GLT Structures, spectral approximation of PDEs, symbol and Preconditioning

Stefano Serra Capizzano

Department of Science and high Technology
Insubria University - Como

Department of Information Technology*
Uppsala University - Uppsala

Grant* via Donation KAW 2013.0341,
Knut & Alice Wallenberg Foundation
with the Royal Swedish Academy of Sciences,
supporting Swedish Research in Mathematics

PRECON2015, Eindhoven, June 17-19 2015

Which messages? Very short answers

a.1) GLT sequences are a subspace of sequences of matrices $\{A_n\}$, A_n of size d_n ($d_k < d_{k+1}$)

a.2) Each GLT sequence $\{A_n\}$ has a symbol ψ

a.3) The singular values of a GLT sequence with symbol ψ are approximately described by equispaced sampling of $|\psi|$

a.4) The spectrum of a (quasi Hermitian) GLT sequence with symbol ψ are approximately described by equispaced sampling of ψ

a.5) The GLT sequences are stable under elementary operations and the symbol is obtained via the same elementary operations

The GLT glasses: a variable coefficient version of the Local Fourier Analysis

b.1) Local methods (including FDs, FEs, IgA, FVs, VEMs) for approximating PDEs, IEs lead to GLT sequences, possibly after proper permutations

b.2) No limitations on variable coefficients and on domains (grids should have some structure at least asymptotically)

b.3) Information on the symbol leads to information on ill-conditioning, on the size of the ill-conditioned subspaces, on the nature of the ill-conditioned subspaces (low frequencies, high frequencies etc)

The GLT glasses.....

c.1) We exploit the symbol for understanding the reason of difficulties of known techniques, w.r.t. finess parameters, problem parameters, approximation parameters

c.2) We exploit the symbol for designing new iterative solvers, new preconditioners or smoothers or prolongation operators, aiming at optimality and robustness.

From continuous to discrete

A continuous infinite-dimensional problem (PDEs, IEs etc) is transformed, via a suitable numerical approximation, into a linear (nonlinear) system of algebraic equations

- Structure inherited from the continuous counterpart
- Large dimensions (e.g. 10^p, $p \geq 10$)
- Spectral features described via a proper Symbol

Goal: solving the resulting linear system by **Optimal Methods** (operation count to obtain the solution of the same order of the matrix-vector multiplication)

In the discrete case

- Large dimensions imply that direct solvers (Gaussian Elimination etc.) have to be avoided
- Iterative solvers: A) operation count per iteration of the same order of the matrix-vector multiplication B) the method is Optimal if the number of iterations $\leq c(\epsilon)$, with ϵ desired precision.

Requirement B) depends on the spectrum of the involved matrices: it depends especially on the possibility of approximating the coefficient matrix in the ill-conditioned subspaces (i.e. associated to the eigenvectors with small eigenvalues).

$$\Downarrow$$

For large classes of matrices coming from continuous problems, **the knowledge of the spectrum** is often compactly represented in a function, called the **symbol**.

Main items

Symbol, preconditioning (and multigrid)

1. Toeplitz structures and symbol
2. Approximation of Differential Operators
3. The GLT algebra and the notion of symbol

Examples

4. FEM of degree p in d dimensions
5. Approximation Q2Q1 of the Linear Elasticity
6. IgA of degree p in d dimensions

Collaborators

Aricó, Bai, Beckermann, Bertaccini, R. Chan, Dell'Acqua, Di Benedetto, Donatelli, Dorostkar, Estatico, Fiorentino, Frangioni, Garoni, Golub, Golinskii, Holmgren, Huckle, Hughes, Kuijlaars, Manni, Mazza, Molteni, Nagy, Neytcheva, Ng, Ngondiep, Noutsos, Pelosi, Pennati, Perrone, Reali, Semplice, Sesana, Speleers, Sundquist, Tablino, Tasche, Tilli, Tyrtyshnikov, Vassalos.

- In blue consolidated collaborations on the themes of the talk;
- In green just started collaborations (mainly on vector PDEs).

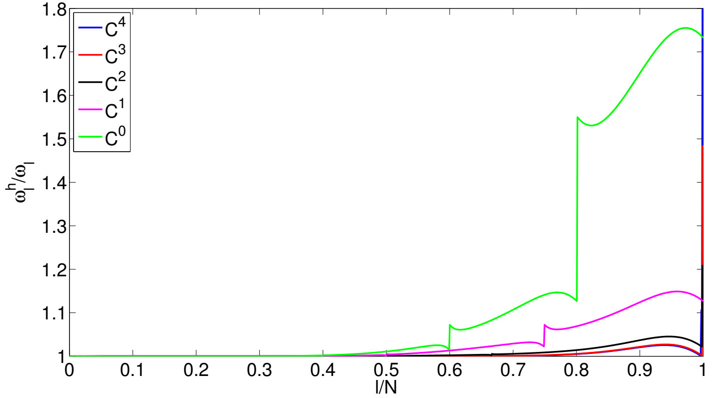
Spectral Distribution: the qualitative idea

- $M_m(\mathbb{C})$ complex matrices of order m,
- $\{A_n\}$, $A_n \in M_{d_n}(\mathbb{C})$, $d_n < d_{n+1}$,
- ψ measurable on $D \subset \mathbb{R}^g$, $g \geq 1$,
- ψ being $M_s(\mathbb{C})$-valued, $s \geq 1$,
- $0 < \mu\{D\} < \infty$, $\mu\{\cdot\}$ can be the Lebesgue measure,

$$\{A_n\} \sim_\lambda (\psi, D).$$

Informal meaning: $s = 1$. If ψ is continuous, then a suitable ordering of the eigenvalues $\{\lambda_j(A_n)\}$, in correspondence with a equispaced gridding on D, reconstructs approximately the surface $t \to \psi(t)$.
Informal meaning: $s > 1$. If ψ is continuous, then a suitable ordering of the eigenvalues $\{\lambda_j(A_n)\}$, in correspondence with a equispaced gridding on D, reconstructs approximately s surfaces, $t \to \lambda_j(\psi(t))$, $j = 1, \ldots, s$.

Comparison IgA-FEM (and furthermore the case of intermediate regularity): $C^0 \rightarrow$FEM$\rightarrow s = p^d$, $C^{p-1} \rightarrow$IgA$\rightarrow s = 1$, $C^k \rightarrow$interm. regularity $\rightarrow s = (p-k)^d$ (figure by A. Reali)



Toeplitz sequences generated by a symbol: $\{T_n(f)\} \sim_\lambda (f, I_d)$ if $f = f^*$

- s, d positive integers, $\mathbf{i}^2 = -1$;
- $f \in L^1(I_d, M_s(\mathbb{C}))$, $I_d = (-\pi, \pi)^d$, $j \in \mathbb{Z}^d$;
- $a_j = \frac{1}{(2\pi)^d} \int_{I_d} f(s) e^{-\mathbf{i} js} \, ds$, $a_j \in M_s(\mathbb{C})$.

