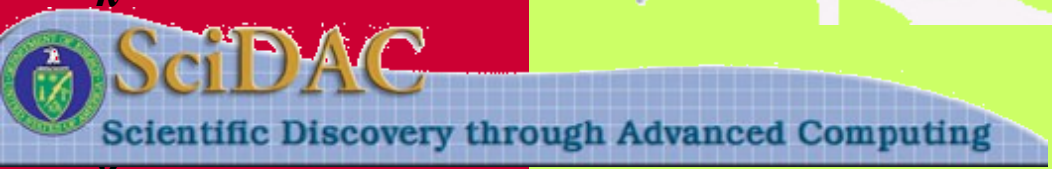
Krylov Deferred Corrections

unJia



Krylov Deferred Corrections

- What is an OPTIMAL stepsize?
- High Order/Spectral Methods for Initial Value Problems? Why? Why NOT?
- MoL or MoL^T?

Fundamentals I

Newton-Krylov Methods

How to solve f(x)=0?

Newton's Method

Newton's method:

$$x_{n+1} = x_n - \delta x$$

where $Jf(x_n)\,\delta x = b$ $\quad (b = f(x_n))$

Assume x has N unknowns, then in general each iteration requires O(N^3) work.

Can we do better?

Krylov Subspace Methods

Suppose Jf(xn)=I-C, and most eigenvalues of C are clustered close to 0, then we can search for the optimal solution in the Krylov subspace defined by

$$K_q(C,b)=\{b, Cb, C^2b, \ldots, C^qb\}$$

This should converge very fast! Consider Neumann series $1/(1-a)=1+a+a^2+a^3+\ldots$

Newton-Krylov Methods

- Newton Method and Krylov methods can be intertwined.

- Efficient implementation requires
  - A formulation such that Jf is close to I (preconditioning)
  - Efficient way to evaluate Jfb (fast matrix vector product). Here, a difference approximation can be used.

$$Jf(xn)b=(f(xn+hb)-f(xn))/h$$

Reference: C. T. Kelley, Solving Nonlinear Equations with Newton's Method, SIAM, 2003

Fundamental II.

Deferred Correction Methods.

Consider ODE initial value problem

$$\begin{cases} \varphi'(t) = f(t, \varphi), \\ \varphi(0) = \varphi_0. \end{cases}$$

Classical Deferred Correction

- Iterated defect corrections, Zadunaisky, 1964;
- Iterated deferred corrections, Peyrera, 1967.

Idea: Iteratively improve the approximate solution using a low order method.

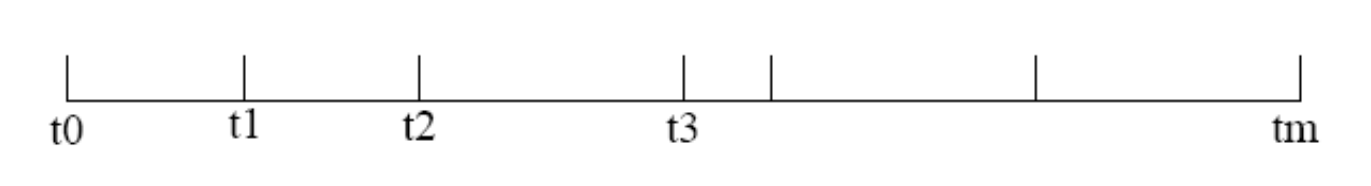
Classical Deferred Correction

Divide one “BIG" time step [Tn, Tn+1] (= [0, T] = [t0, tm]) into m smaller subintervals by using points ti;

Step 0: [Compute initial approximation]

Using a kth order method, compute an approximate solution $\Phi i^{[0]} \sim \Phi(ti)$ at the grid points on the interval [0, T].

Example: Use the backward Euler's method.



Classical Deferred Correction

Step 1: [Compute successive corrections]

1) Compute an interpolating polynomial P(t) of $\Phi i^{[0]}$.

2) Define the error function $\delta(t) = \Phi(t)-P(t)$.

3) Form the error equation

$$\delta'(t) = f(t, \delta(t) + P(t)) - P'(t)$$

$$\delta(0) = 0.$$

4) Using a Kth order method, compute an approximate solution πi at the grid points ti.

5) Define a new approximate solution $\Phi i^{[1]} = \Phi i^{[0]} + \pi$

6) Go back to 1) if necessary.

Classical Deferred Correction

After J corrections, error has order

$$O(h^{\min(K+J\,k,m)}) .$$

- Not recommended for higher order (>8):

-The equal spaced interpolation error may increase without bound near the ends of the interpolation interval due to oscillations as the degree of the polynomial increases;

-Numerical differentiation loses accuracy. (L. N. Trefethen and M. R. Trummer, An instability phenomenon in spectral methods, SIAM J. Numer. Anal., 24 (1987).

- How can we overcome this?

Spectral Deferred Correction

- Gaussian quadrature nodes.

