Functional Calculus for the Spectral Analysis of Iterative Methods

Jan Van lent

BICS (Department of Mathematics)
University of Bath

Wednesday 22 February 2006
Dipartimento di Matematica
Università di Bari

Iterative Methods for Time-Dependent PDEs

- model problem

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

- discretise space and time

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

- iterate $(L = L^+ + L^-)$

$$\frac{u_i^{(\nu)} - u_{i-1}^{(\nu)}}{\Delta t} = L^+ u_i^{(\nu)} + L^- u_i^{(\nu-1)} + f_i$$

Convergence Analysis

- iteration

$$u^{(\nu)} = \mathcal{K} u^{(\nu-1)} + \varphi$$

- spectral radius $\rho(\mathcal{K})$
- convergence if $\rho < 1$, the smaller the better
- for iterative methods for ODEs

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

- matrix-valued function

$$K : \mathbb{C} \to \mathbb{C}^{m \times m} : z \to K(z) = (zI - L^+)^{-1} L^-$$

- complex values

$$\Sigma \subset \mathbb{C} \cup \{\infty\}$$

Convergence Analyses

- many different methods
  - ▶ time stepping, continuous/discrete waveform relaxation
  - ▶ initial value problems on finite/infinite intervals, periodic problems
  - ▶ different iterative methods: Jacobi/Gauss-Seidel, multigrid
- convergence analyses similar
- but different enough to be annoying
- different tools
  - ▶ linear algebra
  - ▶ Volterra convolution operators
  - ▶ Laplace transforms
  - ▶ discrete Volterra convolution operators
  - ▶ discrete Laplace transforms (generating functions)
  - ▶ periodic convolution operators
- convinced that it could fit into one theory

Functional Calculus for Scalar Functions

- function of a matrix
- well known examples: T^2, $\exp(T)$, $\cos(T)$, etc.
- many (equivalent) definitions
- polynomial or series

$$f(z) = c_0 + c_1 z + c_2 z^2 + \ldots$$

- definition 1

$$f(T) = c_0 I + c_1 T + c_2 T^2 + \ldots$$

Calculation Rules

- linearity
$$\alpha f(T) + \beta g(T) = (\alpha f + \beta g)(T)$$

- multiplication
$$f(T) \cdot g(T) = (f \cdot g)(T)$$

- function composition
$$g(f(T)) = (g \circ f)(T)$$

Spectral Mapping Theorem

- eigenvalue decomposition

$$T = V \operatorname{diag} \begin{bmatrix} \lambda_1 & \cdots & \lambda_n \end{bmatrix} V^{-1}$$

- definition 2

$$f(T) = V \operatorname{diag} \begin{bmatrix} f(\lambda_1) & \cdots & f(\lambda_n) \end{bmatrix} V^{-1}$$

- spectral mapping theorem

$$\sigma(f(T)) = \{f(z) : z \in \sigma(T)\} =: f(\sigma(T))$$

- for general matrices Jordan or Schur

Complex Analysis

- Cauchy's integral formula

$$f(t) = \oint \frac{f(z)}{z-t} \tilde{d}z$$

- definition 3

$$f(T) = \oint f(z)(zI - T)^{-1} \tilde{d}z$$

- f holomorphic in a neighborhood of $\sigma(T)$
- works for bounded linear operators
- can be extended to closed linear operator
- $\sigma(T)$ need not be connected

Functional Calculus for Matrix-Valued Functions

- matrix $T \in C^{n \times n}$
- matrix-valued function

$$F : \mathbb{C} \to \mathbb{C}^{m \times p}$$

- F holomorphic in a neighborhood of $\sigma(T)$
- definition $F(T) \in \mathbb{C}^{mn \times pn}$

$$F(T) = \oint F(z) \otimes (zI - T)^{-1} \tilde{d}z$$

- Kronecker or tensor product $A \otimes B = [a_{ij}B]$

Calculation Rules

- linearity

$$\alpha F(T) + \beta G(T) = (\alpha F + \beta G)(T)$$

- multiplication

$$F(T) \cdot G(T) = (F \cdot G)(T)$$

- function composition

$$G(F(T)) = (G \circ F)(T)$$

The Main Theorem

- spectral mapping theorem

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

- corollary

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

- valid for
  - any bounded operator T
  - any matrix-valued function F
    holomorphic in a neighbourhood of $\sigma(T)$