For $d = 1$ the matrix $T_n(f)$ has size ns:

$$T_n(f) = \begin{pmatrix} a_0 & a_{-1} & \cdots & a_{1-n} \\ a_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{-1} \\ a_{n-1} & \cdots & a_1 & a_0 \end{pmatrix}.$$

For $d > 1$ we have a recoursive formula.

Toeplitz sequences generated by a symbol: $\{T_n(f)\} \sim_\lambda (f, I_d)$ if $f = f^*$

For $d > 1$, the d-level Toeplitz matrix $T_n(f)$ has order Ns, $N = \prod n_j$, $n = (n_1, \ldots, n_d)$, and takes the form

$$T_n(f) = \begin{pmatrix} T_0 & T_{-1} & \cdots & T_{1-n_1} \\ T_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & T_{-1} \\ T_{n_1-1} & \cdots & T_1 & T_0 \end{pmatrix},$$

T_j being $(d-1)$-level Toeplitz matrix. If $\otimes$ denotes the Knonecker product

$$T_n(f) = \sum_{|j| \leq n-1} J_n^{[j]}, \qquad J_n^{[j]} = J_{n_1}^{j_1} \otimes \cdots \otimes J_{n_d}^{j_d} \otimes a_j,$$

with $(J_m^r)_{s,t} = 1$ if $s - t = r$ and 0 otherwise.

The symbol: Toeplitz and GLT through an example

Minimalistic example

Assume

$$\begin{cases} -(\kappa_0 u')' + v' & = & g_1(x), \\ u' - \rho v & = & g_2(x), \end{cases}$$

Discretize on a square mesh of stepsize h using bilinear FEM basis functions:

$$\mathcal{A} = \begin{bmatrix} K & B^T \\ B & -\rho M \end{bmatrix}$$

The symbol: Toeplitz and GLT through an example

Minimalistic example, cont.: the arising matrices

$$K = \kappa_0 \cdot \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{bmatrix}, \quad M = \frac{h^2}{6} \begin{bmatrix} 4 & 1 & & & \\ 1 & 4 & 1 & & \\ & 1 & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & 1 & 4 \end{bmatrix},$$

$$B = h \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ & -1 & \ddots & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix}.$$

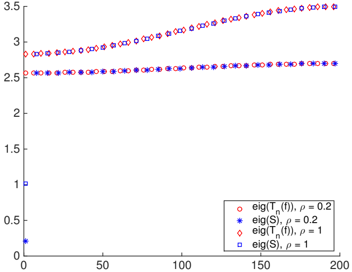
The symbol: Toeplitz and GLT through an example

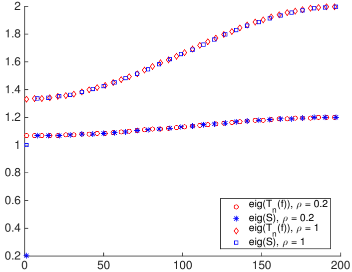
Minimalistic example, cont.: matrices and symbols

$$\begin{array}{rclcrcl} K & = & \kappa_0\, T_n(2-2\cos(\theta)), & & B & = & h\, T_n(1-e^{i\theta}), \\ B^T & = & h\, T_n(1-e^{-i\theta}), & & M & = & \frac{h^2}{3}\, T_n(2+\cos(\theta)). \end{array}$$

$$S = \rho M + B^T K^{-1} B, \quad \left\{ \frac{S}{h^2} \right\} \sim_\lambda \left(\frac{\rho}{3}(2+\cos(\theta)) + \frac{1}{\kappa_0}, (-\pi,\pi) \right).$$

Below $\kappa_0 = 0.4$ and $\kappa_0 = 1$: **if $\kappa_0 = a(x)$ the formula of the symbol (a bivariate function in (x,θ)) is formally exactly the same!**





The symbol: Toeplitz and GLT through an example

$\mathcal{L}_a(u) = -\left(a(x)u^{'}\right)^{'}$ [rod with variable section].

$$K_n = \begin{pmatrix} d_1 & -a_2 & & \\ -a_2 & \ddots & \ddots & \\ & \ddots & \ddots & -a_n \\ & & -a_n & d_n \end{pmatrix}, \quad T_n(e^{i\theta}) = \begin{pmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix},$$

$$T_n(e^{-i\theta}) = T_n^T(e^{i\theta}), \quad D_n(a) = \mathrm{diag}(a(jh)), \quad h = \frac{1}{n+1}.$$

Then

$$\begin{aligned} K_n &= 2D_n(a) - D_n(a)T_n(e^{i\theta}) - T_n(e^{-i\theta})D_n(a) + E_n, \quad \|E_n\| \to 0, \\ \psi(x,\theta) &= 2a(x) - a(x)e^{i\theta} - e^{-i\theta}a(x) + 0 = a(x)(2 - 2\cos(\theta)) \end{aligned}$$

$$\{K_n\} \sim_\lambda (\psi(x,\theta), [0,1] \times (-\pi,\pi)).$$

Spectral Distribution: the definition

$F \in C_0$ (continuous with compact support):

$$\Sigma_\lambda(F, A_n) = \frac{1}{d_n} \sum_{j=1}^{d_n} F[\lambda_j(A_n)].$$

Definition

We write $\{A_n\} \sim_\lambda (\psi, D)$ if $\forall F \in C_0$

$$\lim_{n\to\infty} \Sigma_\lambda(F, A_n) = \frac{1}{s\mu\{D\}} \int_D \mathrm{trace}\left(F\left(\psi(t)\right)\right)\, dt.$$

Moreover, we write $\{A_n\} \sim_\sigma (\psi, D)$ replacing $\lambda_j(A_n)$ by $\sigma_j(A_n)$ (singular values) in $\Sigma_\sigma(F, A_n)$ in place of $\Sigma_\lambda(F, A_n)$ and replacing $\psi(t)$ by $|\psi(t)|$ in the integral. If $s > 1$ then $|\psi(t)| = \left(\psi^*(t)\psi(t)\right)^{1/2}$.

