Matching moments and matrix computations

Jörg Liesen
Technical University of Berlin

and

Petr Tichý
Czech Academy of Sciences

and

Zdeněk Strakoš
Charles University in Prague and Czech Academy of Sciences

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

Householder Symposium XVIII, Granlibakken, June 2011

Thanks

I wish to give my thanks to many authors and collaborators whose work and help has enabled the presented view.

Gene Golub,
Gérard Meurant,
Claude Brezinski,

Jan Papež,
Tomáš Gergelits.

CG $\equiv$ matrix formulation of the Gauss Ch. q.

$$Ax = b\,,\ x_0 \quad \longleftrightarrow \quad \omega(\lambda), \quad \int_\zeta^\xi (\lambda)^{-1}\, d\omega(\lambda)$$

$$\uparrow \qquad\qquad\qquad\qquad \uparrow$$

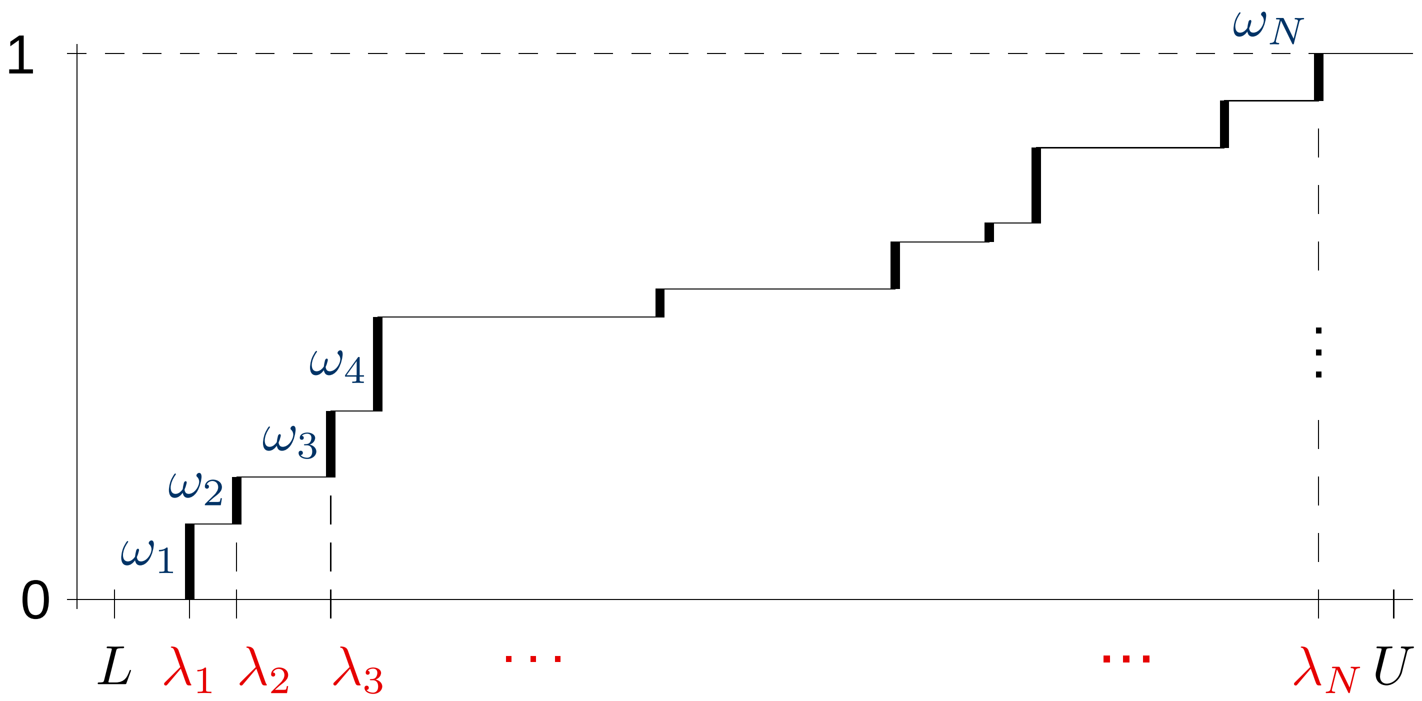
$$T_n\, y_n \ = \ \|r_0\|\, e_1 \quad \longleftrightarrow \quad \omega^{(n)}(\lambda), \quad \sum_{i=1}^{n} \omega_i^{(n)} \left(\theta_i^{(n)}\right)^{-1}$$

