Implicitly Defined High-Order Operator Splittings for Time-Dependent Variable-Coefficient PDE Using Modified Moments

James V. Lambers

Department of Energy Resources Engineering
Institute for Computational and Mathematical Engineering
Stanford University

Bay Area Scientific Computing Day
March 29, 2008

Model Variable-Coefficient Problem

$$\begin{aligned} \frac{\partial u}{\partial t} + L(x, D)u &= 0, \quad 0 < x < 2\pi, \quad t > 0 \\ u(x, 0) &= f(x), \quad 0 < x < 2\pi \\ u(0, t) &= u(2\pi, t), \quad t > 0 \end{aligned}$$

where

$$L(x, D) = -Dp(x)D + q(x), \quad D = \frac{\partial}{\partial x}$$

We assume that $L(x, D)$ is positive semi-definite, and that the coefficients are smooth

Quick-and-Dirty Solution

Let N be an even integer. An approximate solution is

$$\begin{aligned}\tilde{u}(x, t+\Delta t) &= \frac{1}{2\pi}\sum_{\omega=-N/2}^{N/2} e^{i\omega x}\langle e^{i\omega x}, e^{-L(x,D)\Delta t}\tilde{u}(x,t)\rangle \\ &\approx \sum_{\omega=-N/2}^{N/2} e^{i\omega x} e^{-\lambda_\omega \Delta t}\hat{u}(\omega, t)\end{aligned}$$

where $\lambda_\omega = \langle e^{i\omega x}, L(x,D)e^{i\omega x}\rangle$ and $\hat{u}(\omega, t)$ are the coefficients of the Fourier series of $\tilde{u}(x,t)$

This works well if the coefficients of L are nearly constant, but if not, how can we compute $\langle e^{i\omega x}, e^{-L(x,D)\Delta t}\tilde{u}(x,t)\rangle$ as accurately and efficiently as possible?

Elements of Functions of Matrices

If A is $N \times N$ and symmetric, then $u^T e^{-A\Delta t} v$ is given by a Riemann-Stieltjes integral

$$\int_{\lambda_N}^{\lambda_1} e^{-\lambda \Delta t} \, d\alpha(\lambda)$$

where the measure $\alpha(\lambda)$, which is based on the spectral decomposition of A, is positive and increasing if $\mathbf{v} = \mathbf{u}$, or if $\mathbf{v}$ is a small perturbation of $\mathbf{u}$

Gaussian Quadrature

- This integral can be approximated using Gaussian quadrature rules (G. Golub and G. Meurant, '94)
- The nodes and weights are obtained by applying the Lanczos algorithm to A with initial vectors $\mathbf{u}$ and $\mathbf{v}$ to produce T, the tridiagonal matrix of recursion coefficients
- The nodes are the eigenvalues of T, and the weights are obtained from the first components of the left and right eigenvectors
- However, we need to ensure that weights will not be negative

The $\mathbf{u} \neq \mathbf{v}$ Case

For general $\mathbf{u}$ and $\mathbf{v}$, the bilinear form $\mathbf{u}^T e^{-A\Delta t}\mathbf{v}$ can be expressed as the difference quotient

$$\mathbf{u}^T e^{-A\Delta t}\mathbf{v} = \frac{1}{\delta}\left[\mathbf{u}^T e^{-A\Delta t}(\mathbf{u}+\delta\mathbf{v}) - \mathbf{u}^T e^{-A\Delta t}\mathbf{u}\right]$$

where δ is a small constant. Both forms lead to Riemann-Stieltjes integrals with positive, increasing measures

Our goal: to compute Fourier components of the solution by approximating these integrals with $\mathbf{u} = e^{i\omega\mathbf{x}}$, $\mathbf{v} = \tilde{u}(\mathbf{x}, t)$ and A obtained from a discretization of $L(x, D)$

Krylov Subspace Spectral (KSS) Methods[1]

To compute Fourier components of $\tilde{u}(\mathbf{x}, t_{n+1})$, for $n = 0, 1, \ldots$

