Multigrid Methods for Time-Dependent PDEs

Jan Van lent[1]

Katholieke Universteit Leuven
Department of Computer Science

Friday 8 July 2005
BICS, Bath

[1] Research Assistant of the Fund for Scientific Research - Flanders (Belgium) (F.W.O.-Vlaanderen)

Introduction

- parabolic PDEs, model problem: heat equation

$$u_t = u_{xx} + u_{yy} + f$$

- both ODE and PDE aspects
  - PDE perspective: discretise using implicit Euler

    $$u_i - \Delta t L u_i = u_{i-1} + \Delta t f_i$$

    same structure as elliptic PDE
  - ODE perspective: discretise space

    $$\dot{u} = Lu + f$$

    many sophisticated time discretisation schemes
    large, sparse, structured systems
- How to combine best of both worlds?

1. Iterative Methods for Time-Dependent PDEs
   - Classical Iterative Methods
   - Time Stepping
   - Waveform Relaxation
   - Convergence Analysis
2. Multigrid
3. Time Discretisation Schemes
4. Results
   - Heat Equation
   - Varying Coefficients

Classical Iterative Methods

- system $Ax = b$
- splitting $A = A^+ + A^-$
- iteration $A^+ x^{(\nu)} + A^- x^{(\nu-1)} = b$
- A^+ such that simple to solve
  - Jacobi: diagonal of A
  - Gauss-Seidel: lower triangular part of A
- iteration matrix $K = -(A^+)^{-1} A^-$
- spectral radius

$$\rho(K) = \max_{\lambda \in \sigma(K)} |\lambda|$$

- $\rho < 1 \Rightarrow$ convergence, the smaller the better

Time Stepping

- system of ODEs

$$\dot{x} = Ax + b$$

- simple time discretisation scheme: implicit Euler, BDF

$$x_i - \Delta t A x_i = x_{i-1} + \Delta t b$$

- in each time step: system with matrix

$$I - \Delta t A$$

- use same iterative methods as before

Continuous Waveform Relaxation

- iterative method for system of ODEs

$$\dot{x} = Ax + b$$

- same splitting $A = A^+ + A^-$

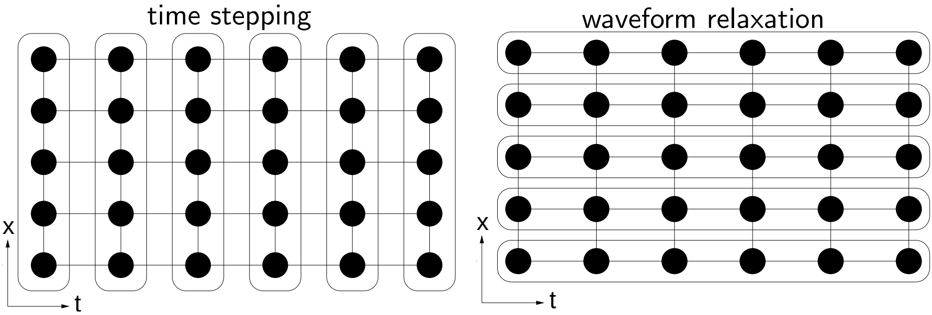
$$\dot{x}^{(\nu)} = A^+ x^{(\nu)} + A^- x^{(\nu-1)} + b$$

- e.g. Jacobi

$$\dot{x}_i^{(\nu)} + a_{ii} x_i^{(\nu)} = b_i - \sum_{j \neq i} a_{ij} x_j^{(\nu-1)}$$

Discrete Waveform Relaxation

- in practice: ODE integrator for each scalar ODE
- many options: LMF, IRK, GLM, BVM, BBVM, CSC
- same equations, but unknowns updated in different order



Convergence Analysis

- system of equations

$$(T \otimes I)x = (I \otimes A)x + b$$

- T : time derivative operator $w = Tv$
  - continuous $\quad w(t) = \frac{dv(t)}{dt}$
  - BDF1 $\quad w_i = \frac{v_i - v_{i-1}}{\Delta t}$
- $A = A^+ + A^- \longrightarrow$ iteration

$$(T \otimes I)x^{(\nu)} = (I \otimes A^+)x^{(\nu)} + (I \otimes A^-)x^{(\nu-1)} + b$$

