Time Parallel Time Integration

Part III

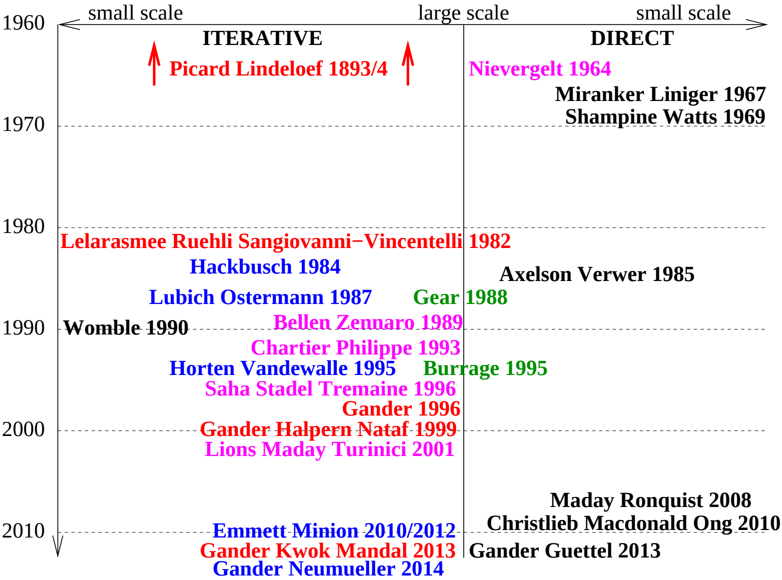
Space-Time Multigrid Methods

Martin J. Gander
martin.gander@unige.ch

University of Geneva

CEMRACS, July 2016

Methods Based on Multigrid



Hackbusch 1984

Parabolic multigrid methods. Computing Methods in Applied Sciences and Engineering, VI, R. Glowinski and J.-L. Lions, Eds. North-Holland, 1984.

Parabolic PDE $u_t + L_h u = f$ discretized by Backward Euler:

$$\frac{1}{\Delta t}(u(t) - u(t - \Delta t)) + L_h u(t) = f(t) \quad (*)$$

> *"The conventional approach is to solve (*) time step by time step; $u(t)$ is computed from $u(t - \Delta t)$, then $u(t + \Delta t)$ from $u(t)$ etc. The following process will be different. Assume that $u(t)$ is already computed or given as an initial state. Simultaneously, we shall solve for $u(t + \Delta t)$, $u(t + 2\Delta t)$, ..., $u(t + k\Delta t)$ in one step of the algorithm."*

Method of Hackbusch

From the problem at each time step

$$Au(t) = -\frac{1}{\Delta t}u(t-\Delta t) + f(t), \quad A := \frac{1}{\Delta t}I + L_h$$

take a Gauss-Seidel smoother S, i.e. $A = L + D + U$

$$\begin{aligned} u_{n+1} &= S(t, u_n, u(t-\Delta t), f(t)) \\ &:= (L+D)^{-1}(-Uu_n - \frac{1}{\Delta t}u(t-\Delta t) + f(t)). \end{aligned}$$

The parabolic multigrid method is a multigrid method in space time with the following smoothing procedure:

for $\tau = t : \Delta t : t + k\Delta t$
  for $j = 1 : \nu$
    $u(\tau) = S(\tau, u(\tau), u(\tau - \Delta t), f(\tau))$;
  end;
end

Results of Hackbusch

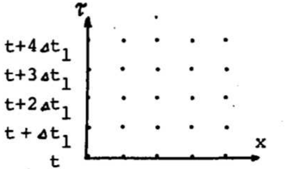


Fig 2.1a: Grid at level l

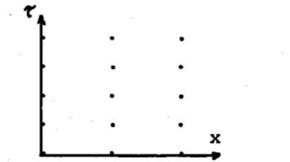


Fig 2.1b: Grid at level l-1

- If one does only coarsen in space, then standard multigrid performance is achieved.
- If one also coarsens in time, one does not obtain standard multigrid performance, and the method can even diverge.
- This is traced back to errors which are smooth in space, but non smooth in time.

Numerical experiments for buoyancy-driven flow with finite difference discretization

Lubich and Ostermann 1987

Multigrid Dynamic Iteration for Parabolic Problems. BIT 27, 1987.

