & C \end{bmatrix}^{-1} &= \begin{bmatrix} I_A & -A^{-1}B^T \\ & I_C \end{bmatrix} \begin{bmatrix} A^{-1} & \\ & S^{-1} \end{bmatrix} \begin{bmatrix} I_A & \\ -BA^{-1} & I_C \end{bmatrix} \\ &= \begin{bmatrix} A^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \Phi S^{-1} \Phi^T,\end{aligned}$$

where

$$\Phi = \begin{bmatrix} -A^{-1}B^T \\ I_C \end{bmatrix}.$$

It is clearly desirable to have a leading block matrix A which is block diagonal, with many small blocks, and also that the order of C and S be small. This can guide us in the design of preconditioners; see the discussion of BDDC and FETI–DP.

The Schwarz Alternating Method



Here $\Omega_1' := \Omega_1 \cup \Gamma_2 \cup \Omega_3$ and $\Omega_2' := \Omega_2 \cup \Gamma_1 \cup \Omega_3$. The algorithm dates back to 1870 and H.A. Schwarz. He used it to show the existence of the solution of elliptic problems for the union of any two domains for which existence is known; the limit of the iterates, shown to converge, solves the elliptic equation on the new domain. He also used recursion to extend the family of domains for which existence could be established.

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:

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

We can also write this algorithm in terms of projections onto subspaces:

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

where $P_i := R_i^T A_i^{-1} R_i A$. This is the basic *multiplicative* Schwarz method.

Details, Following P.-L. Lions, DD1

Consider the first fractional step. We have, with $u^{n+1/2} - u^n \in H_0^1(\Omega_1')$,

$$a(u^{n+1/2} - u^n, \phi) = a(u - u^n, \phi), \quad \forall \phi \in H_0^1(\Omega_1').$$

The right-hand side above represents the negative of the residual from the previous step. We obtain, $u^{n+1/2} - u^n = -P_1(u^n - u)$ and then

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

The multiplicative Schwarz method can be extended immediately to more than two subdomains by recursion.

In the case of two subdomains, we are effectively solving

$$P_{mu}u := (P_1 + P_2 - P_2P_1)u = g,$$

by a Richardson's method. We can simplify by using the two linear terms only. We then get the basic $additive$ (parallel) Schwarz method:

$$P_{ad}u := (P_1 + P_2)u = g_{ad}.$$

This provides a symmetric operator even for more than two subdomains.

There are other symmetric Schwarz methods such as, for three subdomains, given by the Schwarz polynomial

$$(I - P_1)(I - P_2)(I - P_3)(I - P_2)(I - P_1).$$

There are at least three ways of analyzing the Schwarz methods.

Schwarz used a maximum principle; very limited for finite elements.

We can use an abstract Schwarz theory to be discussed shortly.

For two subdomains, one can also argue about Schur complements and show that

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

We view the iteration in terms of an update of the values on Γ_1. The Schur complement $S_{\Gamma_1}^{(1)}$ corresponds to Ω_1', $S_{\Gamma_1}^{(2)}$ to Ω_2, and $S_{\Gamma_1}^{(3)}$ to Ω_3.

We see that $S_{\Gamma_1}^{(2)} + S_{\Gamma_1}^{(3)} > S_{\Gamma_1}^{(1)} + S_{\Gamma_1}^{(2)}$. Q4: *Why?* An increase in the overlap always improves the convergence. Q5: *Why?*

Details

On $\Gamma_1, u^{n+1/2}$ remains equal to u^n in the first fractional step. After the first full step, the residual vanishes in Ω_1' and Ω_2 after the first fractional step. The residual on Γ then equals $(S_{\Gamma_1}^{(1)} + S_{\Gamma_1}^{(2)})e_{\Gamma_1}^{n+1/2} = (S_{\Gamma_1}^{(1)} + S_{\Gamma_1}^{(2)})e_{\Gamma_1}^{n}$, where $e^n = u^n - u$ is the error after n steps, etc.

In the second fractional step, the value of the approximate solution is updated by solving

$$a(u^{n+1} - u^{n+1/2}, \phi) = a(u - u^{n+1/2}, \phi), \ \forall \phi \in V^h \cap H_0^1(\Omega_2');$$

the right-hand side represents the negative of the residual as shown before.

Thus, $u_\Gamma^{n+1} - u_\Gamma^n$ is obtained by multiplying the residual on Γ by

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

Recursion, Three Subdomains

Consider $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$.

Use symmetric, multiplicative Schwarz. Interpret this as first solving exactly on Ω_1, then inexactly on $\Omega_2 \cup \Omega_3$, and then again on Ω_1. The inexact solve is done by using the two subdomain symmetric multiplicative Schwarz method. We find that the Schwarz polynomial is

$$(I - P_1)(I - P_2)(I - P_3)(I - P_2)(I - P_1).$$

This all generalizes easily to more than three subdomains.

The following result is known. It is expressed in terms of the condition numbers for the three-subdomain and two two-subdomain cases:

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

Block Jacobi Preconditioners

Precondition A by

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

Here $A_i = R_i A R_i^T, \quad i = 1, 2$, and the space is split into two subspaces: $V = R_1^T V_1 \oplus R_2^T V_2$. We can write the preconditioned operator as $P_{ad} = A_J^{-1} A$, i.e., as an additive Schwarz operator and also use more than two subdomains. Note that A_J is obtained by a classical splitting, by removing some off-diagonal blocks. We can also introduce overlap to enhance the convergence; the formula then needs to be modified to

$$P_{ad} = R_1 A_1^{-1} R_1^T A + R_2 A_2^{-1} R_2^T A.$$

The convergence rate of the block Jacobi method, without overlap, can be estimated by using a generalized Rayleigh quotient. We find that

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

An estimate

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

now provides the bound

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

and thus a lower bound for the smallest eigenvalue of P_{ad}. Q6: *What are the worst* u? There is an upper bound of 2; P_{ad} is a sum of two projections.

The same type of bounds are equally relevant for overlapping domains. There is an abstract, relatively elementary theory: Estimate C_0^2 such that

$$\sum_0^N a(R_i^T u_i, R_i^T u_i) \leq C_0^2 a(u,u), \ \forall u \ \text{ with } u = \sum_0^N R_i^T u_i, \ u_i \in V_i.$$

The best $C_0^2 = 1/\lambda_{min}(P_{ad})$. Try $u_i = R_i P_i P_{ad}^{-1} u$; use Cauchy–Schwarz:

$$a(u,u) = \sum_0^N a(u, R_i^T u_i) = \sum_0^N a(P_i u, R_i^T u) \leq$$

$$\left(\sum_0^N a(P_i u, P_i u) \right)^{1/2} \left(\sum_0^N a(R_i^T u_i, R_i^T u_i) \right)^{1/2}.$$

Then, since $a(P_i u, P_i u) = a(P_i u, u)$, we find that $a(u,u) \leq C_0^2 a(P_{ad} u, u)$.

