Optimized Schwarz Waveform Relaxation: Roots, Blossoms and Fruits

Laurence Halpern[1]

LAGA, Université Paris XIII, 99 Avenue J-B Clément, 93430 Villetaneuse, France,
halpern@math.univ-paris13.fr

1 Introduction: Parallel Processing of Evolution Problems

There are several ways to solve in parallel the evolution problem

$$P(\partial_t, \partial_1, \cdots, \partial_d)u = f.$$

- *Explicit time discretization* is naturally parallel.
- *Implicit time discretization + spatial domain decomposition*. For the heat equation for instance, with an implicit Euler scheme in time, this amounts to solving at each step the linear problem

  $$\frac{u}{\Delta t} - \Delta u = f.$$

  This gives a very well conditioned problem, and multigrid or domain decomposition algorithms can be used without coarse grid preconditioner (see the reference book of [22] and references therein). Thereafter, improved algorithms were designed (see presentation by F. Nataf in the same minisymposium). However a uniform time-step is needed.

Both procedures imply an exchange of information between processors at every time-step, which can be very penalizing when using a large number of processors. Quoting [1] "*A major obstacle to achieving significant speed-up on parallel machines is the overhead associated with synchronizing the concurrent processes*". One way to overcome this delay problem was invented in the seventies with the concept of **asyncronous algorithms**, [4]. An excellent review can be found in [1]. In that context, amounts of information are sent without waiting for the request. However convergence is weakened, if not destroyed.

- *Time or Space-time multigrid*. Quoting [14], in the previous approach, "*the potential for parallelism is limited to the parallelism of the elliptic solver, since the time dimension is treated strictly sequentially. ... Thus it seems natural to ask whether ... a parallelization strategy for the time-dependent part of the problem*

[can] be found ". The problem has been studied mainly for parabolic operators, leading to the parabolic multigrid method of [11], or the space-time multigrid in [14], or the parareal algorithm of [16].

- *Waveform relaxation*, and more particularly the Schwarz waveform relaxation algorithms, were designed independently in [7, 10]. The main feature of these algorithms is their flexibility. It permits to choose the space and time meshes independently in the subdomains, leading to local space-time refinement with time windows. Different numerical schemes can be used in the subdomains, or even different models can be coupled, and the method adjusts to underlying computing hardware.

2 Roots: Waveform Relaxation for ODEs

The ancestor is Emile Picard, which in the *Journal de mathématiques* introduced the "Méthode des approximations successives" to prove the existence of solutions to ordinary differential equations. The algorithm was reinvented in practical implementations in [15]. We describe below the Picard algorithm for functions y_j in different spaces $\mathbb{R}^{N_j}$:

$$\begin{array}{l|l}
\dfrac{dy_1}{dt} = f_1(t,y_1,y_2,\cdots,y_p), & \dfrac{dy_1^{(k+1)}}{dt} = f_1(t,y_1^{(k)},y_2^{(k)},\cdots,y_p^{(k)}), \\
 & \vdots \\
\dfrac{dy_j}{dt} = f_j(t,y_1,y_2,\cdots,y_p), & \dfrac{dy_j^{(k+1)}}{dt} = f_j(t,y_1^{(k)},y_2^{(k)},\cdots,y_p^{(k)}), \\
 & \vdots \\
\dfrac{dy_p}{dt} = f_p(t,y_1,y_2,\cdots,y_p), & \dfrac{dy_p^{(k+1)}}{dt} = f_p(t,y_1^{(k)},y_2^{(k)},\cdots,y_p^{(k)}), \\
 & \vdots \\
\text{SYSTEM OF ODE'S} & \text{PICARD ITERATES}
\end{array}$$

Note that it is naturally parallel. The error at step $k+1$ on the time interval $[t_0,T]$ is given by:

$$\|y^{(k+1)} - y\|_\infty \leq \frac{L^k (T-t_0)^k}{k!} \|y^{(0)} - y\|_\infty. \tag{1}$$

At first sight, this tends to zero rapidly with k. However, it can be seen using the Stirling formula $n! \approx \sqrt{2\pi n}(\frac{n}{e})^n$ that $\frac{eL(T-t_0)}{k}$ has to be smaller than 1 before the error will start to decrease, which implies many iterations in case of large time intervals, or large Lipschitz constants.