> *"We study the method which is obtained when a multi-grid method (in space) is first applied directly to a parabolic initial-boundary value problem, and discretization in time is done only afterward."*

Laplace transform:

$$u_t + L_h u = f \quad \Longrightarrow \quad A(s)\hat{u} := s\hat{u} + L_h\hat{u} = \hat{f}$$

Multigrid for $A(s)\hat{u} = \hat{f}$: Let $A(s) = L + D + sI + U$,

Initial guess $\hat{u}_0^0(s)$. For $n = 0, 1, 2, \ldots$

for $j = 1 : \nu$

$\hat{u}_n^j(s) := (L + D + sI)^{-1}(-U\hat{u}_n^{j-1}(s) + \hat{f}(s))$

end;

$\hat{u}_{n+1}^0(s) := \hat{u}_n^\nu(s) + EA_c^{-1}R(\hat{f}(s) - A\hat{u}_n^\nu(s))$

smooth again

Algorithm in the Time Domain

The smoothing step

$$(sI + L + D)\hat{u}_n^j(s) = -U\hat{u}_n^{j-1}(s) + \hat{f}(s)$$

becomes in the time domain

$$\partial_t u_n^j + (L + D)u_n^j + Uu_n^{j-1} = f$$

which is a **Gauss Seidel Waveform Relaxation iteration!**

The coarse correction

$$\hat{u}_{n+1}^0(s) := \hat{u}_n^\nu(s) + EA_c^{-1}R(\hat{f} - A\hat{u}_n^\nu(s))$$

becomes

solve $v_t + L_H v = R(f - \partial_t u_n^\nu - L_h u_n^\nu)$
$u_{n+1}^0 = u_n^\nu + Ev$

time continuous parabolic problem on coarse spatial mesh

Results of Lubich and Ostermann

For the heat equation and finite difference discretization:

- Red-black Gauss Seidel smoothing is not as good as for the stationary problem, but still sufficient to give typical multi-grid convergence.
- Damped Jacobi smoothing is as good as for stationary problem.
- Time discretization leads to similar results if no time coarsening is performed.

Numerical experiments with locally adapted time steps for

$$u_t = u_{xx} + e^{-(x-t)^2}$$

$x \in (0, 10)$, $t \in (0, 12)$, $u(x, 0) = 0$.

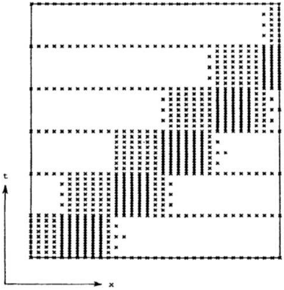


Fig. 1. Grid-points in the (x, t)-plane.

Horton and Vandewalle 1995

A Space-Time Multigrid Method for Parabolic Partial Differential Equations. SIAM J. Sci. Comput, Vol. 16, No. 4

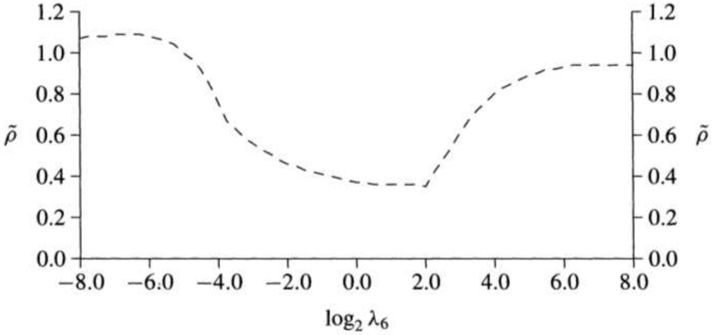
> *"In standard time-stepping techniques multigrid can be used as an iterative solver for the elliptic equations arising at each discrete time step. By contrast, the method presented in this paper treats the whole of the space-time problem simultaneously."*

2D heat equation discretized by centered finite differences in space and backward Euler in time:

$$\begin{pmatrix} A_1 & & & & \\ B_2 & A_2 & & & \\ & B_3 & A_3 & & \\ & & \ddots & \ddots & \\ & & & B_n & A_n \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \\ \vdots \\ \mathbf{u}_n \end{pmatrix} = \begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \mathbf{f}_3 \\ \vdots \\ \mathbf{f}_n \end{pmatrix}$$

Problem of Standard Multigrid in Space-Time

For V-cycle with one Gauss-Seidel red/black pre and post smoothing step, standard coarsening with factor 2 in space and time and full weighting and bilinear prolongation.

