Les méthodes de Schwarz au fil de l'histoire

Martin J. Gander
martin.gander@math.unige.ch

University of Geneva

August 2008

Dirichlet Principle

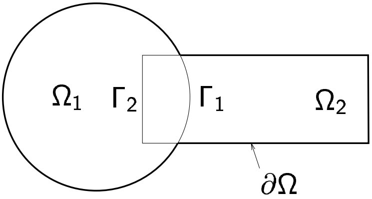
Dirichlet principle: The solution of Laplace's equation $\Delta u = 0$ on a bounded domain Ω with Dirichlet boundary conditions $u = g$ on $\partial\Omega$ is the infimum of the Dirichlet integral $\int_\Omega |\nabla v|^2$ over all functions v satisfying the boundary conditions, $v = g$ on $\partial\Omega$.

H. A. Schwarz 1869:

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

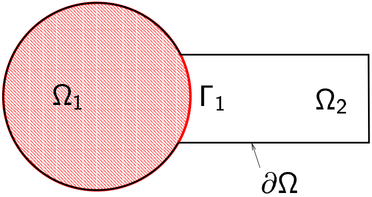
Classical Alternating Schwarz Method

A method to proof that the infimum is attained !



Classical Alternating Schwarz Method

A method to proof that the infimum is attained !

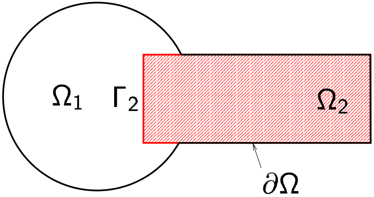


$$
\begin{aligned}
\Delta u_1^1 &= 0 && \text{in } \Omega_1 \\
u_1^1 &= g && \text{on } \partial\Omega \cap \overline{\Omega}_1 \\
u_1^1 &= u_2^0 && \text{on } \Gamma_1
\end{aligned}
$$

disk with known solution

Classical Alternating Schwarz Method

A method to proof that the infimum is attained !

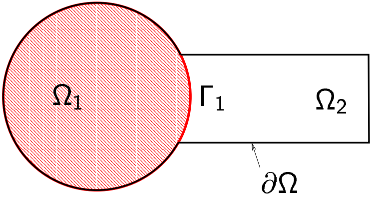


$$
\begin{aligned}
\Delta u_2^1 &= 0 && \text{in } \Omega_2 \\
u_2^1 &= g && \text{on } \partial\Omega \cap \overline{\Omega}_2 \\
u_2^1 &= {\color{red} u_1^1} && \text{on } \Gamma_2
\end{aligned}
$$

rectangle with known solution

Classical Alternating Schwarz Method

A method to proof that the infimum is attained !

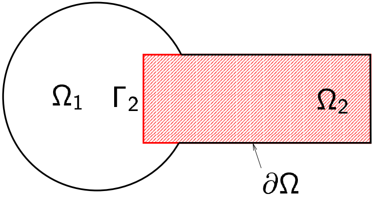


$$\begin{aligned} \Delta u_1^2 &= 0 && \text{in } \Omega_1 \\ u_1^2 &= g && \text{on } \partial\Omega \cap \overline{\Omega}_1 \\ u_1^2 &= u_2^1 && \text{on } \Gamma_1 \end{aligned}$$

disk with known solution

Classical Alternating Schwarz Method

A method to proof that the infimum is attained !

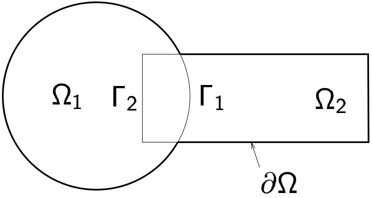


$$
\begin{aligned}
\Delta u_2^2 &= 0 && \text{in } \Omega_2 \\
u_2^2 &= g && \text{on } \partial\Omega \cap \overline{\Omega_2} \\
u_2^2 &= u_1^2 && \text{on } \Gamma_2
\end{aligned}
$$

rectangle with known solution

Classical Alternating Schwarz Method

A method to proof that the infimum is attained !

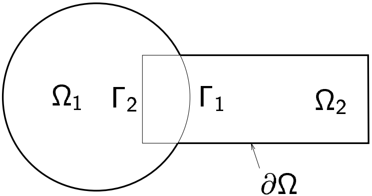


$$
\begin{array}{rlrl}
\Delta u_1^n = 0 & \text{in } \Omega_1 & \Delta u_2^n = 0 & \text{in } \Omega_2 \\
u_1^n = g & \text{on } \partial\Omega \cap \overline{\Omega}_1 & u_2^n = g & \text{on } \partial\Omega \cap \overline{\Omega}_2 \\
u_1^n = u_2^{n-1} & \text{on } \Gamma_1 & u_2^n = u_1^n & \text{on } \Gamma_2
\end{array}
$$

disk with known solution   rectangle with known solution

Classical Alternating Schwarz Method

A method to proof that the infimum is attained !

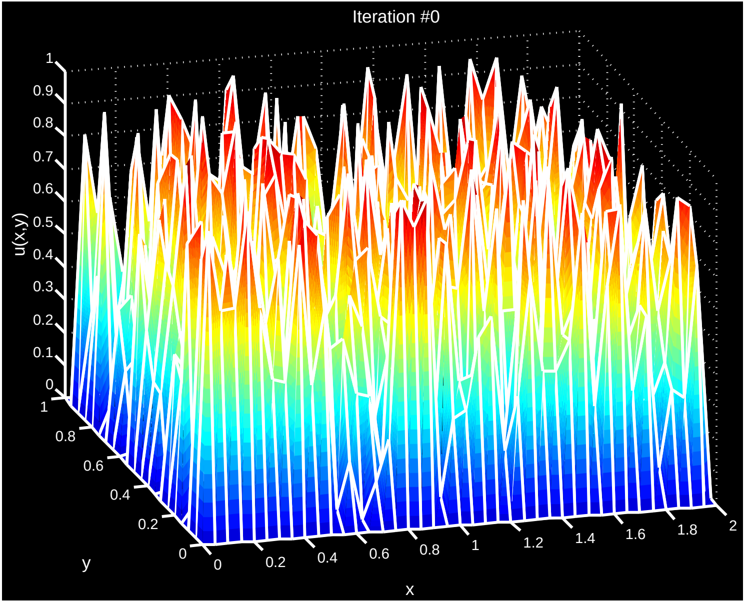


$$\begin{array}{rll} \Delta u_1^n = 0 & \text{in } \Omega_1 \\ u_1^n = g & \text{on } \partial\Omega \cap \overline{\Omega}_1 \\ u_1^n = u_2^{n-1} & \text{on } \Gamma_1 \end{array} \qquad \begin{array}{rll} \Delta u_2^n = 0 & \text{in } \Omega_2 \\ u_2^n = g & \text{on } \partial\Omega \cap \overline{\Omega}_2 \\ u_2^n = u_1^n & \text{on } \Gamma_2 \end{array}$$

disk with known solution   rectangle with known solution

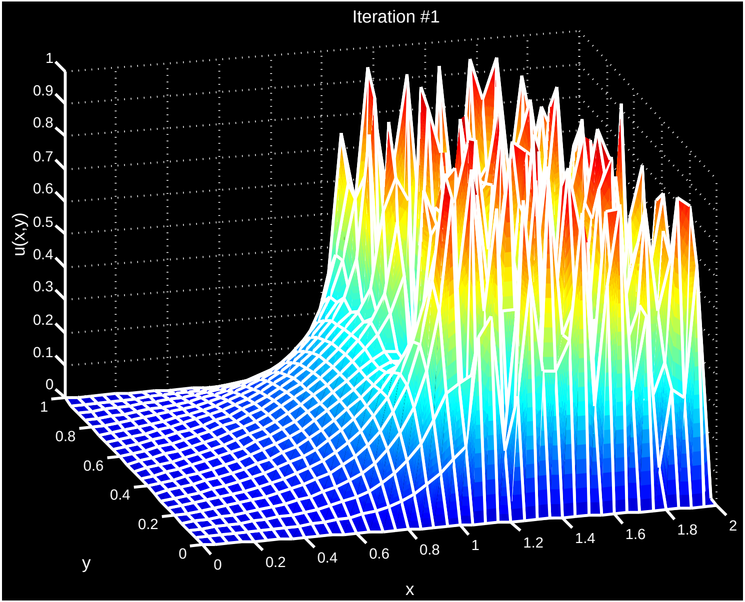
- Schwarz proved convergence in 1869 using the maximum principle.

An Example

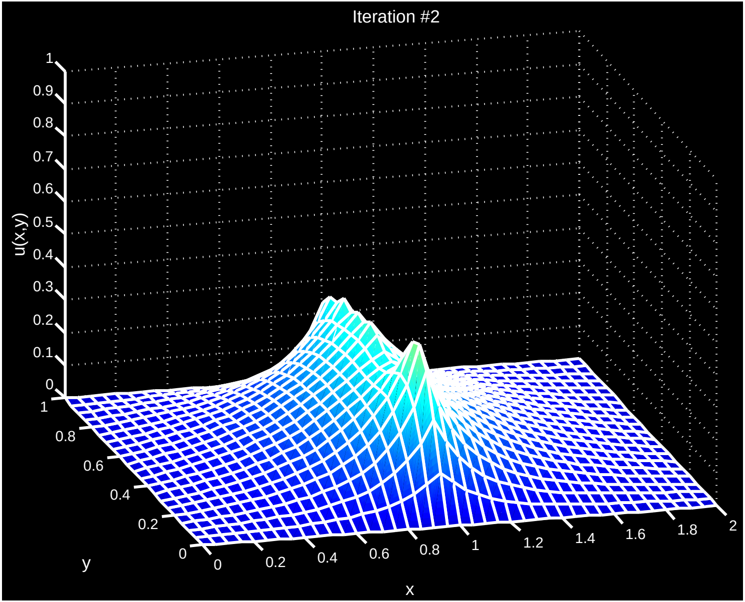


Iteration #0
u(x,y)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
y
1
0.8
0.6
0.4
0.2
0
x
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

Iteration 1

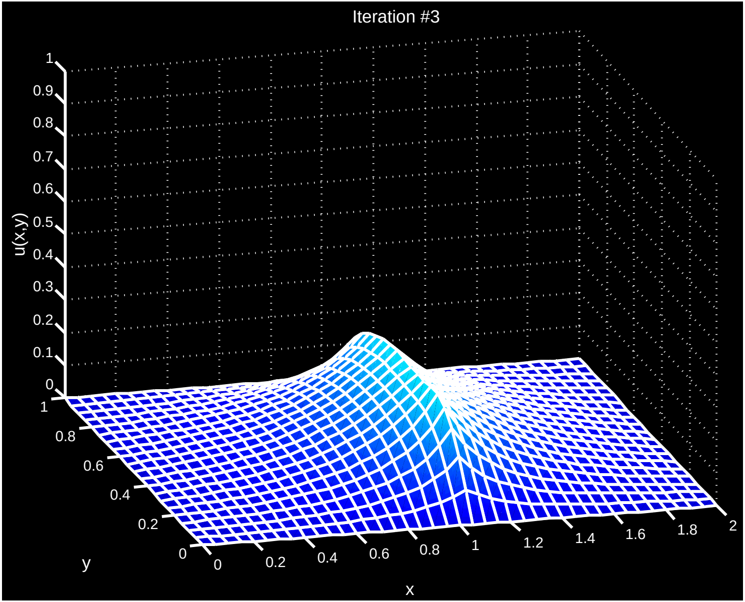

Iteration #1
u(x,y)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
0.8
0.6
0.4
0.2
0
y
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
x