  Remark: Interpolation at Gaussian nodes = Orthogonal polynomial + Gaussian quadrature

- The Picard integral equation:

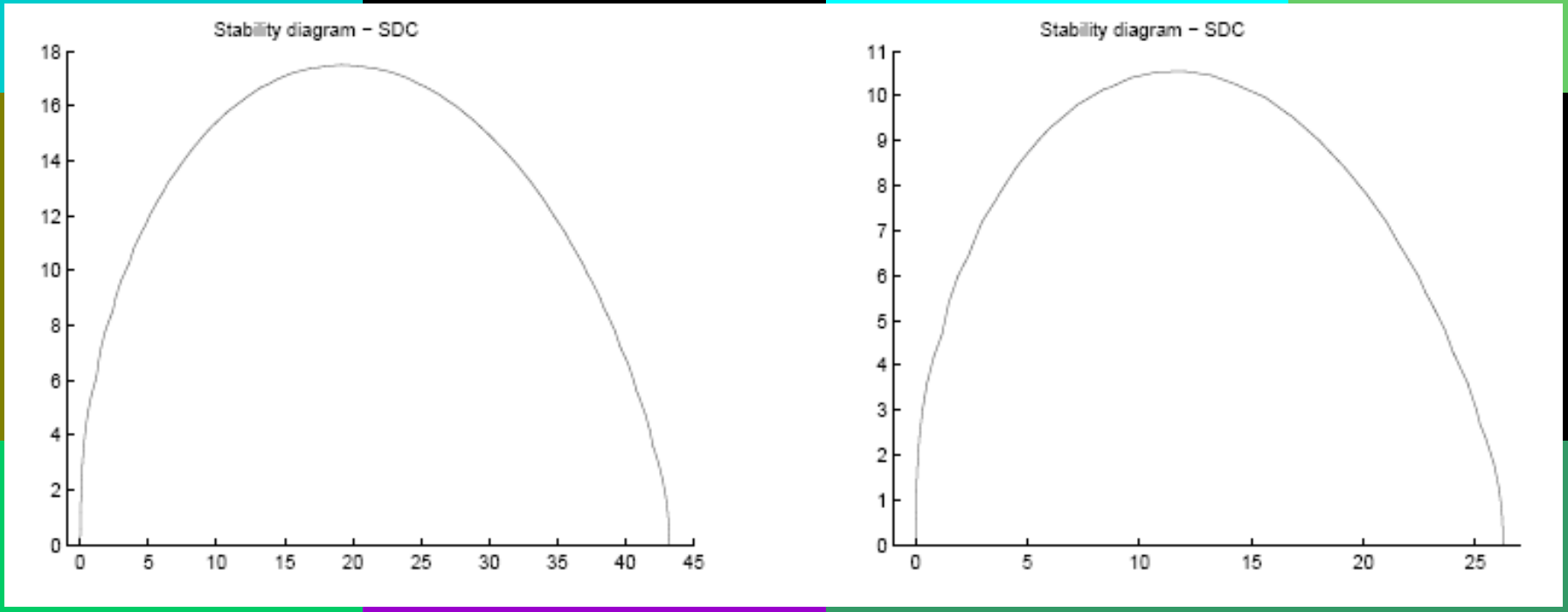
$$\Phi(t) = \Phi(0) + \int f(\tau, \Phi(\tau))d\tau$$

Remark: This avoids numerical differentiation.

- Reference: A. Dutt, L. Greengard, and V. Rokhlin, Spectral deferred correction methods for ordinary differential equations, BIT, 40(2), 2000.12.

SDC works!

- Extremely high order and good stability properties can be derived by implicit methods.
- Left one is for 20th order; right one is for 12th order:



From the Original Paper

- “Our preliminary tests indicate that the schemes with orders between 8 and 20 are roughly competitive with the best existing ones”.
- “If actual accuracies are compared, rather than requested precision, a slightly different picture begins to emerge. The singly implicit deferred correction scheme achieves about eight digits of accuracy at a requested tolerance of 1e-5, using 4, 839 function calls. EULSIM, on the other hand, achieves eight digits of accuracy at a requested tolerance of 1e-10, using 10, 490 function calls”.
- “Another interesting comparison can be made at ten digits of accuracy. The singly implicit deferred correction scheme requires 5, 887 function calls while the RADAU code of Hairer and Wanner [11], which is more efficient in this regime than EULSIM, requires 6, 517”.

Krylov Deferred Correction Methods

Problems with SDC.