- Apply symmetric Lanczos algorithm to L with starting vector $e^{i\omega\mathbf{x}}$ to obtain matrix T_ω of recursion coefficients (note: T_ω is a $K \times K$ matrix, where K is the number of quadrature nodes)
- Apply unsymmetric Lanczos algorithm to L with starting vectors $e^{i\omega\mathbf{x}}$ and $e^{i\omega\mathbf{x}} + \delta\tilde{u}(\mathbf{x}, t_n)$ to obtain matrix $T_\omega(\delta)$ of recursion coefficients
- Use eigenvalues and eigenvectors of T_ω and $T_\omega(\delta)$ to approximate, by Gaussian quadrature,

$$\begin{aligned} y_0 &= \langle e^{i\omega\mathbf{x}}, \exp[-L\Delta t]e^{i\omega\mathbf{x}}\rangle, \\ y_\delta &= \langle e^{i\omega\mathbf{x}}, \exp[-L\Delta t](e^{i\omega\mathbf{x}} + \delta\tilde{u}(\mathbf{x}, t_n))\rangle \end{aligned}$$

- Finally, $\hat{u}(\omega, t_{n+1}) = (y_\delta - y_0)/\delta$

[1]See [1]

Computing the Recursion Coefficients

- We need to apply the Lanczos algorithm to a matrix L derived from the operator $L(x, D)$, with two pairs of starting vectors:
  - Case 1: $e^{i\omega \mathbf{x}}$ and $e^{i\omega \mathbf{x}}$
  - Case 2: $e^{i\omega \mathbf{x}}$ and $e^{i\omega \mathbf{x}} + \delta \tilde{u}(\mathbf{x}, t_n)$
- On an N-point grid, this takes $O(N \log N)$ operations for each ω
- Challenge: how to do this in $O(N \log N)$ operations for all ω?
- For case 1, can compute recursion coefficients analytically, as functions of ω
- For case 2, can perturb coefficients from case 1 using $K - 1$ applications of L to $\tilde{u}(\mathbf{x}, t_n)$ and simple recurrence relations (details in [3])

Example

Let $L(x, D) = -Dp(x)D$, and $K = 2$. Then the entries of

$$T_\omega = \begin{bmatrix} \alpha_1 & \beta_1 \\ \beta_1 & \alpha_2 \end{bmatrix}$$

are

$$\begin{aligned} \alpha_1 &= \overline{p}\omega^2 \\ \beta_1^2 &= \overline{p^2}\omega^4 + \overline{[p']^2}\omega^2 - \alpha_1^2 \\ \alpha_2 &= \beta_1^{-2}[\omega^2(\overline{p^3}\omega^4 - 2\overline{p^2 p''}\omega^2 + 4\overline{p(p')^2}\omega^2 + \overline{p(p'')^2}) + 2\alpha_1\beta_1^2 + \alpha_1^3] \end{aligned}$$

where

$$\overline{f} = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,dx$$

Note: these formulas generalize to higher space dimensions

Modifying Recursion Coefficients

This iteration modifies the entries of T_ω to compute

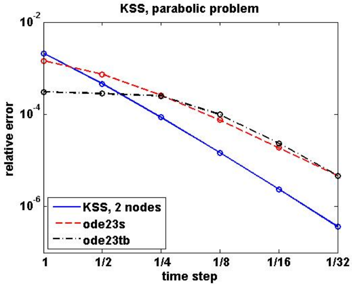
$$T_\omega(\delta) = \begin{bmatrix} \hat{\alpha}_1 & \hat{\beta}_1 \\ \hat{\beta}_1 & \hat{\alpha}_2 \end{bmatrix}$$

produced by the unsymmetric Lanczos algorithm applied to L with starting vectors $\mathbf{r}_0$ and $\mathbf{r}_0 + \mathbf{q}_0$, where the $\mathbf{r}_k$ are the unnormalized Lanczos vectors from the symmetric case

In our case, $\mathbf{r}_0 = e^{i\omega\mathbf{x}}$, so inner products with $\mathbf{r}_k$ can be computed simultaneously for all ω using FFTs