Spectral properties and symbol: the Toeplitz setting

With f real-valued, $m =$ess-inf f, $M =$ess-sup f, $m < M$, and $f(\theta) - m$ asymptotic to $\|\theta - \hat{\theta}\|^\alpha$, $\alpha > 0$, $\theta \in I_d = (-\pi, \pi)^d$:

- $\lambda_j(T_n(f)) \in (m, M)$, for every j, for every n;
- $\lambda_{\min}(T_n(f)) - m \sim n^{-\alpha}$ for $d = 1$;
- $\lambda_{\min}(T_n(f)) - m \sim [N(n)]^{-\alpha/d}$ for $d > 1$;
- the eigenvalues of the matrices $\{T_n(f)\}$ are distributed as the symbol $f \in L^1(I_d)$, i.e., $\{T_n(f)\} \sim_\lambda (f, I_d)$;
- for Hermitian matrix-valued symbols m, M are replace by ess-inf $\lambda_{\min}(f)$, ess-sup $\lambda_{\max}(f)$, respectively;
- for complex/matrix-valued symbols $\{T_n(f)\} \sim_\sigma (f, I_d)$ and $\{T_n(f)\} \sim_\lambda (f, I_d)$ if f is Hermitian-valued.

Example: $T_n(f)$, $f(\theta) = \theta^4$

- f univariate, scalar-valued so that $s = d = 1$;
- $m =$ess-inf $f = 0$, $M =$ess-sup $f = \pi^4$ so that $m < M$ and $\alpha = 4$.

Hence the eigenvalues of the Toeplitz matrix $T_n(f)$, $f(\theta) = \theta^4$, behave as the values $\left(\frac{\pi j}{n+1}\right)^4$ and the minimal eigenvalue tends to zero as n^{-4}, $\alpha = 4$:

$$\Downarrow$$

The Gauss-Seidel method needs $O(n^4)$ steps and also the Conjugate Gradient is far from optimal with loss of orthogonality;

$$\Downarrow$$

Preconditioning (Structured Preconditioning)

Structured Preconditioning: Toeplitz setting

f, g real-valued with g nonnegative, not identically zero, $m=$ess-inf f/g, $M=$ess-sup f/g, $m < M$ and $h(\theta) - m$ asymtotic to $\|\theta - \hat{\theta}\|^{\alpha}$, $h = f/g$, $\alpha > 0$, $\theta \in I_d = (-\pi, \pi)^d$: $T_n(g)$ preconditioner, $P_n = T_n^{-1}(g) T_n(f)$ preconditioned matrix.

- $\lambda_j(P_n) \in (m, M)$, for every j, for every n;
- $\lambda_{\min}(P_n) - m \sim n^{-\alpha}$ for $d = 1$;
- $\lambda_{\min}(P_n) - m \sim [N(n)]^{-\alpha/d}$ for $d > 1$;
- the eigenvalues of the preconditioned matrices $\{P_n\}$ are distributed as the symbol $h = f/g$ measurable over I_d i.e. $\{P_n\} \sim_{\lambda} (h, I_d)$;
- when Hermitian-valued symbols are involved m is replaced by ess-inf $\lambda_{\min}(h)$ and M by ess-sup $\lambda_{\max}(h)$, respectively;
- for complex/matrix-valued symbols $\{P_n\} \sim_{\sigma} (h, I_d)$ and $\{P_n\} \sim_{\lambda} (h, I_d)$ if f, g Hermitian-valued (g positive definite).

Example of Structured Precond.: $T_n(f)$, $f(\theta) = \theta^4$

- Preconditioner $T_n(g)$, $g(\theta) = 16\sin^4(\theta/2)$, $h = f/g$;
- $m =$ess-inf $h = 1$, $M =$ess-sup $h = \pi^4/16$ and hence $m < M$.

Consequently the eigs of the preconditioned matrix $P_n = T_n^{-1}(g)T_n(f)$ behave as the values $h\left(\frac{\pi j}{n+1}\right)$ and they all lie in $(m, M) = (1, \pi^4/16)$:

$$\Downarrow$$

The preconditioner has a pentadiagonal structure and the related PCG is optimal:

- $O(1)$ iterations (available precise bounds since $(m, M) = (1, \pi^4/16)$)
- and optimal cost per iteration $O(n\log(n))$, inherited from the FFT.

IgA, degree p, d dimensions: spectral distribution and classical multigrid

Theorem $\boxed{n^{d-2}A_n^{[p]} \approx T_n(f_p)}$ **and hence** $\boxed{\{n^{d-2}A_n^{[p]}\}_n \sim_\lambda (f_p, I_d)}$

f_p has the expected zero of order 2 at zero, positive elsewhere but it collapses to zero exponentially with p at the boundaries of $I_d = (-\pi, \pi)^d$.

n	$p = 1$	$p = 3$	$p = 5$
16	0.16	0.64	0.96
28	0.17	0.64	0.96
40	0.18	0.64	0.96
52	0.18	0.65	0.96
n	$p = 2$	$p = 4$	$p = 6$
17	0.27	0.88	0.99
29	0.27	0.88	0.99
41	0.29	0.88	0.99
53	0.30	0.88	0.99

Table: spectral radius: standard twogrid, 2D, relaxed GS as smoother

$\approx$ denotes equality up to matrix-sequences with zero symbol.

Multi-iterative idea: either PCG or PGMRES as smoother

A suitable smoother is suggested by the symbol f_p itself: **Structured Preconditioner**.

Take as smoother the PCG or the PGMRES with preconditioner having itself a symbol s_p which 'deletes' the numerical zeros of our symbol f_p, yielding a p-independent preconditioned symbol $s_p^{-1} f_p$

$$\boxed{s_p(\theta_1, \ldots, \theta_d) = h_{p-1}(\theta_1) h_{p-1}(\theta_2) \cdots h_{p-1}(\theta_d)}$$

$$\boxed{s_p^{-1} f_p \sim \text{ symbol of the standard FD Laplacian}}$$

Possible choice of the preconditioner: Toeplitz matrix generated by s_p.

Remark A linear system associated with the preconditioner is easily solvable, since the symbol s_p is a **separable trigonometric polynomial**: as a consequence the preconditioner is a tensor product of banded matrices so that only a linear computational cost is required.