1.Order Reduction

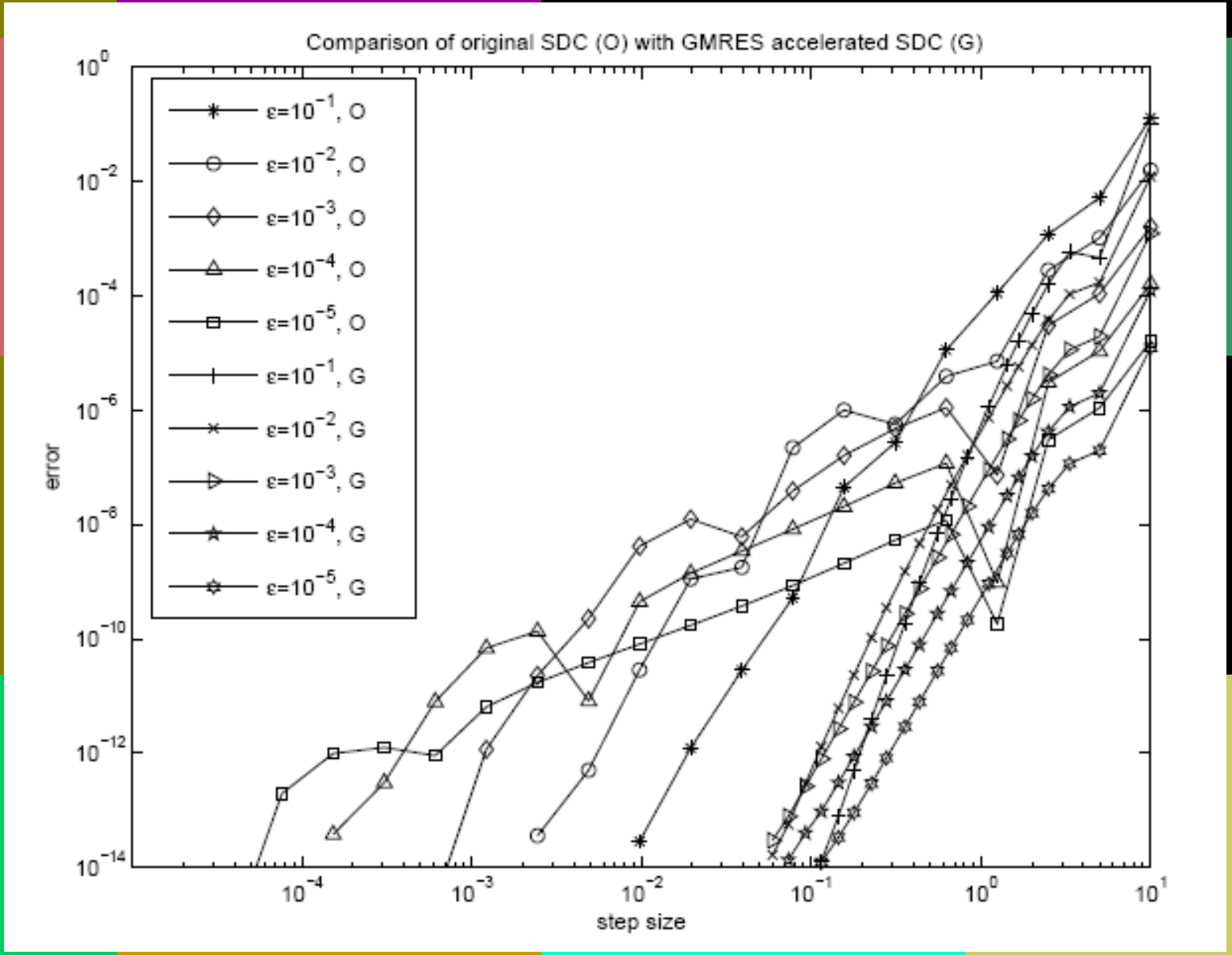
Consider the “stiff” system

$$\varphi'(t) = p'(t) - \frac{1}{\varepsilon}(\varphi(t) - p(t)),$$
$$\varphi(0) = p(0).$$

Stiff: at least two scales in the system.

RK methods suffer from this problem too.

SDC (and other classical solver) results



Problems with SDC.

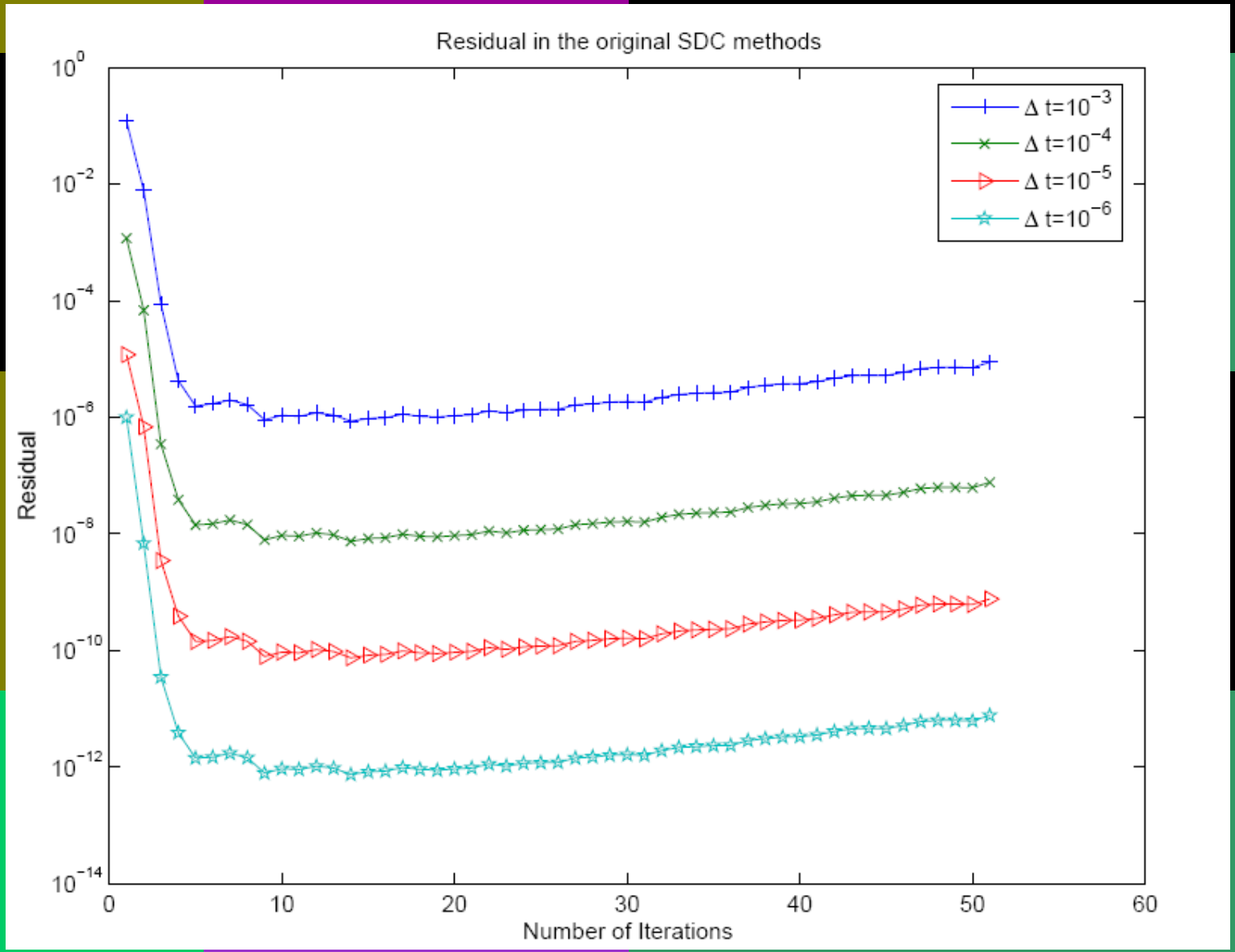
2. For DAE problems, the method may be divergent!!!