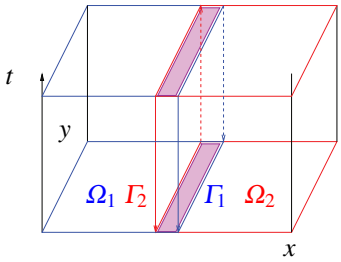
Waveform relaxation algorithms are extensions both of the Picard's "approximations successives" and relaxation methods for algebraic systems. The parallel formulation, of Jacobi type, is obtained by replacing in the "j-" line, $y_j^{(k)}$ by $y_j^{(k+1)}$ in the function f. For the analysis, see [20], [22], for a review [3]. The principal results are:

1. Linear convergence on unbounded time intervals for linear systems with dissipation.
2. Superlinear convergence for finite time.
3. When the ordinary differential equation stems from a discretization in space of a partial differential equation, the convergence rate depends on the discretization parameters, and deteriorates as one refines the mesh.

Later on continuous versions have been developed, like the Schwarz waveform relaxation for partial differential equations.

3 Blossoms: Classical Schwarz Waveform Relaxation for Parabolic Equations

To solve $\mathcal{L}u = f$ in $\Omega \times (0,T)$, with initial condition u_0, with $\mathcal{L}$ the heat operator, it was proposed in [7] to introduce the overlapping domain decomposition algorithm in space-time:



$$\begin{cases} \mathcal{L}u_1^{k+1} = f & \text{in } \Omega_1 \times (0,T), \\ u_1^{k+1}(\cdot,0) = u_0 & \text{in } \Omega_1, \\ u_1^{k+1} = u_2^k & \text{on } \Gamma_1 \times (0,T), \end{cases}$$

$$\begin{cases} \mathcal{L}u_2^{k+1} = f & \text{in } \Omega_2 \times (0,T), \\ u_2^{k+1}(\cdot,0) = u_0 & \text{in } \Omega_2, \\ u_2^{k+1} = u_1^k & \text{on } \Gamma_2 \times (0,T), \end{cases}$$

corresponding to an infinite block Jacobi waveform relaxation. It can be compared to the parallel version of the Schwarz algorithm, introduced in [23], and that is where the name comes from. The algorithm has the same convergence properties 1,2 as the waveform relaxation algorithm. Consider the advection-diffusion-reaction equation in two dimensions,

$$\partial_t u + \mathbf{a} \cdot \nabla u - \nu \Delta u + cu = f.$$

The convergence factor for two semi-infinite subdomains on the Fourier side (ω is the dual variable of time, η is the dual variable of y), is given by $\rho = e^{-\frac{L}{\nu}\sqrt{a_x^2+4\nu(b+i(\omega+a_y\eta)+\nu\eta^2)}}$, and we have for a small overlap L,

$$\sup_{(\omega,\eta)\in\mathbb{R}^{1+1}} |\rho| = e^{-\frac{L}{\nu}\sqrt{a_x^2+4\nu b}} = 1 - \mathcal{O}(L).$$

4 Fruits: Optimized Schwarz Waveform Relaxation for Parabolic Equations

Such algorithms, relying on the theory of absorbing boundary conditions, were announced in [8], for one-dimensional wave and heat equations. For the advection-diffusion-reaction equation in two dimensions, the first thorough analysis was performed in [18]. We proposed in [1] to use a new type of transmission conditions between the subdomains, with or without overlap,

$$\mathcal{B}u := \nu\partial_n u - \frac{a_x}{2}u + \frac{p}{2}u + 2q(\partial_t + a_y u - \nu\Delta_y u + bu), \tag{2}$$

The case $q = 0$ corresponds to Robin transmission conditions, and was explored in [19], with extension to first order conditions (*i.e.* without the Laplace-Beltrami operator in (2)). The convergence factor for two semi-infinite subdomains is given by

$$\begin{aligned} \rho(z,p,q) &= \left(\frac{p+qz-\sqrt{a_x^2+4\nu b+z}}{p+qz+\sqrt{a_x^2+4\nu b+z}}\right)^2 e^{-\frac{L}{\nu}\sqrt{a_x^2+4\nu b+z}}, \\ z &= 4\nu\big(i(\omega+a_y\eta)+\nu|\eta|^2\big). \end{aligned} \tag{3}$$

The terms $O0$ and $O2$ correspond to Robin and second order transmission condition (2), with coefficients determined for two subdomains by minimizing the convergence factor. Note that the use of such transmission conditions leads to problems of the same complexity as the Dirichlet transmission conditions.