The Main Theorem

$\sigma_\infty(T) = \sigma(T) \cup \{\infty\}$

- spectral mapping theorem

$$\sigma(F(T)) = \bigcup_{z \in \sigma_\infty(T)} \sigma(F(z)) =: \sigma(F(\sigma_\infty(T)))$$

- corollary

$$\rho(F(T)) = \max_{z \in \sigma_\infty(T)} \rho(F(z))$$

- valid for
  - ▶ any closed operator T
  - ▶ any matrix-valued function F
    holomorphic in a neighbourhood of $\sigma(T)$
    and at ∞

Iterative Methods for ODEs

- equation

$$(I \otimes T)u = (L \otimes I)u + f$$

- iteration $(L = L^+ + L^-)$

$$(I \otimes T)u^{(\nu)} = (L^+ \otimes I)u^{(\nu)} + (L^- \otimes I)u^{(\nu-1)} + f$$

- iteration operator

$$F(T) = (I \otimes T - L^+ \otimes I)^{-1}(L^- \otimes I)$$

- matrix valued function

$$F(z) = (zI - L^+)^{-1}L^-$$

- analytic in $\mathbb{C} \setminus \sigma(L^+)$ (and at infinity)
- assume $\sigma(L^+) \cap \sigma(T) = \phi$

Continuous Waveform Relaxation

- equation

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

- together with initial condition

$$u(0) = u_0$$

- or periodicity

$$u(0) = u(t_F)$$

- iteration

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

Finite Time Intervals

- Banach space $X = C[0, t_F]$
- operator $(Tx)(t) = \dot{x}(t)$, with domain

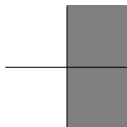$$\mathcal{D}(T) = \{x : \dot{x} \in C[0, t_F], x(0) = 0\}$$

- closed unbounded operator with $\sigma(T) = \phi$
- extended spectrum $\sigma_\infty(T) = \{\infty\}$
- result $\rho(F(T)) = \rho(F(\infty)) = 0$

Infinite Time Intervals

- Banach space $X = L^p([0, \infty], \mathbb{C})$
- operator $(Tx)(t) = \dot{x}(t)$, with domain

  $\mathcal{D}(T) = \{x : x \text{ is absolutely continuous on } [0, a] \text{ for any } a > 0,$
  $\dot{x} \in X, x(0) = 0\}$

- closed unbounded operator with $\sigma(T) = \mathbb{C}^+$
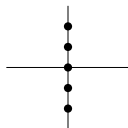- result $\rho(F(T)) = \sup_{z \in \mathbb{C}^+} \rho(F(z))$

Periodic Problems

- Banach space $X = C[0,1]$
- operator $(Tx)(t) = \dot{x}(t)$, with domain

$$\mathcal{D}(T) = \{x : \dot{x} \in C[0,1], x(0) = x(1)\}$$

- closed unbounded operator with $\sigma(T) = 2\pi i \mathbb{Z}$
- result $\rho(F(T)) = \sup_{j \in \mathbb{Z}} \rho(F(2\pi i j))$
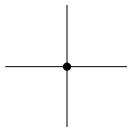
Discrete Waveform Relaxation

- T itself by functional calculus
- implicit Euler $(Tu)_i = \frac{u_i - u_{i-1}}{\Delta t}$
- shift operator $(Sx)_i = x_{i-1}$
- function $g(w) = \frac{1-w}{\Delta t}$
- operator $T = g(S)$

Finite Sequences

- Banach space $X = l^p(n, \mathbb{C})$
- matrix

$$
S = \begin{bmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix}
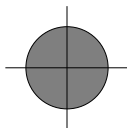$$

