Exponential Integrators

John C. Bowman and Thoma Zoto
Department of Mathematical and Statistical Sciences
University of Alberta

June 24, 2024

www.math.ualberta.ca/∼bowman/talks

Outline

- Motivation
  - Notation
  - Stiffness
- Exponential Integrators
  - Exponential Euler Algorithm
  - History
  - Schur Decomposition
  - Robust Embedded Pairs
- Conclusions

Motivation

- Consider for $y : \mathbb{R} \to \mathbb{R}$ and $L > 0$ the equation

$$\frac{dy}{dt} = -Ly,$$

with the initial condition $y(0) = y_0 \neq 0$.

- We know that the exact solution to this equation is given by

$$y(t) = y_0 e^{-tL}.$$

- Apply Euler's method with time step h:

$$y_{n+1} = (1 - hL) y_n.$$

- For $hL \geq 2$, y_n does not converge to the steady state: if L is too large, the time step is forced to be unreasonably small.

- This phenomenon of linear stiffness manifests itself in general driven systems of ODEs in $\mathbb{R}^n$:

$$\frac{dy}{dt} + Ly = f(y).$$

- When the eigenvalues of L are large compared to the eigenvalues of f', a similar problem will occur.

Notation

$$\frac{dy}{dt} = f(t, y), \qquad y(0) = y_0,$$

- General s-stage Runge–Kutta scheme (scalar case):

$$y_{i+1} = y_0 + h \sum_{j=0}^{i} a_{ij} f(c_j h, y_j), \qquad i = 0, \ldots, s-1.$$

0 is the initial time; h is the time step;

y_s is the approximation to $y(h)$;

a_{ij} are the Runge–Kutta weights;

c_j are the step fractions for stage j.

Butcher Tableau ($s = 3$):

$$c_0 = 0, \qquad c_{i+1} = \sum_{j=0}^{i} a_{ij}.$$

$$\begin{array}{c|ccc}
0 & & & \\
c_1 & a_{00} & & \\
c_2 & a_{10} & a_{11} & \\
\hline
1 & a_{20} & a_{21} & a_{22}
\end{array}$$

Stiffness

Lambert [1991] points out problems with existing notions of stiffness in the literature, either due to the existence of a counterexample or due to their qualitative nature:

- Curtiss and Hirschfelder [1952]: *A system is said to be stiff in a given interval of time if, in that interval, neighbouring solution curves approach the solution curve at a rate which is very large in comparison with the rate at which the solution varies.*
- Lambert [1991]: *If a numerical method with a finite region of absolute stability, applied to a system with any initial conditions, is forced to use in a certain interval of integration a step-length which is excessively small in relation to the smoothness of the exact solution in that interval, then the system is said to be stiff in that interval.*

- Cartwright [1999], Zoto & JCB [2024] : *A system is* stiff *in a given interval if in that interval the most negative local Lyapunov exponent is much larger in absolute value than the curvature of the solution curve.*

- Let $\sigma_i(t)$ be the principal axes of an ellipsoid evolving in phase space.

- In terms of the i^{th} local Lyapunov exponent

$$\gamma_i(h,t) = \lim_{\sigma_i(h)\to 0} \frac{1}{h} \log \frac{\sigma_i(t+h)}{\sigma_i(t)}$$

and the curvature $\kappa = y''(1+y'^2)^{-3/2}$ of the solution y, stiffness may be quantified by the ratio

$$\frac{\left| \min_{1\le i\le n} \gamma_i(h,t) \right|}{\kappa(t)}$$

- This definition recognizes that *stiffness is a local phenomenon.*

Exponential Integrators

- Circumvent linear stiffness by applying a scheme that is exact on the time scale of the linear part of the problem.
- Consider
$$\frac{dy}{dt} + Ly = f(y).$$
- Rewrite the above equation as
$$\frac{d(e^{tL}y)}{dt} = e^{tL}f(y)$$
- and integrate to obtain
$$y(h) = e^{-hL}y(0) + \int_0^h e^{-(h-s)L}f(y(0+s))ds.$$
- A quadrature rule is used to approximate the integral, while treating the exponential term exactly.

Stiff-Order Conditions

$$y_{i+1} = e^{-hL} y_0 + h \sum_{j=0}^{i} a_{ij}(-hL) f(y_j), \quad i = 0, \ldots, s-1.$$

- The weights a_{ij} are constructed from linear combinations of e^x and truncations of its Taylor series:

$$\varphi_0(x) = e^x$$

$$\varphi_{k+1}(x) = \frac{\varphi_k(x) - \frac{1}{k!}}{x} \quad \text{for } k \geq 0,$$

with $\varphi_k(0) = \frac{1}{k!}$.

- Care must be exercised when evaluating φ near 0; see the C++ routines at `www.math.ualberta.ca/~bowman/phi.h`.

- A set of *stiff-order conditions* on the weights were shown by Hochbruck and Ostermann to be *sufficient* to avoid *order reduction* when L has large eigenvalues.

Exponential Euler Algorithm

$$y_{i+1} = e^{-hL} y_i + \frac{1 - e^{-hL}}{L} f(y_i),$$

- Also called Exponentially Fitted Euler, ETD Euler, filtered Euler, Lie–Euler.
- If it has a fixed point, it must satisfy $y = \dfrac{f(y)}{L}$; this is then a fixed point of the ODE.
- In contrast, the popular Integrating Factor method (I-Euler).