Iteration 2



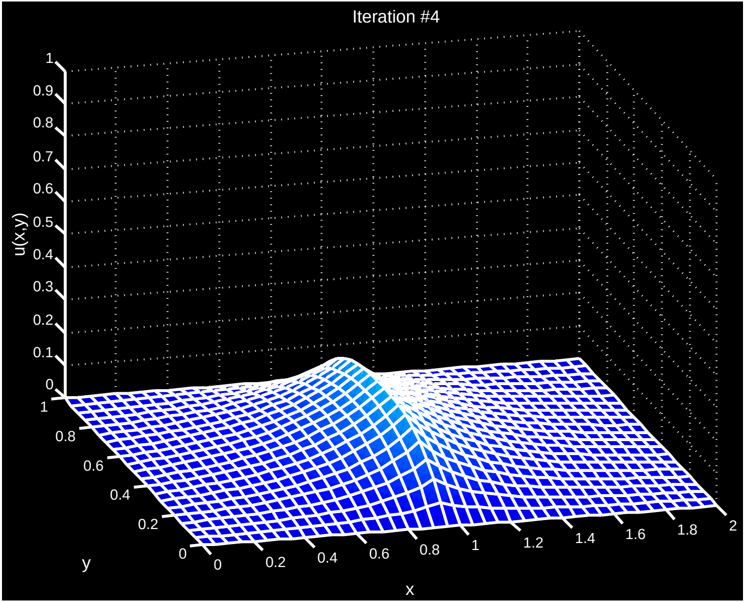
Iteration #2
u(x,y)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
0.8
0.6
0.4
0.2
0
y
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
x

Iteration 3



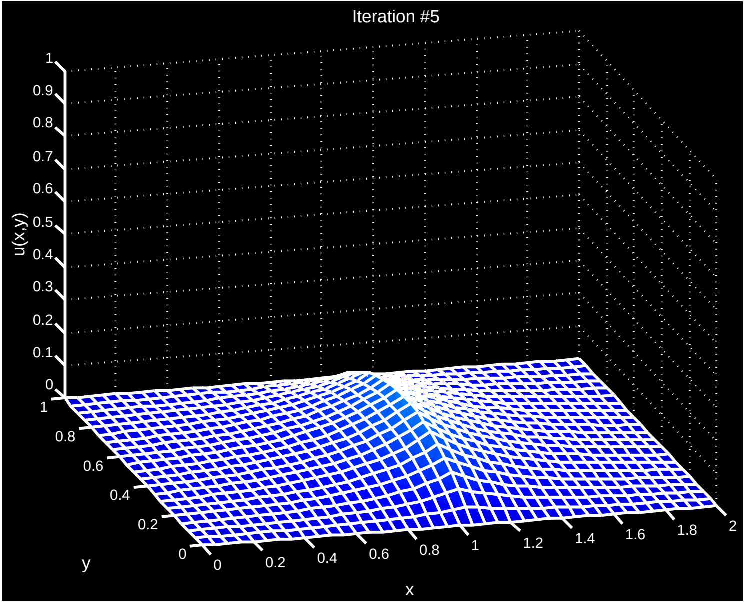
Iteration #3
u(x,y)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
y
1
0.8
0.6
0.4
0.2
0
x
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

Iteration 4



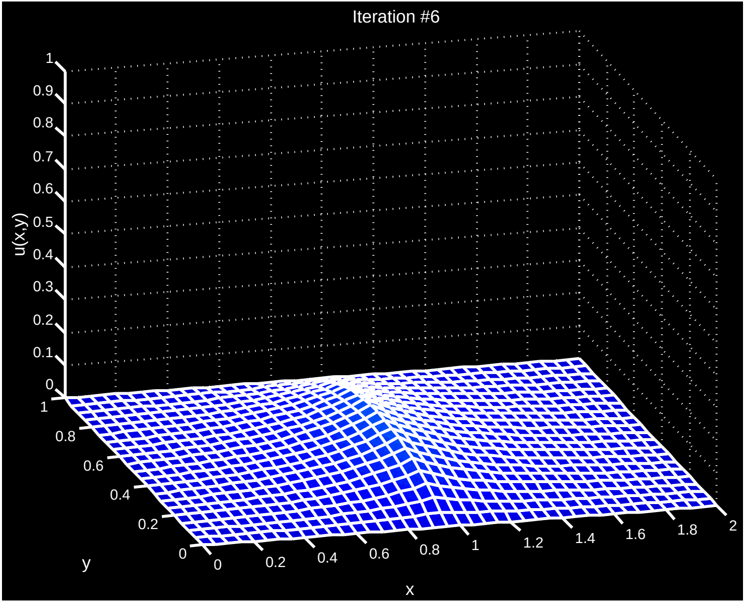
Iteration #4
u(x,y)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
0.8
0.6
0.4
0.2
0
y
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
x

Iteration 5



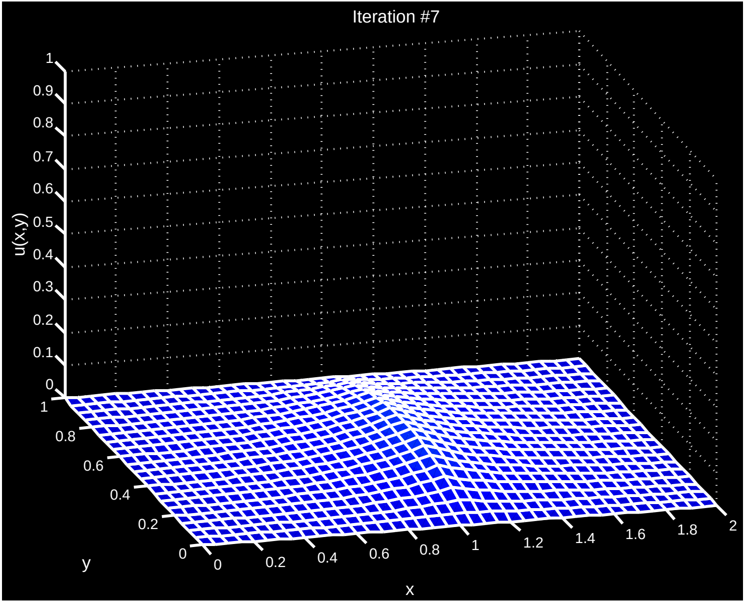
Iteration #5
u(x,y)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
y
1
0.8
0.6
0.4
0.2
0
x
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

Iteration 6

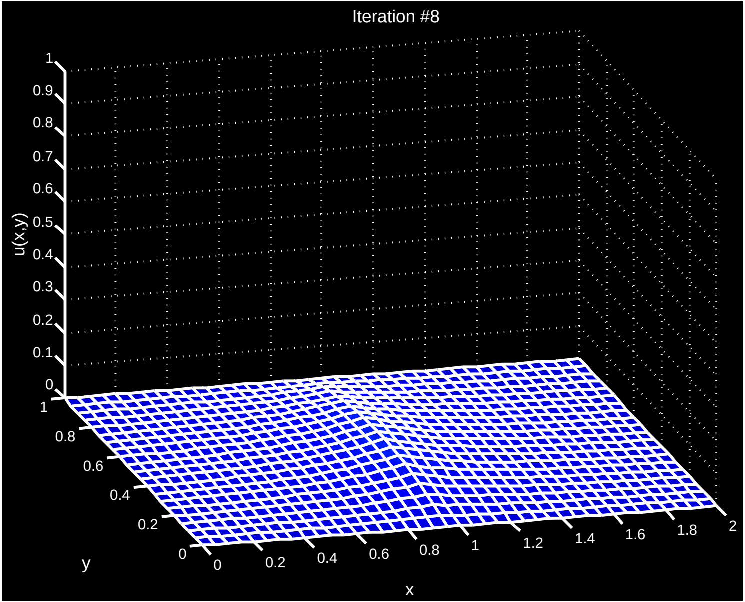


Iteration #6
u(x,y)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
y
0
0.2
0.4
0.6
0.8
1
x
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

Iteration 7

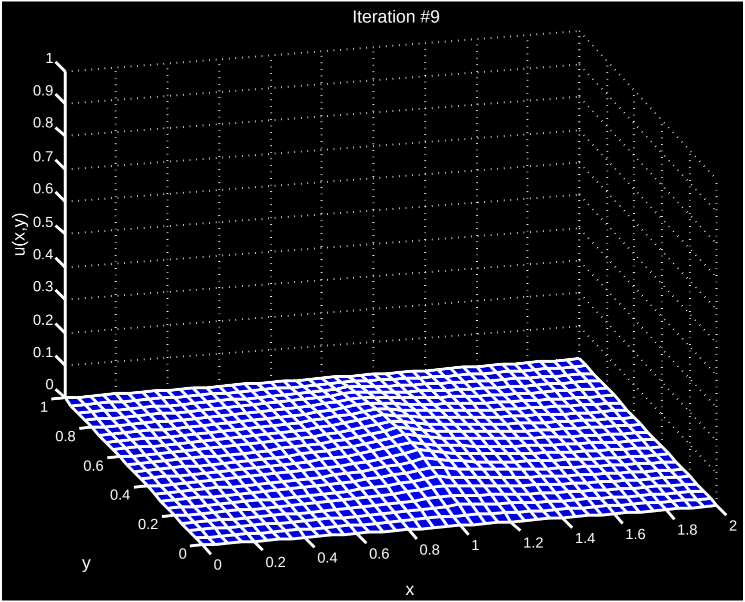

Iteration #7
u(x,y)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
y
0
0.2
0.4
0.6
0.8
1
x
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

Iteration 8



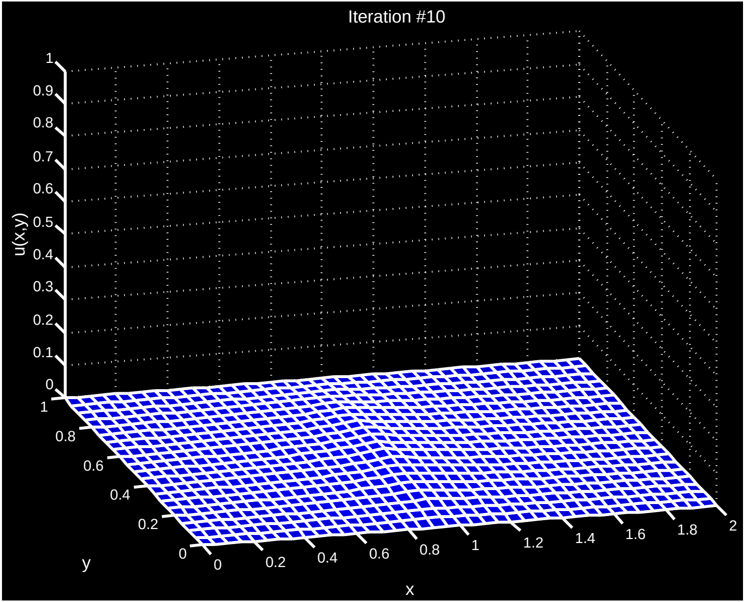
Iteration #8
u(x,y)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
y
1
0.8
0.6
0.4
0.2
0
x
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

