On evaluating computational cost and approximation error in linear algebraic iterative solvers

Jörg Liesen
Technical University of Berlin

and

Zdeněk Strakoš
Charles University in Prague

http://www.karlin.mff.cuni.cz/˜strakos

FoCM, Budapest, July 2011

Main points

- In iterative methods applied to linear algebraic problems, computational cost of finding sufficiently accurate approximation to the exact solution heavily depends on the particular data.

- Any evaluation of cost in iterative computations must take into account effects of rounding errors.

- In mathematical modeling of real world phenomena, the accuracy of the computed approximation must be related to the underlying mathematical model, i.e., its evaluation can not be restricted to algebra.

Conjugate Gradients: A HPD, $x_0,\ r_0,\ p_0 = r_0$

$$\|x - x_n\|_A \;=\; \min_{u \,\in\, x_0 + \mathcal{K}_n(A, r_0)} \|x - u\|_A$$

$$\mathcal{K}_n(A, r_0) \;\equiv\; span\,\{r_0, Ar_0, \cdots, A^{n-1} r_0\}$$

For $n = 1, 2, \ldots$

$$\begin{aligned}
\gamma_{n-1} &= (r_{n-1}, r_{n-1}) / (p_{n-1}, Ap_{n-1}) \\
x_n &= x_{n-1} + \gamma_{n-1}\, p_{n-1} \\
r_n &= r_{n-1} - \gamma_{n-1}\, Ap_{n-1} \\
\delta_n &= (r_n, r_n) / (r_{n-1}, r_{n-1}) \\
p_n &= r_n + \delta_n\, p_{n-1} .
\end{aligned}$$

Hestenes and Stiefel (1952)

CG and Gauss-Christoffel quadrature

$$\int_L^U (\lambda)^{-1}\, d\omega(\lambda) \;=\; \sum_{i=1}^{n} \omega_i^{(n)} \left(\theta_i^{(n)}\right)^{-1} \;+\; R_n(f)$$

$$\frac{\|x - x_0\|_{\mathbf{A}}^2}{\|r_0\|^2} \;=\; n\text{-th Gauss quadrature} \;+\; \frac{\|x - x_n\|_{\mathbf{A}}^2}{\|r_0\|^2}$$

$$\sum_{j=0}^{n-1} \gamma_j \|r_j\|^2$$

CG : model reduction matching $2n$ moments

Golub, Meurant, Reichel, Boley, Gutknecht, Saylor, Smolarski, ... ,
Meurant and S (2006), Golub and Meurant (2010), S and Tichý (2011)

Outline

1. Linear bounds for the nonlinear method?

2. Do we know how to evaluate the computational error?

Thanks

André Gaul,
Jan Papež,
Tomáš Gergelits.

1 Linear bounds for the nonlinear method?

$$\begin{aligned}\|x - x_n\|_A &= \min_{\substack{p(0)=1\\ \deg(p)\le n}} \|A^{1/2} p(A)(x - x_0)\| \\ &= \min_{\substack{p(0)=1\\ \deg(p)\le n}} \|Y p(\Lambda) Y^* A^{1/2}(x - x_0)\| \\ &\le \left(\min_{\substack{p(0)=1\\ \deg(p)\le n}} \max_{1\le j\le N} |p(\lambda_j)| \right) \|x - x_0\|_A\end{aligned}$$

Using the shifted Chebyshev polynomials on the interval $[\lambda_1, \lambda_N]$,

$$\|x - x_n\|_A \le 2 \left(\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1} \right)^n \|x - x_0\|_A \,.$$

1 Linear bounds for the nonlinear method?

This bound should not be used in connection with the behaviour of CG unless $\kappa(A) = \lambda_N / \lambda_1$ is really small or unless the (very special) distribution of eigenvalues makes it relevant.