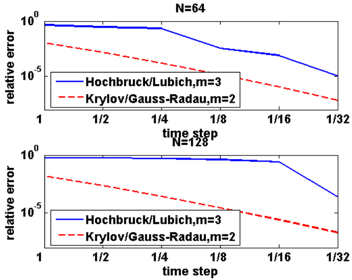
$$\begin{aligned}
\hat{\beta}_0^2 &= \beta_0^2 + \mathbf{r}_0^H \mathbf{q}_0 \\
s_0 &= \frac{\beta_0}{\hat{\beta}_0^2}, \; t_0 = \frac{\beta_0^2}{\hat{\beta}_0^2} \\
\hat{\alpha}_1 &= \alpha_1 + s_0 \mathbf{r}_1^H \mathbf{q}_0 \\
d_1 &= (\alpha_1 - \hat{\alpha}_1) t_0^{1/2} / \hat{\beta}_0 \\
\mathbf{q}_1 &= (L - \hat{\alpha}_1 I)\mathbf{q}_0 \\
\hat{\beta}_1^2 &= t_0 \beta_1^2 + s_0 \mathbf{r}_1^H \mathbf{q}_1 \\
s_1 &= \frac{\beta_1}{\hat{\beta}_1^2} s_0, \; t_1 = \frac{\beta_1^2}{\hat{\beta}_1^2} t_0 \\
\hat{\alpha}_2 &= \alpha_2 + s_1 \mathbf{r}_2^H \mathbf{q}_1 + d_1 \beta_0 t_1^{-1/2}
\end{aligned}$$

Consistency



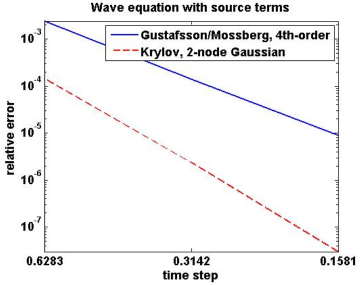
- Gaussian quadrature rules with K nodes are exact for polynomials of degree $2K - 1$ or less
- Quadrature error includes the $2K$-th derivative of the integrand, $e^{-\lambda\Delta t}$
- Therefore, error in each Fourier component is $O(\Delta t^{2K})$!
- Spatial error arises from truncation of Fourier series

Stability



- For parabolic problems, placement of nodes causes damping at higher frequencies (but not too much!)
- Therefore, KSS methods have the stability of implicit methods, even though they are explicit
- Is not restricted by a CFL condition that forces reduction in time step when grid is refined!
- Variation in coefficients, not magnitude, influences stability

The Wave Equation



- The integrands are obtained from the spectral decomposition of the propagator (see [2]):

$$\cos(\sqrt{\lambda}t), \quad \lambda^{\pm 1/2}\sin(\sqrt{\lambda}t)$$

- For each Fourier component, $O(\Delta t^{4K})$ local accuracy!
- For $K = 1$, $p(x)$ constant, $q(x)$ bandlimited, global error is 3rd order in time, and the method is unconditionally stable! (see [4])

In the Limit: Derivatives of Moments!

- Let $T_\omega(\delta)$ be the output of the Lanczos algorithm with starting vectors $e^{i\omega\mathbf{x}}$ and $e^{i\omega\mathbf{x}} + \delta\mathbf{u}^n$, where $\mathbf{u}^n = \tilde{u}(\mathbf{x}, t_n)$
- Each Fourier component of $\mathbf{u}^{n+1}$ approximates the derivative

$$\frac{d}{d\delta}\left[\mathbf{e}_\omega^H(\mathbf{e}_\omega + \delta\mathbf{u}^n)\left(e^{-T_\omega(\delta)t}\right)_{11}\right]_{\delta=0}$$

- Letting $\delta \to 0$ yields

$$[\hat{\mathbf{u}}^{n+1}]_\omega = \sum_{k=1}^{K} w_k e^{-\lambda_k \Delta t}\left([\hat{\mathbf{u}}^n]_\omega + \frac{w_k'}{w_k} - \Delta t \lambda_k'\right)$$

where the λ_k', w_k' are the derivatives of the nodes and weights w.r.t. δ at $\delta = 0$

High-Order Splittings[2]