Convincing comparisons between Dirichlet and optimized transmission conditions where given in [18], the coefficients being computed numerically. Choosing the best coefficients p and q in (2) leads to a new best approximation problem, that was solved by hand for Robin conditions (q=0) in [6] and studied in depth in [1]. Note that various asymptotics come into play: the size of the overlap in terms of the mesh-size in space Δx, the frequencies actually supported by the grid: $\omega \in [\pi/T, \pi/\Delta T]$, $\eta \in [\pi/Y, \pi/\Delta y]$. Asymptotic values of the optimal parameter were given in [6] for Robin transmission conditions in one dimension. We give in Table 1 the asymptotic values in dimension $n > 1$, which were first included in our manuscript for [6], but discarded by the referee who thought it was too complicated. In [1] we give asymptotic values of the parameters of the order 2 method for two subdomains in any dimension. These asymptotics make the algorithm easy to use and fast. Table 2 shows the asymptotic values of the convergence factor for various algorithms.

Note that the value of the optimal parameters depends on the rate between Δt and Δx, while the convergence factor does not. The convergence factor of the optimized of order 2 method is almost independent of the mesh-size, unlike classical methods. Furthermore, the convergence is even better with a small overlap.

We now discuss numerically the dependence with respect to the number of subdomains. In Fig. 1, we consider the advection-diffusion equation in one-dimension on the domain $[0,6]$, with $a = 1$ and $\nu = 1/5$. The final time is $T = 2.5$. It is discretized with implicit Euler, $\Delta x = 0.01$, $\Delta t = 0.0025$. The overlap is 4 gridpoints.

method	parameter p
O0 overlap Δx	$\begin{cases} (\nu(a^2+4\nu b))^{1/3}\Delta x^{-1/3} & \text{if } \beta=1 \\ (2\nu(a^2+4\nu b))^{1/3}\Delta x^{-1/3} & \text{if } \beta=2 \end{cases}$
O0 no overlap	$\begin{cases} (2\pi\nu\sqrt{(n-1)(a^2+4\nu b)})^{1/2}\Delta x^{-1/2} & \text{if } \beta=1 \\ (8\pi\nu(a^2+4\nu b))^{1/4}\Delta x^{-1/2} & \text{if } \beta=2,\ \nu\le\bar{\nu} \\ \left(\frac{8\pi\nu(a^2+4\nu b)((n-1)^2\nu^2\pi^2+1)}{\sqrt{(n-1)^2\nu^2\pi^2+1}+(n-1)\nu\pi}\right)^{1/4}\Delta x^{-1/2} & \text{if } \beta=2,\ \nu>\bar{\nu} \end{cases}$

Table 1. Summary of the asymptotic parameters p in dimension $n>1$ in the Robin transmission conditions, for $\Delta t=\Delta x^{\beta}$, $\beta=1,2$, Δx the discretization step in all spatial directions. $\bar{\nu}\approx 1.5437/(\pi(n-1))$, and $a=a_x$.

	overlap	no overlap
Dirichlet	$1-O(\Delta x)$	
Optimized order 0	$1-O(\Delta x^{1/3})$	$1-O(\Delta x^{1/2})$
Optimized order 2	$1-O(\Delta x^{1/5})$	$1-O(\Delta x^{1/4})$

Table 2. Asymptotic expansion of the convergence factor in dimension $n>1$.

For the optimized Robin (resp. Order 2 optimized) algorithm, the same coefficient is used for all subdomains, computed by the explicit formulas given in [1], that is $p=2.054275607$ (resp. $p=1.366061845, q=0.1363805228$). We compute the number of iterations necessary to reach an error of 10^{-6} (the error is measured as the discrete L^2 error in time on the interface), as a function of the number of subdomains. The convergence history depends strongly on the number of subdomains. However optimized Schwarz waveform relaxation beats the classical Schwarz in two ways: the convergence is much faster (there is a factor 8 to 10), second the influence of the increase of the number of subdomains is much weaker.