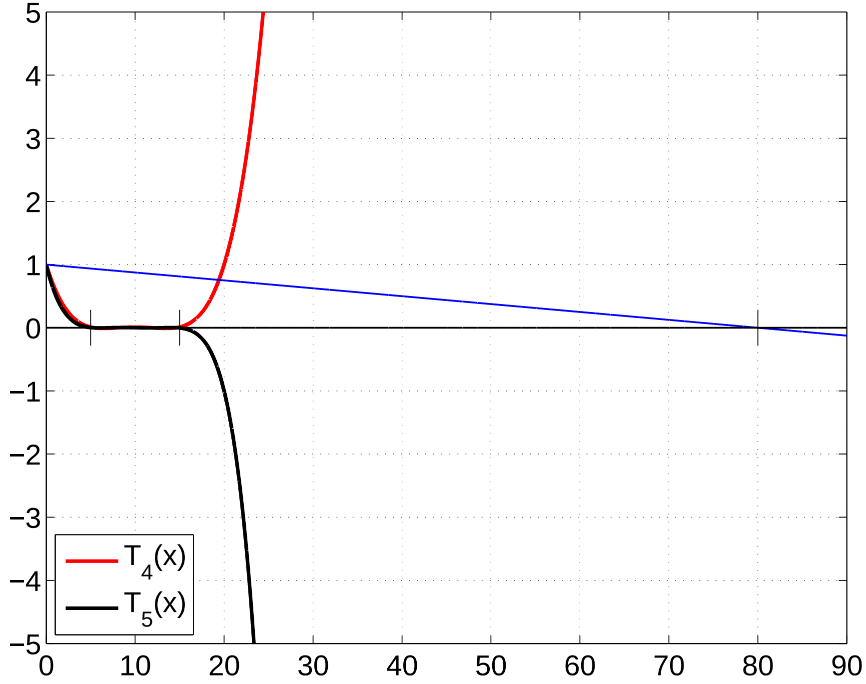
In particular, one should be very careful while using it as a part of a composite bound in the presence of the large outlying eigenvalues

$$\min_{\substack{p(0)=1 \\ \deg(p) \le n-s}} \max_{1 \le j \le N} |\, q_s(\lambda_j)\, p(\lambda_j) \,| \le \max_{1 \le j \le N} |q_s(\lambda_j)| \left| \frac{T_{n-s}(\lambda_j)|}{T_{n-s}(0)} \right|$$

$$< \max_{1 \le j \le N-s} \left| \frac{T_{n-s}(\lambda_j)}{T_{n-s}(0)} \right| .$$

Meinardus (Chebyshev polynomial) bound on the interval $[\lambda_1, \lambda_{N-s}]$ is then valid after s initial steps.

1 The polynomial $q_s(\lambda)$ has desired roots



The Chebyshev polynomials $T_4(\lambda), T_5(\lambda)$, and the polynomial $q_1(\lambda)$, $q_1(0) = 1$ having the single root at the large outlying eigenvalue.

1 Quote (2009, ...): the desired accuracy ϵ

Theorem. After

$$k = s + \left\lceil \frac{\ln(2/\epsilon)}{2} \sqrt{\frac{\lambda_{N-s}}{\lambda_1}} \right\rceil$$

iteration steps the CG will produce the approximate solution x_n satisfying

$$\|x - x_n\|_A \leq \epsilon \, \|x - x_0\|_A \, .$$

This recently republished and used statement is in finite precision arithmetic not true at all.

1 Mathematical model of FP CG

CG in finite precision arithmetic can be seen as the exact arithmetic CG for the problem with the slightly modified distribution function with larger support, i.e., with single eigenvalues replaced by tight clusters.

Paige (1971-80), Greenbaum (1989),
Parlett (1990), S (1991), Greenbaum and S (1992), Notay (1993), ... ,
Druskin, Knizhnermann, Zemke, Wülling, Meurant, ...
Recent review and update in Meurant and S, Acta Numerica (2006).

Fundamental consequence:

In FP computations, the composite convergence bounds eliminating in exact arithmetic large outlying eigenvalues at the cost of one iteration per eigenvalue do not, in general, work.