- Derivatives of nodes and weights w.r.t. δ are Fourier components of applications of pseudodifferential operators applied to $\tilde{u}(\mathbf{x}, t_n)$
- K-node approximate solution has form

$$\tilde{u}(x, t + \Delta t) = \sum_{k=1}^{K} W_k e^{-C_k \Delta t}(I - \Delta t V_k)\tilde{u}(x, t)$$

  where each C_k is a constant-coefficient pseudo-differential operator, of the same order as $L(x, D)$, and positive semi-definite
- Reformulation as splittings facilitates stability analysis
- Unconditionally stable for $K = 1$, solution operator bounded independently of Δt and Δx for $K = 2$

[2] See [5]

Derivatives of the Nodes

There exists a unitary matrix Q_ω^0 such that

$$T_\omega = Q_\omega^0 \Lambda_\omega [Q_\omega^0]^H$$

and the nodes are on the diagonal of Λ_ω. Also,

$$T_\omega(\delta) = Q_\omega(\delta)\Lambda_\omega(\delta)Q_\omega(\delta)^{-1},$$

for suffficiently small δ, where $Q_\omega(0) = Q_\omega^0$

Differentiating with respect to δ and evaluating at $\delta = 0$ yields

$$\mathrm{diag}(\Lambda_\omega') = \mathrm{diag}\left(Q_\omega(0)^H T_\omega' Q_\omega(0)\right),$$

since all other terms arising from differentiation vanish on the diagonal

Derivatives of the Weights

To compute the derivatives of the weights, consider

$$(T_\omega(\delta) - \lambda_j I)\mathbf{w}_j(\delta) = \mathbf{0}, \quad j = 1, \ldots, K,$$

where $\mathbf{w}_j(\delta)$ is a normalized eigenvector of $T_\omega(\delta)$ with eigenvalue λ_j.

- Differentiate with respect to δ, evaluate at $\delta = 0$
- Delete last row and column, using normalization
- We now have a $(K-1) \times (K-1)$ system, where the matrix is tridiagonal plus a rank-one update
- Solve this system, and a similar one for the left eigenvector

We can obtain w_k' from the first components of the two solutions

Derivatives of the Recursion Coefficients

From the expressions for the entries of $T_\omega(\delta)$ in terms of those of T_ω, the derivatives of the recursion coefficients can be obtained by setting $\mathbf{r}_0 = e^{i\omega\mathbf{x}}$ and $\mathbf{q}_0 = \delta\mathbf{u}^n$

By differentiating the recurrence relations with respect to δ and evaluating at $\delta = 0$, we obtain the following algorithm that computes these derivatives

$$[\beta_0^2]' = \mathbf{r}_0^H\mathbf{q}_0$$
$$s_0 = \frac{1}{\beta_0}$$
$$t_0' = -\frac{[\beta_0^2]'}{\beta_0^2}$$
$$d_0' = 0$$

for $j = 1, \dots, K$

$$\alpha_j' = s_{j-1}\mathbf{r}_j^H\mathbf{q}_{j-1} + d_{j-1}'\beta_{j-2}$$

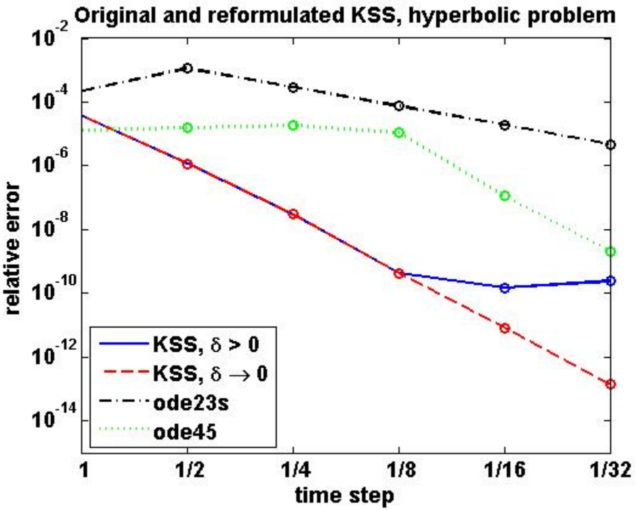
if $j < K$ **then**

$$d_j' = (d_{j-1}'\beta_{j-2} - \alpha_j')/\beta_{j-1}$$
$$\mathbf{q}_j = (L - \alpha_j I)\mathbf{q}_{j-1} - \beta_{j-1}^2\mathbf{q}_{j-2}$$
$$[\beta_j^2]' = t_{j-1}'\beta_j^2 + s_{j-1}\mathbf{r}_j^H\mathbf{q}_j$$
$$s_j = s_{j-1}/\beta_j$$
$$t_j' = t_{j-1}' - \frac{[\beta_j^2]'}{\beta_j^2}$$