We note that V_0 typically plays a special role of a coarse, global space.

An upper bound for P_{ad} can be obtained by coloring the subdomains with different colors for any pair of subdomains that intersect. The sum of the projections of one color is itself a projection and an upper bound of $N_c + 1$ is obtained for P_{ad} : Each color and the coarse space contribute 1.

Coloring can be replaced by strengthened Cauchy-Schwarz inequalities. For $1 \leq i, j \leq N$, let there be constants ϵ_{ij} such that $\forall u_i \in V_i, \forall u_j \in V_j$

$$|a(R_i^T u_i, R_j^T u_j)| \leq \epsilon_{ij} a(R_i^T u_i, R_i^T u_i)^{1/2} a(R_j^T u_j, R_j^T u_j)^{1/2}.$$

We then obtain, with $\rho(\mathcal{E})$ the spectral radius of the matrix with the elements ϵ_{ij},

$$\kappa(P_{ad}) \leq (\rho(\mathcal{E}) + 1) C_0^2.$$

This bound can easily be modified to allow for inexact solvers of the problems on the V_i.

The parameters C_0^2 and $\rho(\mathcal{E})$ are equally relevant for the multiplicative Schwarz methods. Let,

$$E_{mu} := (I - P_N)(I - P_{N-1}) \cdots (I - P_0).$$

Then, it can be shown that

$$a(E_{mu}u, E_{mu}u) \leq (1 - \frac{1}{(2\rho(\mathcal{E})^2 + 1)C_0^2})a(u, u).$$

There is also an estimate for the case of inexact solvers. Then, the multiplicative algorithm requires more care to ensure convergence.

We will encounter additional, interesting Schwarz algorithms later.

Overlapping Schwarz methods, for many subdomains Ω_i', can be improved by introducing a coarse component of the preconditioner defined on a coarse triangulation $\mathcal{T}_H$ and with a coarse space $V_0 = V^H$. This can be done even if $\mathcal{T}_h$ is not a refinement of $\mathcal{T}_H$, at a cost of more complicated programming. The coarse mesh sizes should be locally comparable to the diameters of the subdomains. The basic, sharp result for second order elliptic problems is

$$\kappa(P_{ad}) \leq C \left(1 + \frac{H}{\delta}\right),$$

where δ measures the overlap between the neighboring subdomains. The proof does not work well in 3D if the material properties change a lot.

How to choose $u_0 \in V^H$ in the analysis? We should reproduce constants locally and have a good energy bound. By using (5), it is easy to show, in 2D, that $a(I^H u_h, I^H u_h) \leq C(1 + \log(H/h))a(u, u)$ where $I^H u_h$ is the standard V^H-interpolant.

For 3D, we would get a bound with the factor H/h. If we have constant coefficients, we can avoid these logarithmic and algebraic factors by replacing $I^H u_h$ by a quasi-interpolant $u_0 := \tilde{I}^H u_h$, where we replace the values at any subdomain vertex V by an average, e.g., over the unino of the coarse elements of $\mathcal{T}_H$, which have V as a vertex. We can then obtain quasi-local bounds of the L^2- and H^1-norms of $u_h - \tilde{I}^H u_h$ in terms of $CH|u_h|_{H^1}$ and $C|u_h|_{H^1}$, respectively.

In the decomposition, the local components can be defined by

$$u_i = R_i(I^h(\theta_i w)) \in V_i, \quad 1 \leq i \leq N,$$

where $w = u - u_0$. The $\{\theta_i\}$ form a piecewise linear partition of unity associated with the overlapping partition. We have $\sum_1^N \theta_i(x) = 1$, and $|\nabla \theta_i| \leq C/\delta_i$.

In the proof, we use the Poincaré and Friedrichs inequalities. Without a coarse component, we will have large L_2-terms and a poor convergence rate: Consider $\nabla(\theta_i u_h)$; we obtain a large coefficient in front of one term.

The core to obtaining an optimal $C(1+(H/\delta))$ bound is the following inequality, which holds for all of $H^1(\Omega_i')$:

$$\|u\|^2_{L^2(\Omega_{i,\delta_i})} \leq C\delta_i^2((1+H_i/\delta_i)|u|^2_{H^1(\Omega_i')} + 1/(H_i\delta_i)\|u\|^2_{L^2(\Omega_i')}). \quad (8)$$

Here $\Omega_{i,\delta_i} \subset \Omega_i'$ is the set of points common to more than one overlapping subdomain.

Cover Ω_{i,δ_i} by on the order of H_i/δ_i patches of diameter δ_i. By using Friedrichs' inequality, properly scaled, and after summing over all the patches, we find

$$\|u\|^2_{L^2(\Omega_{i,\delta_i})} \leq C(\delta_i^2 |u|^2_{H^1(\Omega_{i,\delta_i})} + \delta_i \|u\|^2_{L^2(\partial\Omega_i')}).$$

The last term on the right can be estimated by the following trace inequality:

$$\|u\|^2_{L^2(\partial\Omega_i')} \leq C(H_i |u|^2_{H^1(\Omega_i')} + 1/H_i \|u\|^2_{L^2(\Omega_i')}).$$

Now consider $\nabla(I^h(\theta_i w))$. It is easy to show that the operator I^h does not affect our bound. We have $\nabla(\theta_i w) = \nabla\theta_i w + \theta_i \nabla w$. The second term is easy to estimate. Also note that the first term differs from 0 only in Ω_{i,δ_i}. Then use (8) and Poincaré's inequality to complete the proof.

Additional Additive Schwarz Methods

Let us consider a different type of overlapping subdomains for problems in 2D. There is one subdomain Ω_{ij} for each edge Γ_{ij} of the interface. Γ_{ij} is common to Ω_i and Ω_j and it does not include the endpoints of Γ_{ij}. We introduce

$$\Omega_{ij} := \Omega_i \cup \Gamma_{ij} \cup \Omega_j.$$

Several questions can be posed:

Q7: *Does this set of subdomains and the corresponding finite element subspaces give us a convergent additive Schwarz method?*

Q8: *Is there a mechanism for global communication of information and if not, can we add a good coarse space?*

Q9: *Would it then be a good idea just to use the nodal basis functions associated with the subdomain vertices to construct an additional space* V_0*?*

Q10: *Can we derive bounds for the number of colors and for the parameter* C_0*?*

Q11: *What kind of bound can we derive for the condition number of the resulting additive Schwarz method, which has an effective coarse space?*

Linear Elasticity

Find the displacement $\mathbf{u} \in \mathbf{V}$ of the domain Ω, fixed along $\partial\Omega_D$, with a surface force of density $\mathbf{g}$, along $\partial\Omega_N = \partial\Omega \setminus \partial\Omega_D$, and a body force $\mathbf{f}$:

$$2\int_\Omega \mu\ \epsilon(\mathbf{u}) : \epsilon(\mathbf{v})\ dx + \int_\Omega \lambda\ \text{div}\ \boldsymbol{u}\ \text{div}\ \boldsymbol{v}\ dx\ = < \mathbf{F}, \mathbf{v} > \quad \forall \mathbf{v} \in \mathbf{V}.$$