1 Axelsson (1976), quote Jennings (1977)

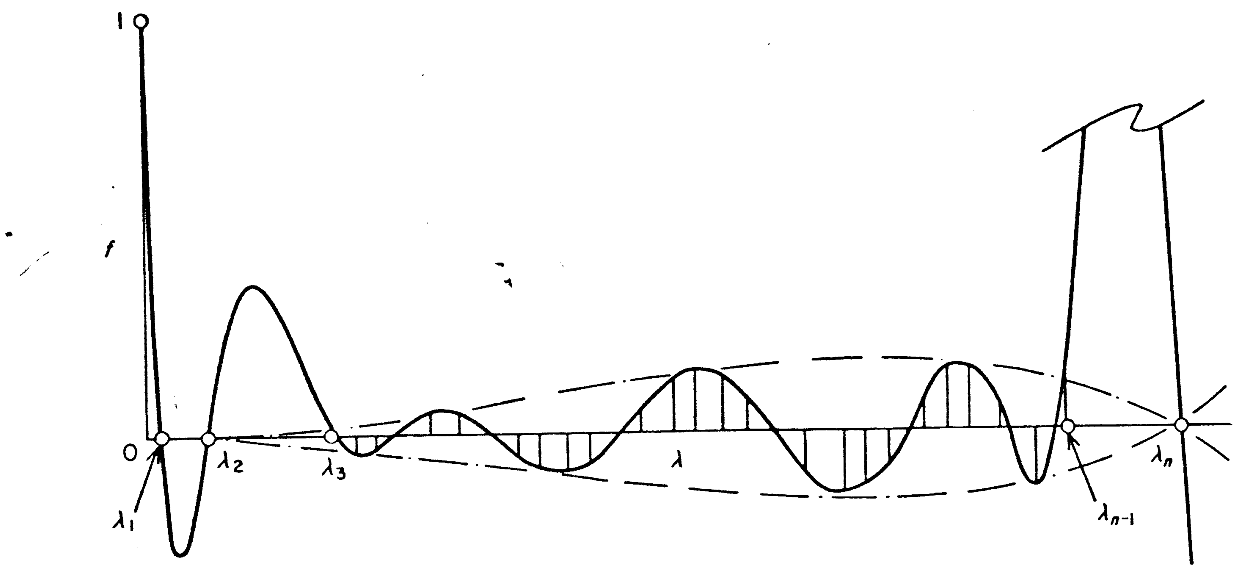
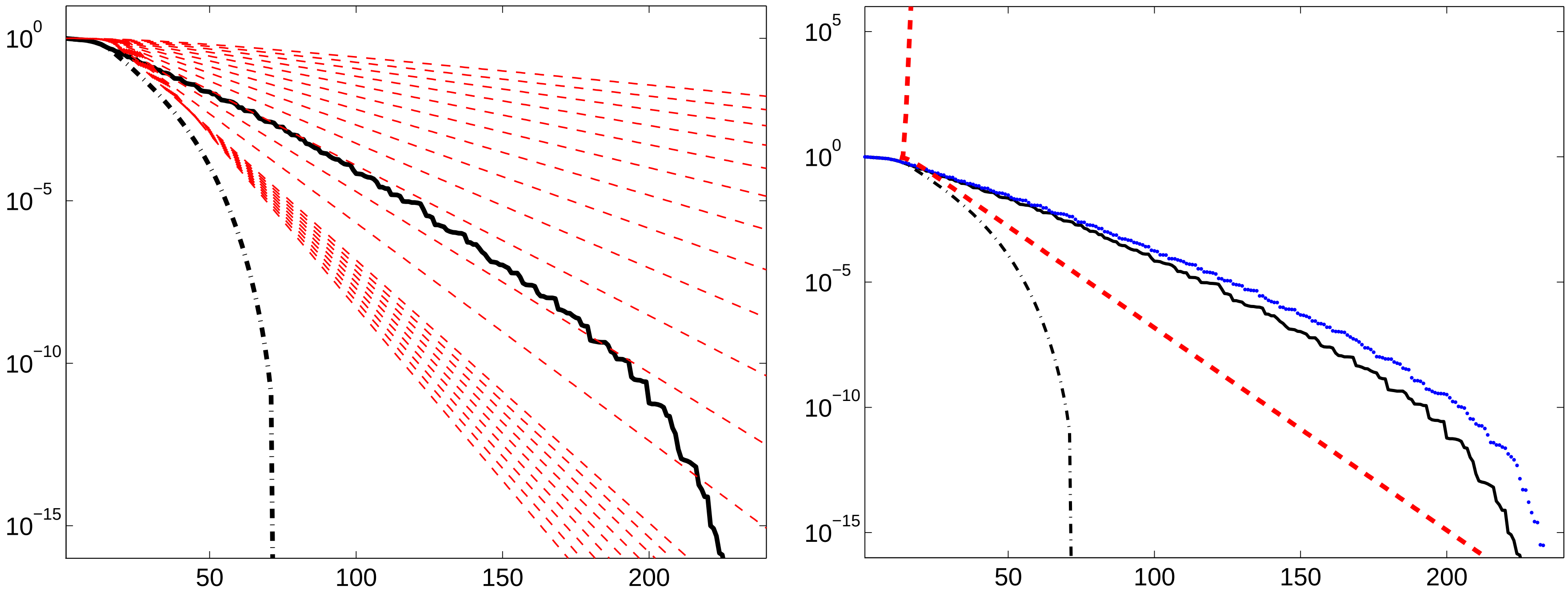