$$x_n \ = \ x_0 + W_n\, y_n$$

$$\boxed{\omega^{(n)}(\lambda) \longrightarrow \omega(\lambda)}$$

Distribution function $\omega(\lambda)$

λ_i, s_i are the eigenpairs of A , $\omega_i = |(s_i, w_1)|^2$, $w_1 = r_0/\|r_0\|$



Hestenes and Stiefel (1952)

CG and Gauss-Ch. quadrature errors

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

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

With $x_0 = 0$,

$$b^* A^{-1} b \;=\; \sum_{j=0}^{n-1} \gamma_j \|r_j\|^2 \;+\; r_n^* A^{-1} r_n \,.$$

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. History

2. Presence

3. Future

1 History

Brouncker and Wallis (1655), Euler (1737, 1748, ...),
Gauss (1814), Jacobi (1826, 1850, 1857), Chebyshev (1855, 1859, ...),
Christoffel (1858, 1877), Heine (1878, 1881), Markov (1884, ...),
Stieltjes (1884–1894)

see

Perron (1913), Szegö (1839), Gantmacher and Krein (1941), Shohat and Tamarkin (1943), Krein (1951), Akhiezer (1965), Gautschi (1981, 2004), Brezinski(1991), Van Assche (1993), Kjeldsen (1993), Gutknecht (1994) ...

Hilbert (1906), F. Riesz (1909), Hellinger and Toeplitz (1910, 1914),
von Neumann (1927, 1932), Wintner (1929),

Krylov (1931), Lanczos (1950, 1952), Hestenes and Stiefel (1952),
Rutishauser (1953), Henrici (1958), Stiefel (1958),
Householder (1953, 1964) ...

1 Continued fraction corresponding to $\omega(\lambda)$

$$\mathcal{F}_N(\lambda) \equiv \cfrac{1}{\lambda - \gamma_1 - \cfrac{\delta_2^2}{\lambda - \gamma_2 - \cfrac{\delta_3^2}{\lambda - \gamma_3 - \dots \cfrac{\ddots}{\lambda - \gamma_{N-1} - \cfrac{\delta_N^2}{\lambda - \gamma_N}}}}}$$

The entries $\gamma_1, \dots, \gamma_N$ and $\delta_2, \dots, \delta_N$ represent coefficients of the Stieltjes recurrence (1893-4), see Chebyshev (1855), Brouncker (1655), Wallis (1656), with the Jacobi matrix T_N (Toeplitz and Hellinger (1914), Simon (2006), S (2011), Liesen and S (201?).)

The nth convergent $\mathcal{F}_n(\lambda)$.

1 Partial fraction decomposition

$$w_1^*(\lambda I - A)^{-1} w_1 \ = \ \int_L^U \frac{d\omega(\mu)}{\lambda - \mu} \ = \ \sum_{j=1}^{N} \frac{\omega_j}{\lambda - \lambda_j} = \ \frac{\mathcal{R}_N(\lambda)}{\mathcal{P}_N(\lambda)} \ = \ \mathcal{F}_N(\lambda) \, ,$$