Here

$$< \mathbf{F}, \mathbf{v} > = \int_\Omega \sum_{i=1}^{3} f_i v_i\ dx + \int_{\partial\Omega_N} \sum_{i=1}^{3} g_i v_i\ dA,$$

$\mu(x)$ and $\lambda(x)$ are the Lamé parameters, the linearized strain tensor

$$\epsilon_{ij}(\mathbf{u}) := \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \quad \text{and} \quad \epsilon(\mathbf{u}) : \epsilon(\mathbf{v}) := \sum_{i=1}^{3}\sum_{j=1}^{3} \epsilon_{ij}(\mathbf{u})\epsilon_{ij}(\mathbf{v}).$$

We also define bilinear forms by

$$a_i(\boldsymbol{u}, \boldsymbol{v}) := 2 \int_{\Omega_i} \epsilon(\mathbf{u}) : \epsilon(\mathbf{v}) \, dx.$$

The Lamé parameters can also be expressed in terms of the Poisson ratio ν and Young's modulus E:

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}, \qquad \mu = \frac{E}{2(1+\nu)}.$$

When $\nu \to 1/2$, we go to the incompressible limit; this is essentially the incompressible Stokes problem.

Rigid Body Modes and Korn's Inequality

For $n = 3$, there are six rigid body modes with zero energy, three translations

$$\boldsymbol{r}_1 := \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{r}_2 := \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{r}_3 := \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

and three rotations

$$\boldsymbol{r}_4 := \frac{1}{H_i} \begin{bmatrix} 0 \\ -x_3 + \hat{x}_3 \\ x_2 - \hat{x}_2 \end{bmatrix}, \boldsymbol{r}_5 := \frac{1}{H_i} \begin{bmatrix} x_3 - \hat{x}_3 \\ 0 \\ -x_1 + \hat{x}_1 \end{bmatrix}, \boldsymbol{r}_6 := \frac{1}{H_i} \begin{bmatrix} -x_2 + \hat{x}_2 \\ x_1 - \hat{x}_1 \\ 0 \end{bmatrix},$$

where $\hat{x}$ is a shift at our disposal and H_i the diameter of Ω_i.

Poincaré's inequality is replaced by Korn's second inequality

$$\|\mathbf{v}\|^2_{\boldsymbol{H}^1(\Omega_i)} \leq C \left(a_i(\mathbf{v},\mathbf{v}) + \frac{1}{H_i^2} \|\mathbf{v}\|^2_{\boldsymbol{L}^2(\Omega_i)} \right).$$

We also have, more importantly, with $\mathcal{RB}$, the space of rigid body modes,

$$\inf_{\mathbf{r}\in\mathcal{RB}} \|\mathbf{v}-\mathbf{r}\|^2_{\boldsymbol{H}^1(\Omega_i)} \leq C a_i(\mathbf{v},\mathbf{v}).$$

Can we successfully extend the two-level additive Schwarz method cases with general subdomains and to compressible elasticity?

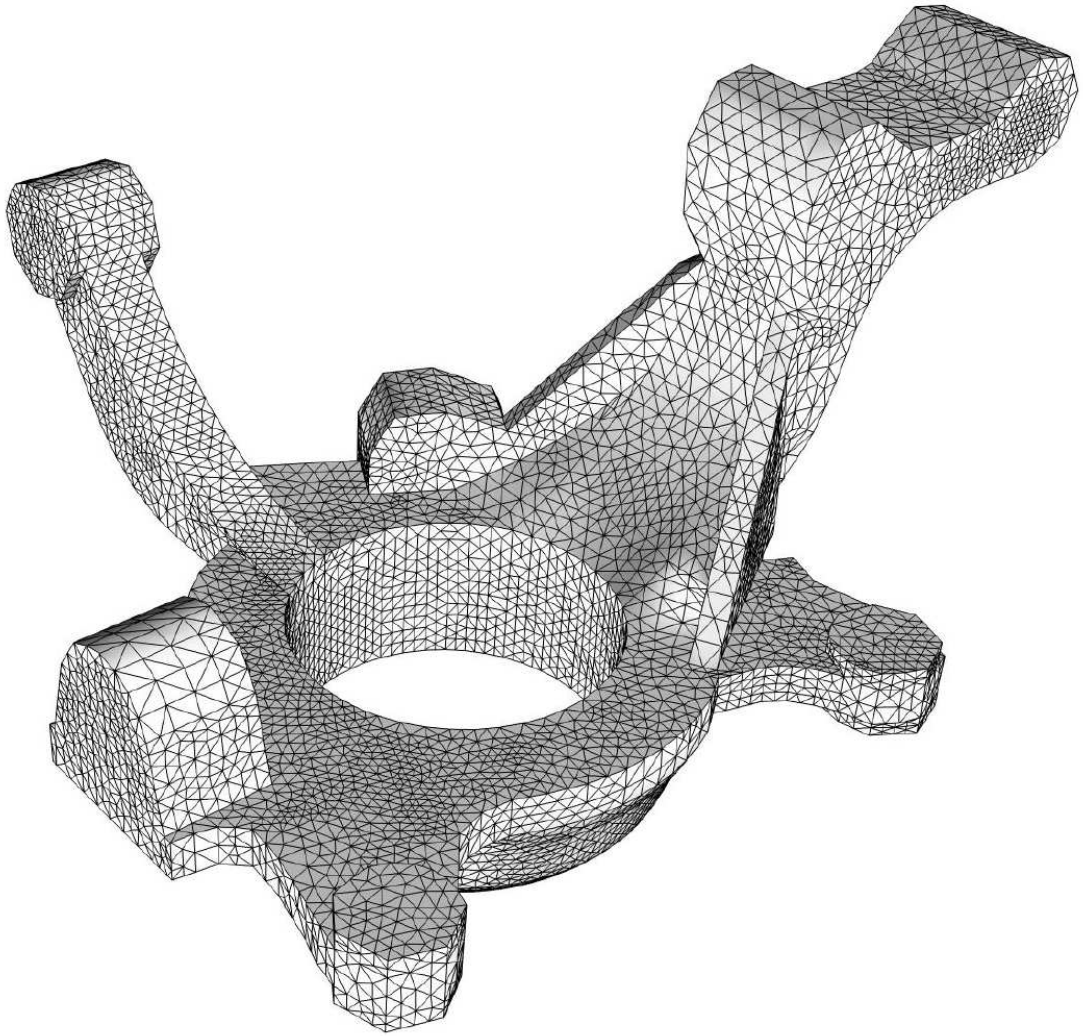


Figure 3: Finite element meshing of a mechanical object.