$$y_{i+1} = e^{-hL}(y_i + hf_i)$$

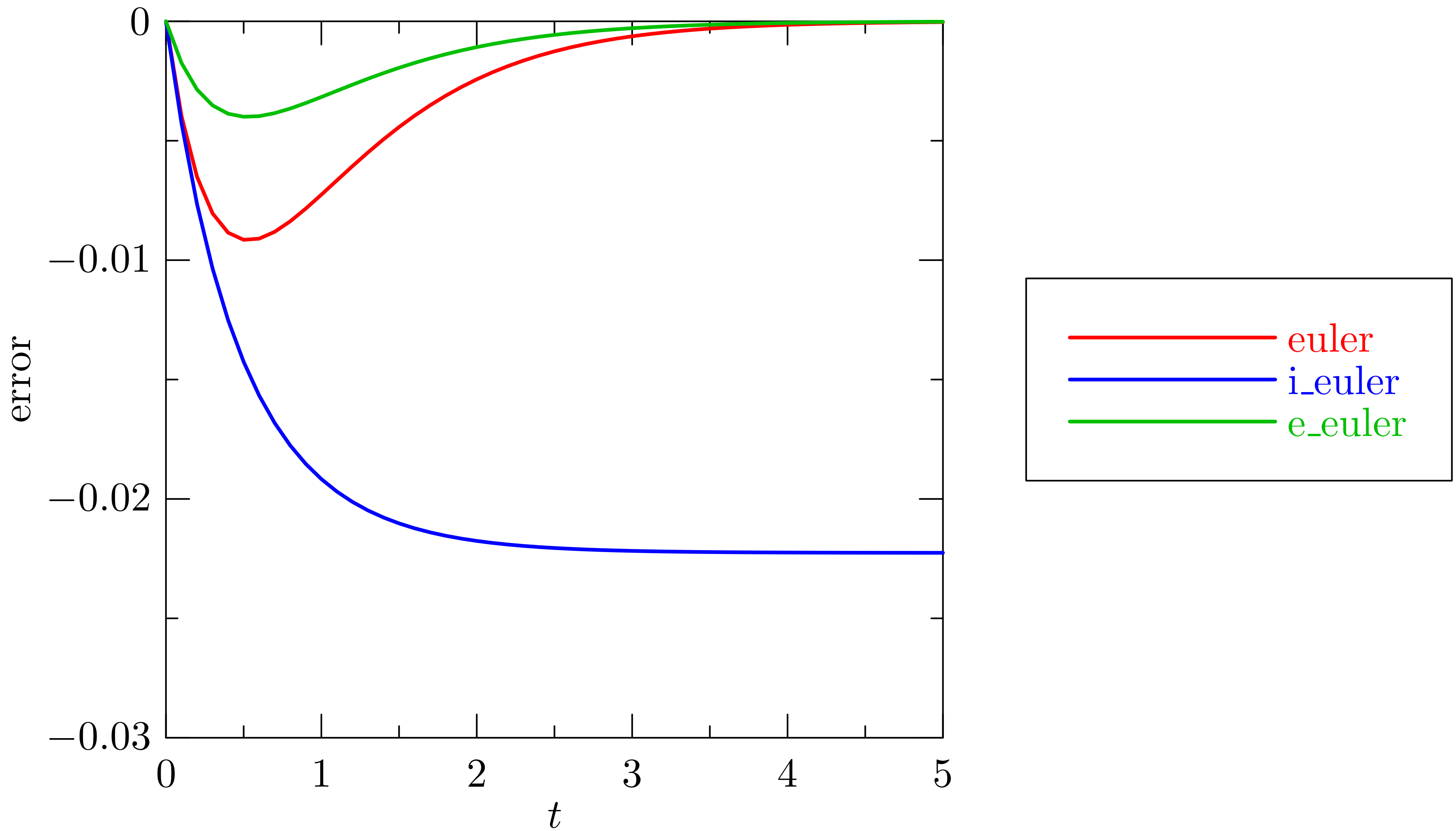
can at best have an incorrect fixed point: $y = \dfrac{hf(y)}{e^{Lh} - 1}$.

- As $h \to 0$ the Euler method is recovered:

$$y_{i+1} = y_i + hf(y_i).$$

Comparison of Euler Integrators

$$\frac{dy}{dt} + y = \cos y, \qquad y(0) = 1.$$



History

- Certaine [1960]: Exponential Adams-Moulton
- Nørsett [1969]: Exponential Adams-Bashforth
- Verwer [1977] and van der Houwen [1977]: Exponential linear multistep method
- Friedli [1978]: Exponential Runge–Kutta
- Hochbruck *et al.* [1998]: Exponential integrators up to order 4
- Beylkin *et al.* [1998]: Exact Linear Part (ELP)
- Cox & Matthews [2002]: ETDRK3, ETDRK4; worst case: stiff order 2
- Lu [2003]: Efficient Matrix Exponential
- Hochbruck & Ostermann [2005a]: Explicit Exponential Runge–Kutta; stiff order conditions.

Schur Decomposition

- When L is a nondiagonal matrix, the matrix exponentials required by exponential integrators are computationally expensive.

- Consider the Schur decomposition of L:

$$L = UTU^{\dagger},$$

where U is a *unitary* matrix and T is an upper triangular matrix.

- Decompose $T = D + S$, where D is a diagonal matrix and S is a strictly upper triangular matrix.

- We obtain

$$\frac{dy}{dt} + U(D + S)U^{\dagger}y = F(t, y).$$

- Multiply by $U^\dagger$ on the left:

$$\frac{d(U^\dagger y)}{dt} + (D+S)U^\dagger y = U^\dagger F(t,y).$$

- In terms of the transformed variable $Y = U^\dagger y$:

$$\frac{dY}{dt} + DY = U^\dagger F(t,UY) - SY.$$

- This transformation allows us to replace exponentials of a full matrix with a diagonal matrix of scalar exponentials.
- Being diagonal, the φ_k functions now require far less storage.
- Although the computation of the Schur decomposition of L is expensive, *it only has to be done once.*
- The explicit treatment of the upper triangular matrix S contributes to the overall error, but *does not contribute to stiffness.*

- Moreover, many matrices encountered in practice are *normal*: they commute with their Hermitian adjoint.
- The following theorem tells us that $S = 0$ for normal matrices:

  Theorem 1: *The triangle matrix in the Schur decomposition of a normal matrix is diagonal.*
- With the optimization afforded by Schur decomposition, embedded ERK methods for step size adjustment becomes computationally viable, *even when L is a nondiagonal matrix.*
- An adaptive exponential method requires re-evaluating the φ_k functions whenever the step size is adjusted.
- However, since these are now functions of diagonal matrices, there is no longer a huge computational cost.