- iteration operator

$$\mathcal{K} = (T \otimes I - I \otimes A^+)^{-1}(I \otimes A^-)$$

- spectral radius

$$\rho(\mathcal{K}) = \max_{z\in\Sigma} \rho(K(z)), \qquad K(z) = (zI - A^+)^{-1}A^-$$

- $\Sigma \subset \mathbb{C} \cup \{\infty\}$
- continuous WR on $[0, T]$: $\Sigma = \{\infty\} \rightarrow \rho = 0$
- $\mathcal{K}$ non-normal operator $\rightarrow$ infinite time domains (alternatives: weighted norms, pseudospectra)
- continuous WR on $[0, \infty)$: $\Sigma = \mathbb{C}^+$
- time stepping $\subset$ discrete WR

Functional Calculus

- traditional analysis based on theory of Volterra convolution equations
- matrix-valued functional calculus gives a meaning to $K(T)$

$$\mathcal{K} = K(T)$$

- spectral mapping theorem

$$\sigma(K(T)) = \bigcup_{z \in \sigma(T)} \sigma(K(z))$$

- spectral radius

$$\rho(K(T)) = \max_{\sigma(T)} \rho(K(z))$$

- which means

$$\Sigma = \sigma(T)$$

Multigrid Principle

- convergence of Jacobi and Gauss-Seidel very slow
- for elliptic PDEs: multigrid
- idea: use calculations on a coarse grid (cheaper) to accelerate iteration on fine grid
  - given approximation y to solution of $Ax = b$
  - write $x = y + e$
  - correction e is solution of $Ae = b - Ay = d$
  - solve this on coarse grid
- can be extended to time dependent PDEs

Two-Grid Algorithm

$\mathrm{mg}(\mathbf{x}^{(0)}, \mathbf{A}, \mathbf{b}) \longrightarrow \mathbf{x}^{(3)}$

- $\mathbf{x}^{(1)} \leftarrow \mathrm{smooth}(\mathbf{x}^{(0)}, \mathbf{A}, \mathbf{b}, \mu_1)$
- $\bar{\mathbf{b}} \leftarrow \mathbf{R}(\mathbf{b} - \mathbf{A}\mathbf{x}^{(1)})$
- $\bar{\mathbf{x}} \leftarrow \mathrm{cgc}(\bar{\mathbf{A}}, \bar{\mathbf{b}})$
- $\mathbf{x}^{(2)} \leftarrow \mathbf{x}^{(1)} + \mathbf{P}\bar{\mathbf{x}}$
- $\mathbf{x}^{(3)} \leftarrow \mathrm{smooth}(\mathbf{x}^{(2)}, \mathbf{A}, \mathbf{b}, \mu_2)$

$\mathrm{cgc}(\bar{\mathbf{A}}, \bar{\mathbf{b}}) \longrightarrow \bar{\mathbf{x}}$

- $\bar{\mathbf{x}} \leftarrow \bar{\mathbf{A}}^{-1}\bar{\mathbf{b}}$

$\mathrm{smooth}(\mathbf{x}^{(0)}, \mathbf{A}, \mathbf{b}, \mu) \longrightarrow \mathbf{x}^{(\mu)}$

- for $\nu = 1, \ldots, \mu$
  solve
  $\mathbf{A}^{+}\mathbf{x}^{(\nu)} = \mathbf{b} - \mathbf{A}^{-}\mathbf{x}^{(\nu-1)}$

Multigrid Algorithm

$\mathrm{mg}(\mathbf{x}^{(0)}, \mathbf{A}, \mathbf{b}) \to \mathbf{x}^{(3)}$

- $\mathbf{x}^{(1)} \leftarrow \mathrm{smooth}(\mathbf{x}^{(0)}, \mathbf{A}, \mathbf{b}, \mu_1)$
- $\bar{\mathbf{b}} \leftarrow \mathbf{R}(\mathbf{b} - \mathbf{A}\mathbf{x}^{(1)})$
- $\bar{\mathbf{x}} \leftarrow \mathrm{mg}(0, \bar{\mathbf{A}}, \bar{\mathbf{b}})$
- $\mathbf{x}^{(2)} \leftarrow \mathbf{x}^{(1)} + \mathbf{P}\bar{\mathbf{x}}$
- $\mathbf{x}^{(3)} \leftarrow \mathrm{smooth}(\mathbf{x}^{(2)}, \mathbf{A}, \mathbf{b}, \mu_2)$