What is DAE: $F(Y,Y',t)=0$

Example: $y'(t)=f(y,z)$

$0=g(y,z)$

SDC is divergent!



Linear Problem Study

Consider

$$f(t, \varphi) = L\varphi + g(t)$$

The collocation formulation:

$$\boxed{\vec{\varphi_0} + hS\vec{f} - \vec{\varphi} = 0.}$$

$$\Updownarrow$$

$$\delta - hSL\delta = \vec{\varphi_0} + hSf(\vec{t}, \varphi^{[0]}) - \varphi^{[0]}.$$

$$\Updownarrow$$

$$\boxed{(I - hSL)\delta = \vec{\epsilon}}$$

Spectral Integration

Given (t_i, g_i) at the Gaussian nodes in $[0, h]$, let $P(x)$ be the interpolating polynomial, evaluate $\int_0^x P(y)dy$ at $t_1, t_2, \cdots t_m$ and denote the values by G_i.

We call the linear mapping from f_i to F_i the spectral integration matrix.

$$\vec{G} = hS\vec{g}$$

The discrete residual is given by

$$\vec{\varepsilon} = \vec{\varphi_0} + hS\vec{f} - \vec{\varphi^{[0]}}.$$

Linear Problem Study

Therefore the low order method solves

$$\boxed{\left(I - h\tilde{S}L\right)\bar{\delta} = \vec{\epsilon}}$$

where

$$h\tilde{S} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ h_1 & 0 & \cdots & 0 & 0 \\ h_1 & h_2 & \cdots & 0 & 0 \\ h_1 & h_2 & \cdots & 0 & 0 \\ . & . & \cdots & 0 & 0 \\ h_1 & h_2 & \cdots & h_m & 0 \end{bmatrix}$$

Linear Problem Study

For backward Euler method:

$$h\tilde{S} = \begin{bmatrix} h_0 & 0 & \cdots & 0 & 0 \\ h_0 & h_1 & \cdots & 0 & 0 \\ h_0 & h_1 & \cdots & 0 & 0 \\ . & . & \cdots & 0 & 0 \\ h_0 & h_1 & \cdots & h_{m-1} & 0 \\ h_0 & h_1 & \cdots & h_{m-1} & h_m \end{bmatrix}.$$

Note: $\tilde{S}$ is a lower triangular approximation of S and the continuous operator $\int_0^x$.

Linear Problem Study

For backward Euler method:

$$h\tilde{S} = \begin{bmatrix} h_0 & 0 & \cdots & 0 & 0 \\ h_0 & h_1 & \cdots & 0 & 0 \\ h_0 & h_1 & \cdots & 0 & 0 \\ . & . & \cdots & 0 & 0 \\ h_0 & h_1 & \cdots & h_{m-1} & 0 \\ h_0 & h_1 & \cdots & h_{m-1} & h_m \end{bmatrix}.$$

Note: $\tilde{S}$ is a lower triangular approximation of S and the continuous operator $\int_0^x$.

What we want vs What we can

This is what we want to solve:

$$\boxed{(I - hSL)\delta = \vec{\epsilon}}$$

This is what we can solve:

$$\boxed{\left(I - h\tilde{S}L\right)\bar{\delta} = \vec{\epsilon}}$$

$S \approx \tilde{S} \approx \int_0^x$

Question: What is SDC?

Low Order Preconditioner

For linear problems, apply the preconditioner

$$\left(I - h\tilde{S}L\right)^{-1}$$

to

$$\boxed{(I - hSL)\delta = \vec{\epsilon},}$$

the resulting equation takes the form

$$(I - C)\delta = \delta^{[1]}.$$

This can be solved using Neumann Series expansion

$$\delta = \sum_{j=1}^{\infty} C^{j-1}\delta^{[1]}.$$

Why Order Reduction

- When there are “bad” eigenvalues in L, Neumann series will converge slowly.

$$C = \left(I - h\tilde{S}L\right)^{-1} h(S - \tilde{S})L$$

SDC: Pros and Cons

- Pro: Low order method is a good preconditioner for high order method.

