# Stability of Krylov Subspace Spectral Methods

James V. Lambers
Department of Energy Resources Engineering
Stanford University

includes joint work with
Patrick Guidotti and Knut Sølna, UC Irvine
Margot Gerritsen, Stanford University

Harrachov '07
August 23, 2007

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

# Quick-and-dirty Solution

Let $\{ \phi_\nu \}$ be a set of orthonormal $2\pi$-periodic functions. Then, an approximate solution is

$$\begin{aligned} u(x,t) &= \sum_\nu \phi_\nu(x) \langle \phi_\nu, e^{-L(x,D)t} f \rangle \\ &\approx \sum_\nu \phi_\nu(x) e^{-\lambda_\nu t} \langle \phi_\nu, f \rangle \end{aligned}$$

where $\lambda_\nu = \langle \phi_\nu, L(x,D)\phi_\nu \rangle$

This works if the $\phi_\nu$ are nearly eigenfunctions, but if not, how can we compute $\langle \phi_\nu, e^{-L(x,D)t} f \rangle$ as accurately and efficiently as possible?

# Elements of Functions of Matrices

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

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

provided the measure $\alpha(\lambda)$, which is based on the spectral decomposition of $A$, is positive and increasing

This is the case if $\mathbf{v} = \mathbf{u}$, or if $\mathbf{v}$ is a small perturbation of $\mathbf{u}$

# The $\mathrm{u} \neq \mathrm{v}$ Case

For general $\mathrm{u}$ and $\mathrm{v}$, the bilinear form $\mathrm{u}^T e^{-At} \mathrm{v}$ can be obtained by writing it as the difference quotient

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

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

How can we approximate these integrals?

# Gaussian Quadrature

- These two integrals 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 starting vectors $\mathrm{u}$ and $\mathrm{v}$ $\mathrm{to\ produce}$ $T$, the tridiagonal matrix of recursion coefficients
- The nodes are the eigenvalues of $T$, and the weights are the products of the first components of the left and right eigenvectors
- We want rules for each Fourier component

# What if $A$ is a differential operator?

- If $\mathrm{u} = \mathrm{v} = \mathrm{e}_\omega = e^{i\omega x}$, then recursion coefficients $\alpha_j$, $\beta_j$ are functions of $\omega$
- Let $A = L(x,D)$ from before, with $q(x) = 0$. The recursion coefficients for a 2-node Gaussian rule are

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

  where $\overline{f} = \mathrm{Avg}\ f$
- Similar formulas apply in higher dimensions

# Updating Coefficients for $\mathrm{u} \neq \mathrm{v}$

To produce modified recursion coefficients generated by $\mathrm{r}_0 = \mathrm{u}$ and $\mathrm{r}_0 + \delta\mathrm{v}$:

$$
\begin{aligned}
&\alpha_0 = \hat{\alpha}_0 = 0,\ \beta_{-1} = 0,\ \mathbf{q}_{-1} = 0 \\
&\mathbf{q}_0 = \delta\mathbf{v} \\
&\hat{\beta}_0^2 = \beta_0^2 + \mathbf{r}_0^H \mathbf{q}_0 \\
&s_0 = \tfrac{\beta_0}{\hat{\beta}_0^2},\ t_0 = \tfrac{\beta_0^2}{\hat{\beta}_0^2},\ d_0 = 0 \ \textbf{for}\ j = 1, \ldots, K \\
&\quad \hat{\alpha}_j = \alpha_j + s_{j-1}\mathbf{r}_j^H \mathbf{q}_{j-1} + d_{j-1}\beta_{j-2} t_{j-1}^{-1/2} \\
&\quad d_j = (d_{j-1}\beta_{j-2} + (\alpha_j - \hat{\alpha}_j) t_{j-1}^{1/2})/\hat{\beta}_{j-1} \\
&\quad \mathbf{q}_j = (A - \hat{\alpha}_j I)\mathbf{q}_{j-1} - \hat{\beta}_{j-1}^2 \mathbf{q}_{j-2} \\
&\quad \hat{\beta}_j^2 = t_{j-1}\beta_j^2 + s_{j-1}\mathbf{r}_j^H \mathbf{q}_j \\
&\quad s_j = \tfrac{\beta_j}{\hat{\beta}_j^2} s_{j-1},\ t_j = \tfrac{\beta_j^2}{\hat{\beta}_j^2} t_{j-1} \\
&\textbf{end}
\end{aligned}
$$