Systematic testing is undergoing to see how relevant these values are for more general geometries and any number of subdomains.

5 Other Fruits: Optimized Schwarz Waveform Relaxation for Other Types of PDEs

The strategy of Schwarz waveform relaxation applies to every type of partial differential equations. We have in each case proposed new algorithms, and compared them to classical Schwarz waveform relaxation, with Dirichlet transmission conditions.

- ♣ We have designed algorithms for the wave equation in one or two dimensions. In two dimensions, due to the presence of evanescent waves, we have been led to choose overlapping algorithms, with optimized of order 2 transmission conditions. The coefficients are given by an exact formula as functions of the overlap and the desired tolerance on the error in [5].
- ♣ For the Schrödinger equation in one dimension, we have tested (complex) Robin conditions, which behave much better with overlap. The coefficient is computed

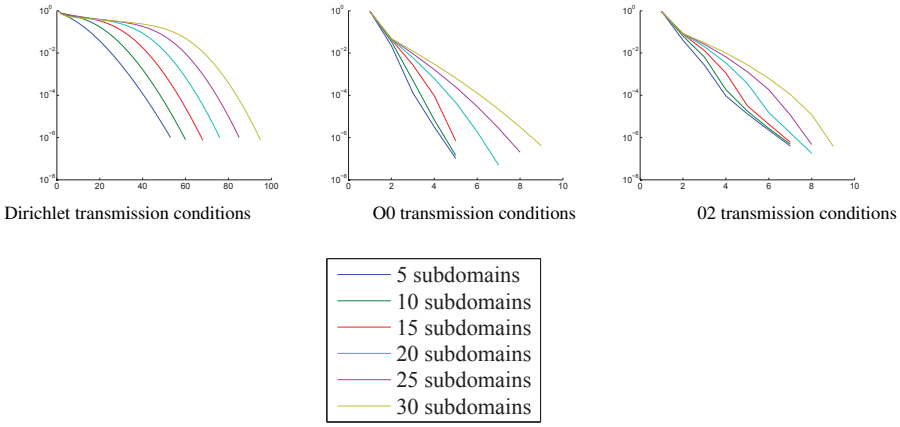


Fig. 1. Convergence history for the three main algorithms for 5 to 30 subdomains. Logaritmic scale.

again by an explicit formula. Moreover, we have suggested to use the exact Dirichlet to Neumann map, designed for the Crank-Nicolson scheme. The convergence is achieved in 2 or 3 iterations, even with non constant potential (linear or parabolic), for which the exact Dirichlet to Neumann map is not known, see [12].

♣ For nonlinear waves, we have designed nonlinear Schwarz waveform relaxation algorithms, which again converge in a few iterations, see [13].

6 New Blossoms: Space-Time Coupling and Refinements

In principle, nothing can prevent us from using different time and space meshes in different subdomains, in any of our algorithms. The question is how to match the discrete approximations on the interfaces of the subdomains. These interfaces have dimension 1 (time) + $N-1$ (space). We designed in [9] an optimal algorithm in one dimension, that we used for the wave equation and space-time refinement. Using a leapfrog scheme in each subdomain, which is of order two in space and time, we were able to keep a CFL number equal to 1 in each subdomain, which minimizes the dispersion and gives convergence in two iterations on time windows, and proved the overall solution to be an order 2 approximation of the wave equation. The extension to the 2-D wave equation is ongoing.

For parabolic problems, we use Discontinuous Galerkin methods in time with Optimized Order 2 Schwarz waveform relaxation (see presentation by Caroline

Japhet in DD18). The projection algorithm in time is coupled with an optimal projection algorithm in space (see contribution by Gander and Japhet in DD18), and mortar projection. Extension to non linear problems is ongoing.

References

[1] Amitai, D., Averbuch, A., Israeli, M., Itzikowitz, S., Turkel, E.: A survey of asynchronous finite-difference methods for parabolic PDEs on multiprocessors. *Appl. Numer. Math.*, 12:27–45, 1993.