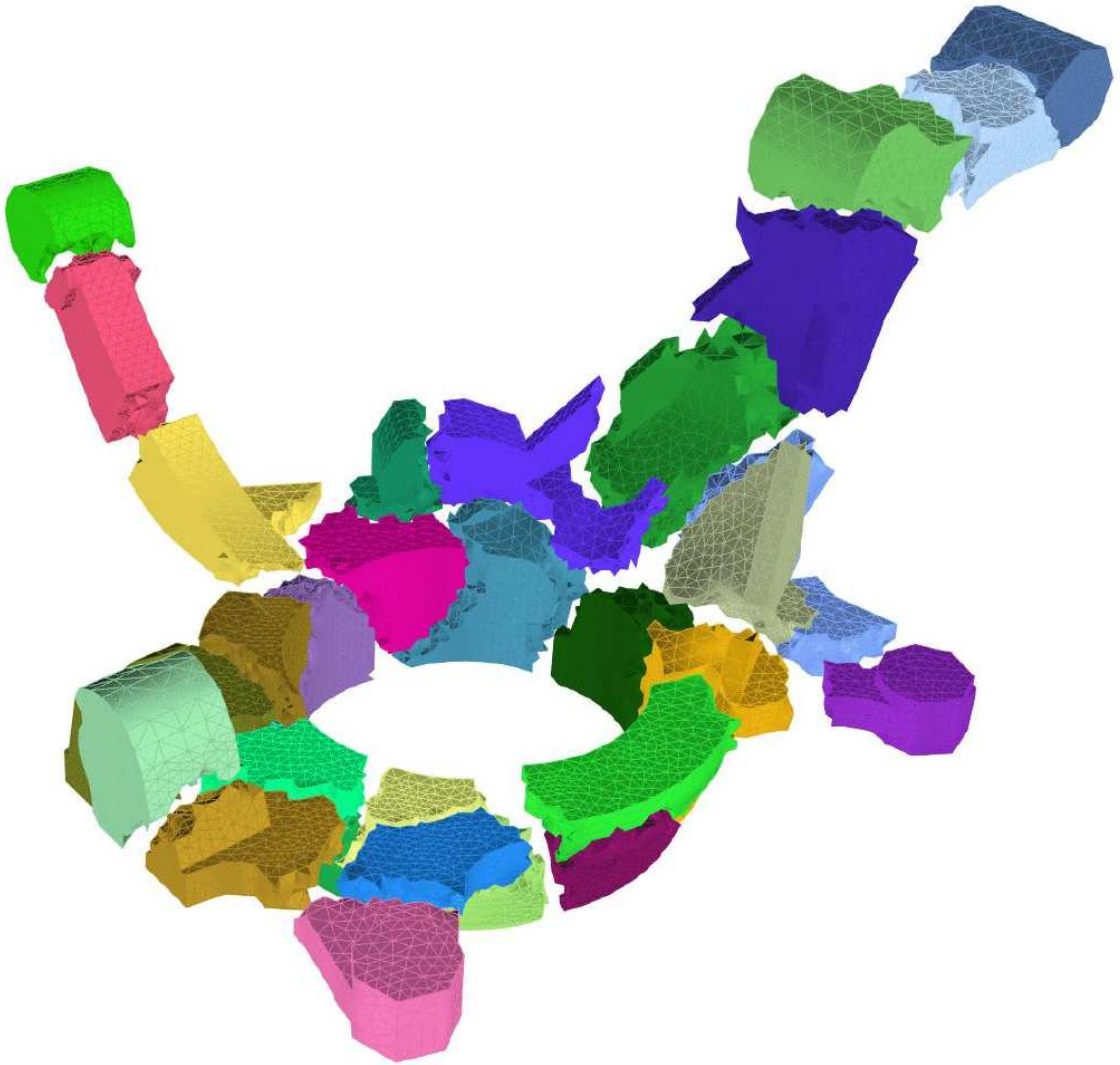


Figure 4: Partition into thirty subdomains. Courtesy Charbel Farhat.

These subdomains, produced by a mesh partitioner, effectively provide our coarse mesh.

Faces, edges, and vertices of quite general subdomains can be defined in terms of certain equivalence classes of finite element nodes. These geometric objects are central in an alternative construction of coarse problems and in the theory. Also highly relevant for parallel computing.

We will use face, edge, and vertex functions, providing a partition of unity on the interface. A face function $\theta_{F^{ij}}$ equals 1 at all nodes of a face common to two subdomains Ω_i and Ω_j and vanishes at all other interface nodes. They are extended as $discrete\ elasto\text{-}harmonic$ functions, i.e., with minimal elastic energy; this determines the values at interior nodes. Similarly, we have edge functions and vertex functions. The restriction of the rigid body modes – all linear functions – to faces and edges, are used for problems of elasticity since the coarse space needs to accommodate all rigid body modes.

Alternative Overlapping Schwarz Methods

Consider a scalar elliptic problem, in three dimensions, and a coarse space which is the range of the interpolation operator

$$I_B^h u(x) = \sum_{V^k \in \Gamma} u(V^k)\theta_{V_k}(x) + \sum_{E^\ell \subset \Gamma} \bar{u}_{E^\ell}\theta_{E^\ell}(x) + \sum_{F^{ij} \subset \Gamma} \bar{u}_{F^{ij}}\theta_{F^{ij}}(x).$$

Here $\bar{u}_{E^i}$ and $\bar{u}_{F^{ij}}$ are averages over edges and faces of the subdomains.

$\theta_{V_k}(x)$ is essentially the standard nodal basis functions of a vertex of the subdomains, $\theta_{E^\ell}(x) = 1$ at the nodes of the edge E^ℓ and vanishes at all other interface nodes, and $\theta_{F^{ij}}(x)$ is a similar function already defined for the face F^{ij}. These functions are extended as discrete harmonic functions into the interior of the subdomains. This interpolation operator I_B^h reproduces constants since the basis functions form a partition of unity.

For nice enough subdomains, we have a genuinely local bound

$$|u - I_B^h u|^2_{H^1(\Omega_i)} \leq C(1 + \log(H_i/h_i))|u|^2_{H^1(\Omega_i)}. \tag{9}$$

We use Cauchy-Schwarz inequality, a trace theorem, bounds on face and edge functions, and finite element Sobolev inequalities; estimate one term at a time. We can now also handle arbitrary coefficient jumps across the interface since the bounds are local. What is needed in the proof? Estimates of the coefficients and of the energy of the basis functions. We have

$$|\theta_{V_k}|^2_{H^1(\Omega_i)} \leq Ch_i,\ |\theta_{E^\ell}|^2_{H^1(\Omega_i)} \leq CH_i,\ |\theta_{F^{ij}}|^2_{H^1(\Omega_i)} \leq CH_i(1+\log(H_i/h_i)).$$

We also need

$$
\begin{aligned}
|u(V_k)|^2 &\leq C/h_i \|u\|^2_{H^1(\Omega_i)}, \\
|\bar{u}_{E^\ell}|^2 &\leq C/H_i \|u\|^2_{L^2(E^\ell)} \leq C/H_i(1+\log(H_i/h_i))\|u\|^2_{H^1(\Omega_i)}, \\
|\bar{u}_{F^{ij}}|^2 &\leq C/H_i^2 \|u\|^2_{L^2(F^{ij})} \leq C/H_i^2(H_i|u|^2_{H^1(\Omega_i)} + 1/H_i\|u\|^2_{L^2(\Omega_i)}).
\end{aligned}
$$

Finally, we use Poincaré's inequality to obtain (9).