- Toeplitz matrix with $\sigma(S) = \{0\}$

Infinite Sequences

- Banach space $X = l^p(\infty, \mathbb{C})$
- operator $[x_1, x_2, \ldots] \to [0, x_1, x_2, \ldots]$

$$S = \begin{bmatrix} 0 & & \\ 1 & \ddots & \\ & \ddots & \end{bmatrix}$$

- Toeplitz operator with $\sigma(S) = \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}$
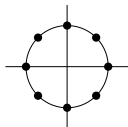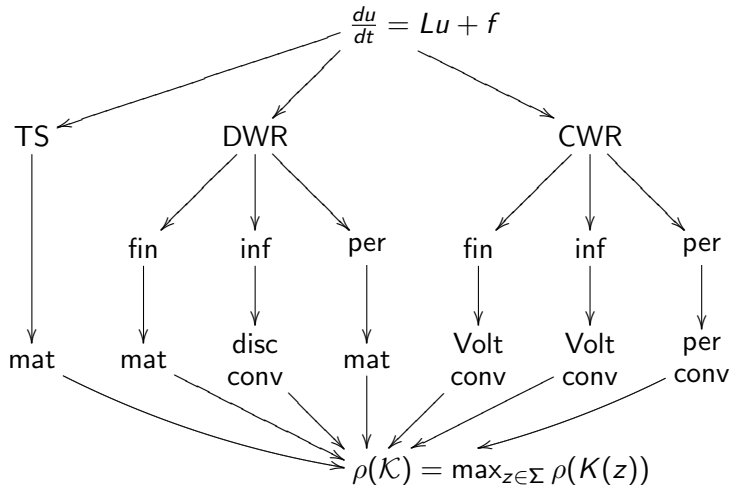
Periodic Sequences

- Banach space $X = l^p(n, \mathbb{C})$
- matrix

$$S = \begin{bmatrix} 0 & \cdots & & 0 & 1 \\ 1 & \ddots & & & 0 \\ & \ddots & \ddots & & \vdots \\ & & & 1 & 0 \end{bmatrix}$$
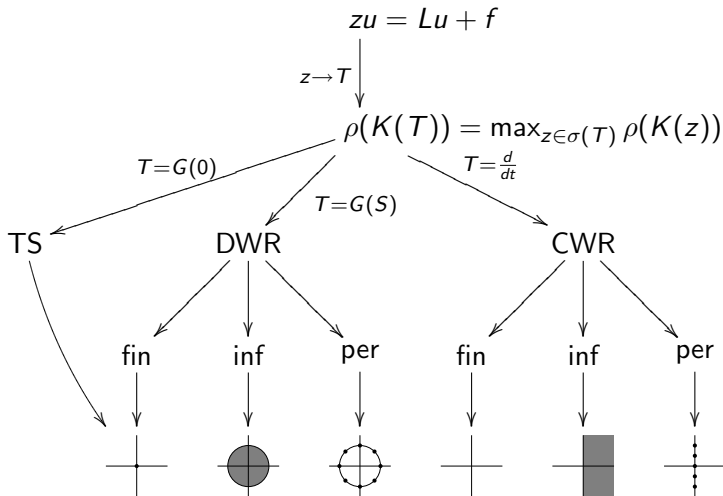
- circulant matrix with $\sigma(S) = \left\{ \exp\left(\frac{2\pi i j}{n} \right), j = 0, \ldots, n-1 \right\}$

Traditional Convergence Analysis Approach

$$\frac{du}{dt} = Lu + f$$

TS   DWR   CWR

fin   inf   per   fin   inf   per

mat   mat   disc conv   mat   Volt conv   Volt conv   per conv

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

Convergence Analysis using Functional Calculus

$$zu = Lu + f$$

$z \to T$

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

$T = G(0)$

$T = G(S)$

$T = \frac{d}{dt}$

TS

DWR

CWR

fin

inf

per

fin

inf

per

Convergence and Stability

- test equation

$$\dot{y} = \lambda y$$

- $y(t) \to 0$ if $\Re(\lambda) < 0$
- time integration: discrete solution $y_k \approx y(k)$
- open stability domain S:
  all λ such that $y_k \to 0$
- T discrete time derivative
- spectrum $\sigma(T)$:
  all λ such that $\lambda I - T$ does not have bounded inverse
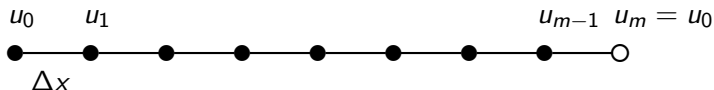- consider $l^p(\infty)$