This is particularly useful if the $\mathrm{r}_j$, $\alpha_j$, $\beta_j$ are parameterized families and $\mathrm{v}$ is a fixed vector

# Krylov Subspace Spectral Methods

To compute Fourier components of $u(x,t^{n+1})$:

- Apply symmetric Lanczos algorithm to $L$ with starting vector $e^{i\omega x}$
- Use fast updating to obtain modified recursion coefficients for starting vectors $e^{i\omega x}$, $e^{i\omega x} + \delta u^n$
- Approximate, by Gaussian quadrature,

$$\begin{aligned} y_0 &= \langle e^{i\omega x}, \exp[-L(x,D)\Delta t] e^{i\omega x} \rangle, \\ y_1 &= \langle e^{i\omega x}, \exp[-L(x,D)\Delta t] (e^{i\omega x} + \delta u^n) \rangle \end{aligned}$$

- Finally, $\hat{u}(\omega, t^{n+1}) = (y_1 - y_0)/\delta$

# Why Do it This Way?

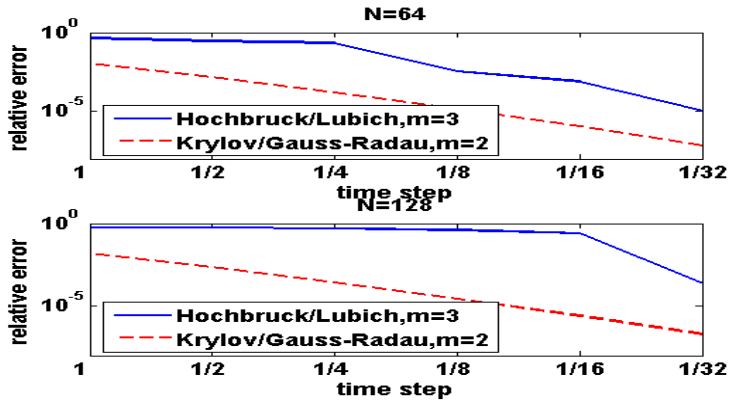
- Other, more general Krylov subspace methods (e.g. M. Hochbruck and C. Lubich, '96) use a single Krylov subspace for each time step, with $\mathrm{u}^n$ as the starting vector, to approximate $\exp(\text{-}Lt)u^n$
- KSS methods obtain Fourier components from ***derivatives of frequency-dependent Krylov subspace bases in the direction of*** $\mathrm{u}^n$
- Thus, each component receives individual attention, enhancing accuracy and stability

# Properties

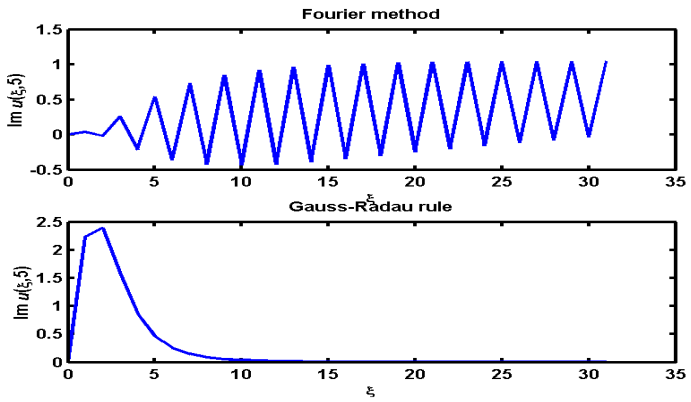
- **They're High-Order Accurate!**
  Each Fourier component of $u(x,t^{n+1})$ is computed with local accuracy of $O(\Delta t^{2K})$, where $K$ is the number of nodes in the Gaussian rule
- **They're Explicit but Very Stable!**
  If $p(x)$ is constant and $q(x)$ is bandlimited, then for $K = 1$, method is ***unconditionally stable!***
  For $K = 2$, solution operator is bounded independently of $\Delta t$ and $N$

# Demonstrating Stability

They do not experience the same difficulties with stiffness as other Krylov subspace methods, or the same weak instability as the unsmoothed Fourier method