FIG. 4. A Chebyshev polynomial modified by a simple third order auxiliary polynomial having zeros at λ_1, λ_2 and λ_n.

p. 72: ... it may be inferred that rounding errors ... affects the convergence rate when large outlying eigenvalues are present.

1 The composite bounds completely fail



Composite bounds with varying number of outliers (left)
and the failure of the composed bounds in FP CG (right),
Gergelits (2011).

2 PDE discretization and matrix computations

$$-\Delta u \;=\; 16\eta_1\eta_2(1-\eta_1)(1-\eta_2)$$

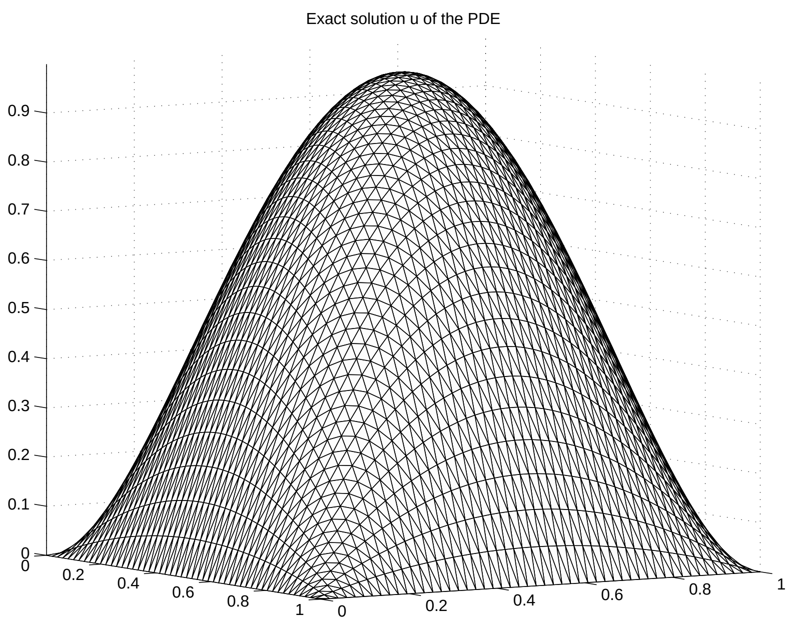
on the unit square with zero Dirichlet boundary conditions. Galerkin finite element method (FEM) discretization with linear basis functions on a regular triangular grid with the mesh size $h = 1/(m+1)$, where m is the number of inner nodes in each direction. Discrete solution

