50 Years of Time Parallel Time Integration

Martin J. Gander
martin.gander@unige.ch

University of Geneva

Jülich, May 2014

Overview

Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

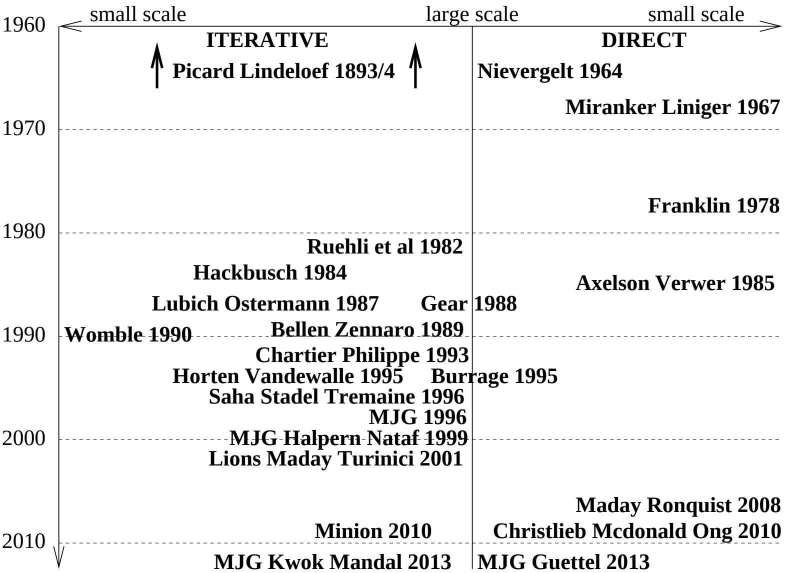
Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary

Time Parallel Methods Over the Course of Time



Overview

Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

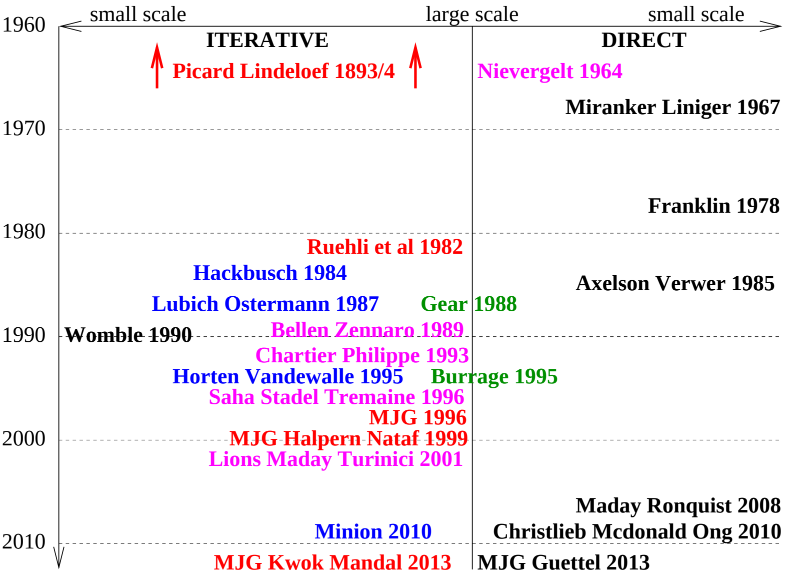
Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary

Methods Based on Multiple Shooting



Overview
Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici
Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR
Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller
Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp
Summary

Jörg Nievergelt 1964

Parallel Methods for Integrating Ordinary Differential Equations. Comm. of the ACM, Vol 7(12), 1964.

> *"For the last 20 years, one has tried to speed up numerical computation mainly by providing ever faster computers.* **Today, as it appears that one is getting closer to the maximal speed of electronic components**, *emphasis is put on allowing operations to be performed in parallel. In the near future, much of numerical analysis will have to be recast in a more 'parallel' form."*

$$y' = f(y), \quad y(x_0) = y_0$$

> *"As an example, a method is proposed for 'parallelizing' the numerical integration of an ordinary differential equation, which process, by all standard methods, is entirely serial"*

Overview

Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

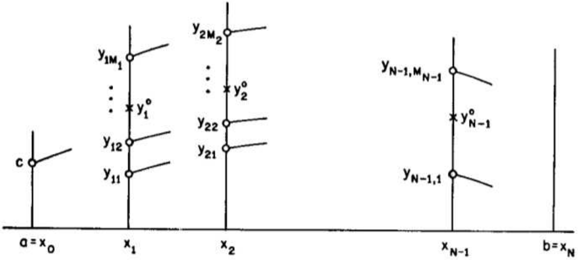
Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary

Nievergelt's Method

The idea is to divide the integration interval [a, b] into N equal subintervals $[x_{i-1}, x_i]$, $x_0 = a$, $x_N = b$, $i = 1, 2, \cdots, N$, to make a rough prediction y_i^0 of the solution $y(x_i)$, to select a certain number M_i of values y_{ij}, $j = 1, 2, \cdots, M_i$ in the vicinity of y_i^0, $i = 1, 2, \cdots, N$, and then to integrate *simultaneously* with an accurate integration method $\mathfrak{M}$ all the initial value problems



. The connection between the branches is now brought about by interpolating the end value of the unique branch in $[x_0, x_1]$

Alfredo Bellen and Marino Zennaro 1989

Parallel algorithms for initial-value problems for difference and differential equations. J. Comp. and Appl. Math. Vol 25.

"...the idea is indeed that of multiple shooting..."

$$y_{n+1} = F_{n+1}(y_n), \quad y_0 \text{ known.}$$

Write all these simultaneously:

$$\mathbf{y} = \mathbf{\Phi}(\mathbf{y}),$$

Apply Steffensen to solve

$$\mathbf{y}^{k+1} = \mathbf{\Phi}(\mathbf{y}^k) + \Delta\mathbf{\Phi}(\mathbf{y}^k)(\mathbf{y}^{k+1} - \mathbf{y}^k)$$

Results of Bellen and Zennaro:

1. each iteration gives one more exact value
2. convergence is locally quadratic
3. corrections can be computed in parallel
4. estimated speedups of 29-53 with 400 steps

Philippe Chartier and Bernard Philippe 1993

A Parallel Shooting Technique for Solving Dissipative ODE's. Computing, Vol 51, 1993.

> *"...die ursprünglich von A. Bellen und M. Zennaro für Differenzengleichungen konzipiert..."*

$$u' = f(u), \quad u(0) = u^0, \qquad t \in [0,1]$$

Multiple shooting: split $[0,1]$ into $[0,\frac{1}{3}]$, $[\frac{1}{3},\frac{2}{3}]$, $[\frac{2}{3},1]$:

$$\begin{array}{rcl} u_0' &=& f(u_0), \\ u_0(0) &=& U_0, \end{array} \qquad \begin{array}{rcl} u_1' &=& f(u_1), \\ u_1(\frac{1}{3}) &=& U_1, \end{array} \qquad \begin{array}{rcl} u_2' &=& f(u_2), \\ u_2(\frac{2}{3}) &=& U_2, \end{array}$$

together with the matching conditions

$$U_0 = u^0, \quad U_1 = u_0(\frac{1}{3}, U_0), \quad U_2 = u_1(\frac{2}{3}, U_1)$$

$$\Longleftrightarrow \quad F(\mathbf{U}) = \left(\begin{array}{c} U_0 - u^0 \\ U_1 - u_0(\frac{1}{3}, U_0) \\ U_2 - u_1(\frac{2}{3}, U_1) \end{array} \right) = 0, \; \mathbf{U} = (U_0, U_1, U_2)^T.$$