As noted, good spaces for elasticity are obtained by multiplying the rigid body modes by the face and edge functions. The interpolation operator can then preserve all rigid body modes. The coefficients built from averages and first order moments. Results in a large coarse space. Estimates are very similar to those of the scalar case, except we need to use Korn's inequality.

For each face, we can use the finite element interpolant of the product of this face cutoff function and the rigid body modes to obtain six linearly independent functions $I^h(\theta_{F^{ij}}\mathbf{r}_k)$; we extend the resulting boundary values into the interior of the subdomains as discrete elasto-harmonic functions. These coarse basis functions can also be obtained by restricting the rigid body modes to the nodes of F^{ij} and setting the values at all other interface nodes to zero.

Similarly, for a straight edge, we obtain five linearly independent rigid body modes since, as is easy to see, a rigid body mode representing a rotation, with the edge as its axis, is invisible on the edge. In earlier work, we have also learned to handle the case of curved edges, for which we use six degrees of freedom. We thus use coarse basis functions associated with the edge, which are given as $I^h(\theta_{E^{ik}}\mathbf{r})$ where $\mathbf{r} \in \mathcal{RB}$.

For each vertex, finally, we have three degrees of freedom representing the displacement at that point.

We can modify this basis to shrink its dimension by replacing vertex, edge and face contributions by fewer terms. The new coarse basis functions are defined as linear combinations of those of the larger space and in terms of simple least squares problems. The dimension of this coarse space can be decreased to be about half of that of the older one.

We consider one face F^{ij} at a time and one rigid body mode $\mathbf{r}_m$, which defines one of the edge coarse basis functions given above. We obtain the corresponding modified edge coarse basis function by extending its values to the faces which have this edge in common. A similar kind of extension will be used for each subdomain vertex. Doing this correctly will produce a coarse space which includes all rigid body modes.

The face contributions to these modified edge and vertex functions are of the form

$$\sum_{\ell=1}^{6} \alpha_{\ell m} I^h(\theta_{F^{ij}} \mathbf{r}_\ell).$$

To determine the coefficients $\alpha_{\ell m}$, for a modified edge basis function, we solve a least squares problem:

$$\min_{\alpha_{\ell m}} \|I^h(\theta_{E^{ik}} \mathbf{r}_m) - \sum_{\ell=1}^{6} \alpha_{\ell m} \mathbf{r}_\ell\|^2_{\mathbf{L}^2(\partial F^{ij})}.$$

Here, $\mathbf{L}^2(\partial F^{ij}) = L^2(\partial F^{ij})^3$. We define the modified vertex basis functions in the same way. It is elementary to show that $\mathbf{L}^2(\partial F^{ij})-$norm of $\sum_1^6 \alpha_{\ell m} \mathbf{r}_\ell$ will be less than or equal to that of $I^h(\theta_{E^{ik}} \mathbf{r}_m)$.

We can prove bounds on the coefficients, which are necessary for showing that the energy of the new basis functions are acceptable. Their energy exceeds those of the original basis functions by a factor $C(1 + \log(H/h))$.

For the almost incompressible case, we need one additional degree of freedom for each face. For a flat face, we choose a $face\ bubble\ function$ $\theta_{F^{ij}}\mathbf{n}_{F^{ij}}$ where $\mathbf{n}_{F^{ij}}$ is a unit normal to the face. We note that this function is linearly independent of the edge and vertex basis functions since it vanishes on the boundary of the face while the modified edge and vertex functions do not. We also know how to handle curved faces.

Submatrices of assembled stiffness matrices can be used to compute the interior values of the basis elements of the coarse space.

The domain is also covered by overlapping subdomains Ω_i'. δ_i/H_i measures the relative overlap between adjacent subdomains, each of which is a union of elements. The local spaces chosen for the Schwarz methods are

$$V_i = V^h \cap H_0^1(\Omega_i'), \quad i > 0.$$

The standard overlapping subdomains Ω_i' are obtained by repeatedly adding layers of elements starting with Ω_i. The bounds for the local components in the Schwarz decomposition require no new ideas in the compressible case.

Another interesting choice is to work with the Ω_i and $\Omega_{i\delta}$. The $\Omega_{i\delta}$ are obtained by adding layers of elements on both sides of $\Gamma_i := \partial\Omega_i \cap \Gamma$. By using a new hybrid Schwarz method, we can make all residuals interior to the Ω_i vanish in each step. In this, we have a situation similar to that for classical iterative substructuring methods.

Schwarz Methods

Schwarz, 1870, Pierre-Louis Lions, 1987: $(I - P_2)(I - P_1)$.

Standard two-level additive, 1988: $P_0 + \sum_{i \geq 1} P_i$.

Standard hybrid, e.g., as in balancing N-N: $P_0 + (I - P_0) \sum_{i \geq 1} P_i (I - P_0)$.

New hybrid: $(I - \sum_{i \geq 1} P_i)(P_0 + \sum_{i \geq 1} P_{i\delta})(I - \sum_{i \geq 1} P_i)$.

We can use that $(I - \sum_{i \geq 1} P_i)$ is a projection; the subdomains Ω_i do not intersect. Therefore, after the first iteration, we need only apply this operator once per step.

Also note that the residuals vanish in the interior of the subdomains, which allows us to save storage.

Result for Overlapping Schwarz Method

Theorem *The condition number of the preconditioned operator* P_{ad} *satisfies*

$$\kappa(P_{ad}) \leq C(1 + H/\delta)(1 + \log(H/h))^q.$$

Here C *is independent of the mesh size, the number of subdomains, the Lamé parameters, as long as the material is compressible.* H/δ *measures the relative overlap between neighboring overlapping sudomains, and* H/h *the maximum number of elements across any sudomain.* $q = 1$ *for the original richer coarse subspace while* $q = 2$ *for the smaller one.*

Numerical experiments indicate that the result is sharp for the larger coarse space but that the bound for the smaller one should hold with $q = 1$.

What needs to be done in the almost incompressible case?

Almost Incompressible Elasticity

In this case, there is locking and very slow convergence of conventional finite element solutions when $h \to 0$.