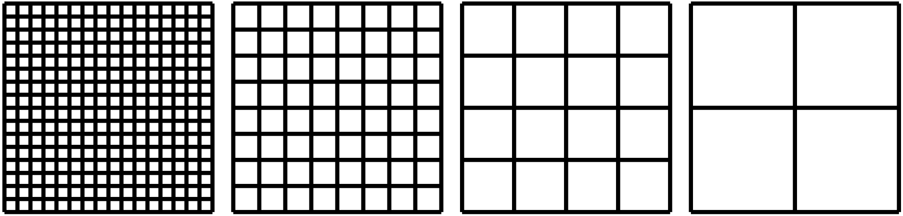
$\mathrm{smooth}(\mathbf{x}^{(0)}, \mathbf{A}, \mathbf{b}, \mu) \to \mathbf{x}^{(\mu)}$

- for $\nu = 1, \ldots, \mu$
  solve
  $\mathbf{A}^{+}\mathbf{x}^{(\nu)} = \mathbf{b} - \mathbf{A}^{-}\mathbf{x}^{(\nu-1)}$

Multigrid Iteration

Many variations

- ν_1 pre- and ν_2 post-smoothing steps
- V-cycle: 1 recursive call
- W-cycle: 2 recursive calls



Multigrid for PDEs

- system of ODEs

$$\dot{u} = Lu + f$$

- continous MG WR: apply multigrid framework with

$$\mathbf{A}^{+} = \frac{\partial}{\partial t} - L^{+}, \qquad \mathbf{A}^{-} = -L^{-}$$

- smoother: continuous WR

$$\dot{u}^{(\nu)} = L^{+} u^{(\nu)} + L^{-} u^{(\nu-1)} + f$$

- restriction and prolongation: linear combinations of functions (instead of scalars)
- discrete MG WR: apply time discretization scheme

Convergence Analysis

- two-grid symbol

$$M(z) = K^{\nu_2}(z)C(z)K^{\nu_1}(z),$$
$$K(z) = (zI - L^+)^{-1}L^-$$
$$C(z) = I - P(zI - \overline{L})^{-1}R(zI - L)$$

- spectral radius of the iteration operator $\mathcal{M}$

$$\rho(\mathcal{M}) = \max_{z \in \Sigma} \rho(M(z))$$

- iteration operator $M(z)$ for MG applied to

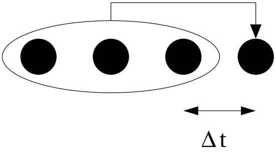
$$zu = Lu + f$$

- each $\rho(M(z))$ by standard two-grid Fourier analysis

Linear Multistep Formula

$$v'(t) = f(v(t))$$

$$v_i \approx v(i\Delta t)$$



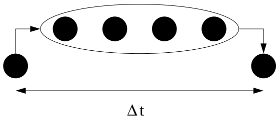
- k-step LMF
- update approximation at point i
  using k previous points $i-k, \ldots, i-1$

$$\sum_{j=-k}^{0} \alpha_{k+j} v_{i+j} = \Delta t \sum_{j=-k}^{0} \beta_{k+j} f(v_{i+j})$$