Applying Newton's Method to

$$F(\mathbf{U}) = \begin{pmatrix} U_0 - u^0 \\ U_1 - u_0(\frac{1}{3}, U_0) \\ U_2 - u_1(\frac{2}{3}, U_1) \end{pmatrix} = 0$$

$$\begin{pmatrix} U_0^{k+1} \\ U_1^{k+1} \\ U_2^{k+1} \end{pmatrix} = \begin{pmatrix} U_0^{k} \\ U_1^{k} \\ U_2^{k} \end{pmatrix} - \begin{bmatrix} 1 & & \\ -\frac{\partial u_0}{\partial U_0}(\frac{1}{3}, U_0^k) & 1 & \\ & -\frac{\partial u_1}{\partial U_1}(\frac{2}{3}, U_1^k) & 1 \end{bmatrix}^{-1} \begin{pmatrix} U_0^k - u^0 \\ U_1^k - u_1(\frac{1}{3}, U_0^k) \\ U_2^k - u_1(\frac{2}{3}, U_1^k) \end{pmatrix}$$

Multiplying through by the matrix, we find the recurrence

$$\begin{aligned} U_0^{k+1} &= u^0, \\ U_1^{k+1} &= u_0(\tfrac{1}{3}, U_0^k) + \tfrac{\partial u_0}{\partial U_0}(\tfrac{1}{3}, U_0^k)(U_0^{k+1} - U_0^k), \\ U_2^{k+1} &= u_1(\tfrac{2}{3}, U_1^k) + \tfrac{\partial u_1}{\partial U_1}(\tfrac{2}{3}, U_1^k)(U_1^{k+1} - U_1^k). \end{aligned}$$

General case with N intervals:

$$U_{n+1}^{k+1} = u_n(t_{n+1}, U_n^k) + \frac{\partial u_n}{\partial U_n}(t_{n+1}, U_n^k)(U_n^{k+1} - U_n^k).$$

Chartier and Philippe prove locally quadratic convergence and global convergence for dissipative systems

Saha, Stadel, Tremaine 1997

A parallel integration method for solar system dynamics. The Astronomical Journal, Vol 114(1).

> *"Our work is related to the so-called* **waveform relaxation methods**. . . *".*

$$\dot p = -H_q,\ \dot q = H_p, \quad H(p,q,t) = H^0(p) + \epsilon H^1(p,q,t).$$

With $y := (p,q)$, $f(y) := (-H_q(y), H_p(y))$, get $\dot y = f(y)$, derive a multiple shooting method as Chartier Philippe:

$$Y_{n+1}^{k+1} = y_n^{\epsilon}(t_{n+1}, Y_n^k) + \frac{\partial y_n^{\epsilon}}{\partial Y_n}(t_{n+1}, Y_n^k)(Y_n^{k+1} - Y_n^k)$$

Approximate derivative by cheap difference:

$$Y_{n+1}^{k+1} = y_n^{\epsilon}(t_{n+1}, Y_n^k) + y_n^0(t_{n+1}, Y_n^{k+1}) - y_n^0(t_{n+1}, Y_n^k).$$

Results: Each step gives additional order ϵ, speedup 400 with 4000 processors for our solar system

Overview
Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici
Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR
Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller
Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp
Summary

Lions, Maday, Turinici 2001

Résolution d'EDP par un schéma en temps "pararéel".
C. R. Acad. Sci. Paris.

> *"Elle a pour principale motivation les problèmes en temps réel, d'où la terminologie proposée de pararéel."*

$$\dot{y} = -ay, \quad \text{on } [0, T], \quad y(0) = y_0.$$

First use Backward Euler on grid T_n with step ΔT

$$Y_{n+1}^1 - Y_n^1 + a\Delta T Y_{n+1}^1 = 0, \quad Y_0^1 = y_0.$$

Then compute on each interval $[T_n, T_{n+1}]$ exactly

$$\dot{y}_n^1 = -ay_n^1, \quad y_n^1(T_n) = Y_n^1.$$

Iteration for $k = 1, 2, \ldots$:

1. Compute jumps $S_n^k := y_{n-1}^k(T_n) - Y_n^k$
2. Propagate jumps $\delta_{n+1}^k - \delta_n^k + a\Delta T \delta_{n+1}^k = S_n^k$, $\delta_0^k = 0$
3. Set $Y_n^{k+1} := y_{n-1}^k(T_n) + \delta_n^k$ and solve in parallel

$$\dot{y}_n^{k+1} = -ay_n^{k+1}, \quad \text{on } [T_n, T_{n+1}], \quad y_n^{k+1}(T_n) = Y_n^{k+1}.$$

"C'est alors un exercice que de montrer la:"

Proposition: The parareal scheme is order k, i.e. there exists c_k s.t.

$$|Y_n^k - y(T_n)| + \max_{t\in[T_n,T_{n+1}]} |y_n^k(t) - y(t)| \leq c_k \Delta T^k.$$

Parareal Algorithm in Modern Notation for $u' = f(u)$

1. $G(t_2, t_1, u_1)$ is a rough approximation to $u(t_2)$ with initial condition $u(t_1) = u_1$,
2. $F(t_2, t_1, u_1)$ is a more accurate approximation of the solution $u(t_2)$ with initial condition $u(t_1) = u_1$.

Starting with a coarse approximation U_n^0 at the time points $t_1, t_2, \ldots, t_N$, parareal performs for $k = 0, 1, \ldots$ the correction iteration

$$U_{n+1}^{k+1} = F(t_{n+1}, t_n, U_n^k) + G(t_{n+1}, t_n, U_n^{k+1}) - G(t_{n+1}, t_n, U_n^k).$$

G, Vandevalle 2007: Parareal is multiple shooting with the Jacobian approximated by differences on a coarser grid

Precise Convergence Estimate for Parareal

Theorem (G, Hairer 2007)

Let $F(t_{n+1}, t_n, U_n^k)$ denote the exact solution at t_{n+1} and $G(t_{n+1}, t_n, U_n^k)$ be a one step method with local truncation error bounded by $C_1 \Delta T^{p+1}$. If

$$|G(t+\Delta T, t, x) - G(t+\Delta T, t, y)| \leq (1 + C_2 \Delta T)|x - y|,$$

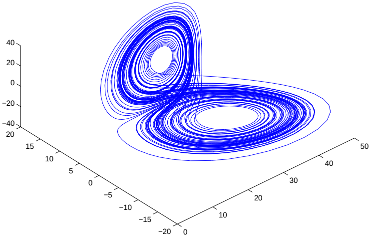
then

$$\begin{aligned}\max_{1\leq n\leq N} |u(t_n) - U_n^k| &\leq \frac{C_1 \Delta T^{k(p+1)}}{k!}(1+C_2\Delta T)^{N-1-k}\prod_{j=1}^{k}(N-j)\max_{1\leq n\leq N}|u(t_n)-U_n^0| \\ &\leq \frac{(C_1 T)^k}{k!} e^{C_2(T-(k+1)\Delta T)} \Delta T^{pk} \max_{1\leq n\leq N}|u(t_n) - U_n^0|.\end{aligned}$$

G and Hairer: Nonlinear Convergence Analysis for the Parareal Algorithm, Domain Decomposition Methods in Science and Engineering XVII, Springer-Verlag, 2007.

Results for the Lorenz Equations

$$\begin{aligned}\dot{x} &= -\sigma x + \sigma y\\ \dot{y} &= -xz + rx - y\\ \dot{z} &= xy - bz\end{aligned}$$



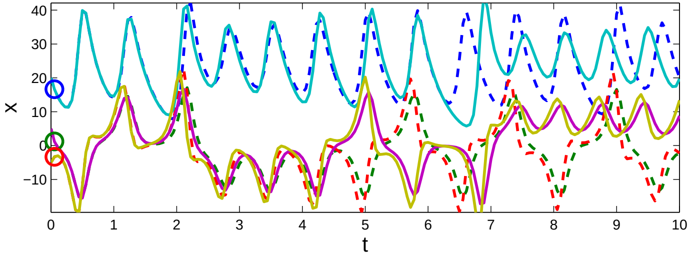
Parameters: $\sigma = 10$, $r = 28$ and $b = \frac{8}{3} \Longrightarrow$ chaotic regime.