Iteration 9

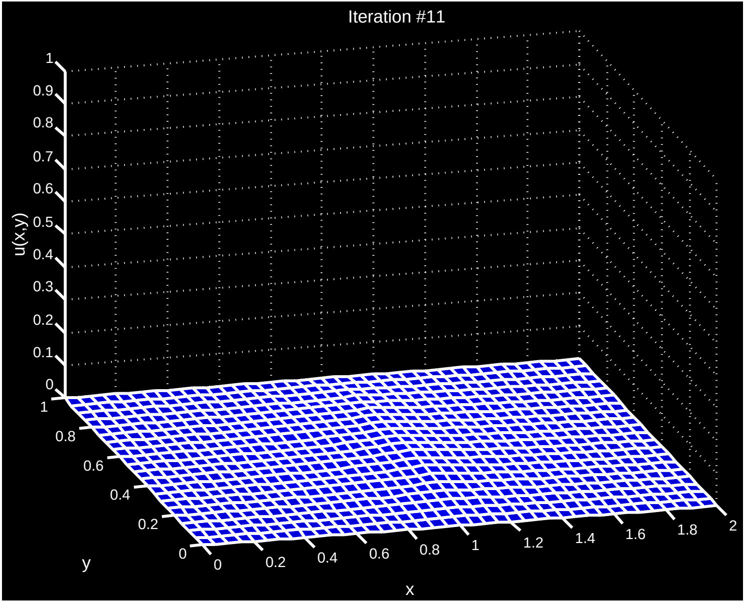


Iteration #9
u(x,y)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
y
1
0.8
0.6
0.4
0.2
0
x
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

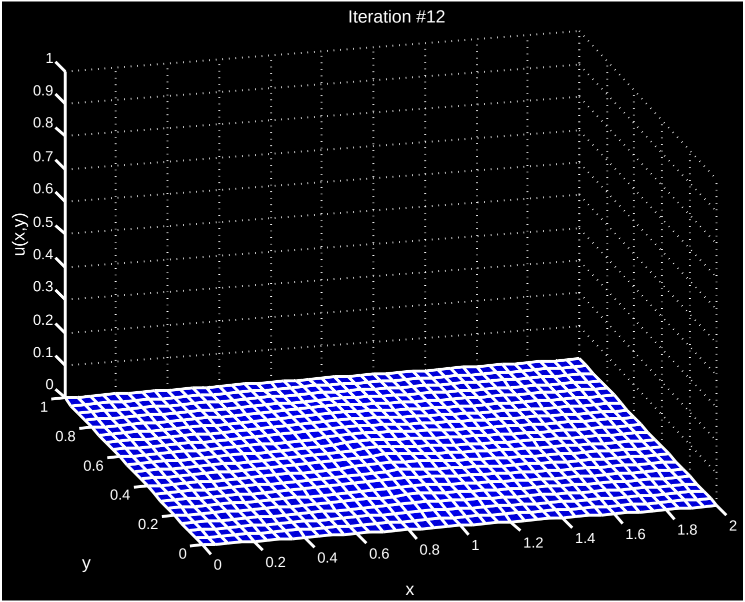
Iteration 10



Iteration 11



Iteration 12



Convergence Result with Fourier Analysis

For the model problem $\mathcal{L}u := (\eta - \Delta)u = 0$ on $\Omega = \mathbb{R}^2$, $\Omega_1 = (-\infty, L) \times \mathbb{R}$ and $\Omega_2 = (0, \infty) \times \mathbb{R}$,

$$\begin{array}{rll} (\eta-\Delta)u_1^n = 0 & \text{in } \Omega_1, \\ u_1^n = u_2^{n-1} & \text{on } x = L, \end{array} \qquad \begin{array}{rll} (\eta-\Delta)u_2^n = 0 & \text{in } \Omega_2, \\ u_2^n = u_1^n & \text{on } x = 0, \end{array}$$

we obtain after a Fourier transform in y

$$\begin{aligned} \hat{u}_j^n(x,k) &= \mathcal{F}(u_j^n) := \int_{-\infty}^{\infty} e^{-iky} u_j^n(x,y)dy, \quad k \in \mathbb{R}, \\ u_j^n(x,y) &= \mathcal{F}^{-1}(\hat{u}_j^n) := \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{iky} \hat{u}_j^n(x,k)dk, \end{aligned}$$

the Schwarz iteration in the Fourier domain (note how derivatives in y become multiplications by ik)

$$\begin{array}{rll} (\eta+k^2-\partial_{xx})\hat{u}_1^n = 0 & \text{in } \Omega_1, \\ \hat{u}_1^n = \hat{u}_2^{n-1} & \text{on } x = L, \end{array} \qquad \begin{array}{rll} (\eta+k^2-\partial_{xx})\hat{u}_2^n = 0 & \text{in } \Omega_2, \\ \hat{u}_2^n = \hat{u}_1^n & \text{on } x = 0. \end{array}$$

Convergence Result with Fourier Analysis

Now the ordinary differential equations

$$(\eta + k^2 - \partial_{xx})\hat{u}_j^n = 0$$

can easily be solved:

$$\hat{u}_j^n(x,k) = A_j^n e^{\sqrt{\eta+k^2}x} + B_j^n e^{-\sqrt{\eta+k^2}x}.$$

On domain Ω_1, solutions must stay bounded at $-\infty$, hence

$$\hat{u}_1^n(x,k) = A_1^n e^{\sqrt{\eta+k^2}x},$$

and on domain Ω_2, solutions must stay bounded at ∞,

$$\hat{u}_2^n(x,k) = B_2^n e^{-\sqrt{\eta+k^2}x}.$$

To determine the constants A_j^n and B_j^n, we use the transmission conditions

$$\hat{u}_1^n(L,k) = \hat{u}_2^{n-1}(L,k), \quad \hat{u}_2^n(0,k) = \hat{u}_1^n(0,k),$$

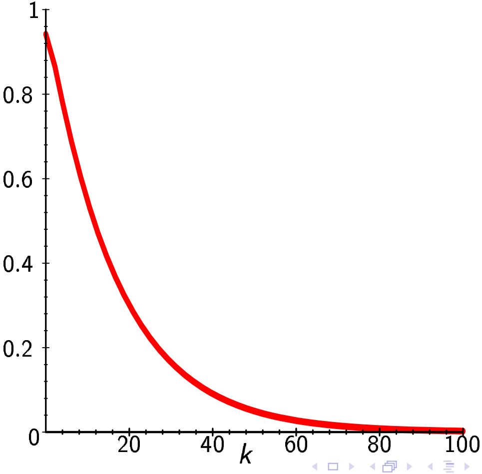
which give

$$A_1^n e^{\sqrt{\eta+k^2}L} = B_2^{n-1} e^{-\sqrt{\eta+k^2}L} = A_1^{n-1} e^{-\sqrt{\eta+k^2}L}.$$

Convergence Result with Fourier Analysis

After one iteration of the alternating Schwarz method, we obtain the convergence factor

$$\rho(\eta, k, L) := \frac{A_1^n}{A_1^{n-1}} = e^{-2\sqrt{\eta+k^2}L}.$$

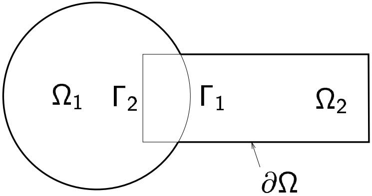


Classical Parallel Schwarz Method

P-L. Lions 1988:

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

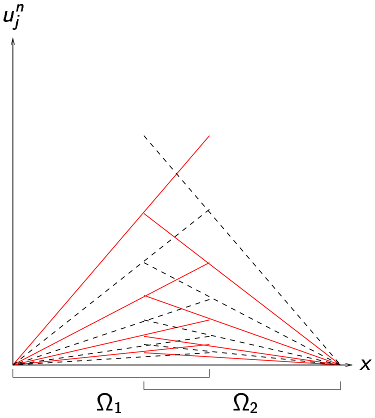
$\mathcal{L}u = f$ in Ω



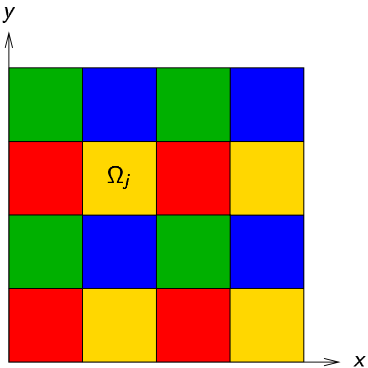
$$\begin{aligned} \mathcal{L}u_1^{n+1} &= f, \text{ in } \Omega_1 \\ u_1^{n+1} &= u_2^n, \text{ on } \Gamma_1 \end{aligned} \qquad \begin{aligned} \mathcal{L}u_2^{n+1} &= f, \text{ in } \Omega_2 \\ u_2^{n+1} &= u_1^n, \text{ on } \Gamma_2 \end{aligned}$$

Comparison of Alternating and Parallel Schwarz

two subdomains



many subdomains



- The alternating Schwarz methods needs a coloring strategy in order to be useful on a parallel computer.
- Sometimes the alternating Schwarz method produces a subset of the iterates of the parallel Schwarz method.

The Multiplicative Schwarz Method (MS)

The discretized PDE leads to the linear system

$$A\mathbf{u} = \mathbf{f}.$$

With the restriction operators

$$R_1 = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \end{bmatrix} R_2 = \begin{bmatrix} & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix}$$

and $A_j = R_j A R_j^T$ the multiplicative Schwarz method is

$$\begin{aligned} \mathbf{u}^{n+\frac{1}{2}} &= \mathbf{u}^n + R_1^T A_1^{-1} R_1(\mathbf{f} - A\mathbf{u}^n) \\ \mathbf{u}^{n+1} &= \mathbf{u}^{n+\frac{1}{2}} + R_2^T A_2^{-1} R_2(\mathbf{f} - A\mathbf{u}^{n+\frac{1}{2}}). \end{aligned}$$

Questions:

- Is MS a discretization of a continuous Schwarz method?
- How is the algebraic overlap related to the physical one?