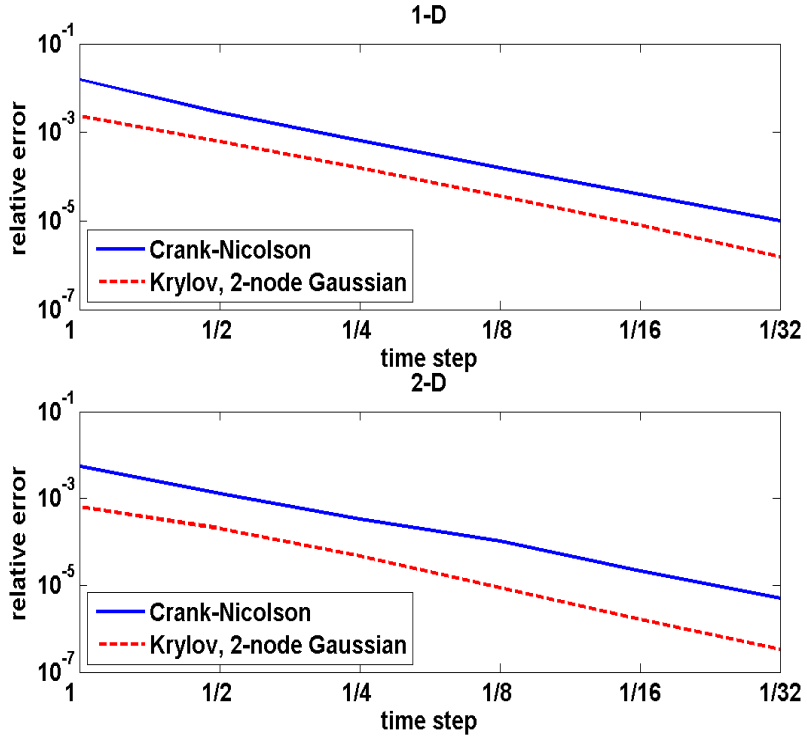

Contrasting Krylov subspace methods applied to parabolic problem on different grids




Fourier, KSS methods applied to $u_t = (\sin(x)u)_x$, smooth initial data until $T = 5$

# Properties, cont'd

- **They're efficient and scalable!**
  - Performance of MATLAB implementation comparable or superior to that of built-in ODE solvers
  - Accuracy and efficiency scale to finer grids or higher spatial dimension



# In the Limit: Derivatives of Moments!

- Each Fourier component approximates the Gâteaux derivative

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

- $K$-node approximate solution has the form

$$\begin{aligned}\hat{u}(\omega, t^{n+1}) &= \hat{u}(\omega, t^n)\sum_{k=1}^{K} e^{-\lambda_k \Delta t} w_k + \\ &\sum_{k=1}^{K} e^{-\lambda_k \Delta t}\left[w_k' - \Delta t w_k \lambda_k'\right]\end{aligned}$$

# The Splitting Perspective

- Derivatives of nodes and weights w.r.t. $\delta$ are Fourier components of applications of pseudodifferential operators applied to $u(x,t^n)$
- $K$-node approximate solution has form

$$u(x,t) = \sum_{k=1}^{K} w_k e^{-C_k t}(I - tV_k)f(x)$$

where and each $C_k$ is a constant-coefficient pseudo-differential operator, of the same order as $L(x,D)$, and positive semi-definite

# Simplest Splittings

- The case $K = 1$ reduces to solving the averaged-coefficient problem exactly, after applying forward Euler with the residual operator
- In the case $K = 2$, with $p(x)$ constant, the operators $V_k$ are second-order, and yet the approximate solution operator $S_N(\Delta t)$ satisfies

$$S_N(\Delta t) = e^{-C_1 \Delta t} + B\Delta t + O(\Delta t^2)$$

where $B$ is a bounded operator!

# General Splittings

- Operators $C_k$ and $V_k$ are defined in terms of derivatives of nodes and weights, which represent pseudo-differential operators
- For each $\omega$, let $T_\omega$ be the $K \times K$ Jacobi matrix output by Lanczos, with initial vector $\mathrm{e}_\omega$
  - Derivatives of nodes: $T_\omega = Q \Lambda_\omega Q^T$, where nodes are on diagonal of $\Lambda_\omega$, and $\Lambda_\omega\text{'} = Q^T T_\omega\text{'} Q$, where $T_\omega\text{'}$ is available from fast updating algorithm
  - Derivatives of weights: from solution of systems of form $A \mathrm{w_j}\text{'} = \mathrm{b}$, where $A = (T_\omega(1{:}K{-}1,1{:}K{-}1)$ - $\lambda_j I) + R_j$, for $j = 1, \ldots, K$, with $R_j$ a rank-one matrix