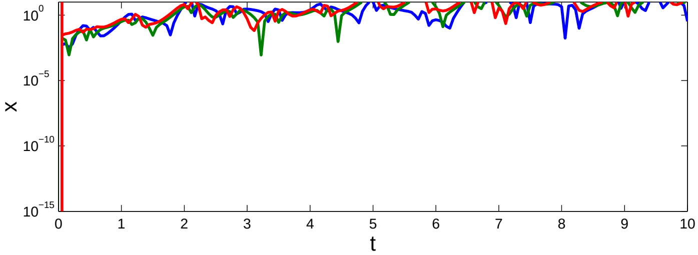
Initial conditions: $(x, y, z)(0) = (20, 5, -5)$

Simulation time: $t \in [0, T = 10]$

Discretization: Fourth order Runge Kutta, $\Delta T = \frac{T}{180}$,

$$\Delta t = \frac{T}{1800}.$$





Overview

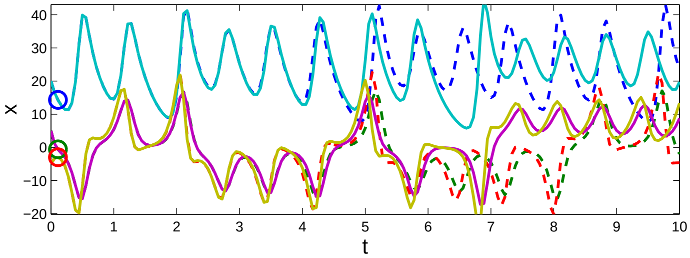
Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

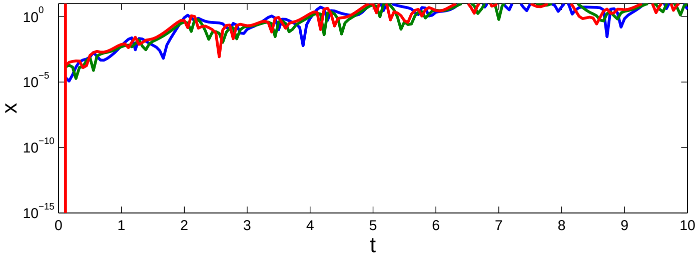
Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary





Overview

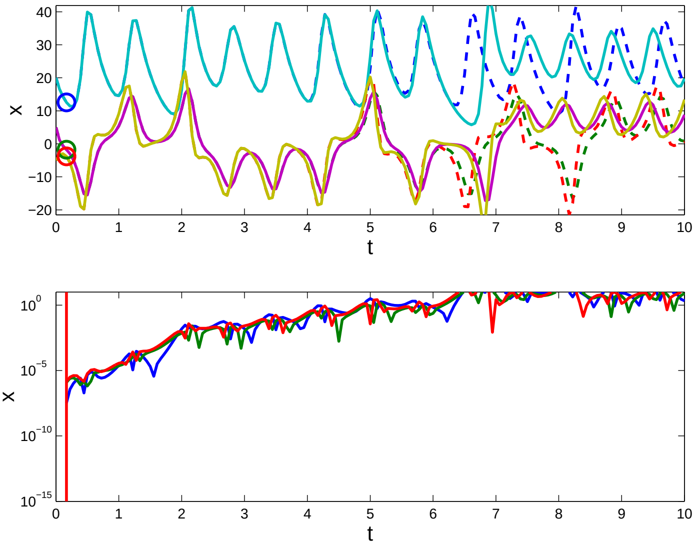
Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

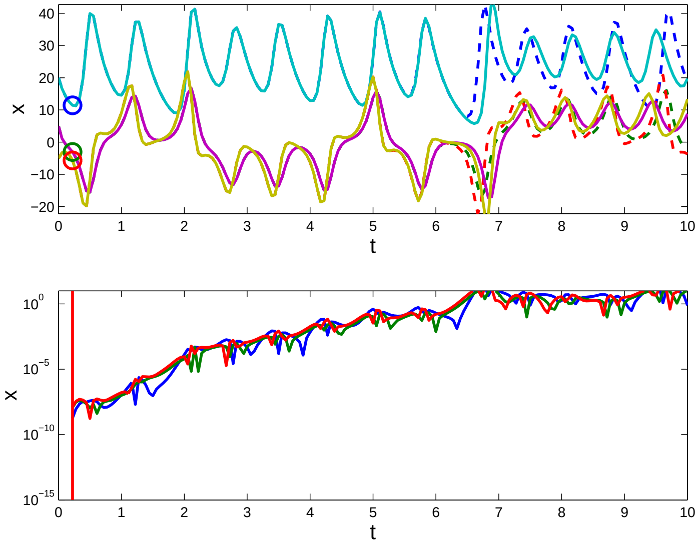
Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary



x
t
0
1
2
3
4
5
6
7
8
9
10
40
30
20
10
0
−10
−20
10^{0}
10^{-5}
10^{-10}
10^{-15}



Overview

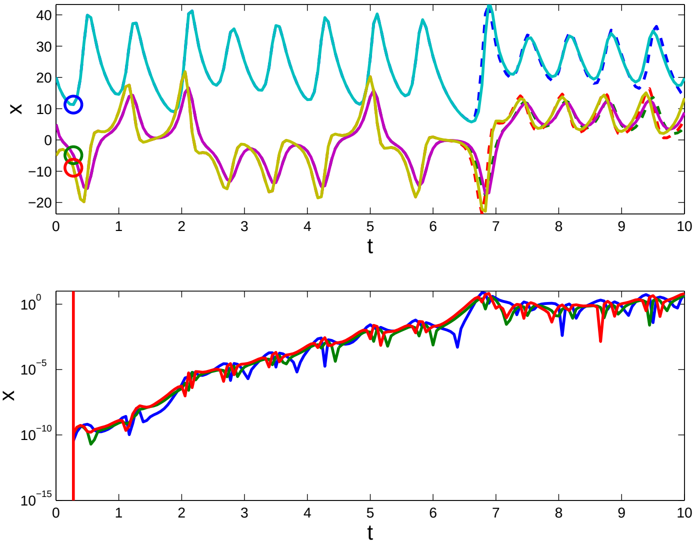
Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

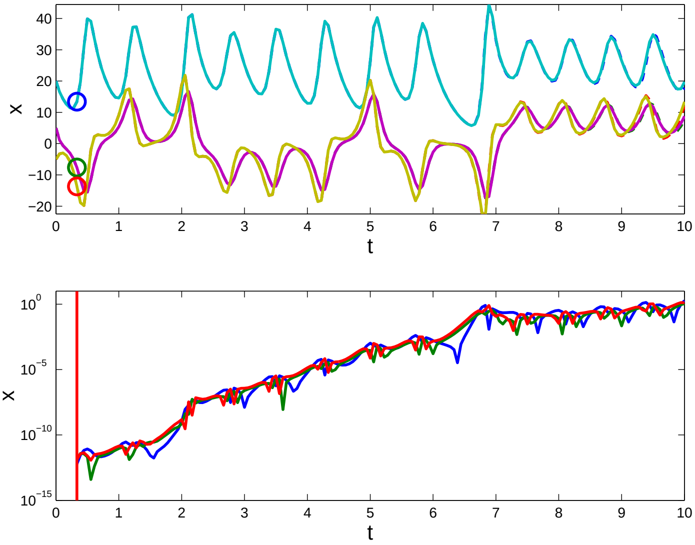
Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary


x
t
40
30
20
10
0
−10
−20
0 1 2 3 4 5 6 7 8 9 10
10^{0}
10^{-5}
10^{-10}
10^{-15}