Relation with Alternating Schwarz

If the R_j are non-overlapping, and we partition accordingly

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

we obtain from the first relation of MS, i.e.

$$\mathbf{u}^{n+\frac{1}{2}} = \mathbf{u}^n + R_1^T A_1^{-1} R_1(\mathbf{f} - A\mathbf{u}^n)$$

an interesting cancellation:

$$\begin{aligned} R_1(\mathbf{f} - A\mathbf{u}^n) &= \mathbf{f}_1 - A_1\mathbf{u}_1^n - A_{12}\mathbf{u}_2^n \\ A_1^{-1}R_1(\mathbf{f} - A\mathbf{u}^n) &= A_1^{-1}(\mathbf{f}_1 - A_{12}\mathbf{u}_2^n) - \mathbf{u}_1^n \\ \begin{pmatrix} \mathbf{u}_1^{n+\frac{1}{2}} \\ \mathbf{u}_2^{n+\frac{1}{2}} \end{pmatrix} &= \begin{pmatrix} \mathbf{u}_1^n \\ \mathbf{u}_2^n \end{pmatrix} + \begin{pmatrix} A_1^{-1}(\mathbf{f}_1 - A_{12}\mathbf{u}_2^n) - \mathbf{u}_1^n \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} A_1^{-1}(\mathbf{f}_1 - A_{12}\mathbf{u}_2^n) \\ \mathbf{u}_2^n \end{pmatrix} \end{aligned}$$

Relation with Alternating Schwarz

Similarly, from the second relation of MS, i.e.

$$\mathbf{u}^{n+1} = \mathbf{u}^{n+\frac{1}{2}} + R_2^T A_2^{-1} R_2(\mathbf{f} - A\mathbf{u}^{n+\frac{1}{2}})$$

we obtain

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

which can be rewritten in the equivalent form

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

and is therefore a discretization of the alternating Schwarz method from 1869,

$$\begin{aligned} \mathcal{L}u_1^{n+1} &= f, \text{ in } \Omega_1 \\ u_1^{n+1} &= u_2^n, \text{ on } \Gamma_1 \end{aligned} \qquad \begin{aligned} \mathcal{L}u_2^{n+1} &= f, \text{ in } \Omega_2 \\ u_2^{n+1} &= u_1^{n+1}, \text{ on } \Gamma_2 \end{aligned}$$

MS is also a block Gauss Seidel method

MS is also equivalent to a block Gauss Seidel method, since

$$A_1\mathbf{u}_1^{n+1} = \mathbf{f}_1 - A_{12}\mathbf{u}_2^n, \quad A_2\mathbf{u}_2^{n+1} = \mathbf{f}_2 - A_{21}\mathbf{u}_1^{n+1}$$

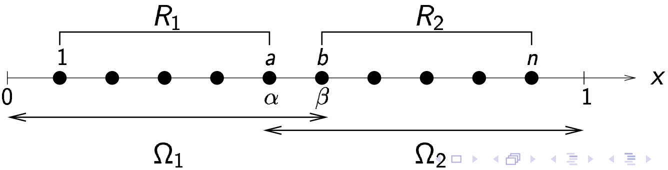
leads in matrix form to the iteration

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

So why the complicated R_j notation ?

- With R_j, one can also use overlapping blocks.
- With R_j, there is a global approximate solution $\mathbf{u}^n$.

Note that even the algebraically non-overlapping case above implies overlap at the PDE level:



The Additive Schwarz Method (AS)

M. Drjya and O. Widlund 1989:

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

Using the same notation as before, the preconditioned system is

$$(R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2) A\mathbf{u} = (R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2)\mathbf{f}$$

Writing this as a stationary iterative method yields

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

Question: Is AS equivalent to a discretization of Lions parallel Schwarz method ?

Algebraically non-overlapping case

If the R_j are non-overlapping, we obtain now

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

which can be rewritten in the equivalent form

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

This is a discretization of Lions' parallel Schwarz method from 1988,

$$\begin{array}{rcl} \mathcal{L}u_1^{n+1} & = & f, \text{ in } \Omega_1 \\ u_1^{n+1} & = & u_2^n, \text{ on } \Gamma_1 \end{array} \qquad \begin{array}{rcl} \mathcal{L}u_2^{n+1} & = & f, \text{ in } \Omega_2 \\ u_2^{n+1} & = & u_1^n, \text{ on } \Gamma_2 \end{array}$$

In the algebraically non-overlapping case, AS is also equivalent to a block Jacobi method,

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

What happens if the R_j overlap ?

If the R_j overlap, the cancellation is more complicated:

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

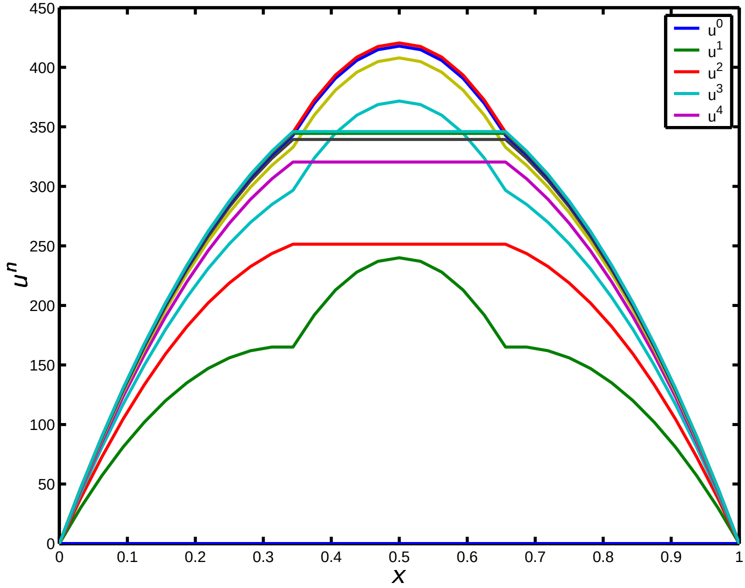
In the overlap, the current iterate is subtracted twice, and a new approximation from the left and right solve is added.

Remarks:

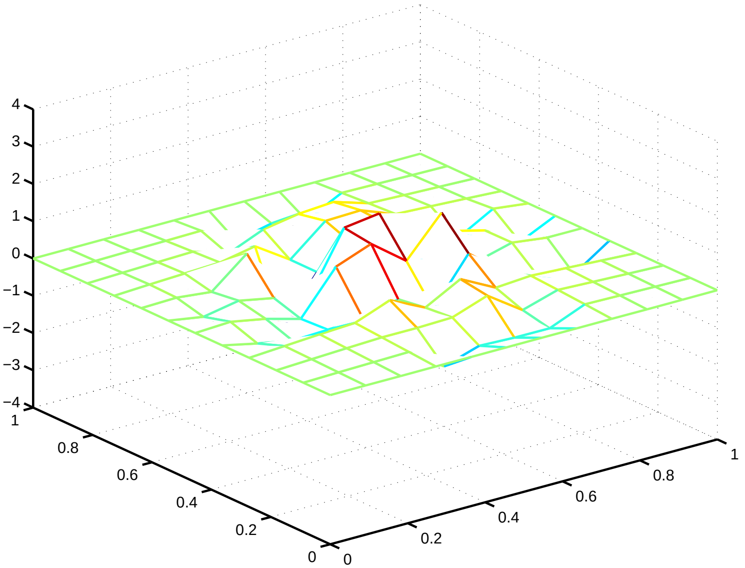
- One can show that the spectral radius of the AS iteration operator equals 1 for two subdomains.
- The method converges outside of the overlap for two subdomains.
- For more than two subdomains with cross points the method diverges everywhere.

AS is thus not equivalent to a discretization of Lions parallel Schwarz method for more than minimal physical overlap.

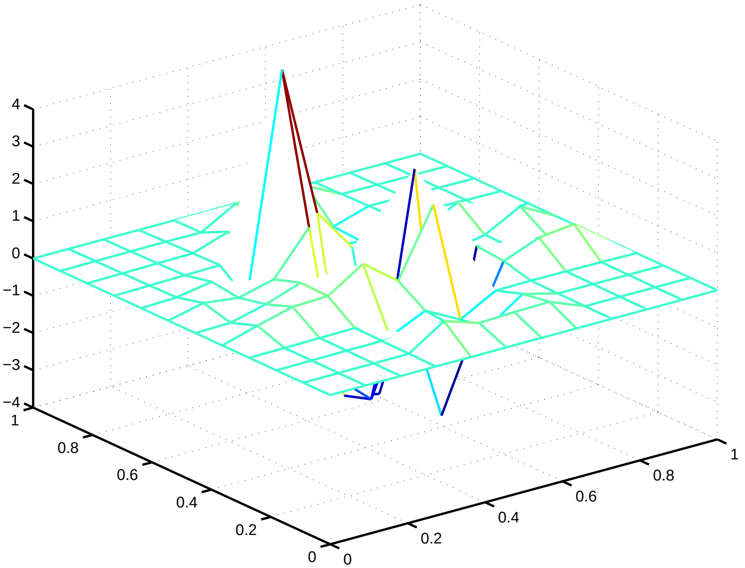
An Example for 2 Subdomains



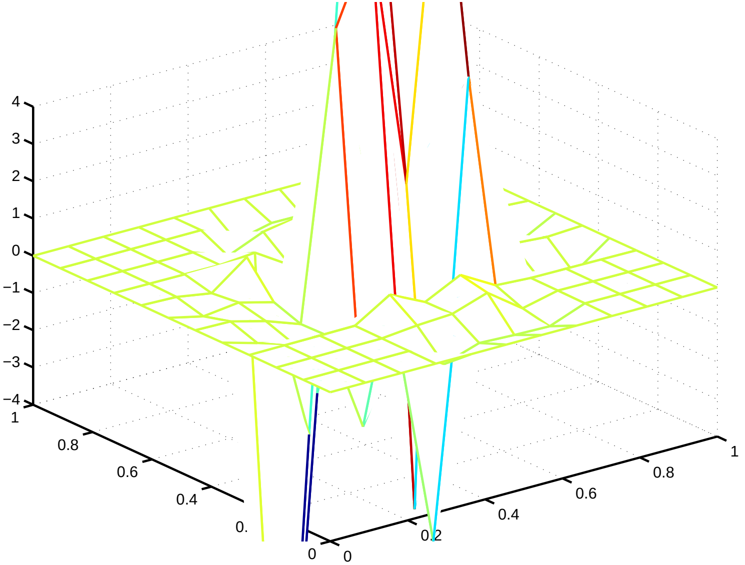
An Example for Many Subdomains



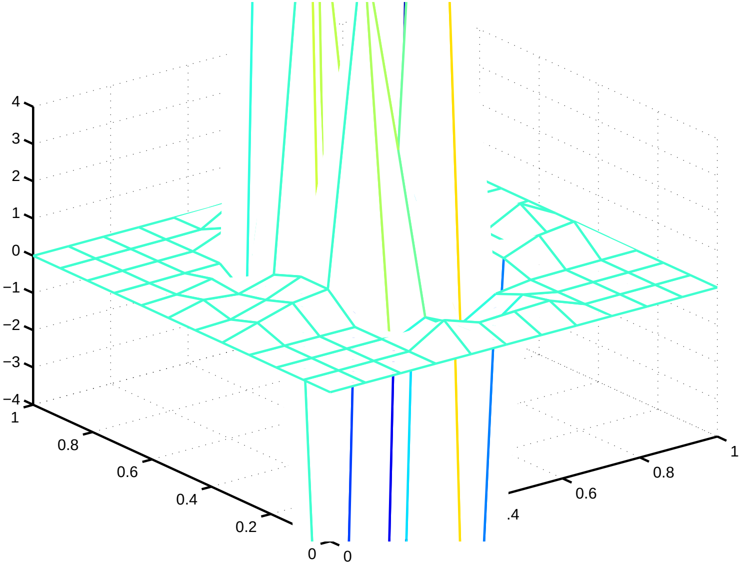
Iteration 2



Iteration 3



Iteration 4



Fundamental Convergence Result

M. Drjya and O. Widlund 1989:

Including a coarse grid correction in the additive Schwarz preconditioner,

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

with characteristic coarse mesh size H and overlap δ, one can show

Theorem

The condition number of the additive Schwarz preconditioned system satisfies

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

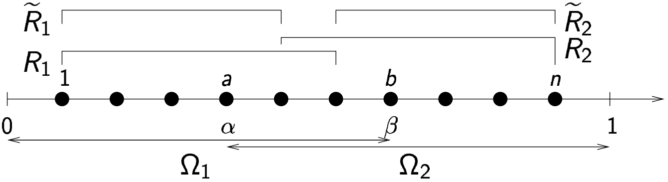
where the constant C is independent of δ and H.