The denominator $\mathcal{P}_n(\lambda)$ corresponding to the nth convergent $\mathcal{F}_n(\lambda)$ of $\mathcal{F}_N(\lambda)$, $n = 1, 2, \ldots$ is the nth orthogonal polynomial in the sequence determined by $\omega(\lambda)$; see Chebyshev (1855). Similarly,

$$e_1^T(\lambda I - T_n)^{-1} e_1 \ = \ \int_L^U \frac{d\omega^{(n)}(\mu)}{\lambda - \mu} \ = \ \sum_{j=1}^{n} \frac{\omega_j^{(n)}}{\lambda - \lambda_j} = \ \frac{\mathcal{R}_n(\lambda)}{\mathcal{P}_n(\lambda)} \ = \ \mathcal{F}_n(\lambda) \, ,$$

1 Minimal partial realization

Consider the expansion

$$w_1^*(\lambda I - A)^{-1} w_1 \;=\; \mathcal{F}_N(\lambda) \;=\; \sum_{\ell=1}^{2n} \frac{\xi_{\ell-1}}{\lambda^\ell} \;+\; \mathcal{O}\left(\frac{1}{\lambda^{2n+1}}\right) .$$

Using the same expansion of the nth convergent $\mathcal{F}_n(\lambda)$ of $\mathcal{F}_N(\lambda)$,

$$e_1^T(\lambda I - T_n)^{-1} e_1 \;=\; \mathcal{F}_n(\lambda) \;=\; \sum_{\ell=1}^{2n} \frac{\xi_{\ell-1}}{\lambda^\ell} \;+\; \mathcal{O}\left(\frac{1}{\lambda^{2n+1}}\right) ,$$

where the moments in the numerators are identical due to the Gauss-Ch. quadrature property. $\mathcal{F}_n(\lambda)$ approximates $\mathcal{F}_N(\lambda)$ with the error proportional to $\lambda^{-(2n+1)}$.

1 Minimal partial realization - roots

Christoffel (1858), Chebyshev (1859), Stieltjes (1894);
Euler (1737), translation published by Wyman and Wyman (1985).

The minimal partial realization was rediscovered in the engineering literature by Kalman (1979).

The works of Hestenes and Stiefel (1952), Vorobyev (1958, 1965) (see Brezinski (1991, ...)), were not studied and recalled.
The links with Chebyshev and Stieltjes were pointed out by
Gragg (1974), Gragg and Lindquist (1983),
Bultheel and Van Barel (1997).

S (2011), S and Tichý (2011), S and Liesen (201?).

2 Presence

A natural shift from the moment matching background to computing, technical issues and algorithmic developments.

Example: The bound

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

should not be used in connection with the behaviour of CG unless $\kappa(A)$ 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 the composed bound in the presence of the large outlying eigenvalues.