Implicit Runge-Kutta Method

$$v'(t) = f(v(t))$$

$$v_i \approx v(i\Delta t)$$

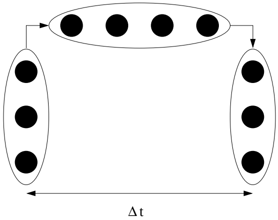


- s stage IRK
- update approximation at point i using s stage values W_i

$$v_i = v_{i-1} + \Delta t b^T f(W_i)$$

$$W_i = v_{i-1} + \Delta t A f(W_i)$$

General Linear Method



$$v'(t) = f(v(t))$$

$$V_i \approx v(i\Delta t),\ v((i-1)\Delta t),\ v'(i\Delta t),\ etc.$$

- update r values V_i
  using s stage values W_i

$$V_i = DV_{i-1} + \Delta t B f(W_i)$$
$$W_i = CV_{i-1} + \Delta t A f(W_i)$$

- IRK: $r = 1,\ B = b^T,\ C = [1, \ldots, 1]^T,\ D = 1$

Remarks

- other methods
  - boundary value methods (BVMs): GBDF, GAM
  - block boundary value methods (BBVMs)
  - Chebyshev spectral collocation (CSC)
- each method has its specific advantages and disadvantages
  - LMF: efficient, but only A-stable for low orders
  - IRK: well established, high order, stable, construction not straightforward
  - GLM: very general framework, construction not straightforward
  - BVM/BBVM: very general framework, straightforward construction
  - CSC: explicit coefficients, A-stable for any order, fast transforms
- all intimately related

Convergence Analysis

$$\max_{z \in \frac{1}{\Delta t}\Sigma} \rho(M(z))$$

Σ	$[0, T]$	$[0, \infty)$
continuous	∞	$\mathbb{C}^+$
LMF	$\frac{\alpha_k}{\beta_k}$	$\frac{a(w)}{b(w)}$
IRK/GLM	$\sigma(A^{-1})$	$\sigma\left((A + C(wI_r - D)^{-1}B)^{-1}\right)$
	$w = \infty$	$\lvert w\rvert \geq 1$

Stability Domains

- test equation : $v' = \lambda v$
- open stability domain S : all $z = \lambda \Delta t$ for which $v \to 0$
- A-stability : $\mathbb{C}^- \subset S$
- $S : z \in \mathbb{C}$ such that
  - LMF : $a(w) - zb(w) = 0, \quad |w| < 1$
  - IRK/GLM : $\rho\big(D + B(z^{-1}I_s - A)^{-1}C\big) < 1$
- $\Sigma : z \in \mathbb{C}$ such that
  - LMF : $z = \frac{a}{b}(w), |w| \geq 1$
  - IRK/GLM : $z \in \sigma\big((A + C(wI_r - D)^{-1}B)^{-1}\big), |w| \geq 1$

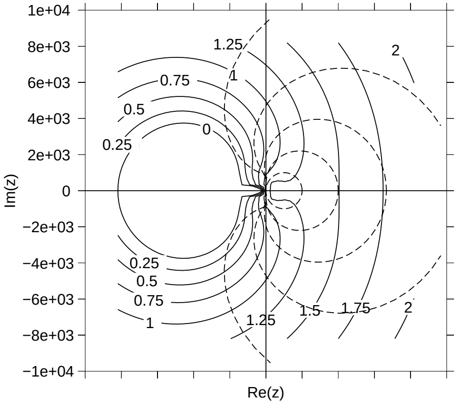
$$\boxed{\Sigma = \mathbb{C} \setminus S}$$

Convergence Analysis

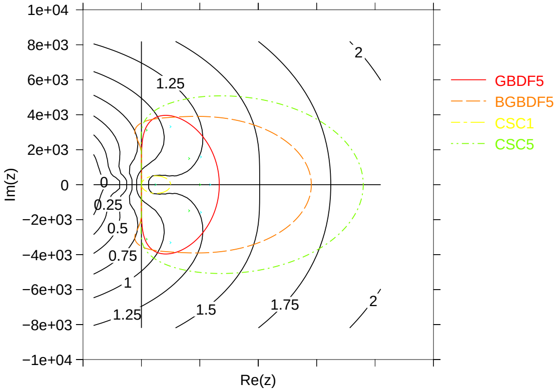
- heat equation

$$u_t = u_{xx} + u_{yy} + f$$

- finite differences, multigrid waveform relaxation $\Delta x = \Delta y = \frac{1}{32}$, $\Delta t = 10^{-3}$
- $\rho(\mathbf{M}(z))$ calculated using standard two-grid Fourier analysis
- contour lines of $R(z) = -\log_{10} \rho(M(z))$
- $\partial\Sigma$ for BDF1–5



Convergence Analysis (cont.)