[1] Bennequin, D., Gander, M.J., Halpern, L.: A homographic best approximation problem with application to optimized Schwarz waveform relaxation. *Math. Comp.*, 78:185–223, 2009.

[3] Burrage, K., Dyke, C., Pohl, B.: On the performance of parallel waveform relaxations for differential systems. *Appl. Numer. Math.*, 20:39–55, 1996.

[4] Chazan, D., Miranker, W.: Chaotic relaxation. *J. Linear Algebra Appl.*, 2: 192–222, 1969.

[5] Gander, M.J., Halpern, L.: Absorbing boundary conditions for the wave equation and parallel computing. *Math. Comp.*, 74(249):153–176, 2004.

[6] Gander, M.J., Halpern, L.: Optimized Schwarz waveform relaxation methods for advection reaction diffusion problems. *SIAM J. Numer. Anal.*, 45(2):666–697, 2007.

[7] Gander, M.J., Stuart, A.M.: Space time continuous analysis of waveform relaxation for the heat equation. *SIAM J. Sci. Comput.*, 19:2014–2031, 1998.

[8] Gander, M.J., Halpern, L., Nataf, F.: Optimal convergence for overlapping and non-overlapping Schwarz waveform relaxation. In C-H. Lai, P. Bjørstad, M. Cross, and O. Widlund, eds., *Eleventh international Conference of Domain Decomposition Methods*, pages 27–36, 1999.

[9] Gander, M.J., Halpern, L., Nataf, F.: Optimal Schwarz waveform relaxation for the one dimensional wave equation. *SIAM J. Numer. Anal.*, 41(5):1643–1681, 2003.

[10] Giladi, E., Keller, H.B.: Space time domain decomposition for parabolic problems. *Numer. Math.*, 93(2):279–313, 2002.

[11] Hackbusch, W.: Parabolic multi-grid methods. In R. Glowinski and J.-L. Lions, eds., *Computing Methods in Applied Sciences and Engineering, VI*, pages 189–197. North-Holland, 1984.

[12] Halpern, L., Szeftel, J.: Optimized and quasi-optimal Schwarz waveform relaxation for the one dimensional Schrödinger equation. In U. Langer, M. Discacciati, D.E. Keyes, O.B. Widlund, and W. Zulehner, eds., *Domain Decomposition Methods in Science and Engineering XVII*, pages 221–228. Springer, 2008.

[13] Halpern, L., Szeftel, J.: Nonlinear Schwarz waveform relaxation for semilinear wave propagation. *Math. Comp.*, 2009.

[14] Horton, G., Vandewalle, S. A space-time multigrid method for parabolic partial differential equations. *SIAM J. Sci. Comput.*, 16(4):848–864, 1995.

[15] Lelarasmee, E., Ruehli, A.E., Sangiovanni-Vincentelli, A.L.: The waveform relaxation method for time-domain analysis of large scale integrated circuits. *IEEE Trans. CAD Integrated Circuits Syst.*, 1:131–145, 1982.

[16] Lions, J.-L., Maday, Y., Turinici, G.: A parareal in time discretization of PDE's. *C. R. Acad. Sci. Paris Sér. I Math.*, 332:661–668, 2001.

[23] Lions, P.-L.: On the Schwarz alternating method. I. In R. Glowinski, G.H. Golub, G.A. Meurant, and J. Périaux, eds., *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 1–42. SIAM, 1988.

[18] Martin, V.: *Méthodes de décomposition de domaines de type relaxation d'ondes pour des équations de l'océnographie*. PhD thesis, Université Paris 13, 2003.

[19] Martin, V.: An optimized Schwarz waveform relaxation method for the unsteady convection diffusion equation. *Appl. Numer. Math.*, 52(4):401–428, 2005.

[20] Miekkala, U., Nevanlinna, O.: Convergence of dynamic iteration methods for initial value problems. *SIAM J. Sci. Statist. Comput.*, 8:459–482, 1987.

[22] Quarteroni, A., Valli, A.: *Domain Decomposition Methods for Partial Differential Equations*. Oxford Science Publications, 1999.

[22] Vandewalle, S., Horton, G.: Fourier mode analysis of the multigrid waveform relaxation and time-parallel multigrid methods. *Computing*, 54(4):317–330, 1995.