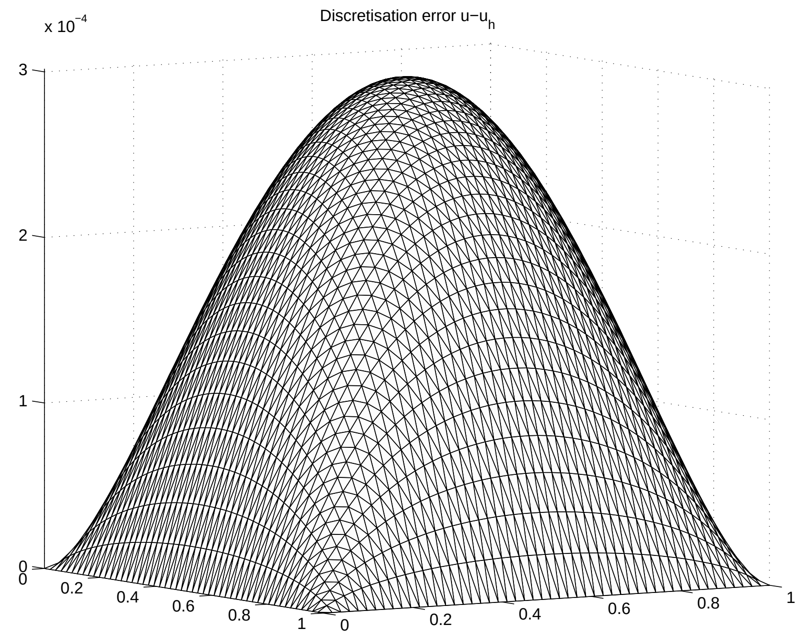
$$u_h \;=\; \sum_{j=1}^{N} \zeta_j \, \phi_j(\eta_1, \eta_2)\,.$$

Up to a small inaccuracy proportional to machine precision,

$$\begin{aligned} \|\nabla(u - u_h^{(n)})\|^2 \;&=\; \|\nabla(u-u_h)\|^2 \;+\; \|\nabla(u_h - u_h^{(n)})\|^2 \\ &=\; \|\nabla(u-u_h)\|^2 \;+\; \|x - x_n\|_A^2\,. \end{aligned}$$