end

end

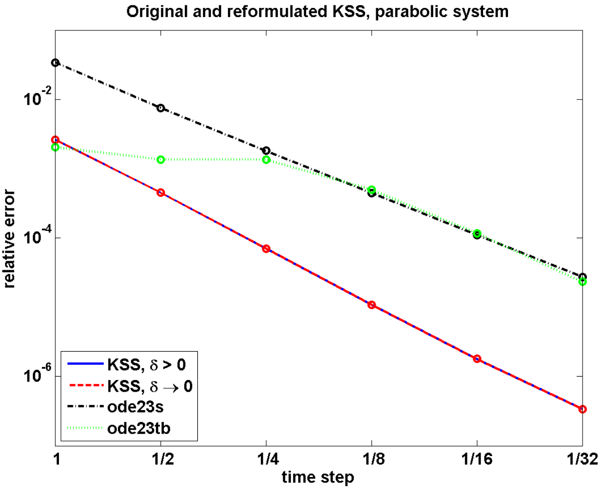
Numerical Stabilization



Systems of Equations

- Generalization to systems of equations is straightforward
- For a system of the form $u_t + A(x, D)u = 0$, where each entry of $A(x, D)$ is a differential operator, we can use trial functions $\mathbf{v}_j \otimes e^{i\omega\mathbf{x}}$ where $\mathbf{v}_j$ is an eigenvector of $\text{Avg}_x A(x, \omega)$
- A corresponding basis of test functions can be used in the same way, except that the left eigenvectors of $\text{Avg}_x A(x, \omega)$ are used instead
- These choices are eigenfunctions of the averaged-coefficient problem

Results


Original and reformulated KSS, parabolic system
relative error
10^-2
10^-4
10^-6
KSS, δ > 0
KSS, δ → 0
ode23s
ode23tb
1
1/2
1/4
1/8
1/16
1/32
time step

Consistency and Stability

- As in the scalar case, a 2-node Gaussian rule yields cubic convergence in time
- Proof is analogous to the scalar case: key is $d^{2K}/d\lambda^{2K}(e^{-\lambda\Delta t}) = (\Delta t)^{2K} e^{-\lambda\Delta t}$ in quadrature error
- Similar stability properties as in scalar case
- However, the Lanczos iteration can break down due to orthogonality of corresponding Lanczos vectors, for example $A(x, D) = [\, 0 \;\; -I; \;\; L(x, D) \; 0 \,]$ where $L(x, D)$ is positive definite

Alternative Bases

- Can use other, orthonormal bases for both trial and test functions
- Instead of using eigenvectors of $\text{Avg}_x A(x, \omega)$, can use Schur vectors, which results in a block diagonalization of the averaged-coefficient operator
- Negligible impact on accuracy, since trial functions are still approximate eigenfunctions
- Fourier coefficients of system components available by solving small orthogonal systems

Time-Dependent Schrödinger Equation

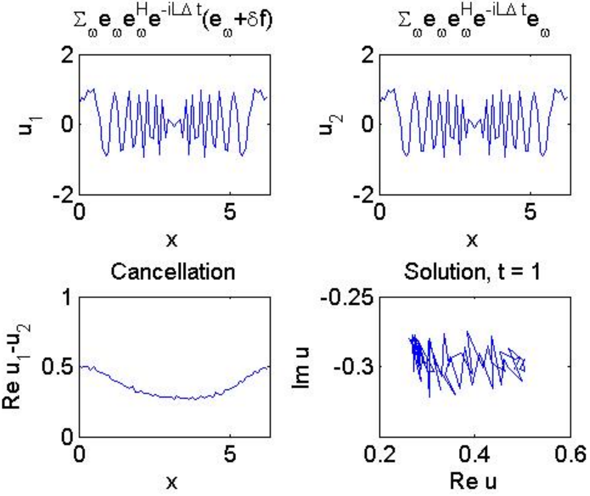
KSS methods can be applied to a problem of the form

$$iu_t = -c\Delta u + Vu$$

simply by using the integrand $e^{-i\lambda\Delta t}$ instead of $e^{-\lambda t}$