A well-known remedy is based on introducing the new variable $p = -\lambda \text{div } \boldsymbol{u} \in U \subset L^2(\Omega)$, called the pressure, and replacing the pure displacement problem with a mixed finite element formulation: find $(\mathbf{u}, p) \in \mathbf{V} \times U$ such that

$$\begin{cases} 2\int_\Omega \mu\ \epsilon(\mathbf{u}) : \epsilon(\mathbf{v})\ dx & - & \int_\Omega \text{div } \boldsymbol{v}\ p\ dx & = & < \mathbf{F}, \mathbf{v} > & \forall \mathbf{v} \in \mathbf{V} \\ -\int_\Omega \text{div } \boldsymbol{u}\ q\ dx & - & \int_\Omega 1/\lambda\ pq\ dx & = & 0 & \forall q \in U; \end{cases}$$

We use a mixed, inf-sup stable finite element method, such as $Q_2 - P_1$, with a discontinuous approximation of $p = -\lambda \operatorname{div} \mathbf{u}$.

The term $\int_\Omega \lambda \operatorname{div} \mathbf{u} \operatorname{div} \mathbf{u} dx$ can dominate the energy norm. Since λ is large but finite, we can eliminate the discontinuous pressure variable on the element level from the saddle point problem. The resulting matrix is then symmetric, positive definite, and very ill conditioned. Will the same domain decomposition algorithms, as for the compressible case, still be fast?

We assume that the problems on the individual subdomains Ω_i are homogeneous, i.e., with slowly varying, or constant, Lamé parameters. We can allow arbitrarily large jumps across the interface Γ.

We will now focus on almost incompressible elasticity; our results also hold for the compressible case and for problems where some subdomain problems are discretized using standard finite element methods and others using mixed finite elements, with discontinuous pressure approximations.

These are the first theoretical results on overlapping Schwarz and saddle point problems. Experimental work and two papers by Klawonn and Pavarino over a decade ago. Our coarse spaces different and more generous; experiments discussed in a DD17 paper, SINUM vol. 47(4), and Internat. J. Numer. Meth. Engng. vol. 82. Algorithms implemented in SALINAS by Dohrmann, which is not in the public domain. They are used extensively by US DOE scientists.

In our proofs, we select an element $\mathbf{u}_0 \in V_0$ which has the same net flux as $\mathbf{u}$ across all individual edges or faces of the interface; for this we use the remaining coarse face degree of freedom representing normal displacements. This makes it possible to have a divergence-free extension of the interface values of $\mathbf{u} - \mathbf{u}_0$.

For the bounds on the local components, we first eliminate the interior parts of $\mathbf{w} := \mathbf{u} - \mathbf{u}_0$ vis a vis the subdomains Ω_i. These components are easy to bound.

We then decompose the remaining discrete, piece-wise, elasto-harmonic part: by construction, the net fluxes of $\mathbf{w}$ vanish across all interior subdomain faces of Γ. The next step involves partitioning the trace of $\mathbf{w}$ *on the interface*, at a cost of two (different) logarithmic factors, and assigning appropriate boundary values for subsets of the Ω_i' or $\Omega_{i\delta}$; each should satisfy the zero net flux condition. The elements of the local components $\mathbf{u}_i$ are then constructed as divergence-free extensions on the relevant subsets. Some of these subsets necessarily have poor aspect ratios and this is reflected in the bound, valid for the method with the richer coarse space:

$$\kappa(P_{ad}) \leq C(1+\log(H/h))(1+\log(H/\delta))(1+H/\delta)^3.$$

A factor $(1 + H/\delta)^2$ originates from the inf-sup constants, (Dobrowolski 2003), which enters when comparing the elastic energy with the square of the norm of $(H^1)^n$. The third factor has the same origin as for the case of Poisson's equations; this is also related to subsets with bad aspect ratios. Experiments reported in the two papers with Dohrmann, previously cited. Dohrmann also has a working variant for the case of continuous pressure spaces, but no theory yet. The results have been extended to H(curl) and edge elements in two dimensions with Jones subdomains; also with Dohrmann, and to H(div) and Reissner-Mindlin plates by Duk-soon Oh and Jong Ho Lee, respectively.

Very good numerical results also for stationary incompressible Navier–Stokes by Clark Dohrmann.

Desirable Properties of Domain Decomposition Solvers

- Should handle arbitrary jumps in material properties between subdomains.
- Use of approximate local solvers should affect iteration count marginally.
- Should work even if stiffness matrix is assembled.
- Should be straightforward to implement in parallel and scalable.
- Should be well supported by theory.
- For elasticity, a seamless transition to the incompressible case.
- Should be well defined for and insensitive to irregularity of subdomains.
- Should handle jumps inside subdomains.

FETI and FETI–DP

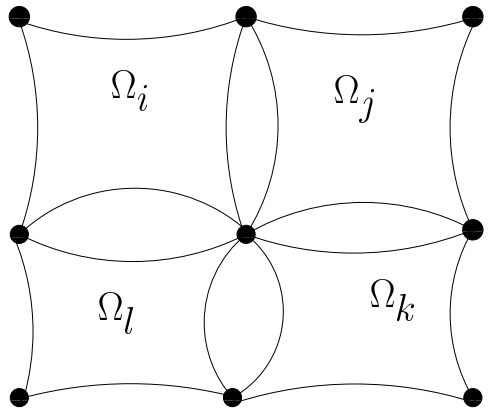
Introduce Lagrange multipliers $\lambda \in U := \mathbf{range}(B_\Gamma)$. B_Γ is a jump operator. Consider the problem:

Find $(u, \lambda) \in W \times U$, such that

$$\left.\begin{array}{rcrcl} Au & + & B_\Gamma^T \lambda & = & f \\ B_\Gamma u & & & = & 0 \end{array}\right\}$$

Eliminate the displacement u by block-Gaussian elimination. Solve the resulting Schur system by PCG. The block diagonal matrix A is, in general, only positive semidefinite. Enforce continuity constraints on $primal$ displacement variables u_Π throughout iterations (as in a primal method); other constraints, on u_Δ, enforced by Lagrange multipliers λ. The local problems then invertible; the primal variables provide a coarse problem.

Two dimensions. Maintain continuity of the primal variables at the vertices (subassemble) and enforce the continuity constraints elsewhere by Lagrange multipliers, which can be interpreted as fluxes.



$$\begin{bmatrix} A_{II}^{(1)} & A_{I\Delta}^{(1)} & & & & \tilde{A}_{\Pi I}^{(1)T} \\ A_{\Delta I}^{(1)} & A_{\Delta\Delta}^{(1)} & & & & \tilde{A}_{\Pi\Delta}^{(1)T} \\ & & \ddots & & & \vdots \\ & & & A_{II}^{(N)} & A_{I\Delta}^{(N)} & \tilde{A}_{\Pi I}^{(N)T} \\ & & & A_{\Delta I}^{(N)} & A_{\Delta\Delta}^{(N)} & \tilde{A}_{\Pi\Delta}^{(N)T} \\ \tilde{A}_{\Pi I}^{(1)} & \tilde{A}_{\Pi\Delta}^{(1)} & \cdots & \tilde{A}_{\Pi I}^{(N)} & \tilde{A}_{\Pi\Delta}^{(N)} & \tilde{A}_{\Pi\Pi} \end{bmatrix}$$

$$\begin{bmatrix} A_{BB} & \tilde{A}_{\Pi B}^T & B^T \\ \tilde{A}_{\Pi B} & \tilde{A}_{\Pi\Pi} & O \\ B & O & O \end{bmatrix} \begin{bmatrix} \mathbf{u}_B \\ \tilde{\mathbf{u}}_\Pi \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_B \\ \tilde{\mathbf{f}}_\Pi \\ 0 \end{bmatrix}$$

FETI–DP in 3D

Good numerical results in 2D; not always very good in 3D. Therefore, in addition to (or instead of) continuity at vertices, constrain certain average values (and moments) of the displacement over individual edges and faces to take common values across the interface.