$$S = \mathbb{C} \setminus \sigma(T)$$

Multigrid Example

- elliptic partial differential equation : 1D Poisson

$$u_{xx} = f, \qquad u, f : [0, 1] \to \mathbb{R}$$

- discretisation : finite differences

$$u_{i-1} - 2u_i + u_{i+1} = \Delta x^{-2} f_i$$
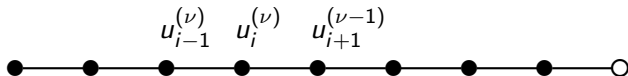
Iterative Solution

- system of equations

$$Lu = f, \qquad u, f \in \mathbb{R}^m$$

- classical iterations

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

- e.g. Jacobi, **Gauss-Seidel**

$$u_i^{(\nu)} = \frac{1}{2}(u_{i-1}^{(\nu)} + u_{i+1}^{(\nu-1)} - \Delta x^{-2} f_i)$$

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 v to solution of $Lu = f$
  - write $u = v + e$
  - correction e is solution of $Le = f - Lv = d$
  - solve this on coarse grid

Multigrid Components

- discretisations at different levels of coarseness
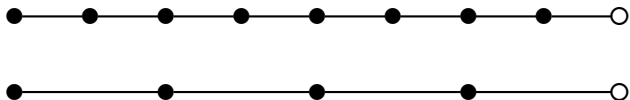
- smoothing : Gauss-Seidel
- restriction : $\bar{f}_i \leftarrow (d_{2i-1} + d_{2i} + d_{2i+1})/4$
- prolongation : $e_{2i} \leftarrow \bar{u}_i, \qquad e_{2i+1} \leftarrow (\bar{u}_i + \bar{u}_{i+1})/2$
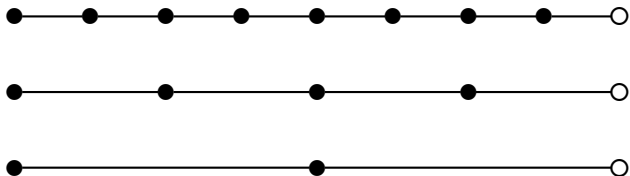
Multigrid Components

- discretisations at different levels of coarseness

- smoothing : Gauss-Seidel
- restriction : $\bar{f}_i \leftarrow (d_{2i-1} + d_{2i} + d_{2i+1})/4$
- prolongation : $e_{2i} \leftarrow \bar{u}_i$, $\qquad e_{2i+1} \leftarrow (\bar{u}_i + \bar{u}_{i+1})/2$

Two-Grid Operator

- multigrid iteration

$$u^{(\nu)} = Mu^{(\nu-1)} + \varphi$$

- two-grid

$$M = S^{\nu_2} C S^{\nu_1}$$
$$S = (L^+)^{-1} L^-$$
$$C = I - P(\bar{L})^{-1} R$$

Multigrid for the 1D Poisson Equation

$$u_{xx} = f$$

- coarse grid with $\bar{m}$ nodes

$$\begin{bmatrix} \bar{u}_0 & \bar{u}_1 & \cdots & \bar{u}_{\bar{m}-1} & \end{bmatrix}$$

- fine grid with $m = 2\bar{m}$ nodes

$$\begin{bmatrix} u_0 & u_1 & u_2 & u_3 & \cdots & u_{m-2} & u_{m-1} \end{bmatrix}$$

- split into even and odd: $u_i^0 = u_{2i}$, $u_i^1 = u_{2i+1}$

$$\begin{bmatrix} u_0^0 & u_0^1 & u_1^0 & u_1^1 & \cdots & u_{\bar{m}-1}^0 & u_{\bar{m}-1}^1 \end{bmatrix}$$