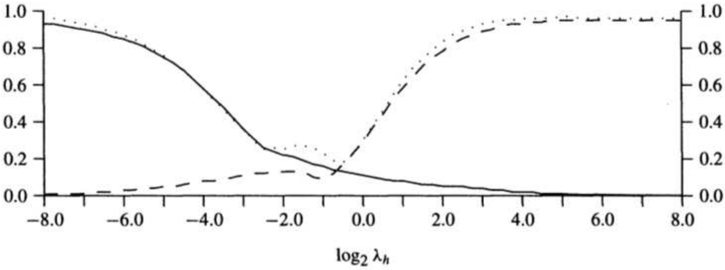


Here $\tilde{\rho}$ is the 2-norm contraction factor, and $\lambda_6 := \Delta t/\Delta x^2$ on the finest level 6, 65×65 points.

- Method only works when $\lambda_6 = \Delta t/\Delta x^2$ is close to 1
- Even then contraction is worse than for spatial multigrid

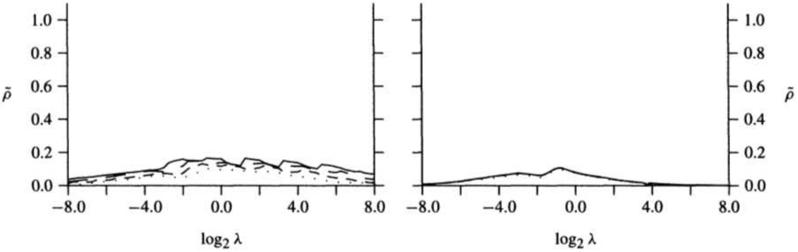
Key New Ideas of Horton and Vandewalle

1. Adaptive semi-coarsening in space or time to keep λ in a good range for the contraction factor.
2. Prolongation operators only forward in time.
3. Restriction operators only backward in time.



Two-grid contraction factor with 2 pre- and postsmoothing steps for $\lambda_h = \Delta t/\Delta x^2$.
Standard coarsening (dotted line), space coarsening (solid line), time coarsening (dashed line)

Results of Horten and Vandewalle



V-cycle and F-cycle contraction factors with 3 smoothing steps. Space time grids of 256×256 (solid), 128×128 (long dashed), 64×64 (short dashed), 32×32 (dotted).

- Analysis based on local Fourier modes
- Good contraction rates for V-cycles, but not quite mesh independent
- Mesh independent convergence rates for F-cycles

Numerical results for 1d, 2d, and 3d heat equations

Minion 2010

A Hybrid Parareal Spectral Deferred Corrections Method, Comm. App. Math. and Comp Sci. Vol. 5, No. 2

> *"This paper investigates a variant of the parareal algorithm first outlined by Minion and Williams in 2008 that utilizes a deferred correction strategy within the parareal iterations."*

Review of deferred (difference or defect) correction:

$$u' = f(u), \quad u(0) = u_0$$

Let $\tilde{u}_m$ be an order one approximation (e.g. FE). If $\tilde{u}(t)$ is its interpolant, the error $e(t) := u(t) - \tilde{u}(t)$ satisfies

$$e'(t) = u'(t) - \tilde{u}'(t) = f(u) - \tilde{u}'(t) = f(e + \tilde{u}) - \tilde{u}'(t)$$

an ODE for $e(t)$ with $e(0) = 0$! Solving with FE gives e_m.

Minion 2010

A Hybrid Parareal Spectral Deferred Corrections Method, Comm. App. Math. and Comp Sci. Vol. 5, No. 2

> *"This paper investigates a variant of the parareal algorithm first outlined by Minion and Williams in 2008 that utilizes a deferred correction strategy within the parareal iterations."*

Review of deferred (difference or defect) correction:

$$u' = f(u), \quad u(0) = u_0$$

Let $\tilde{u}_m$ be an order one approximation (e.g. FE). If $\tilde{u}(t)$ is its interpolant, the error $e(t) := u(t) - \tilde{u}(t)$ satisfies

$$e'(t) = u'(t) - \tilde{u}'(t) = f(u) - \tilde{u}'(t) = f(e + \tilde{u}) - \tilde{u}'(t)$$

an ODE for $e(t)$ with $e(0) = 0$! Solving with FE gives e_m.