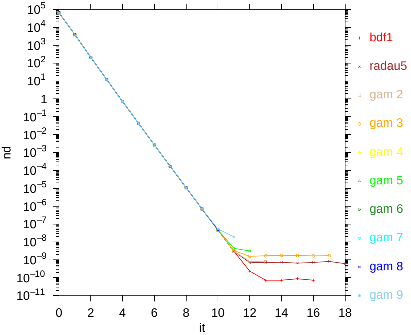


GBDF5	BGBDF5	CSC1	CSC5
0.80	0.74	1.13	0.79

Numerical Results

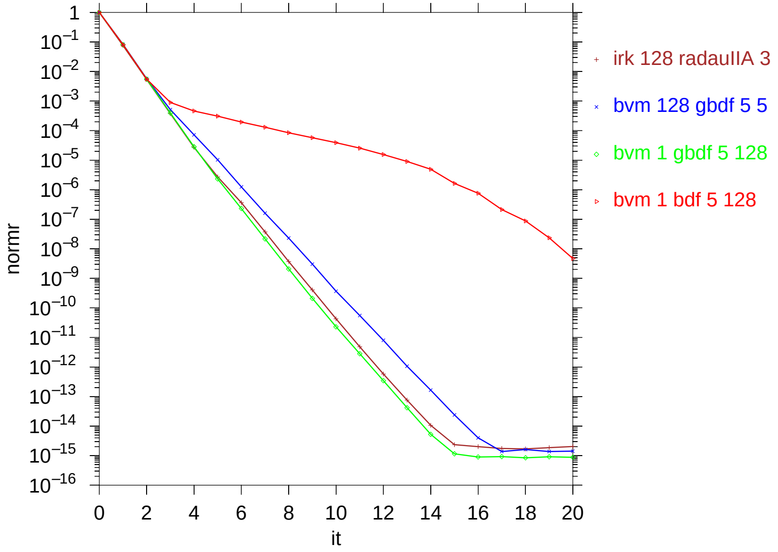
heat equation

- $u_t = u_{xx} + u_{yy} + f$
- $u = 1 + \sin(\frac{\pi}{2}x)\sin(\frac{\pi}{2}y)e^{-\frac{\pi^2}{2}t}$
- $(x, y, t) \in [0, 1]^3$
- $\Delta x = \Delta y = \frac{1}{256}, \Delta t = \frac{1}{128}$
- 8 million unknowns



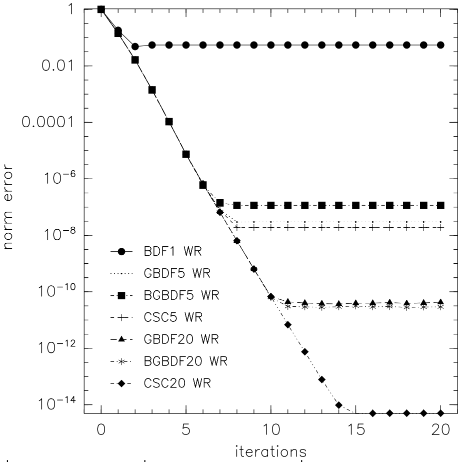
> 1 digit per iteration

Stability Influences Convergence



Numerical Results

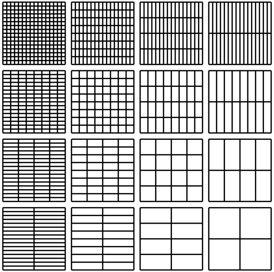
- discretisation error
  - $u(x, y, t) = sin(\frac{2\pi t}{10})$
  - $\Delta x = \Delta y = \frac{1}{8}$
  - time interval and #unknowns fixed
- convergence
  - $u(x, y, t) = 0$
  - $\Delta x = \Delta y = \frac{1}{32}$
  - $\rho^{(\nu)} = \frac{\|u^{(\nu)}\|}{\|u^{(\nu-1)}\|}$
  - $R^{(\nu)} = -\log_{10} \rho^{(\nu)}$



BDF1	GBDF5	BGBDF5	CSC5	GBDF20	BGBDF20	CSC20
0.97	0.87	0.87	0.87	0.11	0.10	0.88

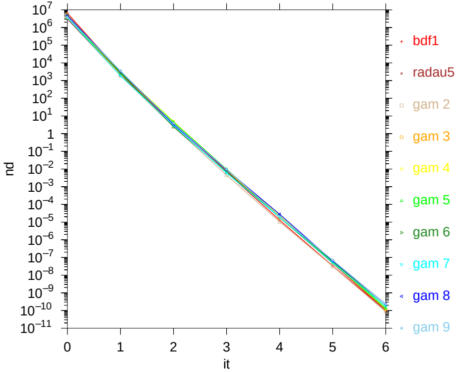
Diffusion Equation with Varying Coefficients