Claim: The term Sy does not incorporate any of the stiffness inherent in the linear term Ly.

Proof:

- On defining the integrating factor $I(t) = e^{tD}$ and $\tilde{y}(t) = I(t)y(t)$, we can transform the autonomous case to

$$\frac{d\tilde{y}}{dt} = I(t)U^{\dagger}F(UI^{-1}(t)\tilde{y}) - \tilde{S}\tilde{y},$$

where $\tilde{S} = I(t)SI^{-1}(t)$ is an $m \times m$ strictly upper triangular matrix.

- If the stiffness only enters through the linear term Ly and not through $F(y)$, the first term on the right-hand side will not contribute any additional stiffness.

- When $F = 0$, we obtain the triangular system of equations

$$\frac{d\tilde{y}_i}{dt} = \sum_{j=i+1}^{m} \tilde{S}_{ij}\tilde{y}_j \text{ for } i = 1, \ldots, m-1 \quad \text{and} \quad \frac{d\tilde{y}_m}{dt} = 0,$$

  which can be solved recursively to obtain the general solution as a polynomial in t.

- Stiffness arises only when nearby solution curves approach the solution curve of interest at exponentially fast rates.

- Thus, the decomposed system of equations is not stiff; it can in fact be solved exactly by a classical Runge–Kutta method whose order is at least the degree of the solution polynomials.

- By linear superposition, the system is not stiff even when F is linear (and, in particular, when F is constant).

- The linear stiffness is thus *entirely contained within the diagonal term* DY.

Hochbruck–Ostermann Test Problems

- For $x \in [0,1]$ and $t \geq 0$:

$$\frac{\partial y}{\partial t}(x,t) - \frac{\partial^2 y}{\partial x^2}(x,t) = H(x,t) + \Phi(x,t).$$

Problem 1:

$$H(x,t) = \int_0^1 y^N(\bar{x},t)\, d\bar{x},$$

where H–O choose $N = 1$.

Problem 2:

$$H(x,t) = \frac{1}{1 + y(x,t)^2}.$$

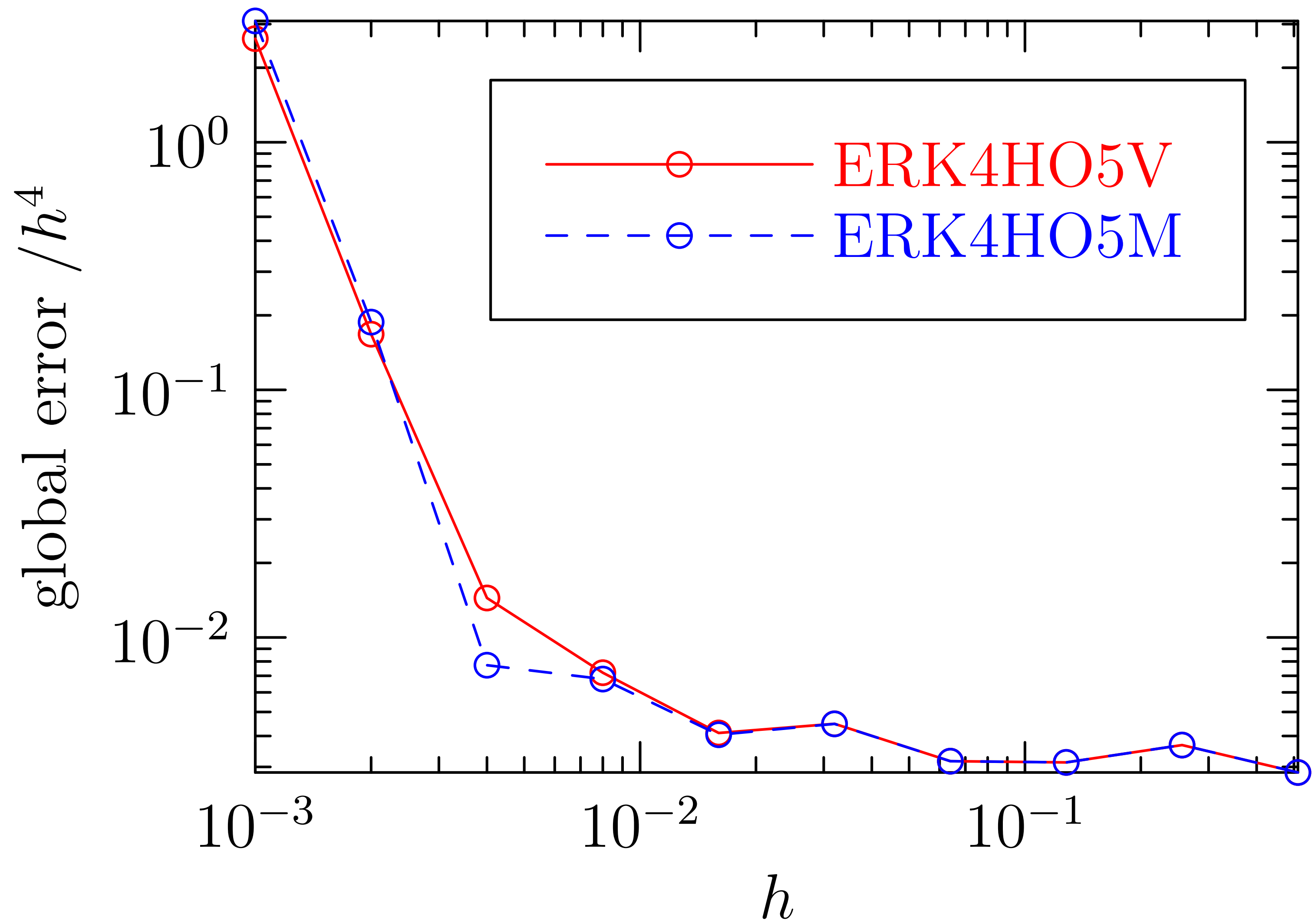
- Φ is chosen so that the exact solution is

$$y(x,t) = x(1-x)e^t.$$

- 200 spatial grid points, evolve from $t = 0$ to $t = 3$.