Overview

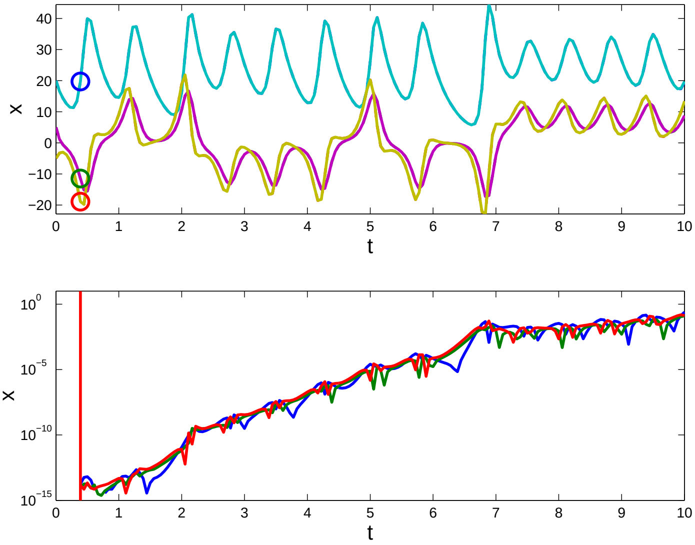
Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary



Overview

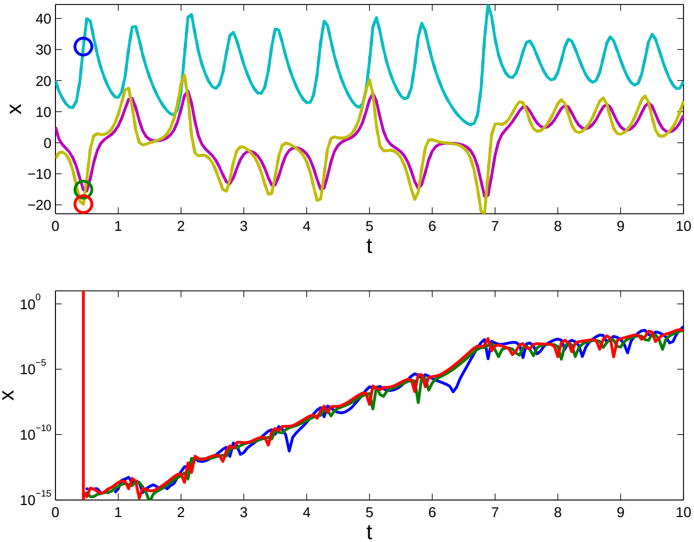
Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary



Overview

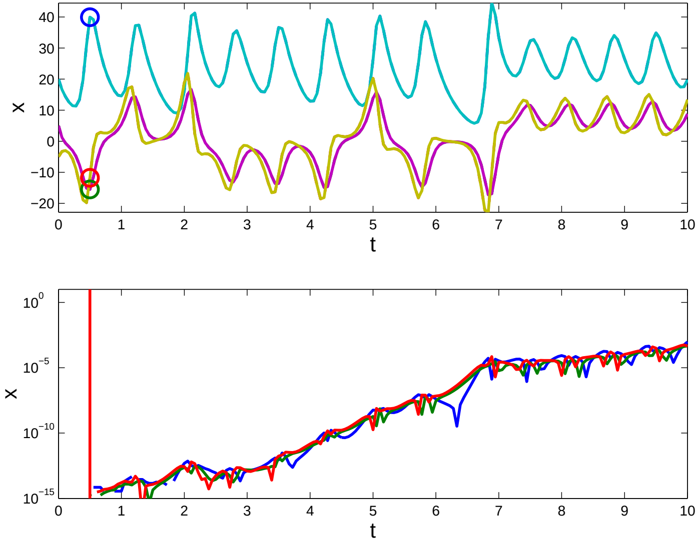
Shooting
- Nievergelt
- Bellen Zennaro
- Chartier Philippe
- Saha et al
- Lions, Maday, Turinici

Schwarz WR
- Picard Lindelöf
- Ruehli et al
- Schwarz WR
- DN and NN WR

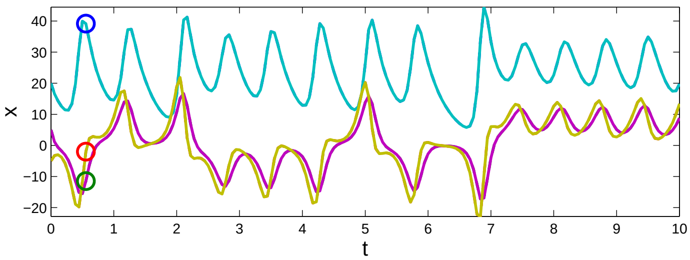
Multigrid
- Hackbusch
- Lubich Ostermann
- Horton Vandewalle
- Minion
- G Neumüller

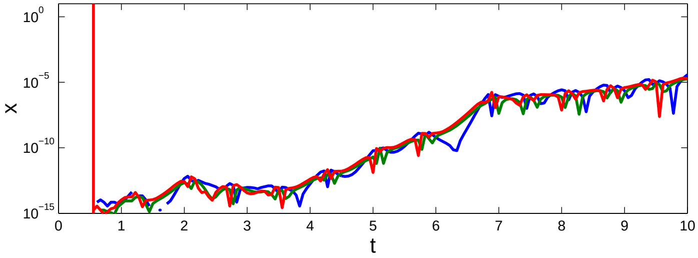
Direct Solvers
- Miranker Liniger
- Axelson Verwer
- Womble
- Maday Ronquist
- RIDC
- ParaExp

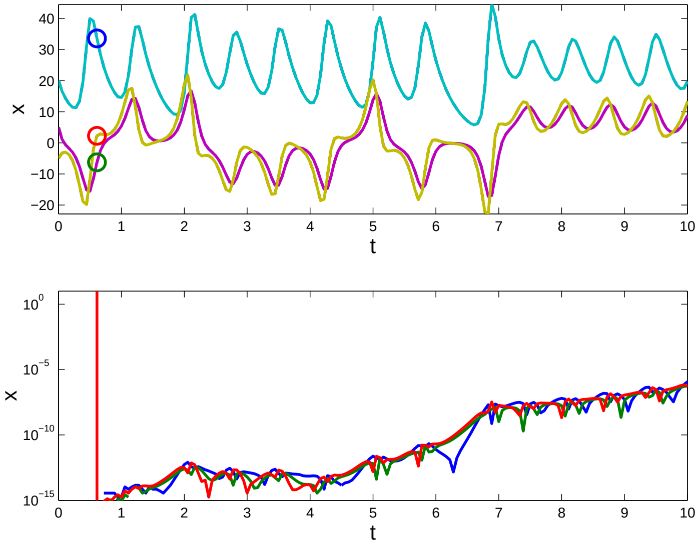
Summary



40
30
20
10
0
−10
−20
x
0 1 2 3 4 5 6 7 8 9 10
t
10^{0}
10^{-5}
10^{-10}
10^{-15}
x
0 1 2 3 4 5 6 7 8 9 10
t


x
t
40
30
20
10
0
−10
−20
0 1 2 3 4 5 6 7 8 9 10


x
t
10^{0}
10^{-5}
10^{-10}
10^{-15}
0 1 2 3 4 5 6 7 8 9 10



Overview

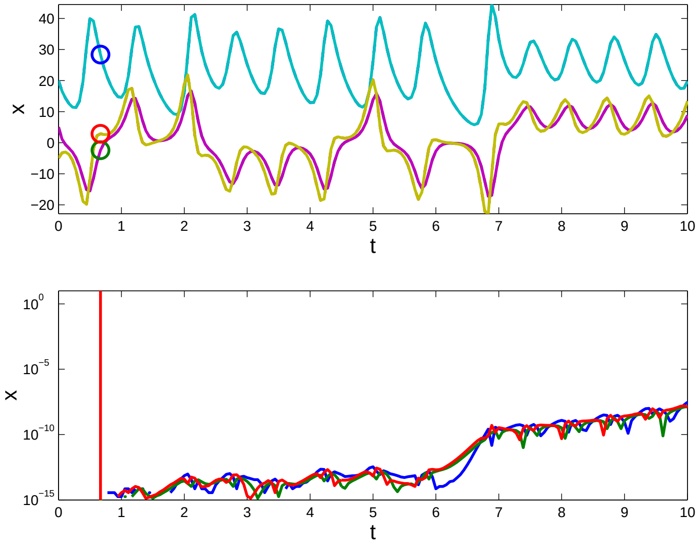
Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