2 Axelsson (1976), Jennings (1977)

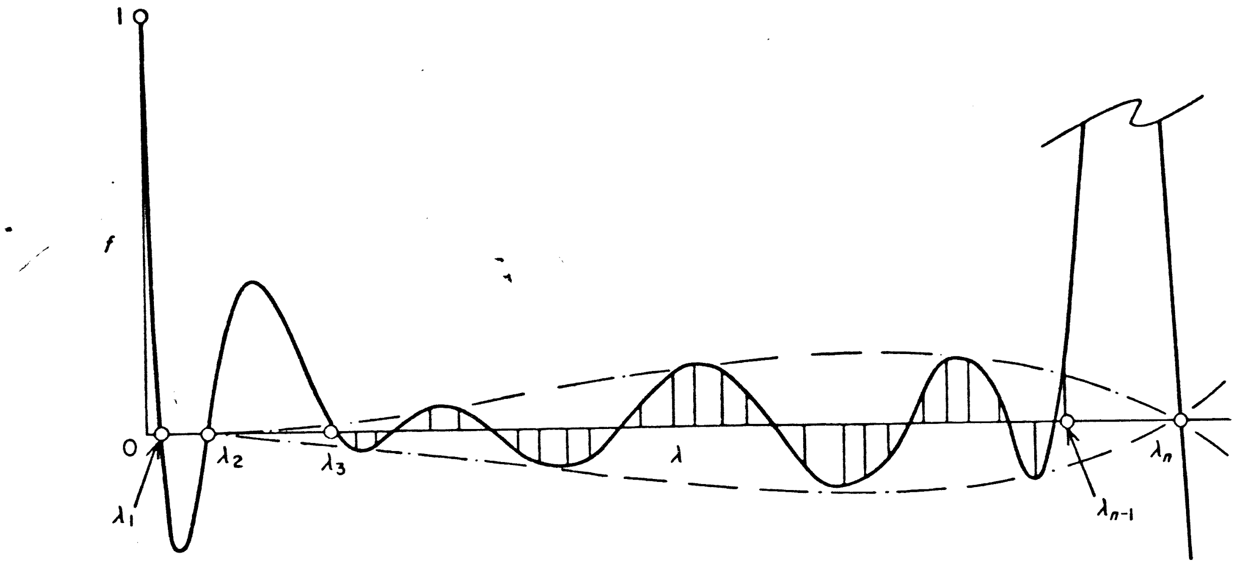


FIG. 4. A Chebyshev polynomial modified by a simple third order auxiliary polynomial having zeros at λ_1, λ_2 and λ_n.

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

2 Mathematical model of FP CG

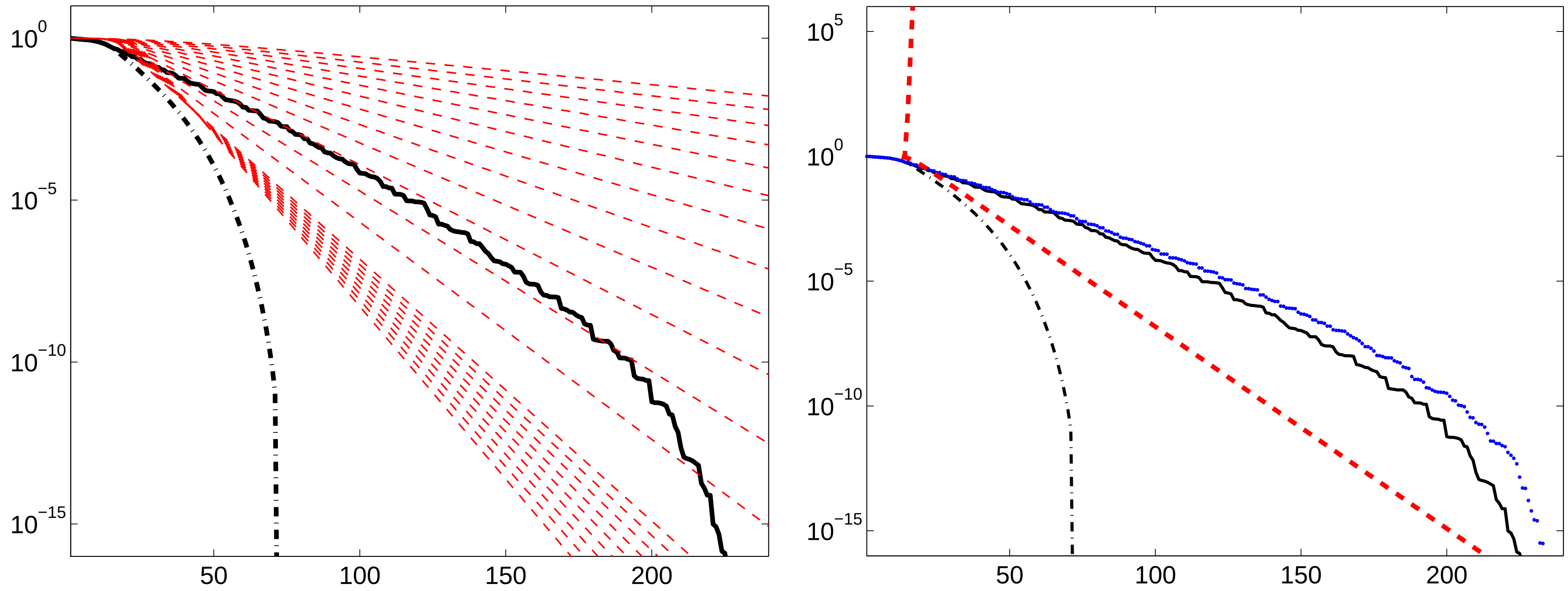
CG in finite precision arithmetic can be seen as the exact arithmetic CG for the problem with the slightly modified distribution function with larger support, i.e., with single eigenvalues replaced by tight clusters.