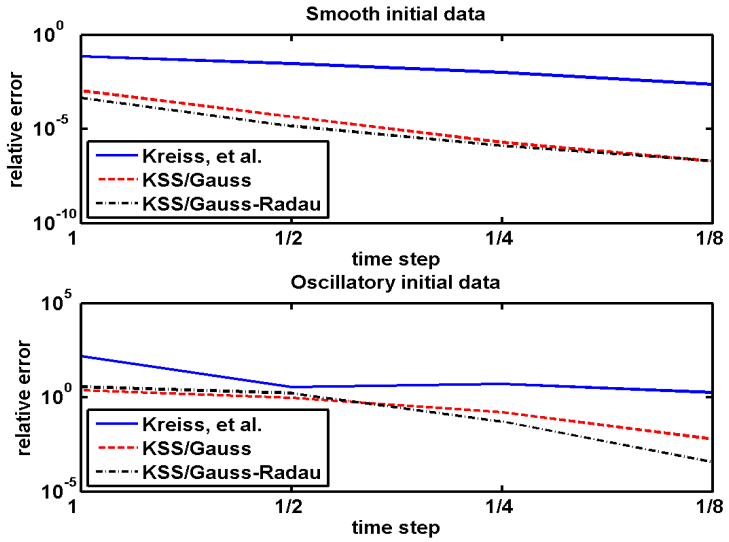
# The Wave Equation

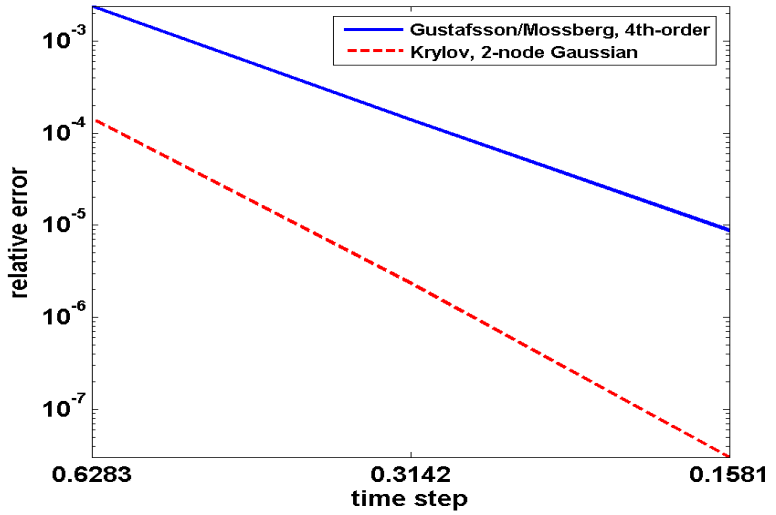
- The integrands are obtained from the spectral decomposition of the propagator (P. Guidotti, JL and K. Sølna '06):