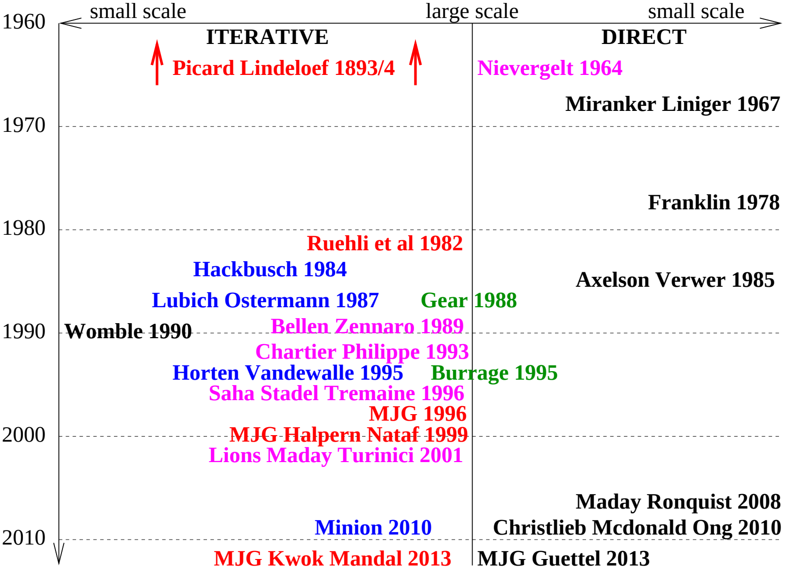
Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary


x
40
30
20
10
0
−10
−20
0 1 2 3 4 5 6 7 8 9 10
t
x
10^{0}
10^{-5}
10^{-10}
10^{-15}
0 1 2 3 4 5 6 7 8 9 10
t

Methods Based on WR and DD



Picard 1893 and Lindelöf 1894

Émile Picard (1893): Sur l'application des méthodes d'approximations successives à l'étude de certaines équations différentielles ordinaires

$$v' = f(v) \quad \Longrightarrow \quad v^n(t) = v(0) + \int_0^t f(v^{n-1}(\tau))d\tau$$

Ernest Lindelöf (1894): Sur l'application des méthodes d'approximations successives à l'étude des intégrales réelles des équations différentielles ordinaires

Theorem (Superlinear Convergence)

On bounded time intervals $t \in [0, T]$, the iterates satisfy the superlinear error bound

$$||\mathbf{v} - \mathbf{v}^n|| \leq \frac{(CT)^n}{n!}||\mathbf{v} - \mathbf{v}^0||,$$

where C is a positive constant.

Overview

Shooting
- Nievergelt
- Bellen Zennaro
- Chartier Philippe
- Saha et al
- Lions, Maday, Turinici

Schwarz WR
- Picard Lindelöf
- Ruehli et al
- Schwarz WR
- DN and NN WR

Multigrid
- Hackbusch
- Lubich Ostermann
- Horton Vandewalle
- Minion
- G Neumüller

Direct Solvers
- Miranker Liniger
- Axelson Verwer
- Womble
- Maday Ronquist
- RIDC
- ParaExp

Summary

Lelarasmee, Ruehli and Sangiovanni-Vincentelli

The Waveform Relaxation Method for Time-Domain Analysis of Large Scale Integrated Circuits. IEEE Trans. on Computer-Aided Design of Int. Circ. a. Sys. 1982

> *"The spectacular growth in the scale of integrated circuits being designed in the VLSI era has generated the need for new methods of circuit simulation. "Standard" circuit simulators, such as SPICE and ASTAP, simply take too much CPU time and too much storage to analyze a VLSI circuit".*

Overview

Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

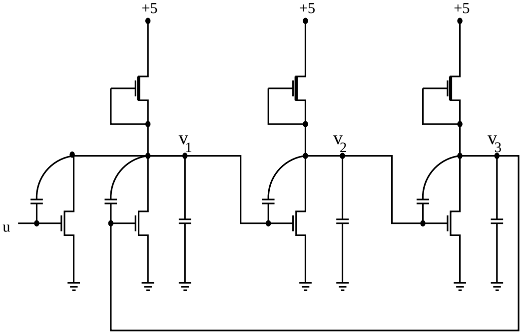
Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary

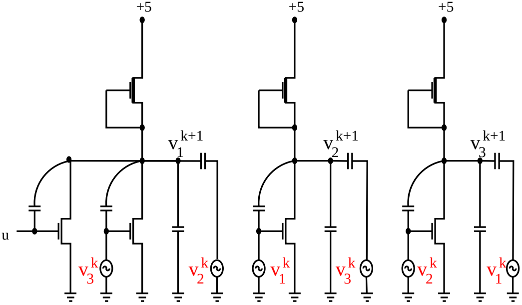
MOS ring oscillator from 1982



Using Kirchhoff's and Ohm's laws gives system of ODEs:

$$\begin{aligned} \frac{\partial \mathbf{v}}{\partial t} &= f(\mathbf{v}), \qquad 0 < t < T \\ \mathbf{v}(0) &= \mathbf{g} \end{aligned}$$

Waveform Relaxation Decomposition



Iteration using sub-circuit solutions only:

$$\begin{aligned}
\partial_t v_1^{k+1} &= f_1(v_1^{k+1}, v_2^k, v_3^k) \\
\partial_t v_2^{k+1} &= f_2(v_1^k, v_2^{k+1}, v_3^k) \\
\partial_t v_3^{k+1} &= f_3(v_1^k, v_2^k, v_3^{k+1})
\end{aligned}$$

Signals along cables are called 'waveforms', which gave the algorithm its name: **Waveform Relaxation**.

Schwarz Waveform Relaxation for PDEs

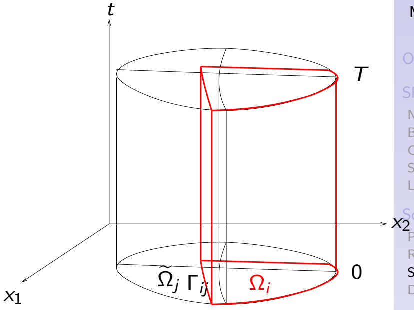
For a given evolution PDE,

$$\partial_t u = \mathcal{L}u + f, \quad \text{in } \Omega \times (0, T),$$

with initial condition

$$u(x, 0) = u_0,$$

the Schwarz waveform relaxation algorithm is:



$$\begin{array}{rcll} \partial_t u_i^k & = & \mathcal{L}u_i^k + f & \text{in } \Omega_i \times (0, T), \\ u_i^k(\cdot,\cdot,0) & = & u_0 & \text{in } \Omega_i, \\ u_i^k & = & u_j^{k-1} & \text{on } \Gamma_{ij} \times (0, T) \end{array}$$

- Many convergence results: heat equation, wave equation, advection reaction diffusion, Maxwell