Restricted Additive Schwarz (RAS)

X. Cai and M. Sarkis 1998:

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

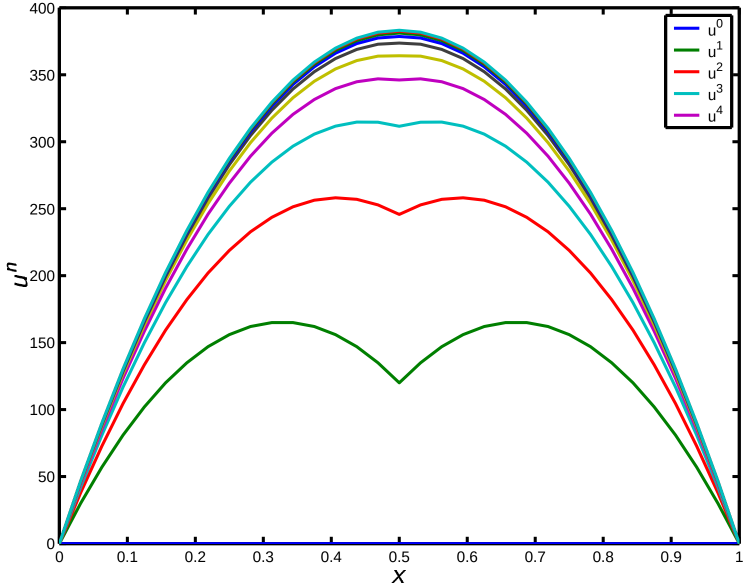
$$u^{n+1} = u^n + (\widetilde{R}_1^T A_1^{-1} R_1 + \widetilde{R}_2^T A_2^{-1} R_2)(f - Au^n)$$



Remarks:

- RAS is equivalent to a discretization of Lions parallel Schwarz method (Efstathiou, G. 2003)
- the preconditioner is **non symmetric**, even if A_j is symmetric

An Example for 2 Subdomains

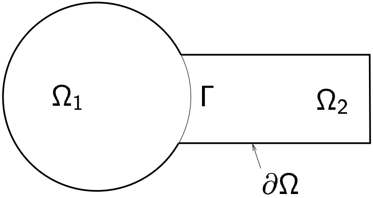


Problems of classical Schwarz: Overlap Necessary

P-L. Lions 1990:

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

$\mathcal{L}u = f$ in Ω

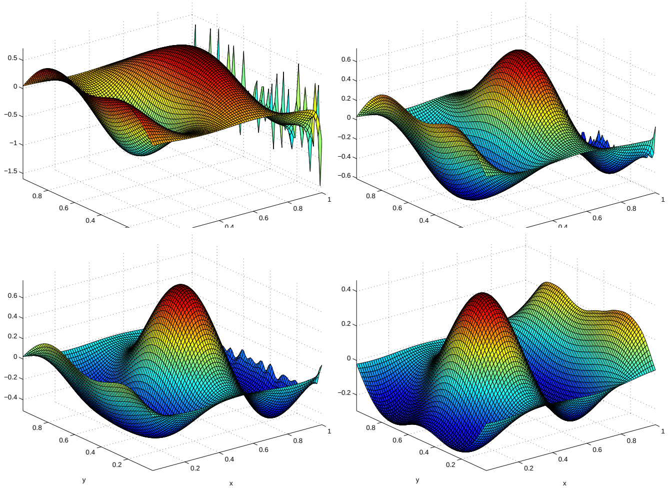


$$
\begin{array}{ll}
\mathcal{L}u_1^n = f \quad \text{in } \Omega_1 & \mathcal{L}u_2^n = f \quad \text{in } \Omega_2 \\
(\partial_{n_1}+p_1)u_1^n = (\partial_{n_1}+p_1)u_2^{n-1} \text{ on } \Gamma & (\partial_{n_2}+p_2)u_2^n = (\partial_{n_2}+p_2)u_1^n \text{ on } \Gamma
\end{array}
$$

P-L. Lions 1990:

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

Other Problem: Lack of Convergence



B. Després 1990:

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

Further Problem: Convergence Speed

T. Hagstrom, R. P. Tewarson and A. Jazcilevich 1988: Numerical experiments on a domain decomposition algorithm for nonlinear elliptic boundary value problems

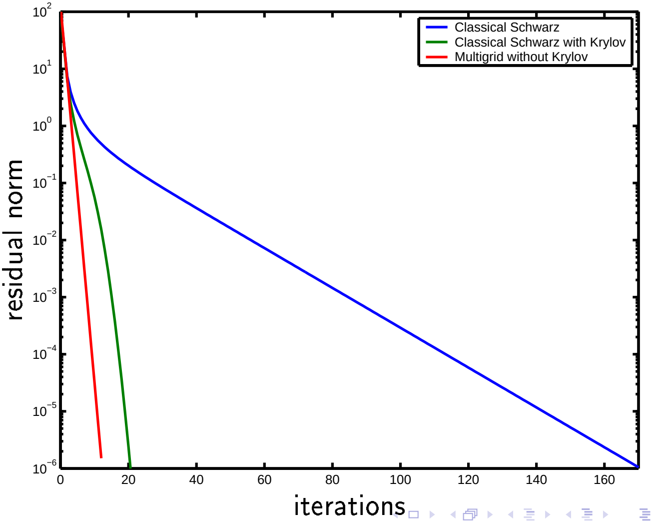
> *In general, [the coefficients in the Robin transmission conditions] may be operators in an appropriate space of function on the boundary. Indeed, we advocate the use of nonlocal conditions.*

W.-P. Tang 1992: Generalized Schwarz Splittings

> *In this paper, a new coupling between the overlap[ping] subregions is identified. If a successful coupling is chosen, a fast convergence of the alternating process can be achieved without a large overlap.*

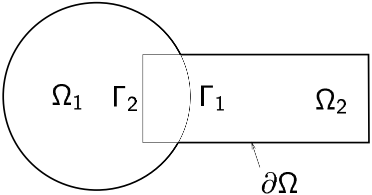
Comparison of Classical Schwarz with Multigrid

Comparison of MS with two subdomains as an iterative solver and a preconditioner for a Krylov method, with a standard multigrid solver:



Continuous Optimized Schwarz Methods

$$\mathcal{L}u = f \text{ in } \Omega$$



Instead of the classical alternating Schwarz method

$$\begin{aligned} \mathcal{L}u_1^n &= f, \text{ in } \Omega_1 \\ u_1^n &= u_2^{n-1}, \text{ on } \Gamma_1 \end{aligned} \qquad \begin{aligned} \mathcal{L}u_2^n &= f, \text{ in } \Omega_2 \\ u_2^n &= u_1^n, \text{ on } \Gamma_2 \end{aligned}$$

one uses transmission conditions adapted to the PDE,

$$\mathcal{B}_1 u_1^n = \mathcal{B}_1 u_2^{n-1}, \text{ on } \Gamma_1 \quad \mathcal{B}_2 u_2^n = \mathcal{B}_2 u_1^n, \text{ on } \Gamma_2$$

Remarks:

- optimal choice for $\mathcal{B}_j$ is $\partial_{n_j} + \mathcal{S}_j$ with $\mathcal{S}_j = DtN_j$
- good approximation is $\mathcal{B}_j = \partial_{n_j} + p_j + r_j\partial_\tau + q_j\partial_{\tau\tau}$
- method can converge even without physical overlap

How to Choose the Parameters

For the model problem $\mathcal{L}u := (\eta - \Delta)u = 0$ on $\Omega = \mathbb{R}^2$, $\Omega_1 = (-\infty, L) \times \mathbb{R}$ and $\Omega_2 = (0, \infty) \times \mathbb{R}$,

$$\begin{array}{rcll} (\eta - \Delta)u_1^n & = & 0 & \text{in } \Omega_1, \\ (\partial_x + \mathcal{S}_1)u_1^n & = & (\partial_x + \mathcal{S}_1)u_2^{n-1} & \text{on } x = L, \\ (\eta - \Delta)u_2^n & = & 0 & \text{in } \Omega_2, \\ (\partial_x - \mathcal{S}_2)u_2^n & = & (\partial_x - \mathcal{S}_2)u_1^n & \text{on } x = 0, \end{array}$$

we obtain as before after a Fourier transform in y the new Schwarz iteration in the Fourier domain

$$\begin{array}{rcll} (\eta + k^2 - \partial_{xx})\hat{u}_1^n & = & 0 & \text{in } \Omega_1, \\ (\partial_x + \sigma_1)\hat{u}_1^n & = & (\partial_x + \sigma_1)\hat{u}_2^{n-1} & \text{on } x = L, \\ (\eta + k^2 - \partial_{xx})\hat{u}_2^n & = & 0 & \text{in } \Omega_2, \\ (\partial_x - \sigma_2)\hat{u}_2^n & = & (\partial_x - \sigma_2)\hat{u}_1^n & \text{on } x = 0, \end{array}$$

where σ_j is the Fourier symbol of the operator $\mathcal{S}_j$.

Convergence Result with Fourier Analysis

As before, the solution of the ordinary differential equations are

$$\hat{u}_1^n(x,k) = A_1^n e^{\sqrt{\eta+k^2}x}, \quad \hat{u}_2^n(x,k) = B_2^n e^{-\sqrt{\eta+k^2}x}.$$

To determine the constants A_j^n and B_j^n, we use the transmission conditions

$$\begin{array}{rcl} (\partial_x + \sigma_1)\hat{u}_1^n(L,k) & = & (\partial_x + \sigma_1)\hat{u}_2^{n-1}(L,k), \\ (\partial_x - \sigma_2)\hat{u}_2^n(0,k) & = & (\partial_x - \sigma_2)\hat{u}_1^n(0,k), \end{array}$$

which give

$$A_1^n(\sqrt{\eta+k^2}+\sigma_1)e^{\sqrt{\eta+k^2}L} = B_2^{n-1}(-\sqrt{\eta+k^2}+\sigma_1)e^{-\sqrt{\eta+k^2}L}$$

and

$$B_2^{n-1}(-\sqrt{\eta+k^2}-\sigma_2) = A_1^{n-1}(\sqrt{\eta+k^2}-\sigma_2).$$

Convergence Result with Fourier Analysis

After one iteration of the optimized Schwarz method, we obtain the convergence factor

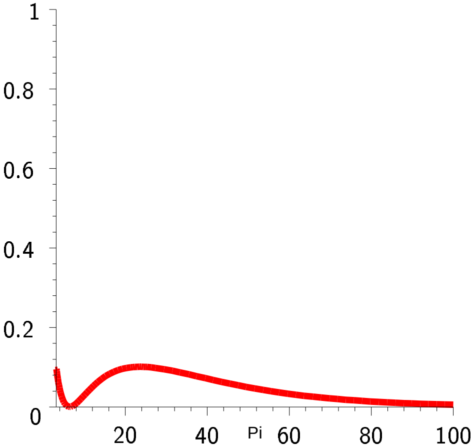
$$\rho(\eta, k, L, \sigma_1, \sigma_2) := \frac{\sqrt{\eta + k^2} - \sigma_1}{\sqrt{\eta + k^2} + \sigma_1} \frac{\sqrt{\eta + k^2} - \sigma_2}{\sqrt{\eta + k^2} + \sigma_2} e^{-2\sqrt{\eta+k^2}L}.$$

- If the symbols are $\sigma_j := \sqrt{\eta + k^2}$, then the convergence factor vanishes identically: convergence after one double step, even without overlap !
- This result can be generalized to convergence after I steps for I subdomains, provided the subdomain connections have no loops.
- This choice is optimal, but not convenient in practice, since the operator associated with the symbol $\sqrt{\eta + k^2}$ is non-local (it represents the DtN operator for the equation)
- One is therefore interested in local approximations.