IgA, degree p, d dimensions: structured PCG/PGMRES and multigrid (const. coeff.... but the technique is equally effective for var. coeff. and singular mappings)

We consider the system $nA_n^{[p]}\mathbf{u} = \mathbf{b}$ coming from the IgA approximation of

$$\begin{cases} -\Delta u = 1 & \text{in } (0,1)^3 \\ u = 0 & \text{on } \partial(0,1)^3 \end{cases}$$

For the solution: V-cycle and W-cycle multigrid

n	$p=1$		n	$p=3$		n	$p=5$	
16	10	7	14	7	6	12	8	8
32	11	7	30	8	6	28	8	7
64	12	7	62	9	6	60	9	6

n	$p=2$		n	$p=4$		n	$p=6$	
15	9	8	13	7	6	11	9	9
31	8	7	29	8	6	27	8	6
63	9	7	61	9	6	59	10	6

Table: number of iterations: 3D with structured PCG/PGMRES

Sequences of diagonal (uniform) sampling matrices

- Let $a(x)$ be Riemann integrable over $(0,1)$ and let us consider the (uniform) diagonal sampling matrix of size n

$$D_n(a) = \begin{pmatrix} a(h) & & & \\ & a(2h) & & \\ & & \ddots & \\ & & & a(nh) \end{pmatrix}, \quad h = \frac{1}{n+1}.$$

It is plain to see that

$$\{D_n(a)\} \sim_{\lambda,\sigma} (a,\Omega), \quad \Omega = (0,1).$$

- The result is also immediate for domains $\Omega \in \mathbb{R}^d$ measurable according to Peano-Jordan (and even for matrix-valued symbols).
- Are we satisfied? ... We have to wait a bit ...

Toeplitz structures and approx. Differential Operators

For simplicity equi-spaced FD, centered and with precision order 2

$$-u^{('')} \Rightarrow \Delta_n^{(2)} = \begin{pmatrix} 2 & -1 & & & \\ -1 & \ddots & \ddots & & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix} = T_n(f), \quad f(\theta) = 4\sin^2(\theta/2),$$

$$u^{iv} \Rightarrow \Delta_n^{(4)} = \begin{pmatrix} 6 & -4 & 1 & & \\ -4 & \ddots & \ddots & \ddots & \\ 1 & \ddots & & \ddots & 1 \\ & \ddots & \ddots & \ddots & -4 \\ & & 1 & -4 & 6 \end{pmatrix} = T_n(f^2) \approx \left[\Delta_n^{(2)}\right]^2.$$

$\approx$ denotes equality up to matrix-sequences with zero symbol. (Using a **'Preconditioning'** terminology: the terms can be viewed as **'low rank'** + **'low norm'**)

Locality and variable coefficients

Let us consider the variable coeff. version of $\mathcal{L}(u) = -u^{('')}$ in divergence form: $\mathcal{L}_a(u) = -\left(a(x)u^{'}\right)^{'}$ [rod with variable section].
For simplicity the same equispaced FD, centered with precision order 2: setting $a_j = a((j-1/2)h), d_j = a_{j-1} + a_j$ we find

$$\mathcal{L}_a(u) \Rightarrow \Delta_n^{(2)}(a) = \begin{pmatrix} d_1 & -a_2 & & \\ -a_2 & \ddots & \ddots & \\ & \ddots & \ddots & -a_n \\ & & -a_n & d_n \end{pmatrix}.$$

$$\Delta_n^{(2)}(1) = T_n(f),$$

but for nonconstant a the Toeplitz character seems to be lost. In fact, we find it again **'in a approximated sense'** and in **'local scale'**.

Locality I: Dyadic representation

Let us consider the variable coeff. version of $\mathcal{L}(u) = -u^{('')}$ in divergence form: $\mathcal{L}_a(u) = -\left(a(x)u^{'}\right)^{'}$.
For simplicity the same equispaced FD, centered with precision order 2: setting $a_j = a((j-1/2)h)$ we find

$$\Delta_n^{(2)}(a) = \sum_{j=1}^{n+1} a_j \Psi_j, \qquad \Delta_n^{(2)}(1) = T_n(f) = \sum_{j=1}^{n+1} \Psi_j,$$

$$\Psi_j = \mathbb{O} \oplus \Theta \oplus \mathbb{O}, \quad \Theta = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \end{pmatrix}.$$

Setting $\Delta_n^{(2)}(1) = T_n(f)$, for nonconstant a, we deduce

$$\begin{aligned} [\min a]\, T_n(f) &\leq \Delta_n^{(2)}(a) \leq [\max a]\, T_n(f), \\ \lambda_{\min}(\Delta_n^{(2)}(a)) &\sim \frac{c_a}{n^2}, \\ \lambda_{\max}(\Delta_n^{(2)}(a)) &< \max_{x,\theta} a(x) f(\theta). \end{aligned}$$

Locality II: Toeplitz + Diagonal representation

$$\Delta_n^{(2)}(a) \approx D_n(a) \begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix} = D_n(a) T_n(f)$$

$$\Delta_n^{(2)}(a) \approx D_n^{1/2}(a) T_n(f) D_n^{1/2}(a), \quad \psi = a(x) f(\theta).$$

$D_n(a)$ uniform diagonal sampling matrix; the decompositions (both Dyadic and Toeplitz + Diagonal) available in the multidimensional:

$$\mathcal{L}_a(u) = -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i} \left(a_{i,j}(x) \frac{\partial u}{\partial x_j} \right) \Rightarrow$$

$$\Delta_{\mathbf{n}}^{(2)}(a) \approx \sum_{i,j=1}^{d} D_{\mathbf{n}}(a_{i,j}) \Delta_{\mathbf{n},i,j}^{(2)}, \quad \Delta_{\mathbf{n},i,j}^{(2)} = T_{\mathbf{n}}(f_{i,j})$$

$$\psi = \sum_{i,j=1}^{d} a_{i,j}(x) f_{i,j}(\theta)$$

Locality III: Toeplitz + Diagonal representation

Variable coefficients with non equispaced grid with $t_j = g(jh)$, $g([0,1]) = [0,1]$, diffeomorfism. Setting

$$\phi(a,g) = \frac{a(g)}{(g^{'})^2}$$

we have

$$-\left(a(x)u^{'}\right)^{'} \Rightarrow_{DF\ su\ \ g(jh)} \Delta_n^{(2)}(a,g) \approx \Delta_n^{(2)}(\phi(a,g)),$$

$$-\left(\phi(a,g)(x)u^{'}\right)^{'} \Rightarrow_{DF\ su\ \ jh} \Delta_n^{(2)}(\phi(a,g)) \approx D_n(\phi(a,g))T_n(f)$$

$$\psi = \phi(a,g)(x)f(\theta)$$

Using a Geometric Map the Toeplitz + Diagonal representation can be extended to the multidimensional: used in the IgA, FEM with 'graded' grids, etc.