Schwarz Waveform Relaxation for PDEs

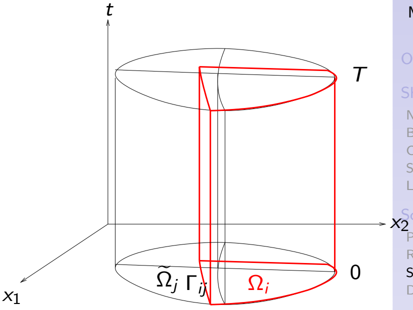
For a given evolution PDE,

$$\partial_t u = \mathcal{L}u + f, \quad \text{in } \Omega \times (0, T),$$

with initial condition

$$u(x, 0) = u_0,$$

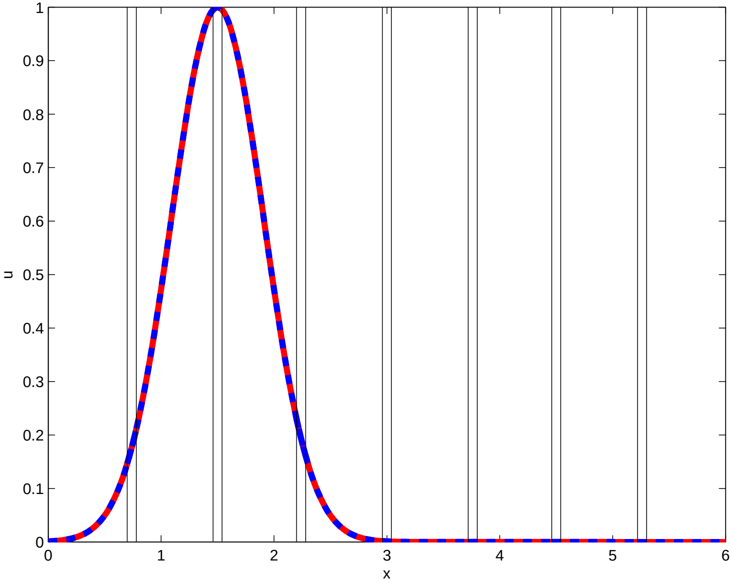
the Schwarz waveform relaxation algorithm is:



$$\begin{array}{rcll} \partial_t u_i^k & = & \mathcal{L}u_i^k + f & \text{in } \Omega_i \times (0, T), \\ u_i^k(\cdot, \cdot, 0) & = & u_0 & \text{in } \Omega_i, \\ \mathcal{B}_{ij} u_i^k & = & \mathcal{B}_{ij} u_j^{k-1} & \text{on } \Gamma_{ij} \times (0, T) \end{array}$$

- Many convergence results: heat equation, wave equation, advection reaction diffusion, Maxwell
- Need to use optimized transmission conditions

An Example with 8 Subdomains



Overview

Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary

Dirichlet-Neumann and Neumann-Neumann WR

Kwok 2014: Neumann-Neumann Waveform Relaxation for the Time-Dependent Heat Equation, DD22

> *"The goal of this paper is to introduce and analyze a new variant of waveform relaxation (WR) methods based on Neumann-Neumann iterations"*

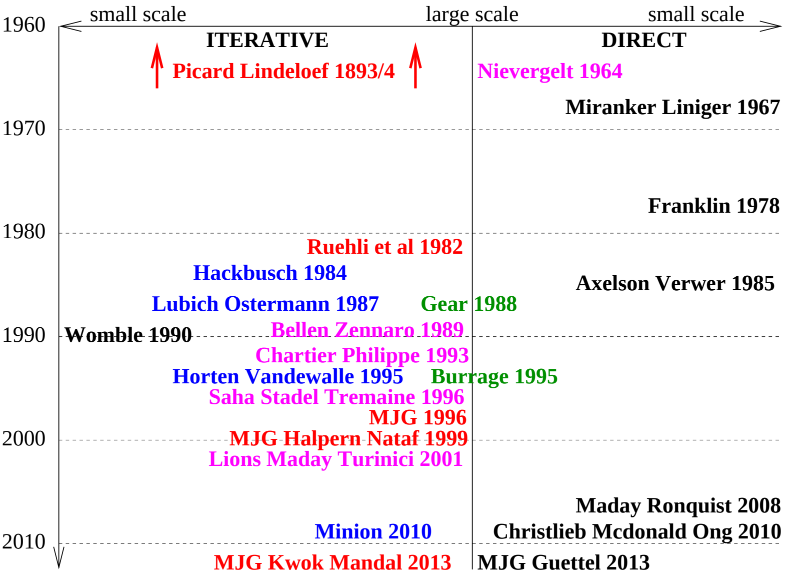
- Apply the Neumann-Neumann algorithm in space-time
- **Theorem:** superlinear convergence on bounded time intervals

Mandal 2014: A Time-Dependent Dirichlet-Neumann Method for the Heat Equation, DD22

> *"We present a waveform relaxation version of the Dirichlet-Neumann method for parabolic problem"*

- Apply the Dirichlet-Neumann algorithm in space-time
- **Theorem:** superlinear convergence on bounded time intervals

Methods Based on Multigrid



Overview
Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici
Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR
Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller
Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp
Summary

Hackbusch 1984

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

Parabolic PDE $u_t + \mathcal{L}u = f$ discretized by Backward Euler:

$$Au_n := (I + \Delta t \mathcal{L})u_n = u_{n-1} + \Delta t f(t_n)$$

Split $A = L + D + U$: smoothing over many time levels

$$\begin{array}{l} \text{for } n = 1 : N \\ \quad \text{for } j = 1 : \nu \\ \quad\quad u_n^j = (L+D)^{-1}(-Uu_n^{j-1} + u_{n-1}^{\nu} + \Delta t f(t_n)) \\ \quad \text{end;} \\ \text{end} \end{array}$$

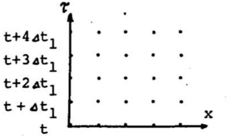


Fig 2.1a: Grid at level l

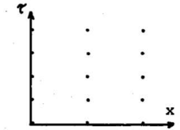


Fig 2.1b: Grid at level l-1

Lubich and Ostermann 1987

Multi-grid Dynamic Iteration for Parabolic Problems. BIT 27.

> *"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:

$$u_t + \mathcal{L}u = f \implies A(s)\hat{u} := s\hat{u} + \mathcal{L}\hat{u} = \hat{f}$$

Let $A(s) = L + D + sI + U$: Two-grid cycle for $A(s)\hat{u} = \hat{f}$:

for $j = 1 : \nu$
  $\hat{u}^j(s) := (L + D + sI)^{-1}(-U\hat{u}^{j-1}(s) + \hat{f}(s))$
end;
$\hat{u}^0_{new}(s) := \hat{u}^\nu(s) + EA_c^{-1}R(\hat{f} - A\hat{u}^\nu(s))$
smooth again

Overview
Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici
Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR
Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller
Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp
Summary

Algorithm in the Time Domain

Smoothing step

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

becomes in the time domain

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

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

The coarse correction

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

becomes

$$\begin{aligned} &\text{solve } v_t + \mathcal{L}_H v = R(f - \partial_t u^\nu - \mathcal{L}u^\nu) \\ &u^0_{new} = u^\nu + Ev \end{aligned}$$

time continuous parabolic problem on coarse spatial mesh

Horton Vandewalle 1995

A Space-Time Multigrid Method for Parabolic Partial Differential Equations. SISC, Vol. 16, No. 4

> *"The method presented in this paper treats the whole of the space-time problem simultaneously."*

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

New ingredients:

1. Adaptive semi-coarsening in space *or* time
2. Prolongation operators only forward in time

Results based on Fourier analysis: good contraction rates for V-cycles, but only mesh independent for F-cycles

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."*

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."*

Overview
Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici
Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR
Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller
Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp
Summary

G and Neumüller 2013

Use block Jacobi smoother for the 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, Steinbach (2013))

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

Overview

Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary

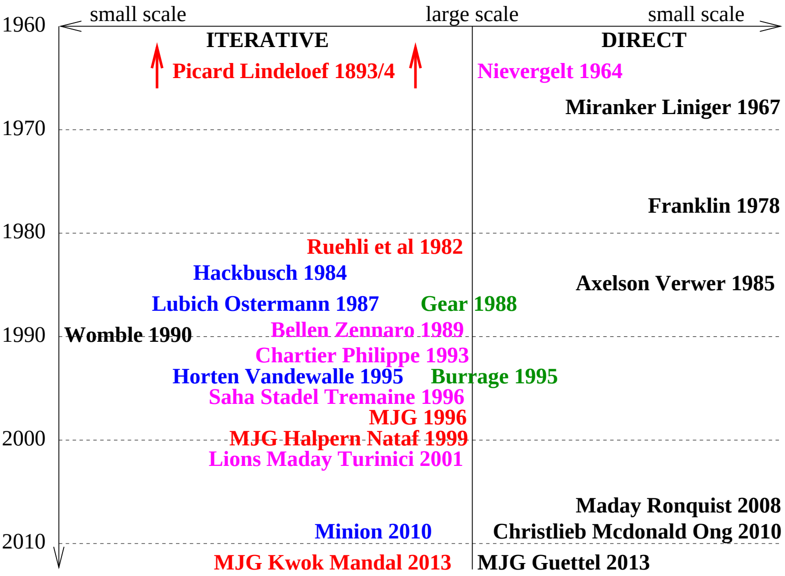
3D Heat Equation Parallelization Results

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)

Direct Time Parallel Methods



Overview

Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

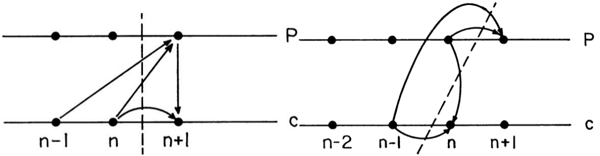
Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary

Miranker Liniger 1967

Parallel Methods for the Numerical Integration of Ordinary Differential Equations. Math. Comp., Vol 21.