- Con: Neumann series type iteration may converge very slowly.

Krylov Deferred Correction Methods

Idea:

1. Use the Good preconditioner
2. Compute the solution using Newton-Krylov methods, instead of Neumann series expansion.

KDC Method for DAEs

- Consider a g DDEs are index 0 DAE)

$$F(y(t), y'(t), t) = 0.$$

Define Y(t)=y'(t) as the new unknown, we have a Picard type equation

$$F(y_0 + \int_0^t Y(\tau)d\tau, Y(t), t) = 0$$

Its discretized version is

$$\mathbf{F}(\mathbf{y}_0 + \Delta t S \otimes \mathbf{Y}, \mathbf{Y}, \mathbf{t}) = \mathbf{0},$$

Krylov Deferred Correction

- Use Newton-Krylov solvers to find the zero of the preconditioned system.
- The function evaluation for each iteration is simplify one SDC correction.

Generalization to PDEs

- Consider PDEs of the form

$$\mathcal{L}\left(u_t, u, u_x, u_{xx}\right) = 0$$

- First discretize in time using Gaussian nodes, and define

$$U_i(x) = U(x, t_i) = u_t(x, t_i)$$

- The discretized high order (pseudo-spectral) formulation is

$$\mathbf{L}\left(\mathbf{U}, \mathbf{u_0} + \Delta t S \otimes \mathbf{U}, \frac{d}{dx}\left(\mathbf{u_0} + \Delta t S \otimes \mathbf{U}\right), \frac{d^2}{dx^2}\left(\mathbf{u_0} + \Delta t S \otimes \mathbf{U}\right)\right) = \mathbf{0}$$

KDC Accelerated MoLT

- The error equation is given by

$$\mathbf{L}(\tilde{\mathbf{U}}+\boldsymbol{\delta}, \mathbf{u_0}+\Delta t S \otimes (\tilde{\mathbf{U}}+\boldsymbol{\delta}), \frac{d}{dx}(\mathbf{u_0}+\Delta t S \otimes (\tilde{\mathbf{U}}+\boldsymbol{\delta})), \frac{d^2}{dx^2}(\mathbf{u_0}+\Delta t S \otimes (\tilde{\mathbf{U}}+\boldsymbol{\delta}))) = \mathbf{0}.$$

- Assume it is easy to derive the low order solution

$$\mathbf{L}(\tilde{\mathbf{U}}+\bar{\boldsymbol{\delta}}, \mathbf{u_0}+\Delta t S \otimes \tilde{\mathbf{U}}+\Delta t \tilde{S} \otimes \bar{\boldsymbol{\delta}}, \frac{d}{dx}(\mathbf{u_0}+\Delta t S \otimes \tilde{\mathbf{U}}+\Delta t \tilde{S} \otimes \bar{\boldsymbol{\delta}}), \frac{d^2}{dx^2}(\mathbf{u_0}+\Delta t S \otimes \tilde{\mathbf{U}}+\Delta t \tilde{S} \otimes \bar{\boldsymbol{\delta}})) = \mathbf{0}$$

- We have a preconditioned system

$$\bar{\boldsymbol{\delta}} = \tilde{\mathbf{H}}(\tilde{\mathbf{U}})$$

KDC Accelerated MoLT: Pros and Cons

Pros:

- High Order in time, with optimized stepsize.
- Can use existing adaptive parallel elliptic equation solvers (MADNESS, FMM accelerated IEM)

Cons:

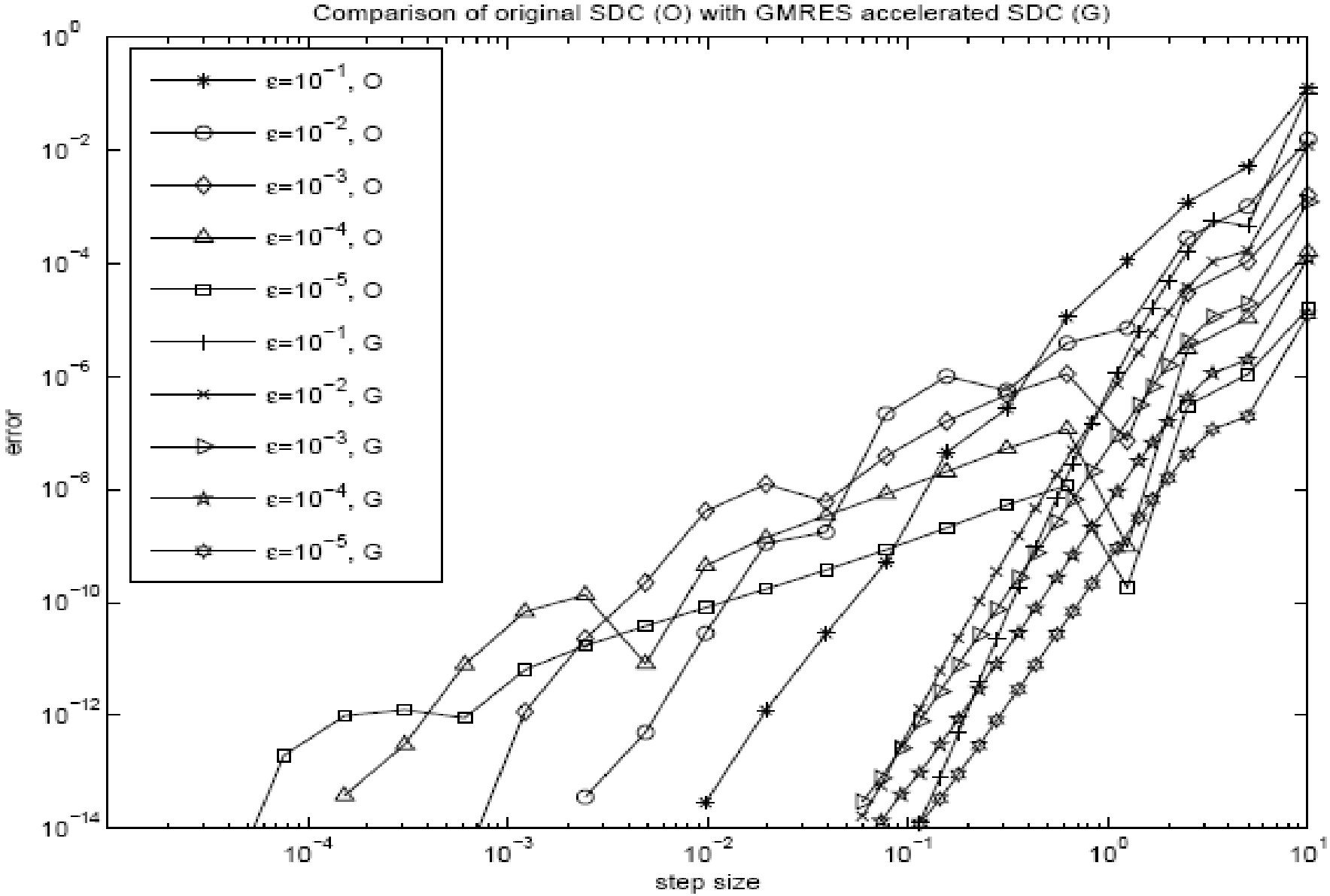
- Further analysis needed.
- Code development.

Analytical Results

Theorem:

For ODE problems, the KDC method using p Gaussian points is order 2p (super convergence), A-stable, B-stable, L-stable, symplectic (structure preserving), and symmetric (time reversible).

Numerical Results



The Ring Modulator Problem

	G-SDC	DASSL	GAMD	MEBDFI	PSIDE	RADAU	VODE
$rtol$	1e-8	1e-12	1e-10	1e-9	1e-10	1e-9	1e-11
$atol$	*	1e-12	1e-10	1e-11	1e-11	1e-10	1e-14
h_0	2.5e-6	*	1e-10	1e-10	*	1e-10	*
$rerr$	3.0e-9	1.1e-9	3.1e-9	2.9e-9	2.1e-9	2.1e-9	1.3e-9
F	1134	2104	4057	2284	3417	2172	2961
$steps$	4	1591	76	669	154	47	2277