The solution has similar smoothness to that of the initial data, but non-smooth or discontinuous coefficients can cause difficulties

Example, Oscillatory Coefficient



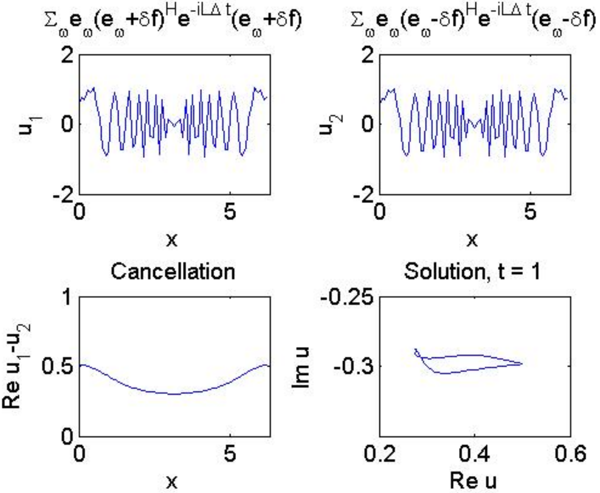
Dealing with Non-Smooth Coefficients

Lack of smoothness in the coefficients can easily be handled by symmetrizing the perturbation, with the polar decomposition

$$\begin{aligned}\mathbf{u}^H e^{-iL\Delta t}\mathbf{v} &= \frac{1}{4}\left[(\mathbf{u}+\mathbf{v})^H e^{-iL\Delta t}(\mathbf{u}+\mathbf{v}) - (\mathbf{u}-\mathbf{v})^H e^{-iL\Delta t}(\mathbf{u}-\mathbf{v})\right] + \\ &\quad \frac{1}{4i}\left[(\mathbf{u}+i\mathbf{v})^H e^{-iL\Delta t}(\mathbf{u}+i\mathbf{v}) - (\mathbf{u}-i\mathbf{v})^H e^{-iL\Delta t}(\mathbf{u}-i\mathbf{v})\right]\end{aligned}$$

If the initial data is not smooth, then $\mathbf{v}$ should be scaled by a parameter δ

Example, Symmetrized Perturbations



Summary and Future Directions

- Krylov subspace spectral methods are showing more promise as their development progresses
  - Applicability to problems with rough behavior, by simply symmetrizing the perturbation!
  - Stability like that of implicit methods
  - Competitive performance and scalability
- Next Steps:
  - Application to other spatial discretizations such as wavelet bases or finite elements
  - Application to more interesting systems of PDE (first results in [5])
  - Generalization of homogenization techniques (see [4]) to higher space dimensions

References

- [1] JL, "Krylov Subspace Spectral Mehods for Variable-Coefficient Initial-Boundary Value Problems", *ETNA* **20** (2005), p. 212-234
- [2] P. Guidotti, JL, K. Sølna, "Analysis of 1-D Wave Propagation in Inhomogeneous Media", *Num. Funct. Anal. Opt.* **27** (2006), p. 25-55
- [3] JL, "Practical Implementation of Krylov Subspace Spectral Methods", *J. Sci. Comput.* **32** (2007), p. 449-476
- [4] JL, "Derivation of High-Order Spectral Methods for Time-Dependent PDE Using Modified Moments", *ETNA* **28** (2008), p. 114-135
- [5] JL, "Implicitly Defined High-Order Operator Splittings for Parabolic and Hyperbolic Variable-Coefficient PDE Using Modified Moments", *Intl. J. Comp. Sci.* (2008), to appear