"Let us consider how we might widen the computation front."



For $y' = f(x, y)$, consider the predictor corrector formulas

$$y^p_{n+1} = y^c_n + \frac{h}{2}(f(y^c_n) - f(y^c_{n-1})), \; y^c_{n+1} = y^c_n + \frac{h}{2}(f(y^p_{n+1}) + f(y^c_n)).$$

This process is sequential. Consider the modified formulas

$$y^p_{n+1} = y^c_{n-1} + 2hf(y^p_n), \quad y^c_n = y^c_{n-1} + \frac{h}{2}(f(y^p_n) + f(y^c_{n-1})).$$

Those two can be evaluated in parallel.

Results: Methods for $2s$ processors with stability and convergence analysis.

Overview
Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici
Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR
Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller
Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp
Summary

Axelson Verwer 1985

Boundary Value Techniques for Initial Value Problems in Ordinary Differential Equations, Math. of Comp, Vol. 45, No. 171

> *"We will consider $\dot{y} = f(x, y)$ as a two point boundary value problem..."*

Discretize $\dot{y} = ay$ with the explicit midpoint rule

$$y_{n+1} - y_{n-1} - 2hay_n = 0, \quad y_0 = y(0), \quad y_1 = ?$$

BE on the left, $y_1 - y_0 - hay_1 = 0$, gives

$$\begin{pmatrix} 1-ah & & & & \\ -2ah & 1 & & & \\ -1 & -2ah & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & -2ah & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} y_0 \\ y_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Overview

Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary

Axelson Verwer 1985

Boundary Value Techniques for Initial Value Problems in Ordinary Differential Equations, Math. of Comp, Vol. 45, No. 171

> *"We will consider $\dot{y} = f(x, y)$ as a two point boundary value problem..."*

Discretize $\dot{y} = ay$ with the explicit midpoint rule

$$y_{n+1} - y_{n-1} - 2hay_n = 0, \quad y_0 = y(0), \quad y_1 = ?$$

BE on the right, $y_N - y_{N-1} - hay_N = 0$ gives

$$\begin{pmatrix} -2ah & 1 & & & \\ -1 & -2ah & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & -2ah & 1 \\ & & & -1 & 1-ah \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} y_0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Overview

Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary

Axelson Verwer 1985

Boundary Value Techniques for Initial Value Problems in Ordinary Differential Equations, Math. of Comp, Vol. 45, No. 171

> *"We will consider $\dot{y} = f(x, y)$ as a two point boundary value problem..."*

Discretize $\dot{y} = ay$ with the explicit midpoint rule

$$y_{n+1} - y_{n-1} - 2hay_n = 0, \quad y_0 = y(0), \quad y_1 = ?$$

BE on the right, $y_N - y_{N-1} - hay_N = 0$ gives

$$\begin{pmatrix} -2ah & 1 & & & \\ -1 & -2ah & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & -2ah & 1 \\ & & & -1 & 1-ah \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} y_0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

This example has already been proposed by Fox in 1954

Overview

Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

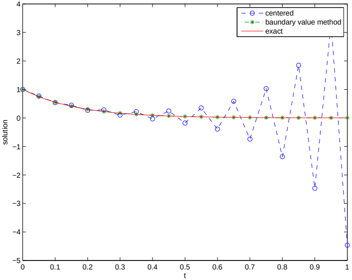
Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

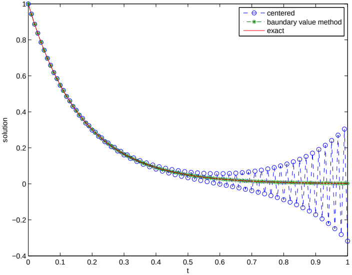
Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary

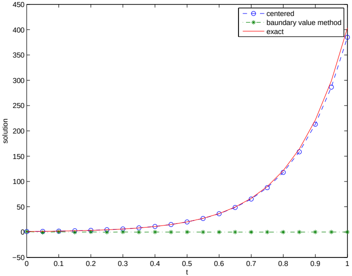
Stability of Boundary Value Methods



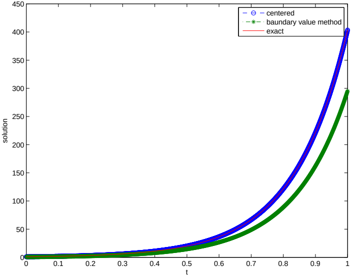
$a = -6,\ h = 1/20$



$a = -6,\ h = 1/100$



$a = 6,\ h = 1/20$



$a = 6,\ h = 1/2000$

Womble 1990

A Time Stepping Algorithm for Parallel Computers, SANDIA REPORT, February 1990

> *"The alternative approach of solving on all time levels simultaneously is not considered practical [9]"*

$$u_t = \mathcal{L}u + f$$

Implicit time discretization and time stepping:

$$A_n u_n = f_n + B_n u_{n-1}.$$

Solution by iteration: $A_n = L_n + D_n + U_n$: for each time level t_n solve for $k = 1, 2, \ldots, K$

$$(L_n + D_n)u_n^k = -U_n u_n^{k-1} + f_n + B_n u_{n-1}^K.$$

Womble 1990

A Time Stepping Algorithm for Parallel Computers, SANDIA REPORT, February 1990

> *"The alternative approach of solving on all time levels simultaneously is not considered practical [9]"*

$$u_t = \mathcal{L}u + f$$

Implicit time discretization and time stepping:

$$A_n u_n = f_n + B_n u_{n-1}.$$

Solution by iteration: $A_n = L_n + D_n + U_n$: for each time level t_n solve for $k = 1, 2, \dots, K$

$$(L_n + D_n)u_n^k = -U_n u_n^{k-1} + f_n + B_n u_{n-1}^{K}.$$

Doing this simultaneously over many time steps $t_1, t_2, \dots, t_N$:

$$(L_n + D_n)u_n^k = -U_n u_n^{k-1} + f_n + B_n u_{n-1}^{k-1}.$$

Results of Womble

Numerical experiment for $u_t = u_{xx} + 3$ on $(0, 1)$, $u(x, 0) = 0$, $u(0, t) = u(1, t) = 3t$, $t \in (0, 5)$ discretized by centered FD and BE, $\Delta x = 1/64$, $\Delta t = 5/200$, using SSOR with $\omega = 1.8$:

P	$\sum_{n=1}^{200} E_n$		$S(P)$		$SS(P)$	
1	12,000	(12,200)	1.00	(1.00)	1.00	(1.00)
2	6,118	(6,121)	1.96	(1.99)	2.00	(2.00)
4	3,168	(3,201)	3.79	(3.81)	3.90	(3.87)
8	1,847	(1,915)	6.49	(6.37)	6.53	(6.78)
16	1,234	(1,319)	9.72	(9.25)	9.50	(10.8)
32	872	(1,137)	13.7	(10.7)	13.9	(15.2)
64	634	(1,137)	18.7	(10.7)	19.1	(20.0)
128	489	(1,137)	24.5	(10.7)	27.6	(23.6)

$E_n = I_n - I_{n-1} + d_n$, I_n iterations needed at level t_n to converge, d_n delay at level t_n before iteration starts, e.g. $d_n = 1$. $S(P)$ speedup, $SS(P)$ "steady state" speedup if all I_n are equal

Maday and Ronquist 2008

Parallelization in time through tensor-product space-time solvers, CRAS, Vol. 346, No. 1.