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

- 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!***

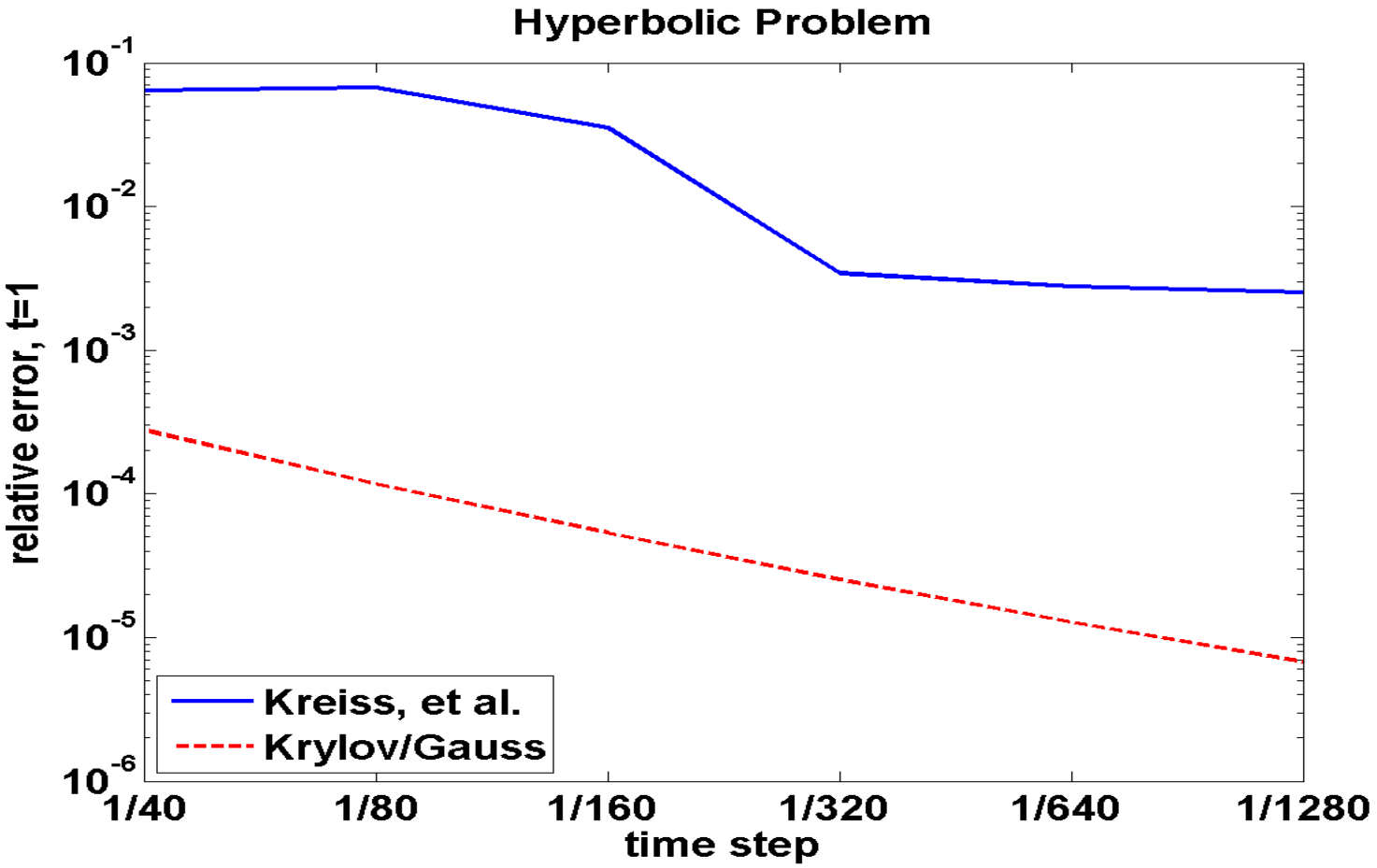
# Results for the Wave Equation




Comparison of 2-node KSS method with semi-implicit 2nd-order method of H.-O. Kreiss, N. Petersson and J. Yström, and 4th-order explicit scheme of B. Gustafsson and E. Mossberg