Theorem (Skeel 1976, see also Fox 1947, Pereyra 1967, Frank and Ueberhuber 1977): The new approximation $\tilde{u}_m + e_m$ is of order two. $\Longrightarrow$ iterated defect correction

Minion 2010

A Hybrid Parareal Spectral Deferred Corrections Method, Comm. App. Math. and Comp Sci. Vol. 5, No. 2

> *"This paper investigates a variant of the parareal algorithm first outlined by Minion and Williams in 2008 that utilizes a deferred correction strategy within the parareal iterations."*

Review of deferred (difference or defect) correction:

$$u' = f(u), \quad u(0) = u_0$$

Let $\tilde{u}_m$ be an order one approximation (e.g. FE). If $\tilde{u}(t)$ is its interpolant, the error $e(t) := u(t) - \tilde{u}(t)$ satisfies

$$e'(t) = u'(t) - \tilde{u}'(t) = f(u) - \tilde{u}'(t) = f(e + \tilde{u}) - \tilde{u}'(t)$$

an ODE for $e(t)$ with $e(0) = 0$! Solving with FE gives e_m.

Theorem (Skeel 1976, see also Fox 1947, Pereyra 1967, Frank and Ueberhuber 1977): The new approximation $\tilde{u}_m + e_m$ is of order two. $\Longrightarrow$ iterated defect correction

Problems: numerical interpolation and differentiation

Dutt, Greengard, Rokhlin 2000

Spectral integral deferred correction:

$$u' = f(u), \quad u(0) = u_0 \quad \Longrightarrow \quad u(t) = u(0) + \int_0^t f(u(\tau))d\tau.$$

Let $\tilde{u}(t)$ be an approximation with residual

$$r(t) := \tilde{u}(0) + \int_0^t f(\tilde{u}(\tau))d\tau - \tilde{u}(t),$$

The error $e(t) := u(t) - \tilde{u}(t)$ satisfies (with $u(0) = \tilde{u}(0)$)

$$\tilde{u}(t) + e(t) = \tilde{u}(0) + \int_0^t f(\tilde{u}(\tau) + e(\tau))d\tau.$$

$$\begin{aligned} \Longrightarrow e(t) &= \tilde{u}(0) + \int_0^t f(\tilde{u}(\tau) + e(\tau))d\tau - \tilde{u}(t) \\ &= r(t) + \int_0^t f(\tilde{u}(\tau) + e(\tau)) - f(\tilde{u}(\tau))d\tau \end{aligned}$$

Using differentiation, we obtain from

$$e(t) = r(t) + \int_0^t f(\tilde{u}(\tau) + e(\tau)) - f(\tilde{u}(\tau))d\tau,$$

the differential equation for the error

$$e'(t) = r'(t) + f(\tilde{u}(t) + e(t)) - f(\tilde{u}(t)).$$

Starting with an order one method, e.g. FE

$$\tilde{u}_{m+1} = \tilde{u}_m + \Delta t f(\tilde{u}_m), \quad \text{for } m = 0, 1, \ldots, M-1.$$

one evaluates with high order quadrature the residual

$$r(t) = \tilde{u}(0) + \int_0^t f(\tilde{u}(\tau))d\tau - \tilde{u}(t),$$

and then computes the approximate error with FE

$$e_{m+1} = e_m + r_{m+1} - r_m + \Delta t(f(\tilde{u}_m + e_m) - f(\tilde{u}_m)).$$

Theorem (Böhmer and Stetter 1984) The new approximation $\tilde{u}_m + e_m$ is of order two.

Spectral Integral Deferred Correction Integrator

Let $\mathbf{u}^0 := (\tilde{u}_0, \tilde{u}_1, \ldots, \tilde{u}_M)^T$ be the initial approximation from forward Euler. Then integral deferred correction is a fixed point iteration

$$\mathbf{u}^k = F(\mathbf{u}^{k-1}, u_0) \qquad (*)$$

Classical approach: partition $[0, T]$ into $[T_{j-1}, T_j]$, $j = 1, 2, \ldots, J$, and then perform $(*)$ sequentially on each interval:

$u_{0,M}^K = u_0$;
for $j = 1 : J$
  compute $\mathbf{u}_j^0$ as Euler approximation on $[T_{j-1}, T_j]$;
  for $k = 1 : K$
    $\mathbf{u}_j^k = F(\mathbf{u}_j^{k-1}, u_{j-1,M}^K)$;
  end;
end;