- $u_t = (au_x)_x + (bu_y)_y + f$
- $a = e^{10(x-y)}$, $b = e^{-10(x-y)}$
- anisotropic
- standard MG does not work
- adapt special MG methods to the parabolic case
- "multigrid as smoother" (MG-S)
- a multiple semi-coarsening method
- same simple smoothers as before, but extended hierarchy of coarse grids



Numerical Results

- $u_t = (au_x)_x + (bu_y)_y + f$
- $a = e^{10(x-y)}$, $b = e^{-10(x-y)}$
- $u^{(0)}$ random
- $(x, y, t) \in [0, 1]^3$
- $\Delta x = \Delta y = \frac{1}{32}$, $\Delta t = \frac{1}{128}$
- multiple semicoarsening, MG-S



$\approx$ 3 digit per iteration

Extensions

same principles have been applied to

- time-Periodic PDEs $u_t = \mathcal{L}u + f, \quad u(t) = u(t + T)$
- delay PDEs $u_t = \mathcal{L}u + \alpha u(t - \tau) + f$
- finite element discretisations, irregular grids
- systems of PDE
- algebraic multigrid (AMG)

e.g. FE + IRK + SAMG for
system of 2 non-linear reaction diffusion equations
modelling the O_2 consumption and CO_2 production in a pear

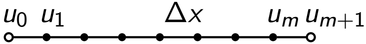
Concluding Remarks

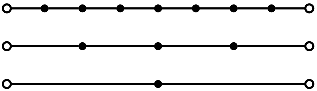
- multigrid methods for elliptic PDEs
  can be extended to parabolic PDEs
- convergence depends on stability of time discretisation
- if continuous WR converges on $[0, \infty)$
  $\longrightarrow$ discrete WR with any A-stable ODE integrator converges
- convergence $\approx$ MG for elliptic case
- complexity $\approx$ MG for elliptic case $\times$ scalar ODE integration

Further Reading

- WR convergence (Miekkala and Nevanlinna '87)
- IRK (Hairer and Wanner; Butcher)
- BVM (Brugnano and Trigiante)
- CSC (Trefethen; Boyd)
- ODE/PDE/WR (Burrage)
- MG/AMG (Briggs, Henson and McCormick; Trottenberg, Oosterlee and Schüller; Stüben)
- MG WR (Lubich and Ostermann '89; Vandewalle)
- MGS (Washio and Oosterlee '95, '98)
- functional calculus (Dunford and Schwartz; Taylor; Haase)

Standard Multigrid Example

- 1D Poisson (elliptic PDE) : $u_{xx} = f$
- finite difference discretisation : 

- system of equations : $u_{i-1} - 2u_i + u_{i+1} = \Delta x^{-2} f_i$
- Gauss-Seidel : $u_i^{(\nu)} = (u_{i-1}^{(\nu)} + u_{i+1}^{(\nu-1)} - \Delta x^{-2} f_i)/2$
- discretisations at different levels : 
- restriction : $\overline{b}_i \leftarrow (r_{2i-1} + 2r_{2i} + r_{2i+1})/4$
- prolongation : $c_{2i} \leftarrow \overline{x}_i$, $\qquad c_{2i+1} \leftarrow (\overline{x}_i + \overline{x}_{i+1})/2$

Functional Calculus

- analytic function $f : \mathbb{C} \rightarrow \mathbb{C}$
- linear operator T in Banach space X
- functional calculus defines $f(T)$ as operator in X

$$f(T) = \sum_{i=0}^{\infty} c_i T^i = \int_C (zI - T)^{-1} f(z) dz$$

- spectral mapping theorem : $\sigma(f(T)) = f(\sigma(T))$
- matrix-valued analytic function $F : \mathbb{C} \rightarrow \mathbb{C}^{m \times m}$
- functional calculus defines $F(T)$ as operator in X^m

$$F(T) = \sum_{i=0}^{\infty} C_i \otimes T^i = \int_C (zI - T)^{-1} \otimes F(z) dz$$

- spectral mapping theorem : $\sigma(F(T)) = \sigma(F(\sigma(T)))$