Zeroth Order Approximation

We approximate the symbols σ_j by a constant, $\sigma_j := p$, $p \in \mathbb{R}$. The transmission conditions are therefore

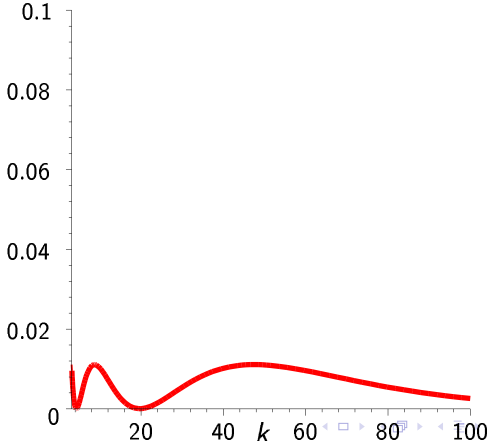
$$\begin{aligned}(\partial_x + p)u_1^n(L,y) &= (\partial_x + p)u_2^{n-1}(L,y),\\ (\partial_x - p)u_2^n(0,y) &= (\partial_x - p)u_1^n(0,y).\end{aligned}$$



Second Order Approximation

We approximate the symbols σ_j by a second degree polynomial in ik, $\sigma_j := p - qk^2$, $p, q \in \mathbb{R}$. The transmission conditions are therefore

$$
\begin{aligned}
(\partial_x + p + q\partial_{yy})u_1^n(L,y) &= (\partial_x + p + q\partial_{yy})u_2^{n-1}(L,y),\\
(\partial_x - p - q\partial_{yy})u_2^n(0,y) &= (\partial_x - p - q\partial_{yy})u_1^n(0,y).
\end{aligned}
$$



How to Choose the Parameters in General

Contraction factor using Fourier analysis:

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

- z related to the Fourier parameters
- s polynomial with coefficients to be optimized.
- f_{PDE} symbol of the DtN of the PDE to be solved.

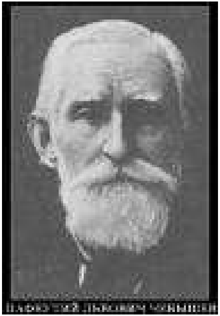
For a fast algorithm, we need to minimize ρ, i.e.

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

- $\mathbb{P}_n$ set of complex polynomials of degree $\leq n$
- K is a bounded or unbounded set in the complex plane

Best Approximation Problems

Chebyshev (1854): Théorie des mécanismes connus sous le nom de parallélogrammes.



Soit $f(x)$ une fonction donnée, U un polynome du degré n avec des coefficients arbitraires. Si l'on choisit ces coefficients de manière à ce que la différence $f(x) - U$, depuis $x = a - h$, jusque à $x = a + h$, reste dans les limites les plus rapprochées de 0, la différence $f(x) - U$ jouira, comme on le sait, de cette propriété:
Parmi les valeurs les plus grandes et les plus petites de la différence $f(x) - U$ entre les limites $x = a - h$, $x = a + h$, on trouve au moins $n+2$ fois la même valeur numérique.

De la Vallée Poussin (1910): Existence, Uniqueness and Equioscillation.

$$\min_{p\in\mathbb{P}_n}\max_{x\in K}|f(x)-p(x)|$$

More General Chebyshev Approximation

Meinardus and Schwedt (1961): Nicht-lineare Approximationen

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

Overview of known results:

1. Linear and finite dimensional: compactness
2. Linear and finite dimensional:
   real: Haar criterion **Haar (1918)**
   complex: **Kolmogoroff (1948)**
3. Real functions of a real variable:
   equioscillation, **Rice (1960)**
   nonlinear, complex rational approximation:
   equioscillation, **Meinardus and Schwedt (1961)**

Homographic Best Approximation

$$\min_{s\in\mathbb{P}_n}\sup_{z\in K}\left|\frac{s(z)-f(z)}{s(z)+f(z)}e^{-Lf(z)}\right|$$

Theorem (Bennequin, G, Halpern 2005)

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

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

Theorem (Local Minima)

If s_n is a strict local minimum, then it is the global minimum.

Case $L > 0$

Without assuming that K is compact, one can show (Bennequin, G, Halpern 2006):

Theorem (Existence)

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

$$\Re f(z) \longrightarrow +\infty \text{ as } z \longrightarrow \infty \text{ in } K.$$

Then for L small enough, there exists a solution.

Theorem (Equioscillation)

Under the same assumptions, if s_n^* is a solution for $L > 0$, then there exist at least $n+2$ points $z_1, \cdots, z_{n+2}$ in K such that

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

Uniqueness, Local Minima and Symmetry

Theorem (Uniqueness)

With the same assumptions, and if K is compact, and L satisfies

$$\delta_n(L)e^{L \sup_{z\in K} \Re f(z)} < 1,$$

where $\delta_n(L)$ is the minimum, then the solution is unique.

Theorem (Local Minima)

If K is compact, and L is small, then if s_n^* is a strict local minimum, then it is the global minimum.

Theorem (Symmetry $\Longrightarrow$ real coefficients)

If K is compact and symmetric with respect to the x-axis, and $f(\bar{z}) = \overline{f(z)}$ in K, then for L small, s_n^* has real coefficients.

Optimized Parameters for a Model Problem

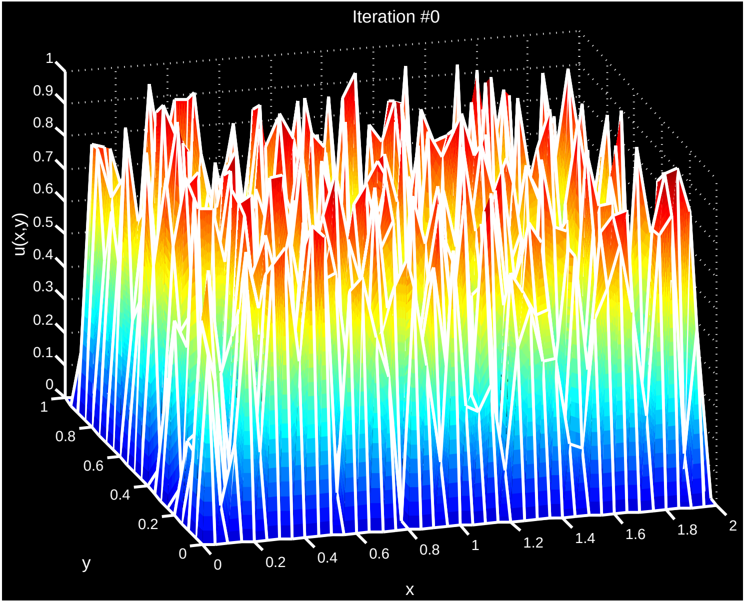
For the self adjoint coercive problem

$$\mathcal{L}u = (\eta - \Delta)u = f$$

the asymptotically optimal parameters are (G 2006)

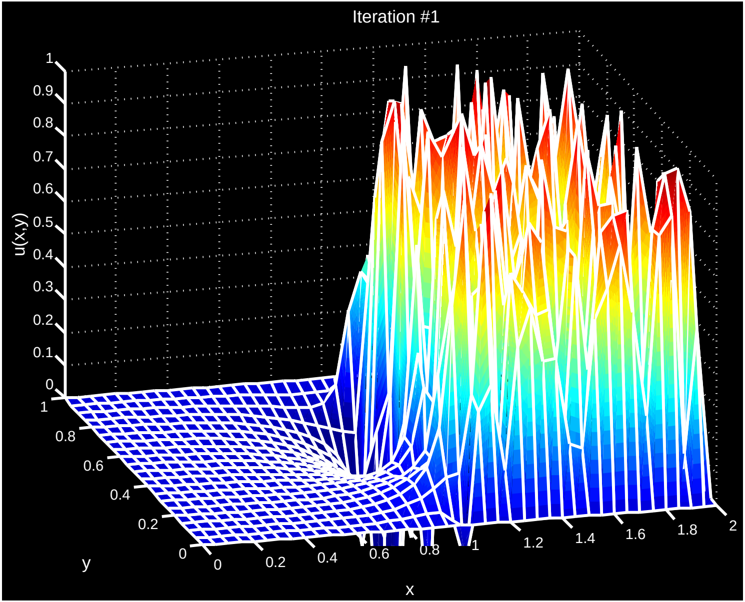
	p	q
OO0	$\frac{\sqrt{\pi}(k_{\min}^2+\eta)^{1/4}}{h^{1/2}}$	0
OO0(Ch)	$\frac{(k_{\min}^2+\eta)^{1/3}}{2^{1/3}(Ch)^{1/3}}$	0
OO2	$\frac{\pi^{1/4}(k_{\min}^2+\eta)^{3/8}}{2^{1/2}h^{1/4}}$	$\frac{h^{3/4}}{2^{1/2}\pi^{3/4}(k_{\min}^2+\eta)^{1/8}}$
OO2(Ch)	$\frac{(k_{\min}^2+\eta)^{2/5}}{2^{3/5}(Ch)^{1/5}}$	$\frac{(Ch)^{3/5}}{2^{1/5}(k_{\min}^2+\eta)^{1/5}}$
TO0	$\sqrt{\eta}$	0
TO2	$\sqrt{\eta}$	$\frac{1}{2\sqrt{\eta}}$

An Example



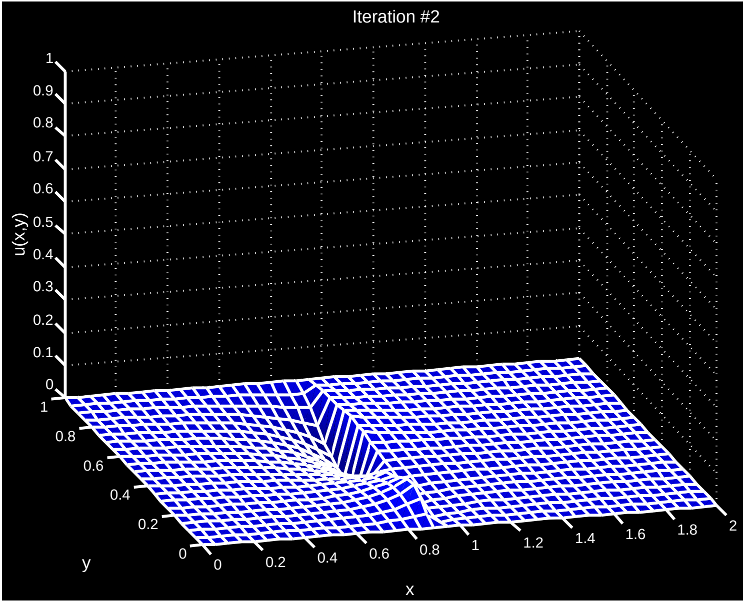
Iteration #0
u(x,y)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
y
0
0.2
0.4
0.6
0.8
1
x
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