# Discontinuous Coefficients and Data

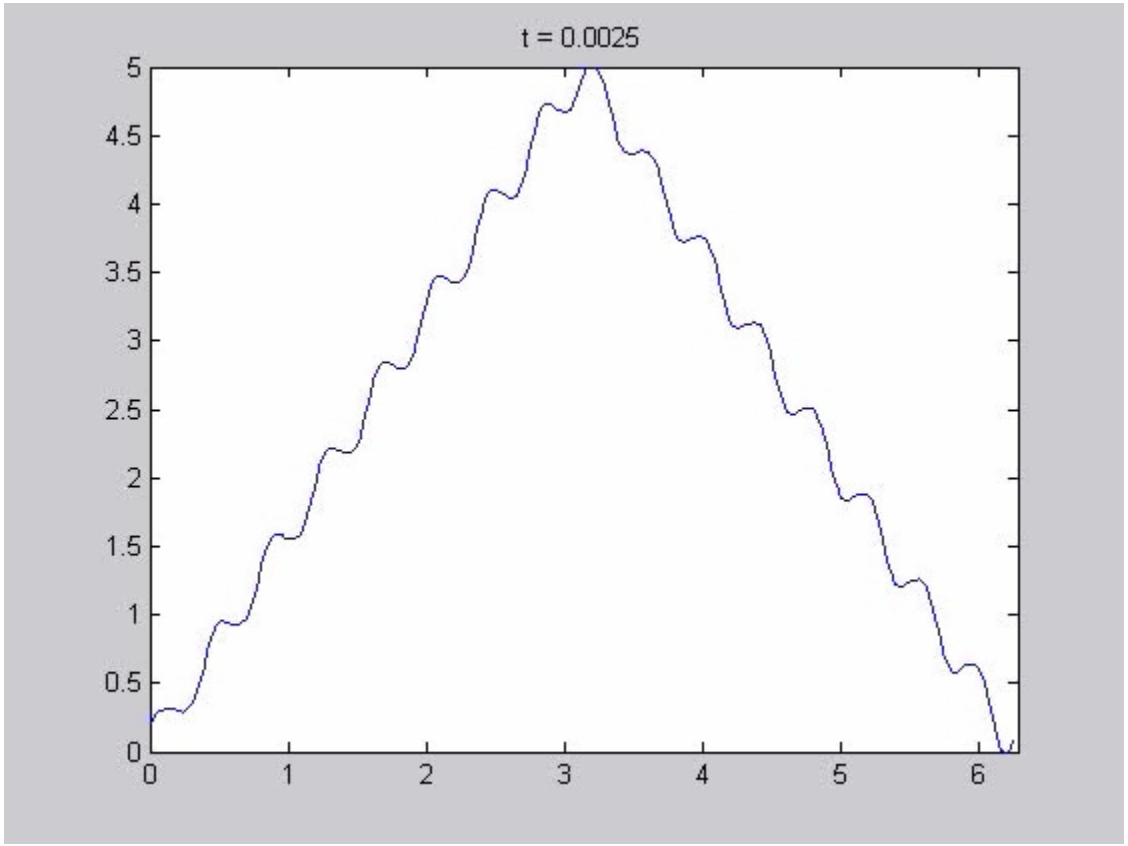


# The Modified Perona-Malik equation

- The Perona-Malik equation is a nonlinear diffusion equation used for image de-noising
- It exhibits both forward and backward diffusion, and so is ill-posed, but surprisingly well-behaved numerically
- A modification proposed by P. Guidotti weakens the nonlinearity slightly, to obtain well-posedness while still de-noising
- KSS methods can limit effects of backward diffusion by truncating recursion coefficients for selected frequencies

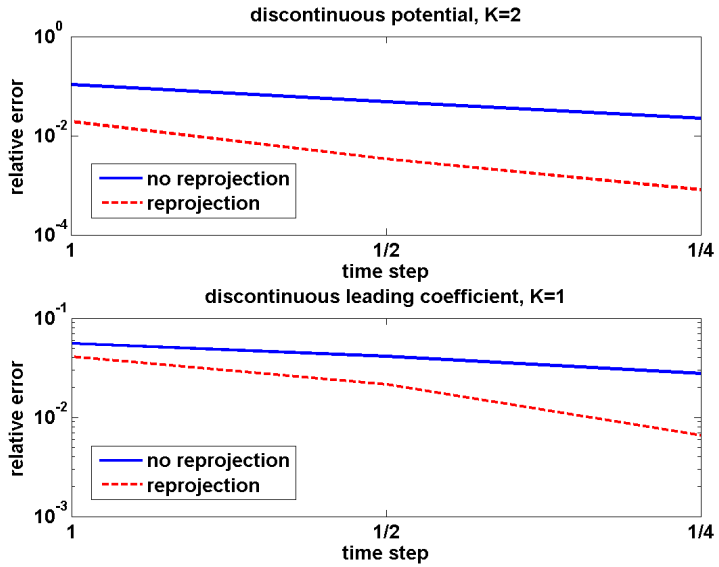
# De-Noising by Modified Perona-Malik

$$u_t - (1 + [(-D^2)^{0.4} u]^2)^{-1} u_{xx} = 0$$



# Handling Discontinuities

- In progress: other bases (e.g. wavelets, multiwavelets) for problems with rough or discontinuous coefficients
- Ideally, trial functions should conform to the geometry of the symbol as much as possible




- Encouraging results: Freud reprojection (A. Gelb & J. Tanner '06; code by A. Nelson) to deal with Gibbs phenomenon
- Improves accuracy for heat equation, stability for wave

# Preconditioning Through Homogenizing Transformations

- KSS methods are most accurate when the coefficients are smooth, so trial functions are also approximate eigenfunctions
- Problem can be preconditioned by applying unitary similarity transformations that homogenize coefficients
- In 1-D, leading coefficient $p(x)$ is easily homogenized by a change of independent variable; to make unitary, add a diagonal transformation