- A Simpson method is used to approximate the integral in Problem 1.
- The resulting matrix-vector multiplication is a linear term that could be combined with the linear term coming from the discretized Laplacian.
- If the linearities were combined, all exponential integrators would solve this problem exactly!
- The choice $N = 4$ prevents the integral from being combined into the linearity, providing a stricter test.
- We calculate the matrix φ_k functions with the help of Padé approximants, along with scaling and squaring.

Prob. 1: Schur Decomposition vs. Full Solution



(4,3) Robust Embedded Pair ERK43ZB

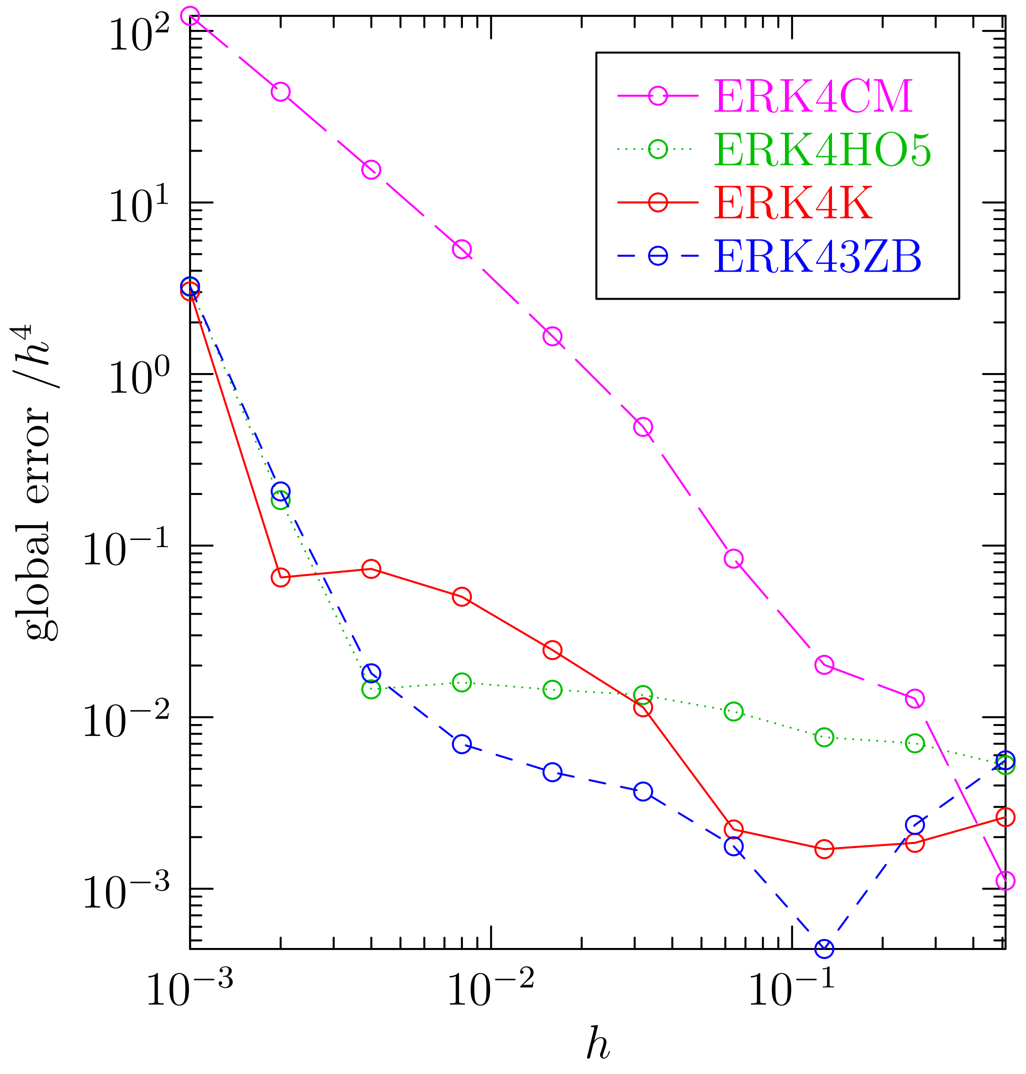
$$
\begin{array}{c|ccccc}
0 & & & & & \\
\frac{1}{6} & \frac{1}{6}\varphi_1\left(-\frac{hL}{6}\right) & & & & \\
\frac{1}{2} & \frac{1}{2}\varphi_1\left(-\frac{hL}{2}\right) - a_{11} & a_{11} & & & \\
\frac{1}{2} & \frac{1}{2}\varphi_1\left(-\frac{hL}{2}\right) - a_{21} - a_{22} & a_{21} & a_{22} & & \\
\hline
1 & \varphi_1 - a_{31} - a_{32} - a_{33} & a_{31} & a_{32} & a_{33} & \\
1 & \varphi_1 - \frac{67}{9}\varphi_2 + \frac{52}{3}\varphi_3 & 8\varphi_2 - 24\varphi_3 & \frac{26}{3}\varphi_3 - \frac{11}{9}\varphi_2 & a_{43} & a_{44},
\end{array}
$$