Iteration 1

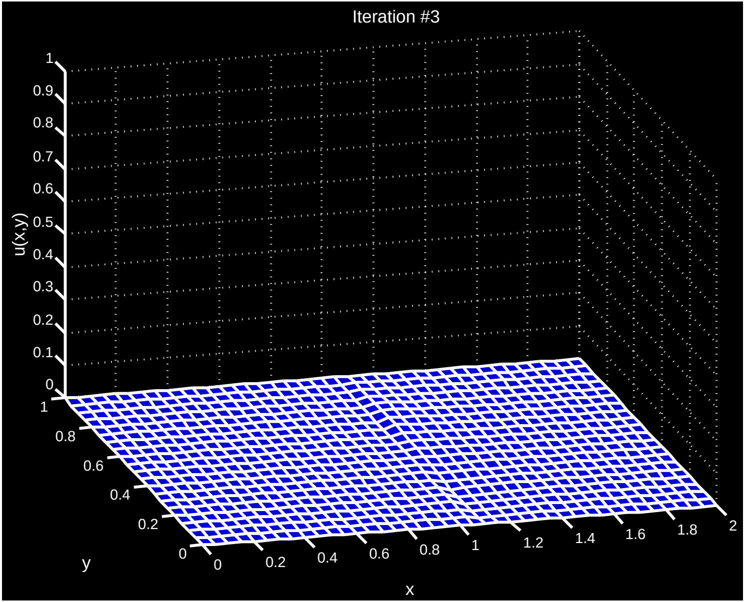


Iteration #1
u(x,y)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
0.8
0.6
0.4
0.2
0
y
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
x

Iteration 2

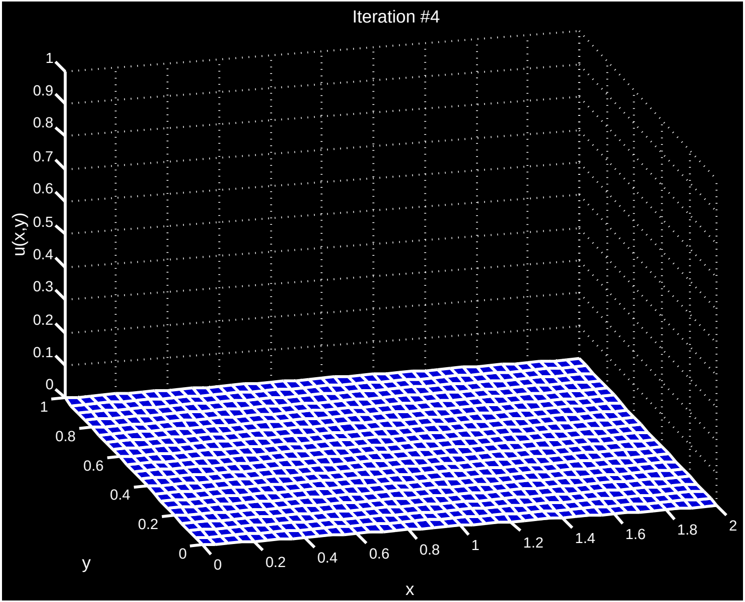

Iteration #2
u(x,y)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
y
1
0.8
0.6
0.4
0.2
0
x
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

Iteration 3



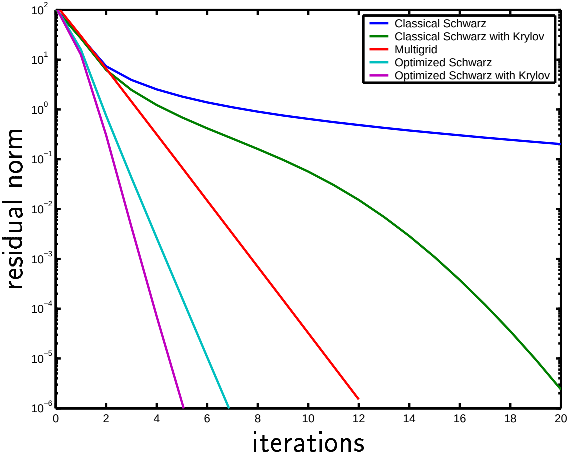
Iteration #3
u(x,y)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
y
1
0.8
0.6
0.4
0.2
0
x
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

Iteration 4


Iteration #4
u(x,y)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
y
1
0.8
0.6
0.4
0.2
0
x
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

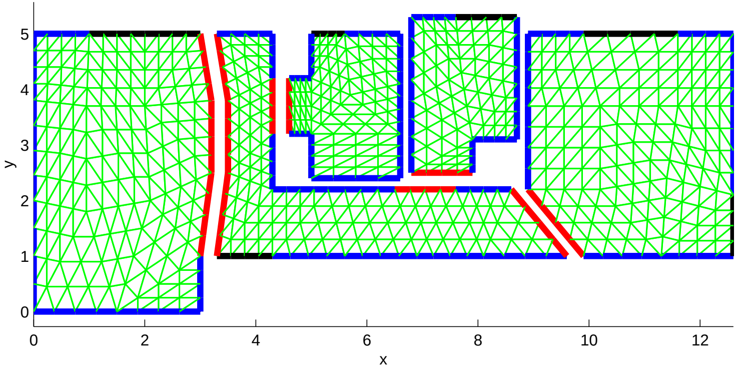
Comparison of Optimized Schwarz with Multigrid

Comparison of MS as an iterative solver, as a preconditioner, multigrid, and an optimized Schwarz methods used iteratively and as a preconditioner:



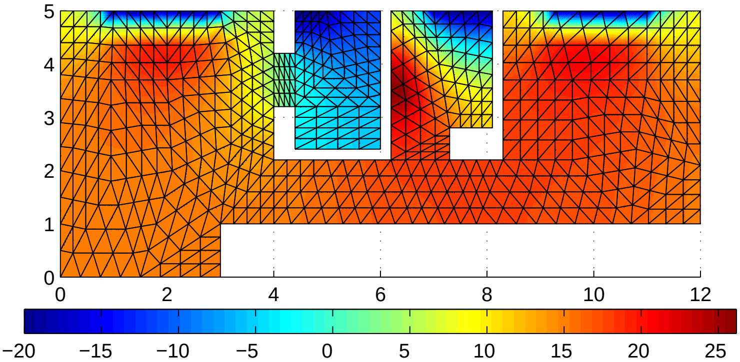
Optimized Schwarz Application

Heating problem in our former apartment in Montreal:



Optimized Schwarz Application

Result of a non-overlapping optimized Schwarz method with Robin transmission conditions:



With the optimal parameter p^* from the two subdomain theory, the convergence factor ratio is in the iterative case $32/25 = 1.28 \approx 2^{1/3} = 1.26$, as predicted by the two subdomain theory.

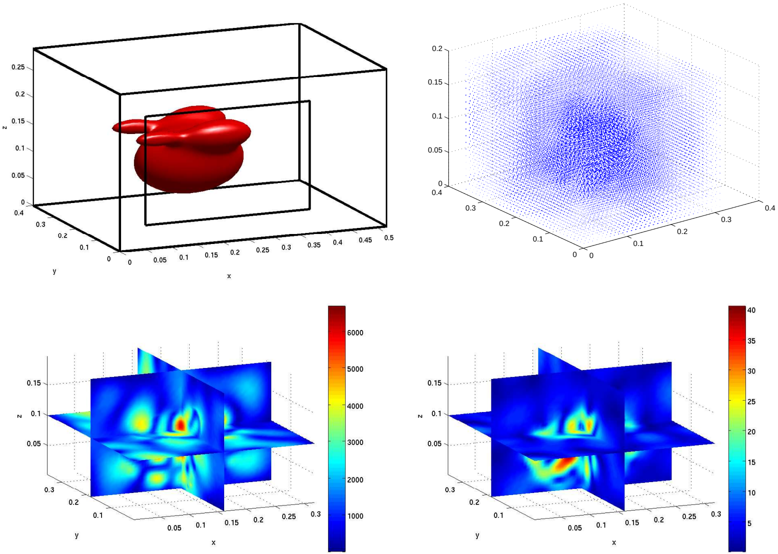
Results for other PDEs

Such formulas are also available for advection-reaction-diffusion (Japhet 1998, Dubois and G. 2007), Helmholtz (Chevalier 1998, G., Magoules, Nataf 2001), Heat (G. Halpern 2003), Wave (G., Halpern, Nataf 2003), Cauchy-Riemann and Maxwell's equations (Dolean, G., Gerardo-Giorda 2007)

	with overlap, $L = h$		without overlap, $L = 0$	
Case	ρ	parameters	ρ	parameters
1	$1-\sqrt{k_+-\omega^2}\,h$	none	1	none
2	$1-2C_\omega^{\frac{1}{6}}h^{\frac{1}{3}}$	$p=\frac{C_\omega^{\frac{1}{3}}}{2\cdot h^{\frac{1}{3}}}$	$1-\frac{\sqrt{2}C_\omega^{\frac{1}{4}}}{\sqrt{C}}\sqrt{h}$	$p=\frac{\sqrt{C}C_\omega^{\frac{1}{4}}}{\sqrt{2}\sqrt{h}}$
3	$1-2(k_+^2-\omega^2)^{\frac{1}{6}}h^{\frac{1}{3}}$	$p=\frac{(k_+^2-\omega^2)^{\frac{1}{3}}}{2\cdot h^{\frac{1}{3}}}$	$1-\frac{\sqrt{2}(k_+^2-\omega^2)^{\frac{1}{4}}}{\sqrt{C}}\sqrt{h}$	$p=\frac{\sqrt{C}(k_+^2-\omega^2)^{\frac{1}{4}}}{\sqrt{2}\sqrt{h}}$
4	$1-2^{\frac{2}{5}}C_\omega^{\frac{1}{10}}h^{\frac{1}{5}}$	$\begin{cases} p_1=\frac{C_\omega^{\frac{2}{5}}}{2^{\frac{7}{5}}\cdot h^{\frac{1}{5}}}, \\ p_2=\frac{C_\omega^{\frac{1}{5}}}{2^{\frac{6}{5}}\cdot h^{\frac{3}{5}}} \end{cases}$	$1-\frac{C_\omega^{\frac{1}{8}}}{C^{\frac{1}{4}}}h^{\frac{1}{4}}$	$\begin{cases} p_1=\frac{C_\omega^{\frac{3}{8}}\cdot C^{\frac{1}{4}}}{2\cdot h^{\frac{1}{4}}}, \\ p_2=\frac{C_\omega^{\frac{1}{8}}\cdot C^{\frac{3}{4}}}{h^{\frac{3}{4}}} \end{cases}$
5	$1-2^{\frac{2}{5}}(k_+^2-\omega^2)^{\frac{1}{10}}h^{\frac{1}{5}}$	$\begin{cases} p_1=\frac{(k_+^2-\omega^2)^{\frac{2}{5}}}{2^{\frac{7}{5}}\cdot h^{\frac{1}{5}}}, \\ p_2=\frac{(k_+^2-\omega^2)^{\frac{1}{5}}}{2^{\frac{6}{5}}\cdot h^{\frac{3}{5}}} \end{cases}$	$1-\frac{(k_+^2-\omega^2)^{\frac{1}{8}}}{C^{\frac{1}{4}}}h^{\frac{1}{4}}$	$\begin{cases} p_1=\frac{(k_+^2-\omega^2)^{\frac{3}{8}}\cdot C^{\frac{1}{4}}}{2\cdot h^{\frac{1}{4}}}, \\ p_2=\frac{(k_+^2-\omega^2)^{\frac{1}{8}}\cdot C^{\frac{3}{4}}}{h^{\frac{3}{4}}} \end{cases}$

Large Scale Optimized Schwarz Computations

- Heating a chicken in our Whirlpool Talent Combi 4 microwave oven (with V. Dolean):



Discrete Optimized Schwarz Methods

How does one have to change the RAS

$$M_{RAS}^{-1} = (\tilde{R}_1^T A_1^{-1} R_1 + \tilde{R}_2^T A_2^{-1} R_2)$$

and the MS preconditioner

$$M_{MS}^{-1} = \left[I - \prod_{j=1}^{J} \left(I - R_j^T A_j^{-1} R_j A \right) \right] A^{-1},$$

to obtain an optimized method ?