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

2 : Composed bounds



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

3 PDE discretization and matrix computations

$$-\Delta u \; = \; 32(\eta_1 - \eta_1^2 + \eta_2 - \eta_2^2)$$

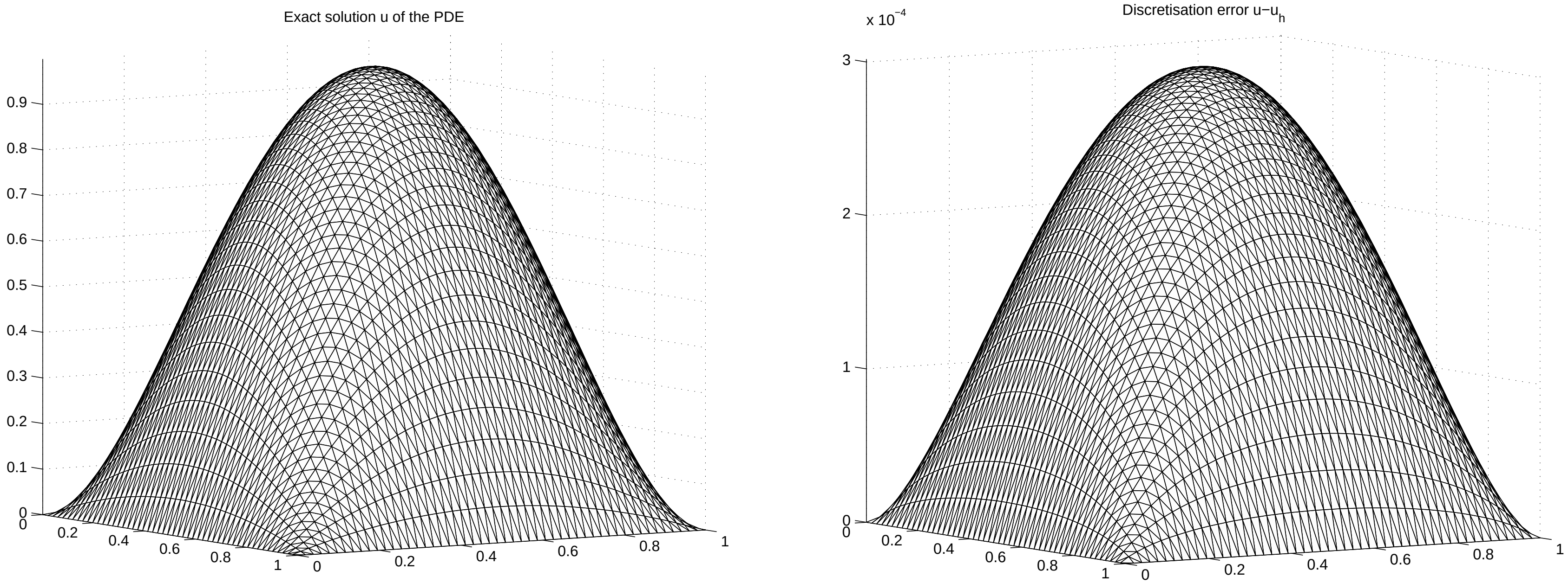
on the unit square with zero Dirichlet boundary conditions. Galerkin finite element method (FEM) discretization with linear basis functions on a regular triangular grid with the mesh size $h = 1/(m+1)$, where m is the number of inner nodes in each direction. Discrete solution