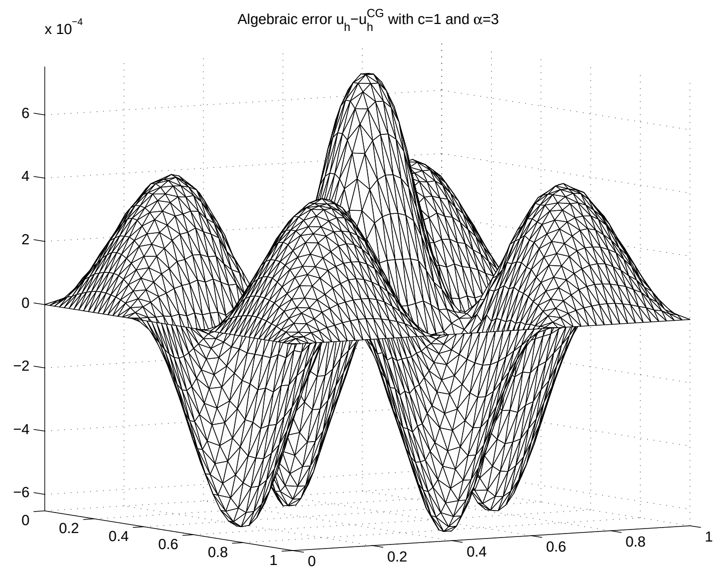
2 Solution and the discretization error

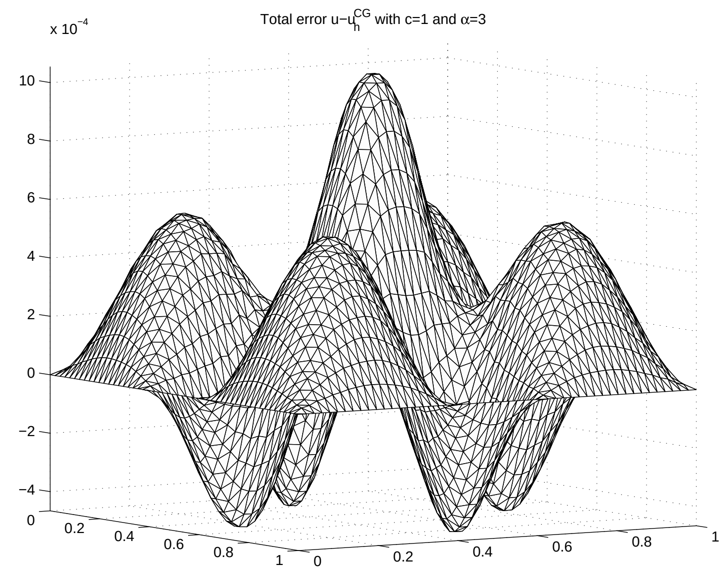




Exact solution u of the Poisson model problem (left)
and the MATLAB `trisurf` plot of the discretization error $u - u_h$ (right).

2 Algebraic and total errors

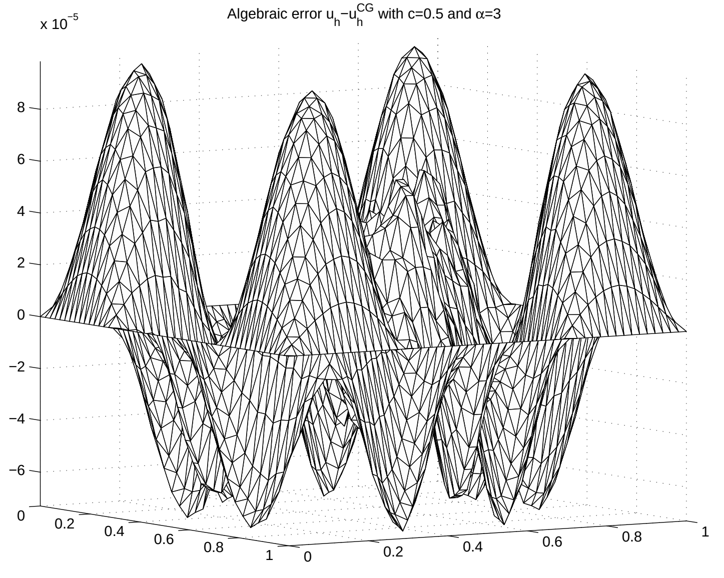


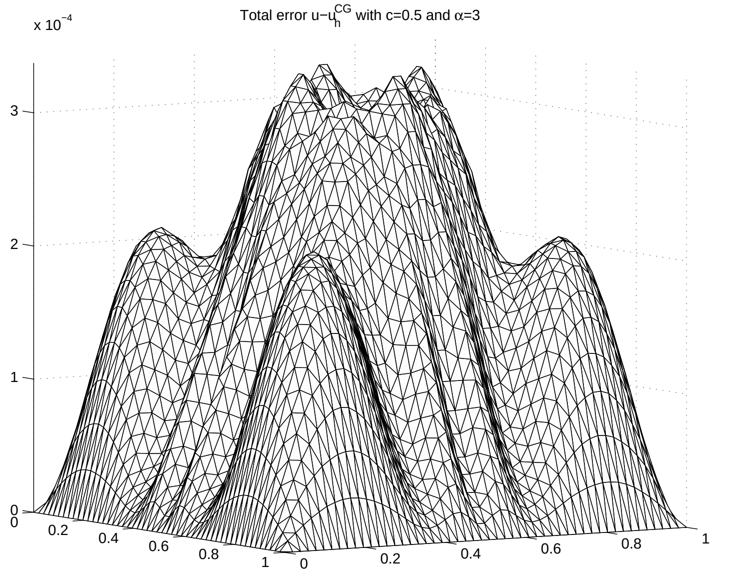


Algebraic error $u_h - u_h^{(n)}$ (left) and the MATLAB `trisurf` plot of the total error $u - u_h^{(n)}$ (right)

$$\|\nabla(u - u_h^{(n)})\|^2 = \|\nabla(u - u_h)\|^2 + \|x - x_n\|_A^2$$
$$= 5.8444e-03 + 1.4503e-05\,.$$