For scalar second order elliptic equations, this approach yields a condition number estimate $C(1 + \log(H/h))^2$ for certain choices of the primal constraints. Results are independent of jumps in coefficients, if the scaling is chosen carefully. There are good algorithms with quite small coarse problems, i.e., relatively few primal constraints.

Reliable recipes exist for selecting sets of primal constraints for elasticity in 3D which primarily use edge averages and first order moments as primal constraints. High quality PETSc-based codes have been developed and successfully tested on very large parallel computing systems.

These algorithms can be described in terms of three product spaces of finite element functions/vectors defined by their interface nodal values:

$$\widehat{W}_\Gamma \subset \widetilde{W}_\Gamma \subset W_\Gamma.$$

W_Γ: no constraints; $\widehat{W}_\Gamma$: continuity at every point on Γ; $\widetilde{W}_\Gamma$: common values of the primal variables.

We change variables, explicitly introducing primal variables and complementary sets of dual displacement variables. This also appears to make the methods more robust. We can then write the subdomain Schur complements in the form

$$S^{(i)} = \begin{pmatrix} S^{(i)}_{\Delta\Delta} & S^{(i)}_{\Delta\Pi} \\ S^{(i)}_{\Pi\Delta} & S^{(i)}_{\Pi\Pi} \end{pmatrix}.$$

N-N Methods of Same Flavor: BDDC

Work with $\widetilde{W}_\Gamma$, i.e., with a set of primal constraints. At the end of each iterative step, the approximate solution will be continuous at all nodal points of the interface; continuity is restored by applying a weighted average operator E_D, which maps $\widetilde{W}_\Gamma$ into $\widehat{W}_\Gamma$.

In each iteration, we first compute the residual of the fully assembled Schur complement. We then apply E_D^T to obtain the right-hand side of the partially subassembled Schur complement. Solve this system and then apply E_D. This will change the values on Γ, unless the iteration already has converged, and it gives rise to non-zero residuals at nodes next to those on Γ. In a final step of the step of the iteration, these residuals are eliminated by solving Dirichlet problems on each of the subdomains. Accelerate with the preconditioned conjugate gradient method.

The theory can be focused on an estimate of the norm of E_D.

BDDC matrices

Let $\tilde{S}_\Gamma$ denote the partially assembled Schur complement. In practice, we work with interior variables as well when solving linear systems with $\tilde{S}_\Gamma$.

For the BDDC method, we use the fully assembled Schur complement $\tilde{R}_\Gamma^T \tilde{S}_\Gamma \tilde{R}_\Gamma$, $\tilde{R}_\Gamma : \widehat{W} \rightarrow \widetilde{W}$, when computing the residual. Using the preconditioner involves solving a system with the partially subassembled matrix $\tilde{S}_\Gamma$:

$$M_{BDDC}^{-1} := \tilde{R}_{D\Gamma}^T \tilde{S}_\Gamma^{-1} \tilde{R}_{D\Gamma},$$

where $\tilde{R}_{D\Gamma}$ is a scaled variant of $\tilde{R}_\Gamma$ with scale factors computed from the PDE coefficients. $\tilde{R}_{D\Gamma}$ and $\tilde{R}_\Gamma$ have the same sparsity pattern.

This scaling is chosen so that $E_D := \tilde{R}_\Gamma \tilde{R}_{D\Gamma}^T$ is a projection, i.e., $E_D^2 = E_D$. More details later.

FETI–DP Matrices

The basic operator is now $B_\Delta \tilde{S}^{-1} B_\Delta^T$. $\tilde{S}$ is a Schur complement of $\tilde{S}_\Gamma$ obtained after eliminating all primal variables. It is elementary to show that $\tilde{S}^{-1} = R_{\Gamma\Delta} \tilde{S}_\Gamma^{-1} R_{\Gamma\Delta}^T$, where $R_{\Gamma\Delta}$ removes the primal part of any vector defined on Γ. We can therefore write the basic operator as $B_\Gamma \tilde{S}_\Gamma^{-1} B_\Gamma^T$, where $B_\Gamma := B_\Delta R_{\Gamma\Delta}$.

The preconditioner is now

$$M_{FETI}^{-1} := B_{D\Delta} S_{\Delta\Delta} B_{D\Delta}^T = B_{D\Gamma} \tilde{S}_\Gamma B_{D\Gamma}^T,$$

where $S_{\Delta\Delta} = R_{\Gamma\Delta} \tilde{S}_\Gamma R_{\Gamma\Delta}^T$ is the Δ block of $\tilde{S}_\Gamma$ and $B_{D\Delta}$ is a scaled jump operator. We have $B_{D\Gamma} := B_{D\Delta} R_{\Gamma\Delta}$. The scale factors should depend on material parameters and are chosen so that $P_D := B_{D\Gamma}^T B_\Gamma$ is a projection.

We choose the scale factors so that

$$E_D w_\Gamma(x) := \sum_{j \in \mathcal{N}_x} \delta_j^\dagger(x) w^{(j)}(x), \quad x \in \Gamma_i, \tag{10}$$

where $\mathcal{N}_x$ is the set of indices such that $x \in \Gamma_j \cap \Gamma_i$ and $\sum_j {\delta_j}^\dagger(x) = 1$. Similarly,

$$P_D w_\Gamma(x) := \sum_{j \in \mathcal{N}_x} \delta_j^\dagger(x)(w^{(i)}(x) - w^{(j)}(x)), \quad x \in \Gamma_i. \tag{11}$$

We then find that $E_D + P_D = I$ and we can also prove that $E_D P_D = P_D E_D = 0$. Thus, E_D and P_D are complementary projections.