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

Up to a small inaccuracy proportional to machine precision,

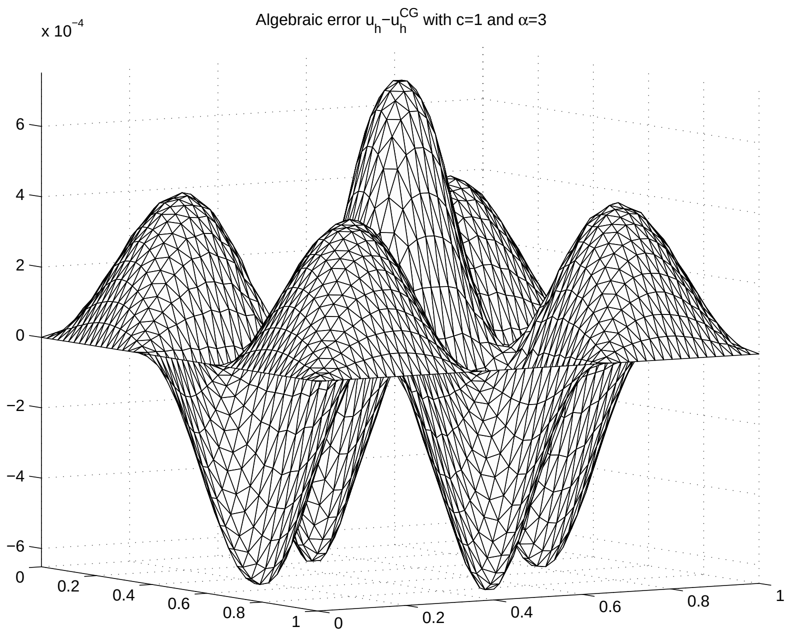
$$\begin{aligned} \|\nabla(u - u_h^{(n)})\|^2 \; &= \; \|\nabla(u - u_h)\|^2 \; + \; \|\nabla(u_h - u_h^{(n)})\|^2 \\ &= \; \|\nabla(u - u_h)\|^2 \; + \; \|x - x_n\|_A^2 \, . \end{aligned}$$

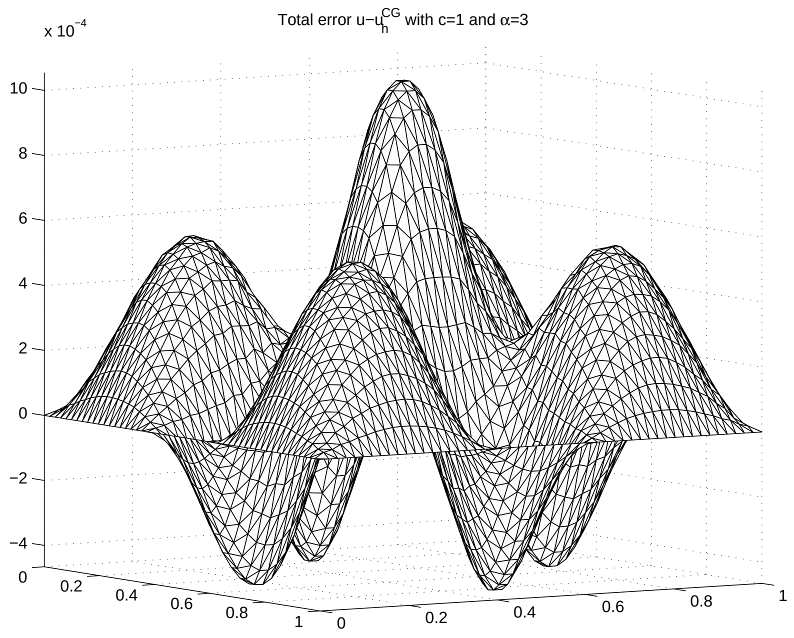
3 Solution and discretization error



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

3 Algebraic and total error





Algebraic error $u_h - u_h^{(n)}$ (left) and the total error $u - u_h^{(n)}$ (right) when CG is stopped with $\|b - Ax_n\| / (\|b\| + \|A\|\,\|x_n\|) < 1e - 03\,.$