where $\varphi_i = \varphi_i(-hL)$ and

$$
\begin{aligned}
a_{11} &= \frac{3}{2}\varphi_2\left(-\frac{hL}{2}\right) + \frac{1}{2}\varphi_2\left(-\frac{hL}{6}\right)\\
a_{21} &= \frac{19}{60}\varphi_1 + \frac{1}{2}\varphi_1\left(-\frac{hL}{2}\right) + \frac{1}{2}\varphi_1\left(-\frac{hL}{6}\right)\\
&\quad + 2\varphi_2\left(-\frac{hL}{2}\right) + \frac{13}{6}\varphi_2\left(-\frac{hL}{6}\right) + \frac{3}{5}\varphi_3\left(-\frac{hL}{2}\right)\\
a_{22} &= -\frac{19}{180}\varphi_1 - \frac{1}{6}\varphi_1\left(-\frac{hL}{2}\right) - \frac{1}{6}\varphi_1\left(-\frac{hL}{6}\right)\\
&\quad - \frac{1}{6}\varphi_2\left(-\frac{hL}{2}\right) + \frac{1}{9}\varphi_2\left(-\frac{hL}{6}\right) - \frac{1}{5}\varphi_3\left(-\frac{hL}{2}\right)\\
a_{33} &= \varphi_2 + \varphi_2\left(-\frac{hL}{2}\right) - 6\varphi_3 - 3\varphi_3\left(-\frac{hL}{2}\right)\\
a_{31} &= 3\varphi_2 - \frac{9}{2}\varphi_2\left(-\frac{hL}{2}\right) - \frac{5}{2}\varphi_2\left(-\frac{hL}{6}\right) + 6a_{33} + a_{21}\\
a_{32} &= 6\varphi_3 + 3\varphi_3\left(-\frac{hL}{2}\right) - 2a_{33} + a_{22}\\
a_{43} &= \frac{7}{9}\varphi_2 - \frac{10}{3}\varphi_3, \qquad a_{44} = \frac{4}{3}\varphi_3 - \frac{1}{9}\varphi_2.
\end{aligned}
$$

Prob. 1: Fixed-Timestep Methods

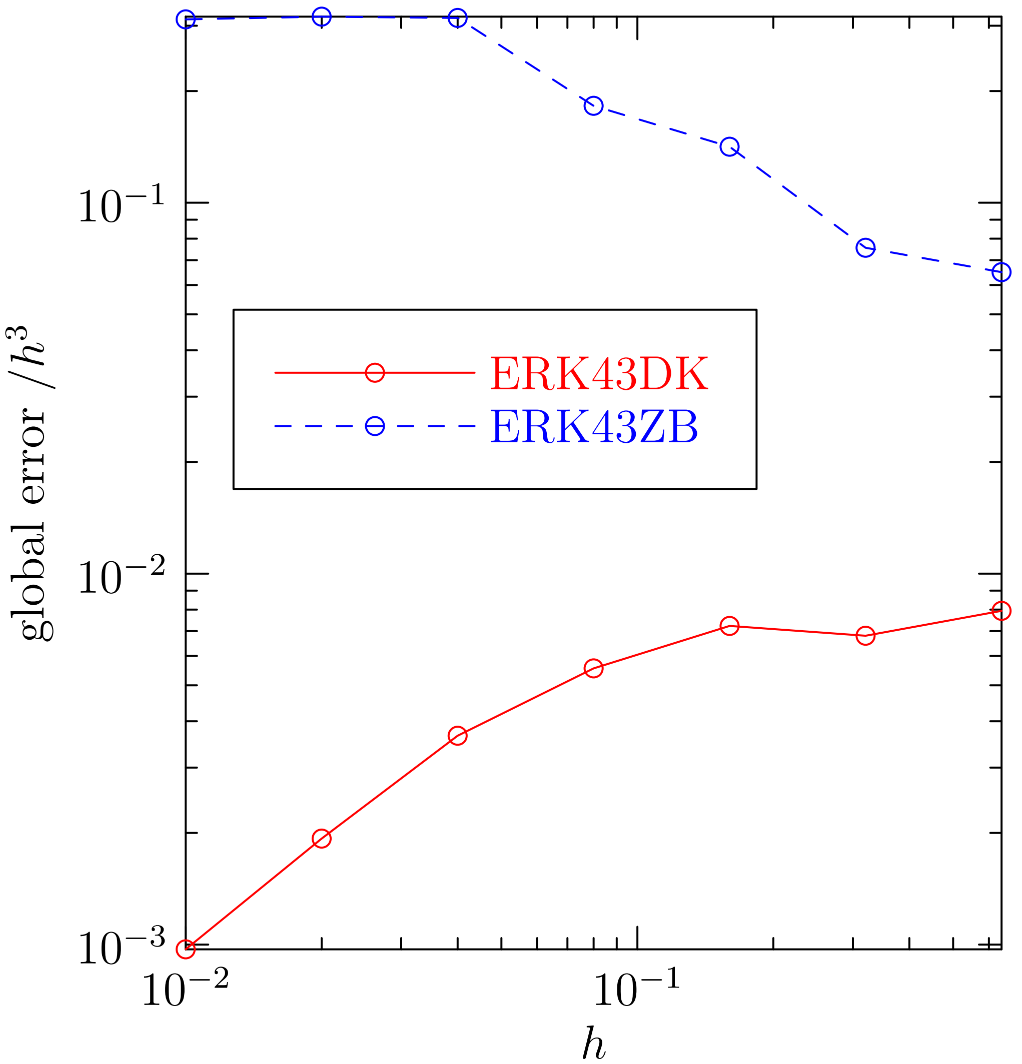
- ERK4CM and ERK4K exhibit stiff-order reduction:



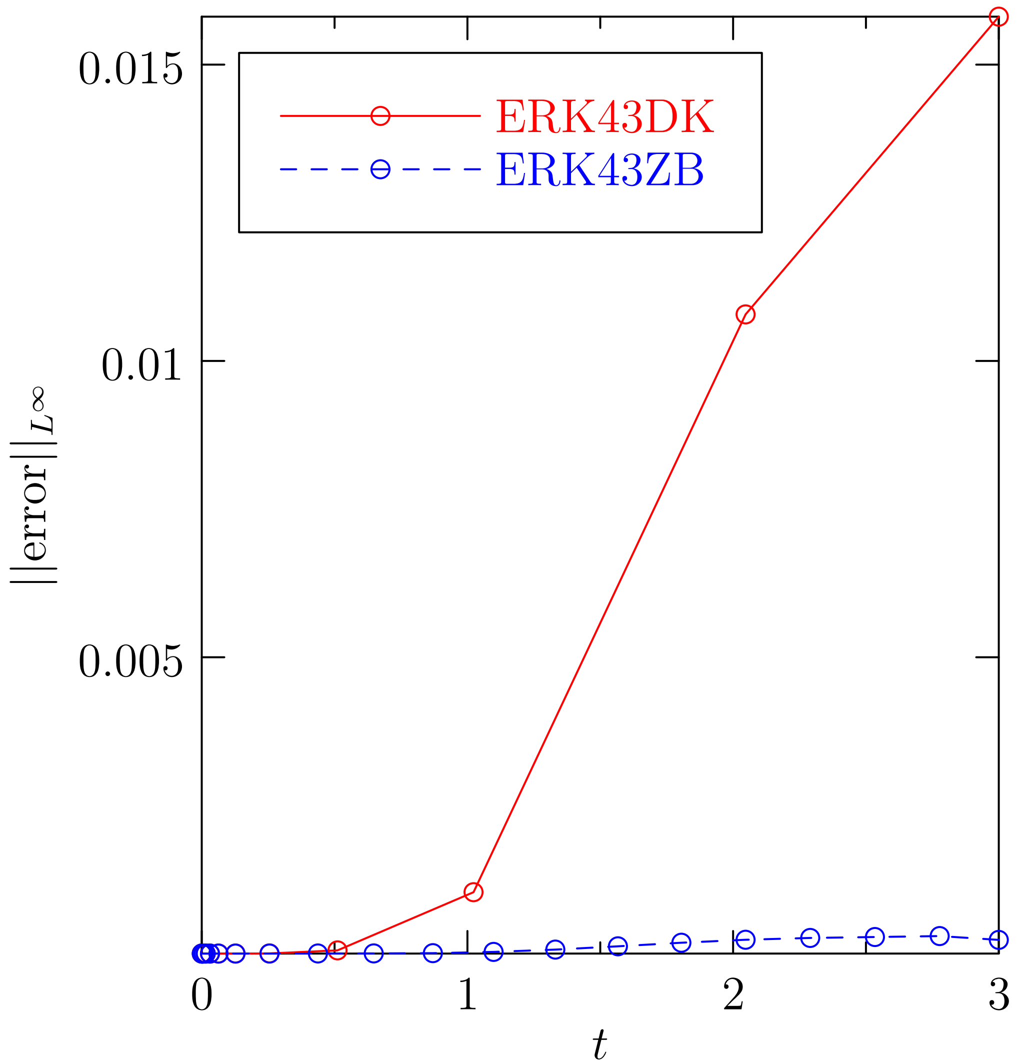
Robustness

- An adaptive pair is *robust* if the order of the low-order method is never equal to the order n of the high-order method for any source function $G(t) = F(t, y(t))$ with a nonzero derivative of order less than n.
- A nonrobust method can mislead the time step adjustment algorithm into adopting too large a time step, leading to catastrophic loss of accuracy.
- We illustrate nonrobustness with Hochbruck–Ostermann Test Problem 2.