Same Eigenvalues

The preconditioned FETI–DP operator is

$$B_{D\Gamma}\tilde{S}_\Gamma B_{D\Gamma}^T * B_\Gamma \tilde{S}_\Gamma^{-1} B_\Gamma^T$$

Multiply by B_Γ^T on the left and remove the same factor on the right to obtain

$$P_D^T \tilde{S}_\Gamma P_D \tilde{S}_\Gamma^{-1}.$$

The BDDC preconditioned operator is

$$\tilde{R}_{D\Gamma}^T \tilde{S}_\Gamma^{-1} \tilde{R}_{D\Gamma} * \tilde{R}_\Gamma^T \tilde{S}_\Gamma \tilde{R}_\Gamma.$$

Multiply by $\tilde{R}_\Gamma$ on the left and remove same factor on the right to obtain

$$E_D \tilde{S}_\Gamma^{-1} E_D^T \tilde{S}_\Gamma.$$

Let φ be an eigenvector of $P_D^T \tilde{S}_\Gamma P_D \tilde{S}_\Gamma^{-1}$ with the eigenvalue λ. Let $\psi = E_D \tilde{S}_\Gamma^{-1} \varphi$. Then,

$$E_D \tilde{S}_\Gamma^{-1} E_D^T \tilde{S}_\Gamma * E_D \tilde{S}_\Gamma^{-1} \varphi = E_D \tilde{S}_\Gamma^{-1} (I - P_D^T) \tilde{S}_\Gamma (I - P_D) \tilde{S}_\Gamma^{-1} \varphi.$$

This gives us three terms,

$$E_D \tilde{S}_\Gamma^{-1} P_D^T \tilde{S}_\Gamma P_D \tilde{S}_\Gamma^{-1} \varphi = \lambda E_D \tilde{S}_\Gamma^{-1} \varphi \tag{12}$$

and

$$-E_D P_D \tilde{S}_\Gamma^{-1} \varphi + E_D \tilde{S}_\Gamma^{-1} (I - P_D^T) \varphi. \tag{13}$$

$E_D P_D = 0$. Also $(I - P_D^T)\varphi = 0$ since $\varphi \in \mathbf{range}(P_D^T)$. Similarly, any eigenvalue of the BDDC operator is an eigenvalue of the FETI–DP operator.

What we just did is not quite correct. BDDC always has an eigenvalue equal to 1; FETI–DP does not always. There is no real problem modifying the argument. It is in fact correct for any eigenvalue different from 0 and 1.

The analysis of BDDC requires a bound of the $\widetilde{S}-$norm of the average operator E_D. Interestingly enough, a main role, in 2D, is played by the special edge functions θ_E and by the finite element extension theorem; both were previously discussed. In 3D, the face functions θ_F also come into play. As we will soon show, in a special case, we can prove a $C(1+\log(H/h))^2$ condition number estimate if the primal constraints and the scale factors are chosen carefully. Sue Brenner has shown that this bound is sharp.

Role of P_D in Condition Number Bound

Consider the preconditioned FETI–DP operator $B_{D\Gamma}\tilde{S}_\Gamma B_{D\Gamma}^T * B_\Gamma \tilde{S}_\Gamma^{-1} B_\Gamma^T$. Then, for λ in the range of $B_{D\Gamma}$, to which the Krylov space belongs, we have $F\lambda = FB_{D\Gamma}B_\Gamma\lambda$, where $F = B_\Gamma \tilde{S}_\Gamma^{-1} B_\Gamma^T$.

Then, using the matrix F to define an inner products, we find,

$$\lambda^T F\lambda = \lambda^T F B_{D\Gamma} B_\Gamma^T \lambda = \lambda^T F B_{D\Gamma} \tilde{S}_\Gamma^{1/2} \tilde{S}_\Gamma^{-1/2} B_\Gamma^T \lambda \leq$$
$$((\tilde{S}_\Gamma^{1/2} B_{D\Gamma}^T F\lambda)^T \tilde{S}_\Gamma^{1/2} B_{D\Gamma}^T F\lambda)^{1/2} ((\tilde{S}_\Gamma^{-1/2} B_\Gamma \lambda)^T \tilde{S}_\Gamma^{-1/2} B_\Gamma \lambda)^{1/2} =$$
$$(\lambda^T F M^{-1} F\lambda)^{1/2} (\lambda^T F\lambda)^{1/2}$$

from which follows that all eigenvalues of $M^{-1}F$ are ≥ 1.

We also need an upper bound for the eigenvalues of $M^{-1}F$. We find that

$$\lambda^T F M^{-1} F \lambda = (B_{D\Gamma}^T B_\Gamma \tilde{S}_\Gamma^{-1} B_\Gamma \lambda)^T \tilde{S}_\Gamma B_{D\Gamma}^T B_\Gamma \tilde{S}_\Gamma^{-1} B_\Gamma \lambda \leq$$
$$(P_D \tilde{S}_\Gamma^{-1} B_\Gamma \lambda)^T \tilde{S}_\Gamma P_D \tilde{S}_\Gamma^{-1} B_\Gamma \lambda \leq |P_D|^2_{\tilde{S}_\Gamma} \lambda^T F \lambda.$$

Thus, all eigenvalues of $M^{-1}F$ are $\leq |P_D|^2_{\tilde{S}_\Gamma}$ and $\kappa(M^{-1}F) \leq |P_D|^2_{\tilde{S}_\Gamma}$.

We can establish similar bounds for BDDC, with $|E_D|^2_{\widetilde{S}_\Gamma}$ replacing $|P_D|^2_{\widetilde{S}_\Gamma}$, or we can rely on the close connection of the spectra of the two algorithms to obtain an estimate of the condition number of the preconditioned BDDC operator.

A Final Bound for FETI–DP

Consider the simplest interesting case, which is the scalar problem in 2D with the bilinear form

$$a(u,v) := \sum_{1}^{N} \rho_i \int_{\Omega_i} \nabla u \cdot \nabla v dx.$$

Let all subdomain vertices be primal. Choose $\delta_i^\dagger = \rho_i/(\rho_i + \rho_j)$ in (11) for $x \in E^{ij}$, the edge which is the interior of $\Gamma_i \cap \Gamma_j$.

When estimating $|P_D|^2_{\tilde{S}_\Gamma}$, we will use formula (11) and split this sum into terms representing individual subdomain edges.

We also note that $\rho_i(\delta_j^\dagger)^2 \leq \min(\rho_i, \rho_j)$.

Thus, to estimate the $\widetilde{S}_\Gamma-$norm of P_D, we have to consider

$$\rho_i|\mathcal{H}(\theta_{E^{ij}}\delta_j^\dagger(w^{(i)}-w^{(j)}))|^2_{H^1(\Omega_i)}.$$

Write

$$w^{(i)}-w^{(j)} = w^{(i)}-\bar{w}^{(i)}_{\Omega_i}-(w^{(j)}-\bar{w}^{(j)}_{\Omega_j})+\bar{w}^{(i)}_{\Omega_i}-w^{(i)}(V)-(\bar{w}^{(j)}_{\Omega_j}-w^{(j)}(V)).$$

Here $\bar{w}^{(i)}_{\Omega_i}$ is the average of $w^{(i)}$ over Ω_i, etc., and $w^{(i)}(V) = w^{(j)}(V)$ the value at a subdomain vertex, i.e., at one of the endpoints of E^{ij}.

The resulting terms can all be estimated by $\rho_i|w^{(i)}|^2_{H^1(\Omega_i)}$ or $\rho_j|w^{(j)}|^2_{H^1(\Omega_j)}$ by using tools previously developed.