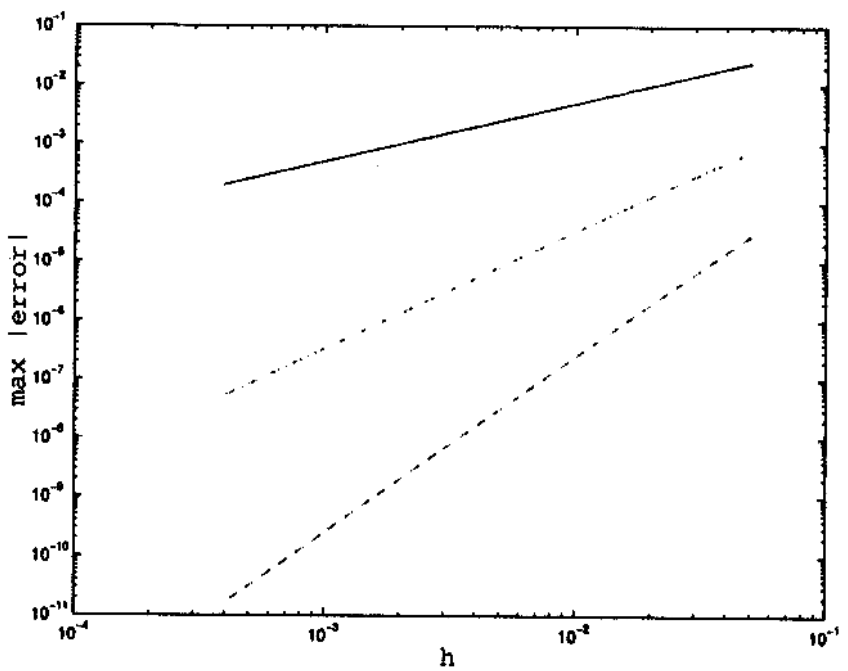
DAE Problem

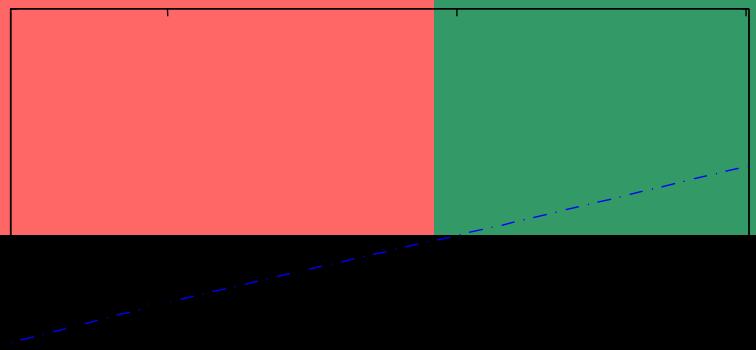
- Consider

$$My' = \begin{pmatrix} 10 - \frac{1}{2+t} & 0 & 10(2+t) \\ \frac{9}{2+t} & -1 & 9 \\ t+2 & t^2-4 & 0 \end{pmatrix} y + \begin{pmatrix} e^t \frac{3+t}{2+t} \\ 2e^t \\ e^t(2-t-t^2) \end{pmatrix},$$

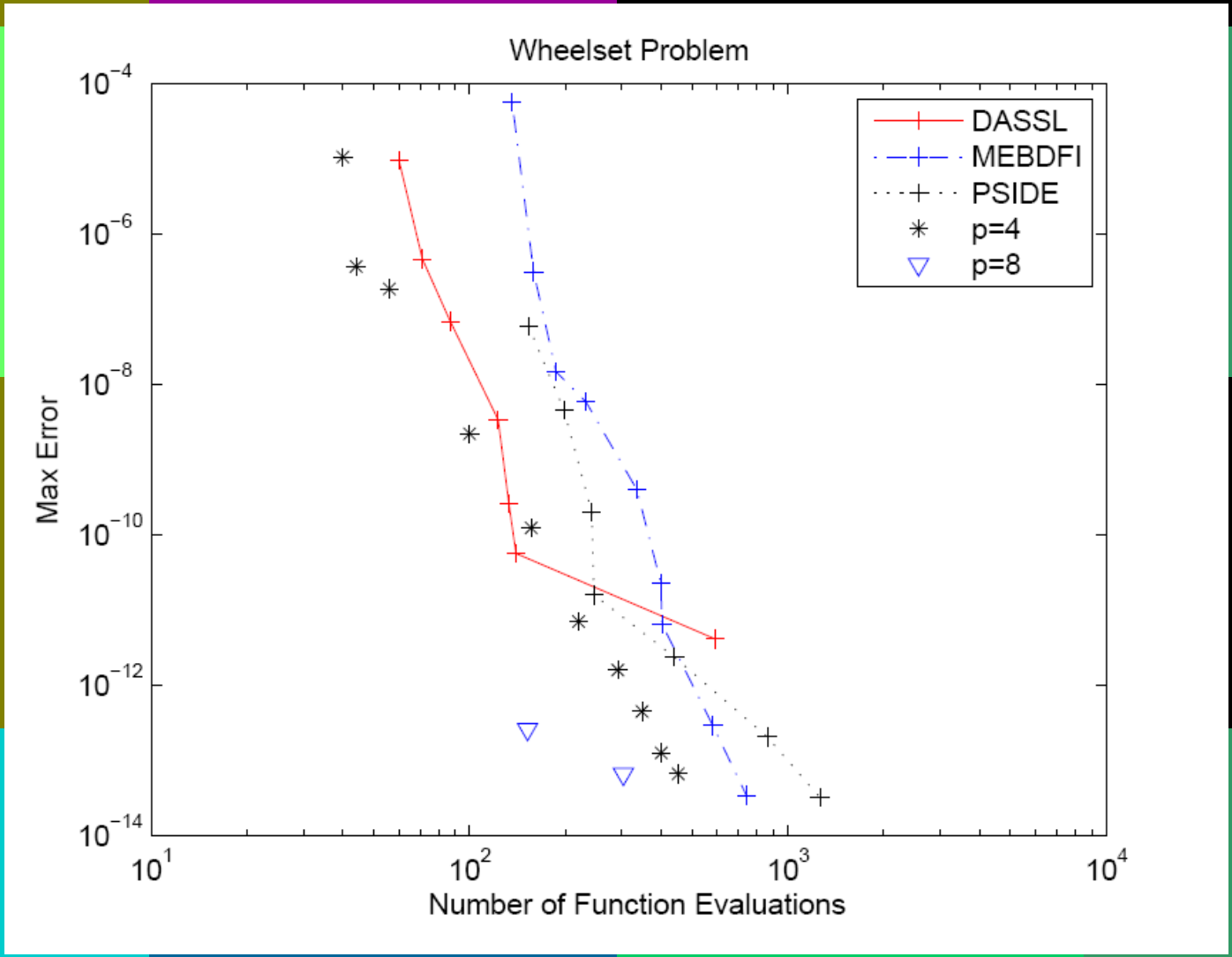
where $M = diag(1, 1, 0)$ and the real solution is $y(t) = (e^t, e^t, -e^t/(2+t))$.