Simply replace A_j by a slightly modified $\tilde{A}_j$!

(St-Cyr, G and Thomas, 2007)

An Example

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

Finite volume discretization leads to

$$A\mathbf{u} = \mathbf{f}$$

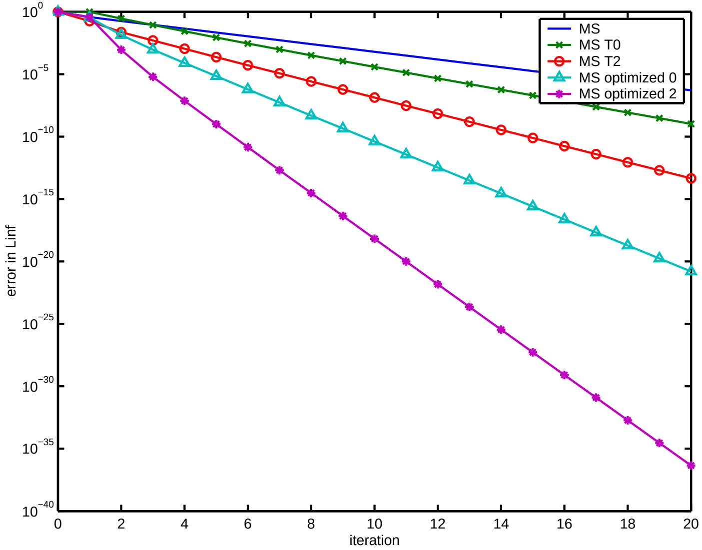
$$A = \frac{1}{h^2}\begin{bmatrix} T_\eta & -I & \\ -I & T_\eta & \ddots \\ & \ddots & \ddots \end{bmatrix}, \quad T_\eta = \begin{bmatrix} \eta h^2 + 4 & -1 & \\ -1 & \eta h^2 + 4 & \ddots \\ & \ddots & \ddots \end{bmatrix}$$

The classical subdomain matrices are $A_j = R_j A R_j^T$.
The optimized $\tilde{A}_j$ are obtained from A_j by simply replacing the interface diagonal block T_η by

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

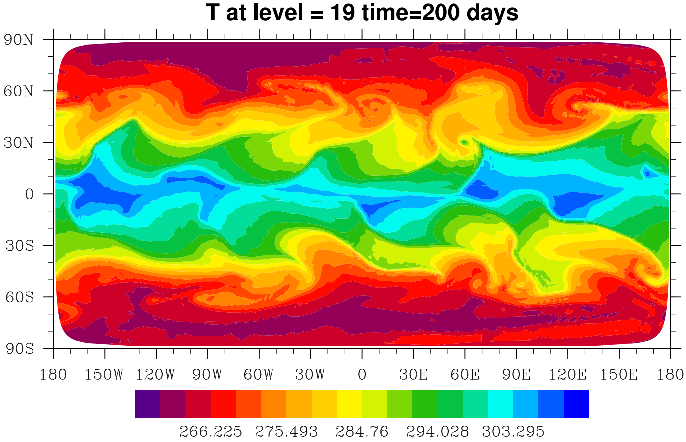
where p and q are solutions of the associated min-max problem.

Result for the Example



Atmospheric General Circulation Model (GCM)

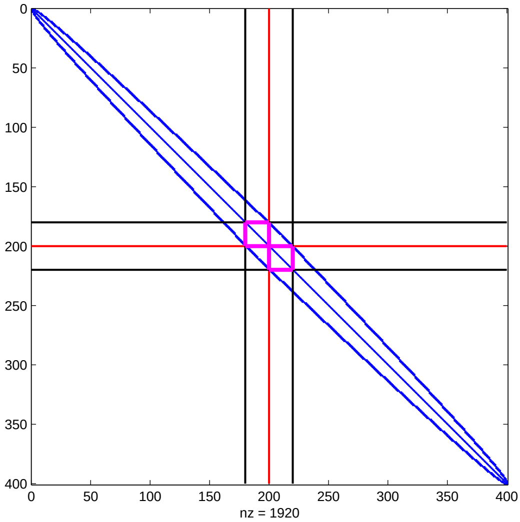
Held Suarez test, temperature field, at the surface of the planet after 200 days of simulation (with A. St-Cyr)



(6144 spectral subdomains on the IBM Blue Gene)

What if one does not know the PDE ?

After a bandwidth minimization reordering, we partition the matrix algebraically and identify interface blocks:



Algebraic Optimization

Using the notation

$$A = \left[\begin{array}{cc|c|c|cc} \ddots & \ddots & \ddots & & & \\ \hline 0 & B_2 & E_2 & X & 0 & \\ \hline & 0 & X & E_1 & B_1 & 0 \\ \hline & & & \ddots & \ddots & \ddots \end{array}\right]$$

the iteration operator squared has the structure

$$T^2 = \left[\begin{array}{c|c|c|c} & T_1 & & \\ \hline & & T_2 & \end{array}\right]$$

Analysis

If the small matrices E_j are replaced by D_j in the new subdomain matrices $\tilde{A}_j$, an algebraic analysis reveals

$$T_1 = Y(D_1^{-1} - Y_n)^{-1}(Z_2 + D_1^{-1}B_1Z_3)(D_2^{-1} - Z_3^{-1})^{-1}[D_2^{-1}B_2e_{n-2}^T + e_{n-1}^T]$$

$$T_2 = Z(D_2^{-1} - Z_1)^{-1}(Y_{n-1} + D_2^{-1}B_2Y_{n-2})(D_1^{-1} - Y_n^{-1})^{-1}[D_1^{-1}B_1e_2^T + e_1^T]$$

where

$$A_1^{-1} = \begin{bmatrix} X & \dots & X & Y_1 \\ \vdots & & \vdots & \vdots \\ & & & Y_{n-2} \\ \vdots & & \vdots & Y_{n-1} \\ X & \dots & X & Y_n \end{bmatrix}, \quad A_2^{-2} = \begin{bmatrix} Z_1 & X & \dots & X \\ Z_2 & \vdots & & \vdots \\ Z_3 & & & \\ \vdots & \vdots & & \vdots \\ Z_n & X & \dots & X \end{bmatrix}$$

Hence we can make T^2 small in norm, if we can make

$$D_1Z_2 + B_1Z_3 \qquad D_2Y_{n-1} + B_2Y_{n-2}$$

small in norm.

Algebraic Optimization Problems

Choosing the Frobenius norm, we need for a given sparsity pattern $\mathcal{P}_1$ and $\mathcal{P}_2$ to solve the minimization problems

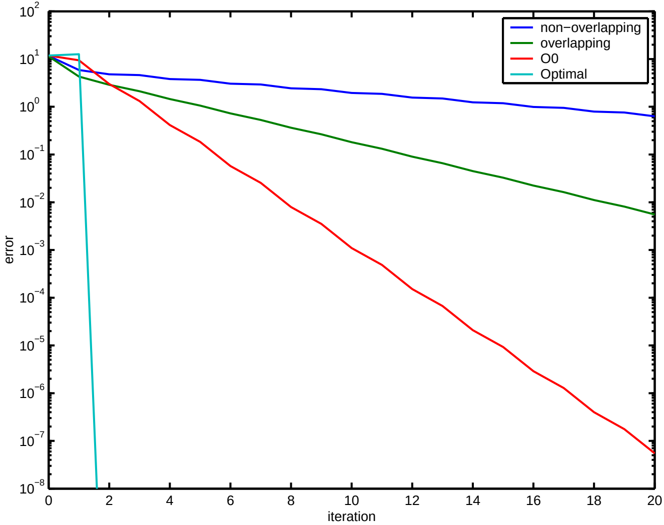
$$\min_{D_1 \in \mathcal{P}_1} ||D_1 Z_2 + B_1 Z_3||_F, \quad \min_{D_2 \in \mathcal{P}_2} ||D_2 Y_{n-1} + B_2 Y_{n-2}||_F,$$

which decouples into smaller least squares problems, for example

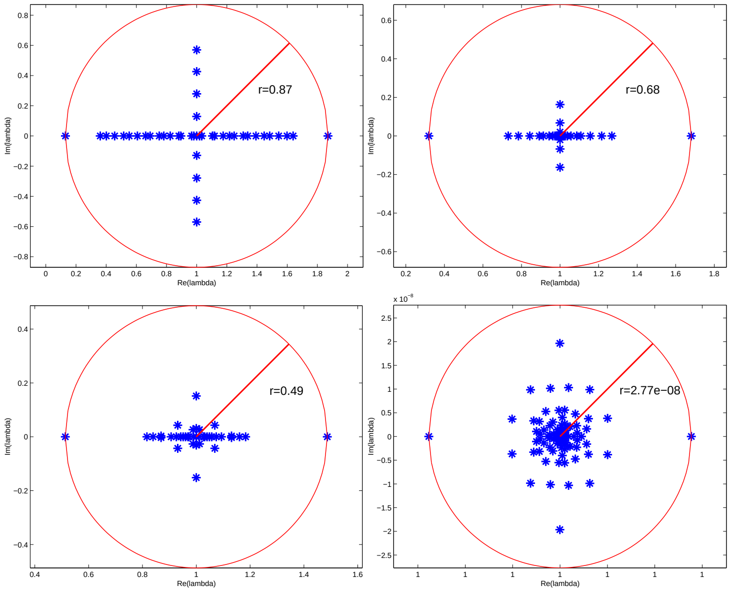
- If the sparsity pattern is diagonal, we get scalar least squares problems.
- If the sparsity pattern is tridiagonal, we get least squares problems with 3 unknowns each.
- If we allow complete fill-in in D_j, we can obtain $T_j = 0$

Performance with Modified Matrix Entries

We obtain for the unknown example (which was an advection diffusion equation with rotating velocity) the convergence curves



Clustering of the Spectra of T



Conclusions

- Discrete Schwarz methods are in most cases discretizations of continuous Schwarz methods (exception: AS with overlap!)
- Optimized Schwarz Methods use transmission conditions adapted to the underlying PDE, which can greatly improve their convergence rate
- Replacing classical subdomain matrices A_j by optimized ones, leads to optimized MS, RAS and AS (on an augmented system)

Important current problems: (2008)

- Algebraically optimized $\tilde{A}_j$
- Coarse grid corrections for optimized Schwarz
- General convergence proof for overlapping optimized Schwarz