3 The point

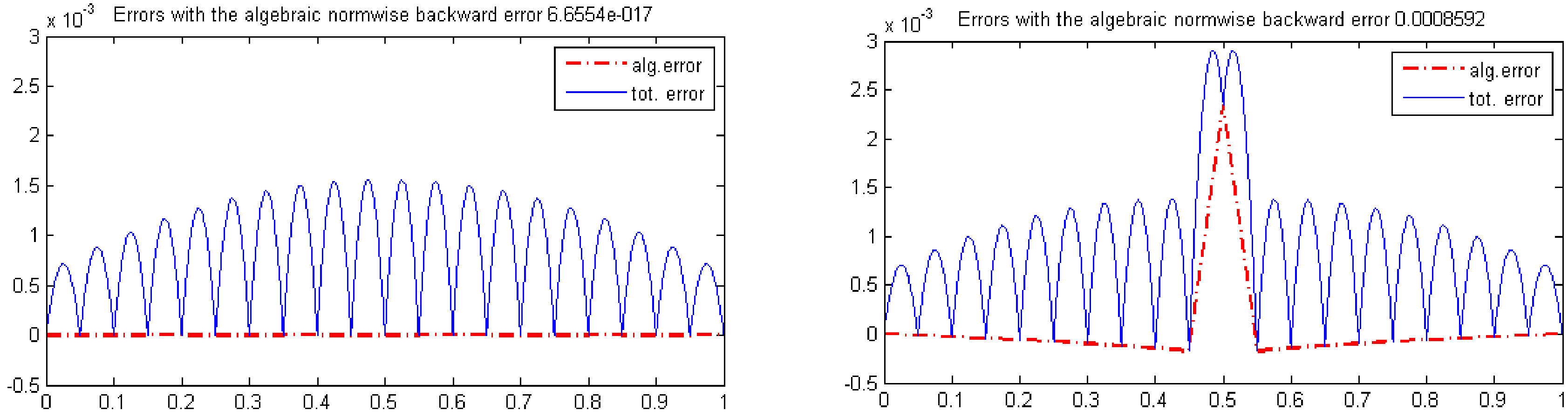
$$\|\nabla(u - u_h^{(n)})\|^2 = \|\nabla(u - u_h)\|^2 + \|x - x_n\|_A^2$$
$$= 5.8444e-03 + 1.4503e-05\,.$$

Message?

- PDE: Matrix computations do not provide exact results.
- Algebra: Error should be evaluated in the function space.
- Both: Local distribution of the discretization and the algebraic error can be very different.

Liesen and S (2011), Liesen and S (201?).

3 One can see the 1D analogy



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

Ideas and people

- Euclid, ... , Brouncker and Wallis (1655-56): Three term recurences (for numbers)
- Euler (1737, 1748), ... , Brezinski (1991), Khrushchev (2008)
- Gauss (1814), Jacobi (1826), Christoffel (1858, 1857), ... , Gautschi (1981, 2004)
- Chebyshev (1855, 1859), Markov (1884), Stieltjes (1884, 1893-94): Orthogonal polynomials
- Hilbert (1906), ... , Von Neumann (1927, 1932), Wintner (1929)
- Krylov (1931), Lanczos (1950), Hestenes and Stiefel (1952), Rutishauser (1953), Henrici (1958), ... , Vorobyev (1958, 1965), Golub and Welsh (1968), ... , Laurie (1991 - 2001), ...
- Gordon (1968), Schlesinger and Schwartz (1966), Reinhard (1979), ...
- Paige (1971), Reid (1971), Greenbaum (1989), ...
- Gragg (1974), Kalman (1979), Gragg, Lindquist (1983), Gallivan, Grimme, Van Dooren (1994), ...

$\longrightarrow$ Euler, Christoffel, Chebyshev (Markov), Stieltjes !

Thank you for your kind patience