- $\bar{u}$, u^0, $u^1 \in \mathbb{R}^{\bar{m}}$

Differential Operator on the Coarse Grid

- grid spacing $\overline{\Delta x} = 2$
- operator $\overline{L} : \mathbb{R}^{\bar{m}} \to \mathbb{R}^{\bar{m}} : \bar{u} \to \bar{v} = \overline{L}\bar{u}$

$$\bar{v}_i = (\bar{u}_{i-1} - 2\bar{u}_i + \bar{u}_{i+1})/4$$

- factor 4 corresponds to $\overline{\Delta x}^2$
- function $\overline{L} : \mathbb{C} \to \mathbb{C}$

$$\overline{L}(w) = (w - 2 + w^{-1})/4$$

- $\overline{L} = \overline{L}(S)$ with S the periodic shift operator

Differential Operator on the Fine Grid

- grid spacing $\Delta x = 1$
- operator $L : \mathbb{R}^{2\bar{m}} \to \mathbb{R}^{2\bar{m}} : u \to v = Lu$

$$v_i = u_{i-1} - 2u_i + u_{i+1}$$

- even points $\begin{bmatrix} \cdots & u_{i-1}^1 & \boxed{u_i^0} & u_i^1 & \cdots \end{bmatrix}$

$$v_i^0 = u_{i-1}^1 - 2u_i^0 + u_i^1$$

- odd points $\begin{bmatrix} \cdots & u_i^0 & \boxed{u_i^1} & u_{i+1}^0 & \cdots \end{bmatrix}$

$$v_i^1 = u_i^0 - 2u_i^1 + u_{i+1}^0$$

- matrix-valued function $L : \mathbb{C} \to \mathbb{C}^{2\times 2}$

$$L(w) = \begin{bmatrix} -2 & w+1 \\ 1 + w^{-1} & -2 \end{bmatrix}$$

Full Weighting Restriction

- operator $R : \mathbb{R}^{2\bar{m}} \to \mathbb{R}^{\bar{m}} : u \to \bar{v} = Ru$

$$\bar{v}_i = (u_{2i-1} + 2u_{2i} + u_{2i+1})/4$$

- even-odd

$$\bar{v}_i = (u_{i-1}^1 + 2u_i^0 + u_i^1)/4$$

- matrix-valued function $R : \mathbb{C} \to \mathbb{C}^{2\times 1}$

$$R(w) = \begin{bmatrix} 2 & w+1 \end{bmatrix}/4$$

Bilinear Interpolation

- operator $P : \mathbb{R}^{\bar{m}} \to \mathbb{R}^{2\bar{m}} : \bar{u} \to v = P\bar{u}$

$$v_i^0 = \bar{u}_i$$
$$v_i^1 = (\bar{u}_i + \bar{u}_{i+1})/2$$

- matrix-valued function $P : \mathbb{C} \to \mathbb{C}^{1 \times 2}$

$$P(w) = \begin{bmatrix} 1 \\ (1 + w^{-1})/2 \end{bmatrix}$$

Smoothing

- Jacobi
$$S_{\mathrm{J}} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 0 & w+1 \\ 1+w^{-1} & 0 \end{bmatrix}$$

- red-black Gauss-Seidel $S_{\mathrm{RB}} = S_{\mathrm{B}} S_{\mathrm{R}}$

$$S_{\mathrm{R}} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & w+1 \\ 0 & 1 \end{bmatrix}$$

$$S_{\mathrm{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 1+w^{-1} & 0 \end{bmatrix}$$

Multigrid for the 2D Poisson Equation

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

- differential operator on the fine grid $\Delta x = \Delta y = 1$
- split fine grid in even-even, even-odd, odd-even, odd-odd
- matrix-valued function $L : \mathbb{C}^2 \to \mathbb{C}^{4 \times 4}$

$$L(w_x, w_y) = L(w_x) \otimes I_2 + I_2 \otimes L(w_y)$$
$$= \begin{bmatrix} -4 & w_y + 1 & w_x + 1 & 0 \\ 1 + w_y^{-1} & -4 & 0 & w_x + 1 \\ 1 + w_x^{-1} & 0 & -4 & w_y + 1 \\ 0 & 1 + w_x^{-1} & 1 + w_y^{-1} & -4 \end{bmatrix}$$