Stepsize Comparison





Wheelset Problem



A Simple PDE Example

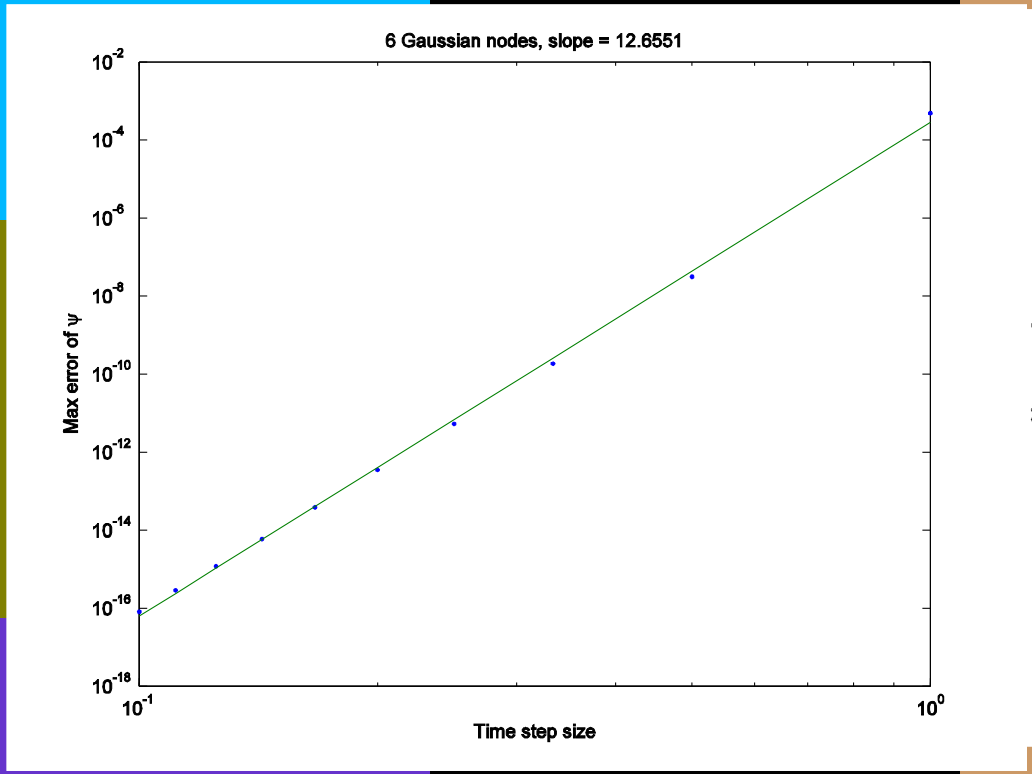
- We consider the PDE

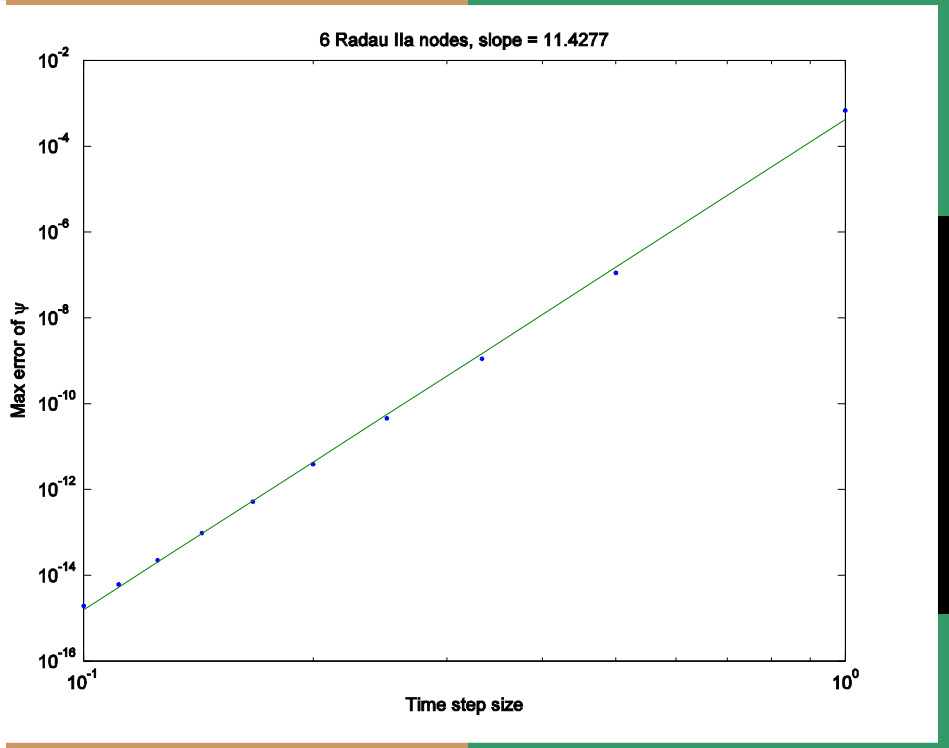
$$M(x)\frac{\partial \psi}{\partial t} = N(x)\frac{\partial^2 \psi}{\partial x^2} + f(x,t).$$

With Periodic boundary condition and exact solution

$$\psi(x,t) = e^{\cos(2\pi(x+t^2))-kt}$$

Order of the Method





Summary

- Lower order methods are good preconditioners for higher order methods.
- The resulting system can be solved efficiently using Newton-Krylov methods.

- Based on this, it is possible to design "optimal" time stepping strategies for initial value problems

Thanks