Approximation Theory for matrix-sequences I

Definition [S. LAA 01]
For $\{A_n\}$ con $d_n < d_{n+1}$
$\{\{B_{n,m}\}\}_m$, $m \in \mathbb{N}$ is an a.c.s. (approximating class) if

$$A_n = B_{n,m} + R_{n,m} + N_{n,m}, \quad \forall n > n_m, \ \forall m \in \mathbb{N},$$

$$\mathrm{rank}\ R_{n,m} \le d_n\, c(m), \quad \|N_{n,m}\| \le \omega(m).$$

The quantities n_m, $c(m)$ and $\omega(m)$ are functions of m and

$$\lim_{m\to\infty} \omega(m) = 0, \quad \lim_{m\to\infty} c(m) = 0.$$

Approximation Theory for matrix-sequences II

Theorem [S. LAA 01-06]
Assuming the following

- $\{\{B_{n,m}\}\}_m$, $m \in \hat{\mathbb{N}} \subset \mathbb{N}$, $\#\hat{\mathbb{N}} = \infty$ a.c.s. per $\{A_n\}$,
- $\{B_{n,m}\} \sim_\sigma (\psi_m, D)$ $(\{B_{n,m}\} \sim_\lambda (\psi_m, D))$,
- $\psi_m \to_\mu \psi$,

we obtain

$$\{A_n\} \sim_\sigma (\psi, D) \ \ (\{A_n\} \sim_\lambda (\psi, D)).$$

In the case of eigenvalues every involved matrix-sequence has to be Hermitian (or the non-Hermitian perturbation has to satisfy a trace norm condition; trace norm = sum of all singular values).

GLT: Generalized Locally Toeplitz [S. et al, 03-15]

We know a lot on spectral features of either Toeplitz or Diagonal matrix-sequences: exploiting these 'two ingredients' we build up a class of matrix-sequences called Generalized Locally Toeplitz (GLT):

- for a Riemann integrable over $[0,1]$ and f being $L^1(-\pi,\pi)$, we define $LT_n^m(a,f) = D_m(a) \otimes T_{n/m}(f)$;
- a sequence $\{A_n\}$ is sLT if $\{\{LT_n^m(a,f)\}\}_m$ is an a.c.s. for $\{A_n\}$: in that case $a(x)f(\theta)$, $(x,\theta) \in [0,1] \times (-\pi,\pi)$, is the symbol of the sequence of matrices $\{A_n\}$;
- a sequence $\{A_n\}$ is GLT with respect to the measurable function $\kappa(x,\theta)$ if for every $\epsilon > 0$, there exist $\{A_n^{(j,\epsilon)}\}$ sequences sLT with symbol $a_{(j,\epsilon)}(x)f_{(j,\epsilon)}(\theta)$, $N_\epsilon \geq j \geq 1$ such that
  - $\sum_{j=1}^{N_\epsilon} a_{(j,\epsilon)}(x)f_{(j,\epsilon)}(\theta)$ converges a.e. to $\psi(x,\theta)$;
  - $\{\{\sum_{j=1}^{N_\epsilon} A_n^{(j,\epsilon)}\}\}_{m,\epsilon}$ is a.c.s. for $\{A_n\}$.
- a GLT sequence $\{A_n\}$ GLT with respect to the measurable function $\psi(x,\theta)$ has $\psi(x,\theta)$ as symbol (extension for $d > 1$).

GLT as super-algebra containing $\{D_n(a)\}$, $\{T_n(f)\}$, a Riemann integrable, f belonging to L^1

- Any linear combination of products (and inverses) involving uniform sampling diagonal matrix-sequences and Toeplitz sequences is GLT and has as symbol the function obtained by the same operations on the symbols...... sequences distributed in the singular value sense as the zero function are GLT with $\psi \equiv 0$.
- Surprisingly enough, we prove formally that any 'reasonable' approximation by local methods (Finite Differences, Finite Elements, IgA etc.) of PDEs leads to GLT sequences, i.e., to matrices that can be approximated by linear combinations of products involving uniform sampling diagonal matrix-sequences and Toeplitz sequences.
- A_n (from Finite Differences on a convection-diff Pb) is approximated by $D_n(a)T_n(4-2\cos(\theta_1)-2\cos(\theta_2))$ and this explains why its eigs are an approximated uniform sampling of

$$\psi(x,\theta) = a(x)(4-2\cos(\theta_1)-2\cos(\theta_2)), \quad x \in \Omega, \theta \in (-\pi,\pi)^2.$$

Concrete Examples in a differential context

- FEM approximation of degree p in d dimensions
- Q2Q1 approximation of the Linear Elasticity
- IgA approximation of degree p in d dimensions

Elasticity - as a part of the 'rebound' analysis in the Glacial Isostatic Adjustment (GIA) models, Donation KAW 2013.0341, Knut & Alice Wallenberg Foundation, in collaboration with the Royal Swedish Academy of Sciences

FEM: of degree p on a d dimensional domain

We consider the Laplacian over $[0,1]^d$ and we denote by $A_n^{(p)}$ the degree p FEM matrix on quadrilaterals.

- There exists a permutation matrix Π such that

$$\Pi A_n^{(p)} \Pi^T \approx T_n(f);$$

- f is defined over $I_d = (-\pi, \pi)^d$ and Hermitian matrix-valued with size p^d (anyy comment is redundant!);
- hence, the eigs of $A_n^{(p)}$ are divided into p^d branches (of the same cardinality), each of them represented by a different real-valued eigenvalue of f: $\lambda_1(f) \leq \ldots \leq \lambda_{p^d}$;
- the spreading of the spectrum, measured by the ratio

$$\frac{\max(\lambda_{p^d})}{\max(\lambda_1),}$$

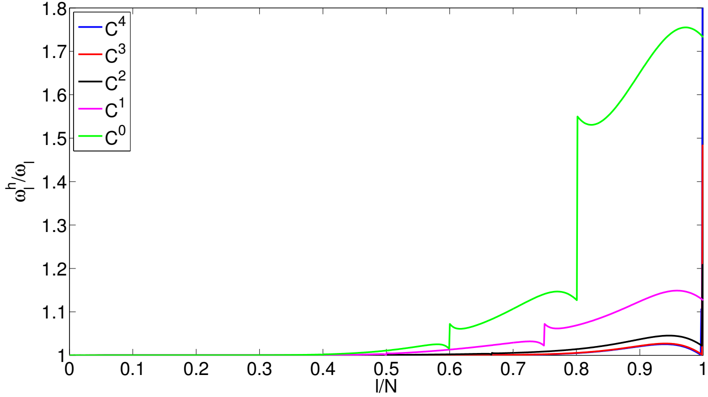
depends on the choice of the basis (Legendre, integrated Legendre, Bernstein etc).

IgA: of degree p on a d dimensional domain

We consider the Laplacian over $[0,1]^d$ and we denote by $A_n^{(p)}$ the spline-degree p IgA matrix.