> *"Pour briser la nature intrinsèquement séquentielle de cette résolution, on utilise l'algorithme de produit tensoriel rapide."*

Suppose we discretize $u_t = Lu$ using Backward Euler:

$$\begin{pmatrix} \frac{1}{\Delta t_1} - L & & & \\ -\frac{1}{\Delta t_2} & \frac{1}{\Delta t_2} - L & & \\ & \ddots & \ddots & \\ & & -\frac{1}{\Delta t_N} & \frac{1}{\Delta t_N} - L \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{pmatrix} = \begin{pmatrix} f_1 + \frac{1}{\Delta t_1} u_0 \\ f_2 \\ \vdots \\ f_N \end{pmatrix}$$

Using the Kronecker symbol, we get $(B \otimes I_x - I_t \otimes L)\mathbf{u} = \mathbf{f}$,

$$B := \begin{pmatrix} \frac{1}{\Delta t_1} & & & \\ -\frac{1}{\Delta t_2} & \frac{1}{\Delta t_2} & & \\ & \ddots & \ddots & \\ & & -\frac{1}{\Delta t_N} & \frac{1}{\Delta t_N} \end{pmatrix}$$

If B is diagonalizable, $B = SDS^{-1}$

$$(B \otimes I_x - I_t \otimes L)\mathbf{u} = \mathbf{f} \quad \Longleftrightarrow \quad (S \otimes I_x)(\mathrm{diag}(D - L))(S^{-1} \otimes I_x)\mathbf{u} = \mathbf{f}$$

and we can hence solve in 3 steps:

$$\begin{array}{lrcl} (a) & (S \otimes I_x)\mathbf{g} & = & \mathbf{f}, \\ (b) & (\frac{1}{\Delta t_n} - L)\mathbf{w}^n & = & \mathbf{g}^n, \quad 1 \leq n \leq N, \\ (c) & (S^{-1} \otimes I_x)\mathbf{u} & = & \mathbf{w}. \end{array}$$

Results by Maday and Ronquist: 1d heat equation:

- close to perfect speedup
- use geometric time mesh $\Delta t_k = \rho^{k-1}\Delta t_1$, with $\rho = 1.2$
- "choosing ρ much closer to 1 may lead to instabilities"

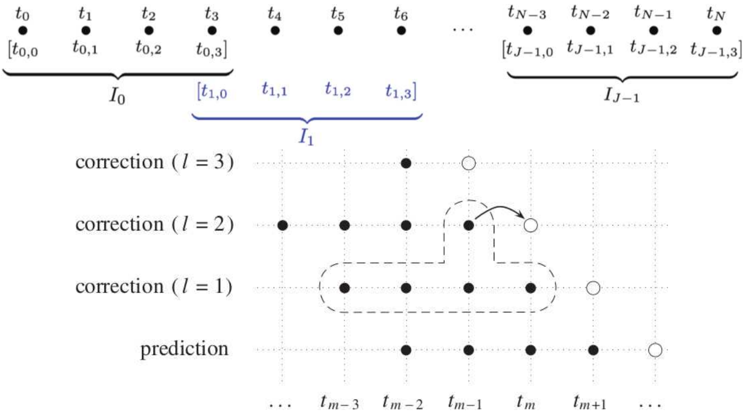
See also **A direct solver for time parallelization** G., Halpern, Ryan and Tran, DD22

- Truncation error analysis for geometric time grid
- Round-off error analysis due to diagonalization
- Error estimates based on trade-off between the two

Christlieb Macdonald Ong 2010

Parallel High-Order Integrators, SISC, Vol. 32, No. 2.

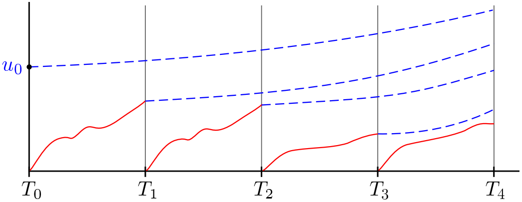
> *"...we discuss a class of defect correction methods which is easily adapted to create parallel time integrators for multicore architectures"*



G and Güttel 2012

For linear problems $\mathbf{u}'(t) = A\mathbf{u}(t) + \mathbf{g}(t), \quad \mathbf{u}(0) = \mathbf{u}_0$

ParaExp: use overlapping decomposition



Solve first non-overlapping inhomogeneous problems

$$\mathbf{v}_j'(t) = A\mathbf{v}_j(t) + \mathbf{g}(t), \quad \mathbf{v}_j(T_{j-1}) = \mathbf{0}, \quad t \in [T_{j-1}, T_j],$$

and then overlapping homogeneous problems

$$\mathbf{w}_j'(t) = A\mathbf{w}_j(t), \quad \mathbf{w}_j(T_{j-1}) = \mathbf{v}_{j-1}(T_{j-1}), \quad t \in [T_{j-1}, T]$$

The solution is then obtained by summation:

$$\mathbf{u}(t) = \mathbf{v}_k(t) + \sum_{j=1}^{k} \mathbf{w}_j(t) \quad \text{with } k \text{ such that} \quad t \in [T_{k-1}, T_k].$$

Overview

Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary

Wave Equation Experiment

$$\partial_{tt} u(t,x) = \alpha^2 \partial_{xx} u(t,x) + \text{hat}(x)\sin(2\pi f t) \qquad x,t \in (0,1)$$
$$u(t,0) = u(t,1) = u(0,x) = u'(0,x) = 0$$

α^2	f	serial τ_0	serial error	parallel $\max(\tau_1)$	parallel $\max(\tau_2)$	parallel error	effi-ciency
0.1	1	2.54e−01	3.64e−04	4.04e−02	1.48e−02	2.64e−04	58 %
0.1	5	1.20e+00	1.31e−04	1.99e−01	1.39e−02	1.47e−04	71 %
0.1	25	6.03e+00	4.70e−05	9.83e−01	1.38e−02	7.61e−05	76 %
1	1	7.30e−01	1.56e−04	1.19e−01	2.70e−02	1.02e−04	63 %
1	5	1.21e+00	4.09e−04	1.97e−01	2.70e−02	3.33e−04	68 %
1	25	6.08e+00	1.76e−04	9.85e−01	2.68e−02	1.15e−04	75 %
10	1	2.34e+00	6.12e−05	3.75e−01	6.31e−02	2.57e−05	67 %
10	5	2.31e+00	4.27e−04	3.73e−01	6.29e−02	2.40e−04	66 %
10	25	6.09e+00	4.98e−04	9.82e−01	6.22e−02	3.01e−04	73 %

$\Delta x = \frac{1}{101}$, $\Delta t_0 = \min\{5 \cdot 10^{-4}/\alpha, 1.5 \cdot 10^{-3}/f\}$, RK45 and Chebyshev exponential integrator, 8 processors

Summary

Four Main Time Parallelization Strategies:

- Multiple Shooting Methods (parareal)
- Schwarz, Dirichlet-Neumann and Neumann-Neumann Waveform Relaxation (also for modeling and local adaptivity)
- Space-Time Multigrid Methods (Need semi-coarsening and special prolongation, or Block smoothers, PFASST)
- Direct Time Parallel Solvers (Tensor methods, RIDC, ParaExp)

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

Overview

Shooting
Nievergelt
Bellen Zennaro
Chartier Philippe
Saha et al
Lions, Maday, Turinici

Schwarz WR
Picard Lindelöf
Ruehli et al
Schwarz WR
DN and NN WR

Multigrid
Hackbusch
Lubich Ostermann
Horton Vandewalle
Minion
G Neumüller

Direct Solvers
Miranker Liniger
Axelson Verwer
Womble
Maday Ronquist
RIDC
ParaExp

Summary