Spectral Integral Deferred Correction Integrator

Let $\mathbf{u}^0 := (\tilde{u}_0, \tilde{u}_1, \ldots, \tilde{u}_M)^T$ be the initial approximation from forward Euler. Then integral deferred correction is a fixed point iteration

$$\mathbf{u}^k = F(\mathbf{u}^{k-1}, u_0) \qquad (*)$$

Classical approach: partition $[0, T]$ into $[T_{j-1}, T_j]$, $j = 1, 2, \ldots, J$, and then perform $(*)$ sequentially on each interval:

$u^K_{0,M} = u_0;$
for $j = 1 : J$
  compute $\mathbf{u}^0_j$ as Euler approximation on $[T_{j-1}, T_j]$;
  for $k = 1 : K$
    $\mathbf{u}^k_j = F(\mathbf{u}^{k-1}_j, u^k_{j-1,M});$
  end;
end;

Idea of Minion (2010): replace K by k (see also Womble later), and use this as fine propagator in parareal.

Emmett and Minion 2012

Toward an efficient parallel in time method for partial differential equations, Comm. App. Math. and Comp Sci. Vol. 7.

> *"A new method for the parallelization of numerical methods for partial differential equations (PDEs) in the temporal direction is presented. The method is iterative with each iteration consisting of deferred correction sweeps performed alternately on fine and coarse space-time discretizations. The coarse grid problems are formulated using a space-time analog of the full approximation scheme popular in multigrid methods for nonlinear equations."*

This now called PFASST algorithm uses the parallel spectral deferred correction iteration of Minion (2010) as a smoother in a multigrid full approximation scheme in space-time for non-linear problems.

Gander and Neumüller 2014

Use block Jacobi smoother for a multi-grid method applied to the entire space-time system

$$\begin{pmatrix} A_1 & & & & \\ B_2 & A_2 & & & \\ & B_3 & A_3 & & \\ & & \ddots & \ddots & \\ & & & B_n & A_n \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \\ \vdots \\ \mathbf{u}_n \end{pmatrix} = \begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \mathbf{f}_3 \\ \vdots \\ \mathbf{f}_n \end{pmatrix}$$

Theorem (G, Neumüller (2014))

For the heat equation, and block Jacobi smoother, we have:

- *the optimal relaxation parameter is* $\omega = \frac{1}{2}$
- *always good smoothing in time (semi-coarsening is always possible)*
- *for* $\frac{\Delta t}{\Delta h^2} \geq C$ *also good smoothing in space*
- *one V-cycle in space suffices to invert* A_n

3D Heat Equation Parallelization Results

Scaling results on the Vienna Scientific Cluster VSC-2

Weak Scaling					Strong Scaling			
cores	$\frac{1}{\Delta T}$	dof	iter	time	$\frac{1}{\Delta T}$	dof	iter	time
1	4	59768	9	6.8	4096	61202432	9	6960.7
2	8	119536	9	8.1	4096	61202432	9	3964.8
4	16	239072	9	9.2	4096	61202432	9	2106.2
8	32	478144	9	9.2	4096	61202432	9	1056.0
16	64	956288	9	9.2	4096	61202432	9	530.4
32	128	1912576	9	9.3	4096	61202432	9	269.5
64	256	3825152	9	9.4	4096	61202432	9	135.2
128	512	7650304	9	9.4	4096	61202432	9	68.2
256	1024	15300608	9	9.4	4096	61202432	9	34.7
512	2048	30601216	9	9.4	4096	61202432	9	17.9
1024	4096	61202432	9	9.4	4096	61202432	9	9.4
2048	8192	122404864	9	9.5	4096	61202432	9	5.4

(all simulations performed by M. Neumüller)

Conclusions Part III: Space-Time Multigrid

- Hackbusch was the first in 1984 to try multigrid for parabolic problems.
- Lubich and Ostermann invented multigrid waveform relaxation in 1989.
- First fully functioning space-time multigrid method by Horton and Vandewalle in 1995.
- PFASST by Emmett and Minion in 2012 from a combination of integral spectral deferred correction and parareal.
- Fully scalable space-time multigrid method by Gander and Neumüller in 2014

Preprints are available at www.unige.ch/∼gander