- It holds $A_n^{[p]} \approx T_n(f)$ so that $\{A_n^{(p)}\} \sim_\lambda (f, I_d)$, $I_d = (-\pi, \pi)^d$;
- f is defined over $I_d = (-\pi, \pi)^d$, is scalar-valued, nonnegative with a unique zero at zero (as in the FD case: it is somehow the revenge of the smoothness);
- the function f tends exponentially to zero as p in every point of the type $\theta = (\theta_1, \ldots, \theta_d)$ for which $\theta_j = \pi$ for some j;
- the latter property induces a bad conditioning in the high frequency subspace, growing exponentially with p and which is not expected for a differential problem: the knowledge of the symbol is an essential guide for finding the right preconditioner.

Comparison IgA-FEM (and furthermore the case of intermediate regularity): **the picture has a clear interpretation as the revenge of the smoothness** (thanks to A. Reali)

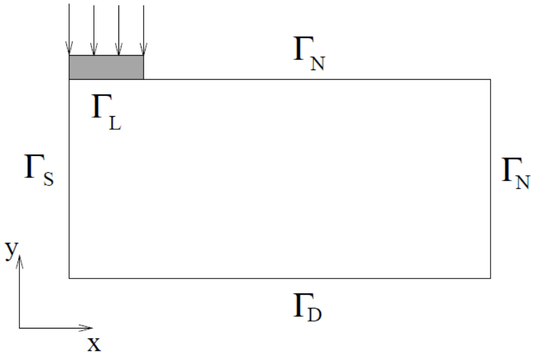


Elasticity Pb: some simplifications

To enable fully incompressible models, i.e. $\lambda = \infty$, we write

$$\begin{aligned} -\nabla \cdot (2\mu\varepsilon(\mathbf{u})) - \nabla(\mathbf{u} \cdot \nabla p_0) - \mu \nabla p &= \mathbf{f} \text{ in } \Omega \\ \mu \nabla \cdot \mathbf{u} - \frac{\mu^2}{\lambda} p &= 0 \text{ in } \Omega \end{aligned}$$

- p_0 is the pre-stress,
- $p = \frac{\lambda}{\mu} \nabla \cdot \mathbf{u}$ is the kinematic pressure.



Q2Q1 for the Elasticity Pb: two-by-two block structure (Stokes, NS, Cahn-Hilliard, PDE constrained opt. etc)

We use the stable pair of spaces Q2-Q1; we obtain a block structure $\mathcal{A}$ with a block factorization

$$\mathcal{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ A_{21} & S_{\mathcal{A}} \end{bmatrix} \begin{bmatrix} I_1 & A_{11}^{-1} A_{12} \\ 0 & I_2 \end{bmatrix},$$
$$S_{\mathcal{A}} = A_{22} - A_{21} A_{11}^{-1} A_{12}.$$

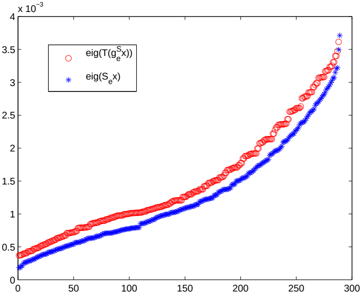
Several possible preconditioners. One option is

$$\mathcal{D} = \begin{bmatrix} D_{11} & 0 \\ A_{21} & S \end{bmatrix},$$

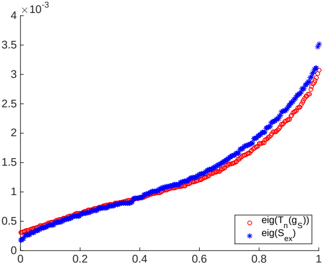
$$D_{11} \approx A_{11}, \quad \text{or} \quad D_{11} \approx A_{11}^{-1} \quad \text{and} \quad S \approx S_{\mathcal{A}}$$

To be computationally efficient, in the majority of the cases, the Schur complement must be approximated.

Approximation of the exact Schur compl $S_{\mathcal{A}}$ (case Q1-Q1): eig(S) in blue vs the GLT symbol $g_S(\theta_1, \theta_2)$ in red



Approximation of the exact Schur compl $S_{\mathcal{A}}$ (case Q2-Q1): eig(S) in blue vs the GLT symbol $g_S(\theta_1, \theta_2)$ in red



IgA, degree p, d dimensions: spectral distribution

Theorem $\boxed{n^{d-2} A_n^{[p]} \approx T_n(f_p)}$ **and hence** $\boxed{\{n^{d-2} A_n^{[p]}\}_n \sim_\lambda (f_p, I_d)}$

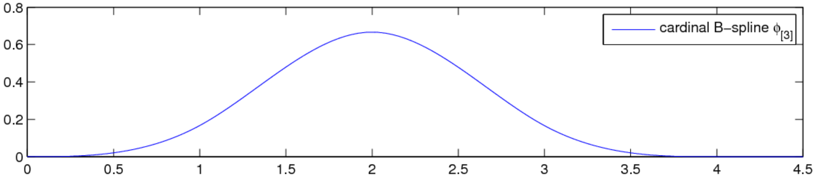
(GLT matrix sequence with symbol $f_p : I_d \to \mathbb{R}$, $I_d = (-\pi, \pi)^d$)

$$f_p(\theta_1, \ldots, \theta_d) = \sum_{k=1}^{d} h_p(\theta_1) \cdots h_p(\theta_{k-1}) f_p(\theta_k) h_p(\theta_{k+1}) \cdots h_p(\theta_d)$$

$$f_p(\theta) = (2 - 2\cos\theta) h_{p-1}(\theta)$$

$$h_p(\theta) = \phi_{[2p+1]} + 2 \sum_{k=1}^{p} \phi_{[2p+1]}(p + 1 - k) \cos(k\theta)$$

$\phi_{[q]}$ = cardinal B-spline of degree q on the nodes $0, 1, \ldots, q + 1$



Properties of the symbol I

For $d = 1$ the symbol is $f_p(\theta) = (2 - 2\cos\theta)h_{p-1}(\theta)$

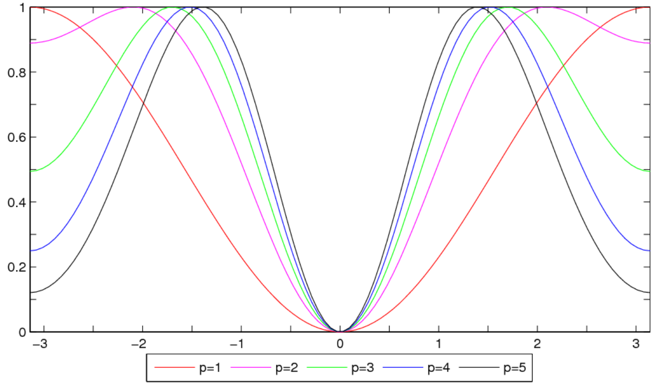


Figure: graph of the normalized symbol f_p/M_{f_p}

- $\lim_{\theta\to 0} \frac{f_p(\theta)}{\theta^2} = 1$, $f_p(\theta) > 0$ for $\theta \neq 0 \quad \Rightarrow \quad \theta = 0$ unique zero of f_p (order 2)
- setting $M_{f_p} = \max_\theta f_p(\theta)$, $\frac{f_p(\pi)}{M_{f_p}} \leq \frac{f_p(\pi)}{f_p\left(\frac{\pi}{2}\right)} = \frac{1}{2^{p-2}} \to 0$ exponentially

Properties of the symbol II

The normalized symbol f_p/M_{f_p} has a numerical zero at $\theta = \pi$ for large p!