Multigrid for the 2D Heat Equation

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

- combine time-dependent and multigrid
- matrix-valued function $M : \mathbb{C}^3 \to \mathbb{C}^{4\times 4}$

$$M(w_t, w_x, w_y) =$$
$$\begin{bmatrix} \times\times\times\times \\ \times\times\times\times \\ \times\times\times\times \\ \times\times\times\times \end{bmatrix}^{\nu_2} \left(\begin{bmatrix} \times \\ \times \\ \times \\ \times \end{bmatrix} - \begin{bmatrix} \times \\ \times \\ \times \end{bmatrix} [\times]^{-1} [\times\times\times\times] \right) \begin{bmatrix} \times\times\times\times \\ \times\times\times\times \\ \times\times\times\times \\ \times\times\times\times \end{bmatrix}^{\nu_1}$$
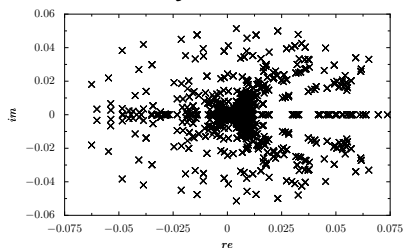
- convergence for infinite grids

$$\rho(M) = \max_{|w_t|=|w_x|=|w_y|=1} \rho(M(w_t, w_x, w_y))$$

Pictures

- heat equation, two-grid, red-black Gauss-Seidel, full weighting restriction, bilinear interpolation, $\nu_1 = \nu_2 = 1$
- periodic grids $\bar{n}_x = \bar{n}_y = 16$
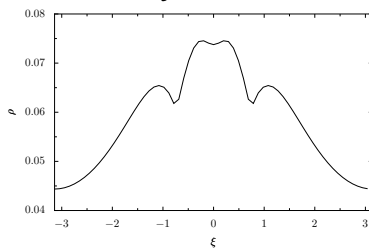- implicit Euler $\Delta t = 1$
- $\rho = 0.074$

$$\sigma(M(w_t, w_x, w_y))$$
$$n_t = 16$$

$$\max_{w_x, w_y} \rho(M(e^{i\xi}, w_x, w_y))$$
$$n_t = 64$$

Ideas for the Future

- extend to differential algebraic equations (DAEs)
  could be straightforward

$$M\dot{u} = Lu + f, \qquad M \text{ singular}$$

- extend to delay differential equations (DDEs)

$$\dot{u}(t) = Lu(t) + u(t - \tau)$$

- multigrid for curl-curl
- operator-valued functions
  (Banach algebras and Gelfand theory)

$$T \in X, \qquad F : \mathbb{C} \to Y, \qquad F(T)$$

Delay Equations

$$\dot{u}(t) = Lu(t) + u(t - \tau)$$

- derivative operator

$$T_1 u(t) = \dot{u}(t)$$

- delay operator (continuous shift)

$$T_2 u(t) = u(t - \tau)$$

- we need
  - ▶ functional calculus for functions not analytic at infinity
    $T_2 = \exp(T_1)$ (formally)
    semigroup theory, Hille-Phillips functional calculus
  - ▶ functional calculus for commuting operators
    T_1 and T_2 commute
    joint spectrum
    $\sigma(F(T_1, T_2)) = \sigma(F(\sigma(T_1, T_2)))$

Further Reading

- WR for IVPs (Miekkala and Nevanlinna '87)
- WR for periodic ODEs (Vandewalle and Piessens '92, '93)
- WR for finite elements (Janssen and Vandewalle '96)
- functional calculus (Dunford and Schwartz; Taylor; Haase; . . .)
- MG (Briggs, Henson and McCormick; Trottenberg, Oosterlee and Schüller; Stüben)
- MG WR (Lubich and Ostermann '89; Vandewalle)
- Banach algebras (Rudin; Lax)
- tensor products (Reed and Simon '73; Ichinose '78)
- semigroup theory (Hille and Philips)
- joint spectra (Taylor '70; Vasilescu)