# Continuing the Process

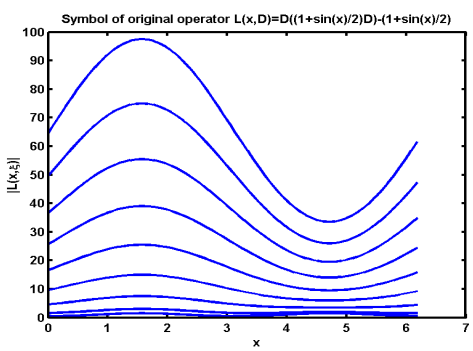
- To homogenize $q(x)$, we use a transformation of the form $L_1(x,D) = U(x,D)^* L(x,D) U(x,D)$:

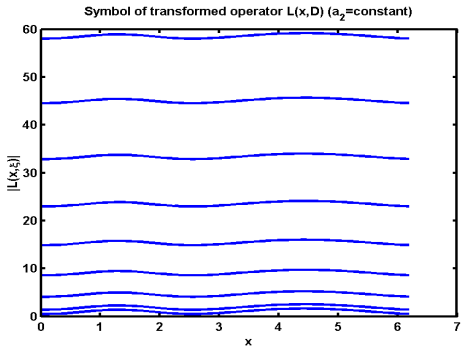
$$\begin{aligned} U(x,D) &= \exp[\phi(x,D) - \phi(x,D)^*], \\ \phi(x,D) &= -\frac{1}{4p}[D^+ q(x)] D^+, \end{aligned}$$

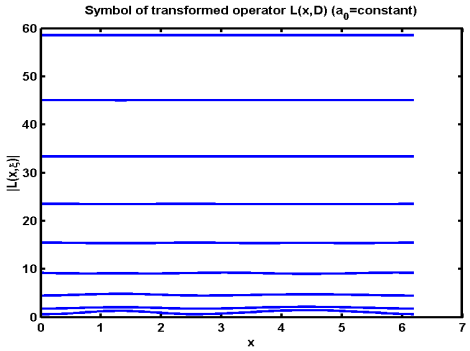
  where $D^+$ is the pseudo-inverse of $D$
- This introduces variable coefficients of order $-2$, but process can be repeated
- Generalizes to higher dimensions, can be applied efficiently using Fast FIO algorithms (Candes, Demanet and Ying '06)

# Smoothing

These transformations yield an operator that is nearly constant-coefficient.




The gain in accuracy is comparable to that achieved by deferred correction with an analytically ***exact*** residual, which KSS methods naturally provide!

# In Progress: Systems of Equations

- Generalization to systems of equations is straightforward
- For a system of the form $\mathrm{u}_t = A(x,D)\mathrm{u}$, where each entry of $A(x,D)$ is a differential operator, we can use trial functions $\mathrm{v}_j \oplus e^{i\omega \mathrm{x}}$ where $\mathrm{v}_j$ is an eigenvector of $\mathrm{Avg}_x\, A(x,\omega)$
- For systems arising from acoustics, in which the solution operator can be expressed in terms of products of entries of $A(x,D)$, order $O(\Delta t^{4K})$ accuracy is possible
- More on this topic at ICNAAM '07

# Conclusions

- Krylov subspace spectral methods are showing more promise as their development progresses
  - Applicability to problems with rough behavior
  - Stability like that of implicit methods
  - Competitive performance and scalability
- Through the splitting perspective, they provide an effective means of *stably* extending other solution methods to variable-coefficient problems
- Still much to do! Especially implementation with other bases of trial functions

# References


- JL, “Krylov Subspace Spectral Methods for Variable-Coefficient Initial-Boundary Value Problems”, *Electr. Trans. Num. Anal.* **20** (2005), p. 212-234
- P. Guidotti, JL, K. Sølna, “Analysis of 1D Wave Propagation in Inhomogeneous Media”, *Num. Funct. Anal. Opt.* **27** (2006), p. 25-55
- JL, “Practical Implementation of Krylov Subspace Spectral Methods”, *J. Sci. Comput.* **32** (2007), p. 449-476
- JL, “Derivation of High-Order Spectral Methods for Time-Dependent PDE Using Modified Moments”, submitted

# Contact Info

James Lambers

lambers@stanford.edu

http://www.stanford.edu/~lambers/