Besides the canonical zero $\theta = 0$, when p is large the normalized symbol has a non-canonical numerical zero at $\theta = \pi$.

In the d-variate case, the situation is even worse.

Besides the canonical zero $(\theta_1, \ldots, \theta_d) = (0, \ldots, 0)$, when p is large the normalized symbol has **infinitely many** non-canonical numerical zeros located at the π-edge points

$$\{(\theta_1, \ldots, \theta_d) : \ \theta_j = \pi \text{ for some } j\}$$

Design of fast iterative solvers: use the symbol

From the properties of the symbol:

- standard multigrid methods for the IgA matrices, which take care of the actual zero $(\theta_1, \ldots, \theta_d) = (0, \ldots, 0)$, will be optimal, i.e. with convergence rate independent of the discretization parameter n
- for large p, standard multigrid methods, which do not take care of the numerical zeros at the π-edge points, will have a bad convergence rate

$$\Downarrow$$

multi-iterative idea (Grenbaum SINUM 84, S. CMA 93) to be fully considered for designing optimal and robust solvers

Target: **use carefully the symbol** to design fast (optimal and robust) multi-iterative solvers for the IgA matrices

Multi-iterative methods: the multi-iterative idea

n	$p = 1$	$p = 3$	$p = 5$
16	0.16	0.64	0.96
28	0.17	0.64	0.96
40	0.18	0.64	0.96
52	0.18	0.65	0.96
n	$p = 2$	$p = 4$	$p = 6$
17	0.27	0.88	0.99
29	0.27	0.88	0.99
41	0.29	0.88	0.99
53	0.30	0.88	0.99

Table: spectral radius: standard twogrid, 2D, with relaxed GS as smoother

Multi-iterative idea: keep the classical full-weighting projector for dealing with the zero $(\theta_1, \ldots, \theta_d) = (0, \ldots, 0)$ and replace the relaxed Gauss-Seidel smoother with another smoother that takes care of the numerical zeros of the symbol: **PCG with structured precond.**

Multi-iterative methods: PCG or PGMRES as smoother

A suitable smoother is suggested by the symbol f_p itself.

Take as smoother the PCG or the PGMRES with preconditioner having itself a symbol s_p which 'deletes' the numerical zeros of our symbol f_p, yielding a p-independent preconditioned symbol $s_p^{-1} f_p$

$$\boxed{s_p(\theta_1, \ldots, \theta_d) = h_{p-1}(\theta_1) h_{p-1}(\theta_2) \cdots h_{p-1}(\theta_d)}$$

$$\boxed{s_p^{-1} f_p \sim \text{ symbol of the standard FD Laplacian}}$$

Possible choice of the preconditioner: Toeplitz matrix generated by s_p.

Remark A linear system associated with the preconditioner is easily solvable, since the symbol s_p is a **separable trigonometric polynomial**: as a consequence the preconditioner is a tensor product of banded matrices so that only a linear computational cost is required.

Multi-iterative methods: twogrid-PCG experiments, 2D

Here we consider the system $A_n^{[p]}\mathbf{u} = \mathbf{b}$ coming from the IgA approximation of

$$\begin{cases} -\Delta u + u = 1 & \text{in } (0,1)^2 \\ u = 0 & \text{on } \partial(0,1)^2 \end{cases}$$

n	$p = 1$	$p = 3$	$p = 5$
40	6	6	6
60	6	6	6
80	5	6	6
100	5	6	6
120	5	6	6
n	$p = 2$	$p = 4$	$p = 6$
41	6	6	6
61	6	6	5
81	6	6	5
101	6	6	5
121	6	6	5

Table: number of iterations

Multi-iterative methods: multigrid-PCG experiments in 2D

Here we consider the system $A_n^{[p]}\mathbf{u} = \mathbf{b}$ coming from the IgA approximation of

$$\begin{cases} -\Delta u = 1 & \text{in } (0,1)^2 \\ u = 0 & \text{on } \partial(0,1)^2 \end{cases}$$

For the solution: V-cycle and W-cycle multigrid

n	$p = 1$		n	$p = 3$		n	$p = 5$	
16	10	7	14	7	6	12	7	7
32	11	7	30	9	6	28	8	6
64	12	7	62	9	6	60	10	6
128	13	7	126	10	6	124	11	6
256	13	7	254	11	6	252	12	6

n	$p = 2$		n	$p = 4$		n	$p = 6$	
15	8	6	13	7	6	11	7	7
31	9	6	29	8	6	27	8	6
63	10	6	61	10	6	59	10	6
127	11	6	125	11	6	123	11	6
255	12	7	253	12	6	251	12	6

Table: number of iterations

Multi-iterative methods: multigrid-PCG experiments in 3D

Here we consider the system $nA_n^{[p]}\mathbf{u} = \mathbf{b}$ coming from the IgA approximation of

$$\begin{cases} -\Delta u = 1 & \text{in } (0,1)^3 \\ u = 0 & \text{on } \partial(0,1)^3 \end{cases}$$

For the solution: V-cycle and W-cycle multigrid

n	$p = 1$		n	$p = 3$		n	$p = 5$	
16	10	7	14	7	6	12	8	8
32	11	7	30	8	6	28	8	7
64	12	7	62	9	6	60	9	6

n	$p = 2$		n	$p = 4$		n	$p = 6$	
15	9	8	13	7	6	11	9	9
31	8	7	29	8	6	27	8	6
63	9	7	61	9	6	59	10	6

Table: number of iterations

Conclusion: We have obtained **optimal and robust multi-iterative multigrid methods for the IgA Laplacian $A_n^{[p]}$.**