2 Algebraic and total errors

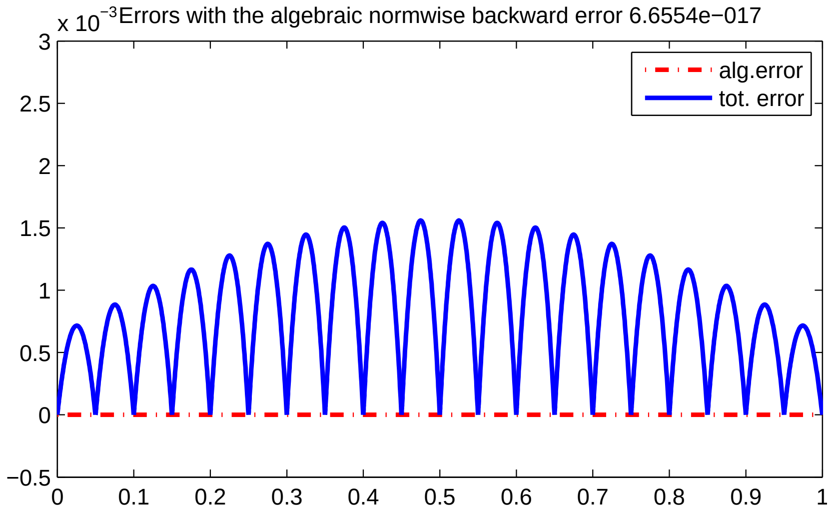


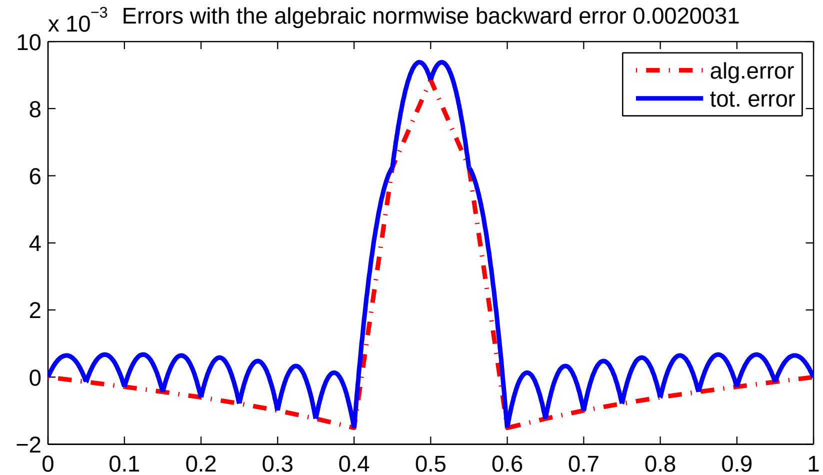


Algebraic error $u_h - u_h^{(n)}$ (left) and the MATLAB `trisurf` plot of the total error $u - u_h^{(n)}$ (right)

$$\begin{aligned} \|\nabla(u - u_h^{(n)})\|^2 &= \|\nabla(u - u_h)\|^2 + \|x - x_n\|_A^2 \\ &= 5.8444e-03 + 5.6043e-07\,. \end{aligned}$$

2 One can see 1D analogy





The discretization error (left),
the algebraic and the total error (right),
Papež (2011).

Challenges

- Using a formula from literature requires understanding of the whole context. Rounding errors can not be ignored.
- Numerical PDE: Matrix computations do not provide exact results. Verification in scientific and engineering computing should take this into account. Whenever possible, one should aim at the local distribution of the total error. Norms can hide important things.
- Algebra: Error should be evaluated in the function space. The backward error analysis and perturbation theory seems not sufficient.
- Both: Local distribution of the discretization and the algebraic errors can be very different. The algebraic computation can not be considered a black box part of the whole solution process. It must be integrated (from both sides) into it.

Liesen and S (2011), Liesen and S (201X)

Thank you for your kind patience