Robust vs. Non-Robust Third-Order Estimate

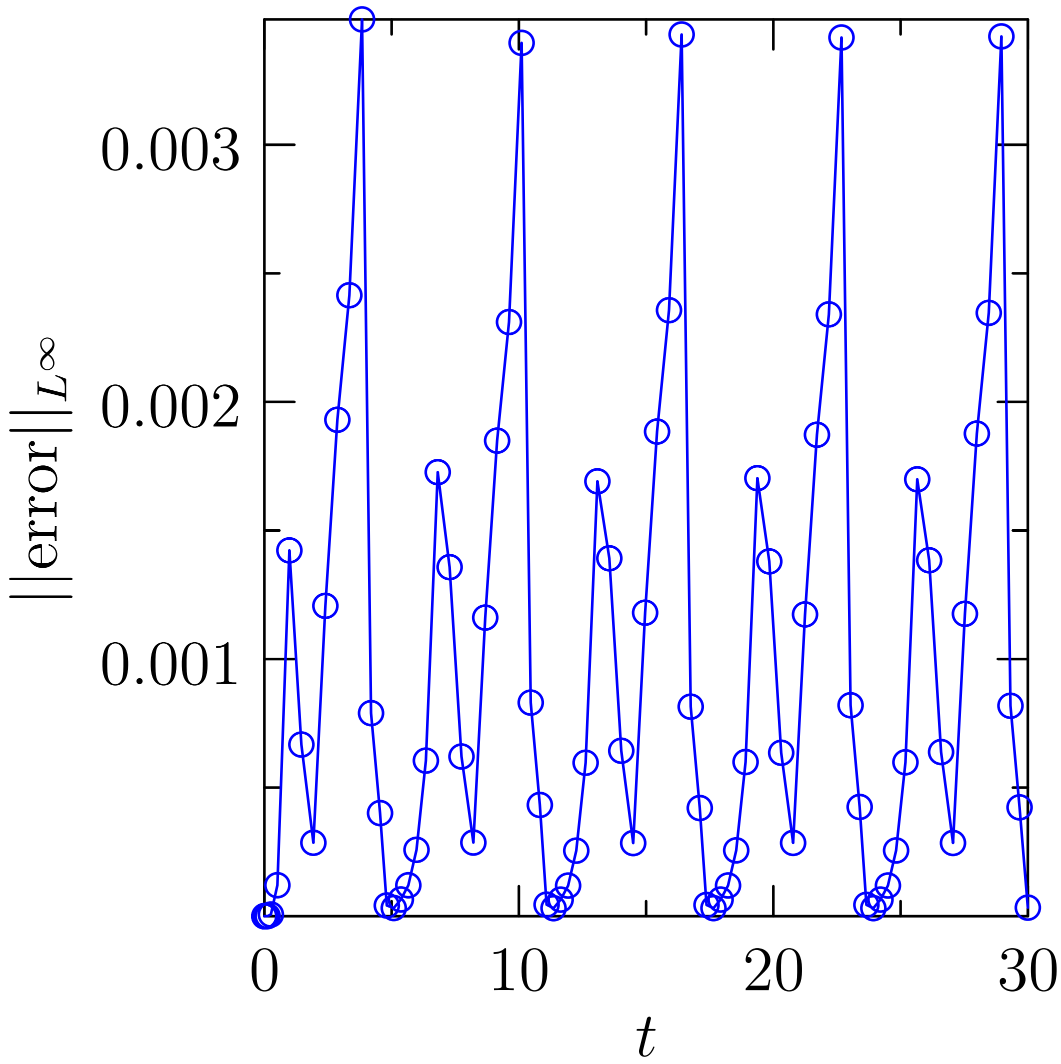


Robust vs. Non-Robust Time Evolution

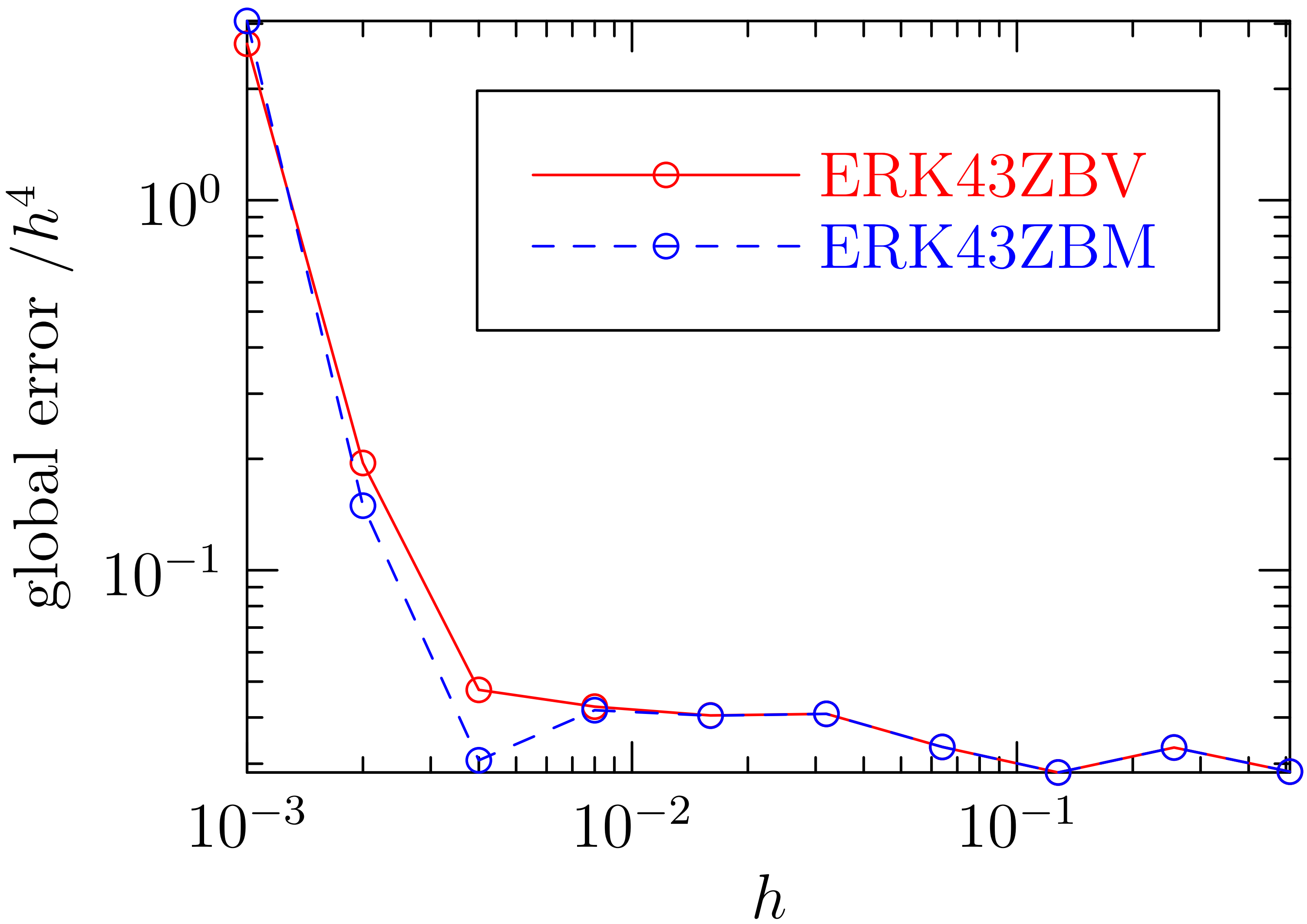


Adaptive Performance of ERK43ZB

- Choose Φ such that $y(x,t) = 10(1-x)x(1+\sin t)+2$:



Prob. 1: Schur Decomposition vs. Full Solution



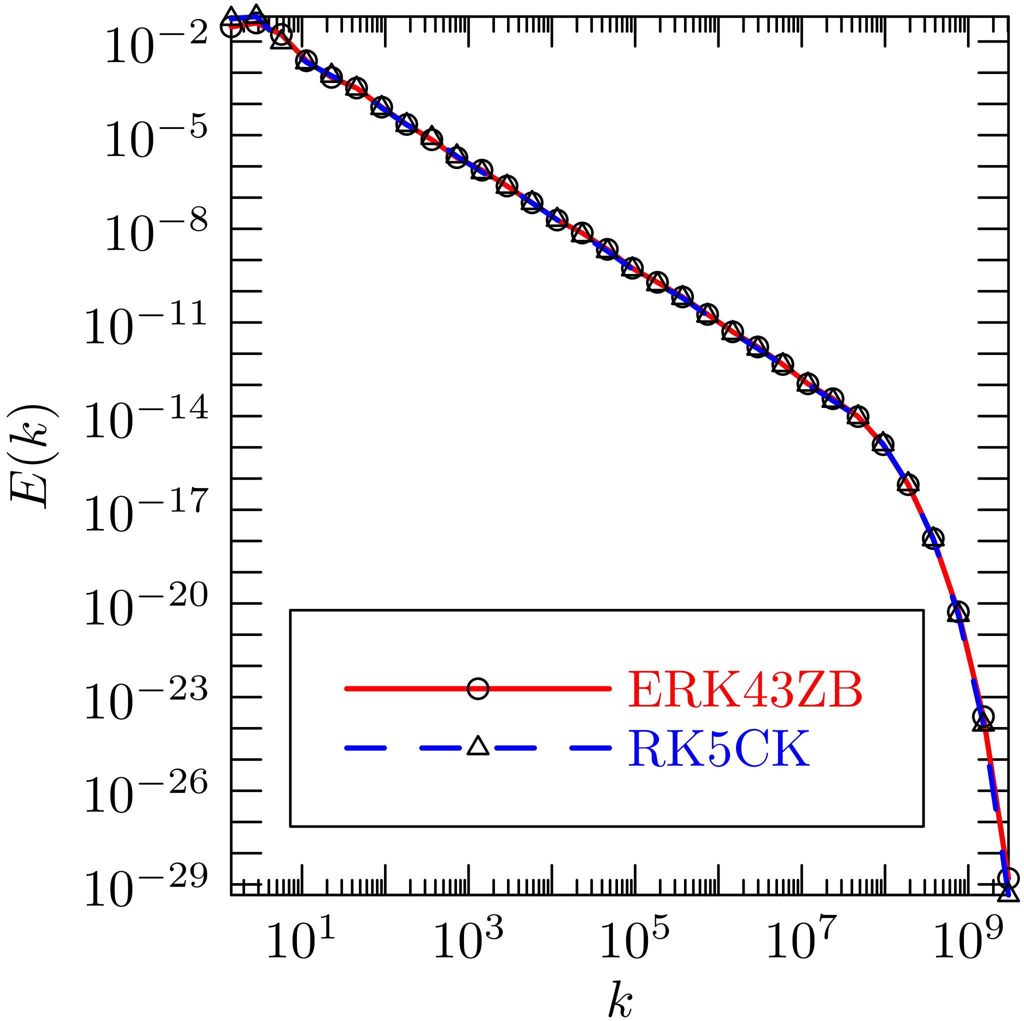
Performance: Schur Decomposition with an Embedded Pair

- Using Schur decomposition to integrate H–O Problem 2 from $t = 0$ to $t = 200$ with ERK43ZB was 117 times faster than the full matrix formulation, even after taking account of the cost of the Schur decomposition.

- Since the spatially discretized Laplacian matrix L is normal, both methods produced identical results.

GOY Shell Model of 3D Turbulence

- ERK43ZB runs over 3 times faster than the classical Cash–Karp (5,4) pair on a shell model of 3D turbulence exhibiting both linear and nonlinear stiffness:



Conclusions

- Unlike integrating factor methods, exponential integrators have the correct fixed point behaviour.
- A Schur decomposition avoids the need for computing matrix exponentials, while still circumventing linear stiffness.
- This technical advance makes adaptive exponential integration for general matrix linearities practical.
- We derived adaptive ERK pairs by symbolically solving the Hochbruck–Ostermann stiff-order conditions.
- A key requirement is that the pair be robust: if the nonlinear source function has nonzero total time derivatives, the order of the low-order estimate should never exceed its design value.
- We have derived robust exponential Runge–Kutta (3,2) and (4,3) embedded pairs well-suited to initial value problems with a dominant linearity.

References

[Beylkin *et al.* 1998] G. Beylkin, J. M. Keiser, & L. Vozovoi, J. Comp. Phys., **147**:362, 1998.

[Cartwright 1999] J. H. Cartwright, Physics Letters A, **264**:298, 1999.

[Certaine 1960] J. Certaine, Mathematical methods for digital computers, **1**:128, 1960.

[Cox & Matthews 2002] S. Cox & P. Matthews, J. Comp. Phys., **176**:430, 2002.

[Curtiss & Hirschfelder 1952] C. F. Curtiss & J. O. Hirschfelder, Proceedings of the National Academy of Sciences, **38**:235, 1952.

[Friedli 1978] A. Friedli, Lecture Notes in Mathematics, **631**:214, 1978.

[Hochbruck & Ostermann 2005a] M. Hochbruck & A. Ostermann, SIAM J. Numer. Anal., **43**:1069, 2005.

[Hochbruck & Ostermann 2005b] M. Hochbruck & A. Ostermann, Appl. Numer. Math., **53**:323, 2005.

[Hochbruck *et al.* 1998] M. Hochbruck, C. Lubich, & H. Selfhofer, SIAM J. Sci. Comput., **19**:1552, 1998.

[Lambert 1991] J. D. Lambert, *Numerical methods for ordinary differential systems: the initial value problem*, John Wiley & Sons, Inc., 1991.

[Lu 2003] Y. Y. Lu, J. Comput. Appl. Math., **161**:203, 2003.

[Nørsett 1969] S. Nørsett, Lecture Notes in Mathematics, **109**:214, 1969.

[van der Houwen 1977] P. J. van der Houwen, *Construction of integration formulas for initial value problems*, North-Holland Publishing Co., Amsterdam, 1977, North-Holland Series in Applied Mathematics and Mechanics, Vol. 19.

[Verwer 1977] J. Verwer, Numer. Math., **27**:143, 1977.

[Zoto & Bowman 2024] T. Zoto & J. C. Bowman, 2024, to be submitted.