Variable coefficients: symbol of the IgA matrix-sequences associated to a full elliptic Pb (a GLT sequence)

Full elliptic problem:

$$\begin{cases} -\nabla \cdot K \nabla u + \boldsymbol{\beta} \cdot \nabla u + \gamma u = f & \text{on } \Omega \subset \mathbb{R}^d \\ u = 0 & \text{on } \partial\Omega \end{cases}$$

IgA approximation: take a geometry map $\mathbf{G} : [0,1]^d \to \overline{\Omega}$ to transfer the problem from Ω to $[0,1]^d$; on $[0,1]^d$ use again splines of deg. p.

$$\mathcal{A}_n^{[p]} = \text{resulting IgA approximation matrix}$$

Theorem $\{n^{d-2}\mathcal{A}_n^{[p]}\}$ sequence of matrices in the GLT super-algebra

$$\boxed{\{n^{d-2}\mathcal{A}_n^{[p]}\} \sim_\lambda \mathbf{1}\big(|\det(J_{\mathbf{G}}(x_1,\ldots,x_d))| K_{\mathbf{G}}(x_1,\ldots,x_d) \circ H_p(\theta_1,\ldots,\theta_d)\big)\mathbf{1}^T}$$

$$K_{\mathbf{G}} = (J_{\mathbf{G}})^{-1} K(\mathbf{G}) (J_{\mathbf{G}})^{-T}, \quad J_{\mathbf{G}} = \text{Jacobian matrix of } \mathbf{G}$$

$H_p =$ symmetric $d \times d$ matrix whose (i,j) entry represents the 'formula' used to approximate $\partial^2/\partial x_i \partial x_j$

Variable coefficients: fast multi-iterative solver for IgA matrices in full elliptic problems

So far, we have seen fast multi-iterative solvers for $A_n^{[p]}$ = the IgA Laplacian over the hypercube $(0,1)^d$

What about $\mathcal{A}_n^{[p]}$?

The good news: the IgA Laplacian $A_n^{[p]}$, with proper diagonal scalings in the case of degeneracy of the coefficients, is an optimal and robust CG/GMRES preconditioner for $\mathcal{A}_n^{[p]}$

Example of P-GMRES optimality and robustness I

Let us take

$$\Omega = \{(x,y) \in \mathbb{R}^2 :\ r^2 < x^2 + y^2 < R^2,\ x > 0,\ y > 0\} \qquad (r = 1,\ R = 4)$$

$$K(x,y) = \begin{bmatrix} (2+\cos x)(1+y) & \cos(x+y)\sin(x+y) \\ \cos(x+y)\sin(x+y) & (2+\sin y)(1+x) \end{bmatrix}$$

$$\beta(x,y) = \sqrt{x^2+y^2} \begin{bmatrix} \cos \frac{x}{\sqrt{x^2+y^2}} \\ \sin \frac{y}{\sqrt{x^2+y^2}} \end{bmatrix}$$

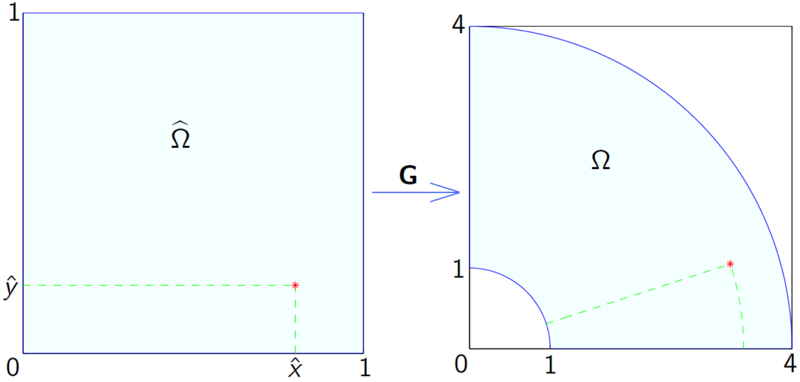
$$\gamma(x,y) = xy$$

$$f(x,y) = x\cos y + y\sin x$$

and...

Example of P-GMRES optimality and robustness II

$\mathbf{G} : \widehat{\Omega} := [0,1]^2 \to \overline{\Omega}$ $\qquad \mathbf{G}(\hat{x}, \hat{y}) = (x, y)$ $\qquad \begin{cases} x = [r + \hat{x}(R - r)]\cos(\frac{\pi}{2}\hat{y}) \\ y = [r + \hat{x}(R - r)]\sin(\frac{\pi}{2}\hat{y}) \end{cases}$



then...

Example of P-GMRES optimality and robustness III

n	$p = 1$	$p = 2$	$p = 3$	$p = 4$	$p = 5$	$p = 6$
10	16	18	19	21	22	23
20	20	21	22	23	24	25
30	22	23	23	24	25	26
40	23	24	24	25	26	26
50	24	24	25	26	26	27

Table: number of P-GMRES iterations (Collocation IgA)

Further Technical Insights

- The symbol can be recovered in the IgA Collocation/Galerkin setting with variable coefficient PDEs, general physical domain, general geometrical mapping.

- The symbol can be recovered in the FEM setting with variable coefficient PDEs, general physical domain, general graded griddings.

- Concerning the numerical methods, the dimensionality d is not an issue and singular mappings are not an issue.

- We are now completing the analysis when the model space is given by NURBS.

Conclusions

- In the case of constant coefficients PDEs the GLT approach and the Local Fourier Analysis lead to the same conclusions and to the same tools.
- The GLT tool has to be considered as an extension of the Local Fourier Analysis (for variable coefficients, irregular domains etc) and indeed the symbol analysis via GLT is more general and includes also integral problems, preconditioning, involved iteration matrices (PHSS), variable coefficients.
- Future work: Navier-Stokes and other vector problems to be considered, with the idea of using the spectral information and the symbol, in order to obtain faster and more robust (preconditioned) iterative solvers.

References

- GLT: Tilli LAA 98, S. LAA 03 e 06 (previous results by Kac, Parter, Widom etc)
- a.c.s: S. LAA 01, 03, 06 (previous results by Tilli)
- Spectral Tools (Bottcher, Grudky, Silbermann, Tilli, Tyrtyshnikov, Golinskii, Kuijlaars, S. + coauthors)
- FEM: Beckermann, S. SINUM 07, Garoni, S., Sesana, SIMAX 15 (to appear)
- IgA: Garoni, Manni, Pelosi, S., Speleers NM 14, Donatelli, Garoni, Manni, S., Speleers MC 15, CMAME 14, CMAME 15
- Vector Problems: collaborations with Donatelli, Dorostkar, Garoni, Hughes, Manni, Mazza, Neytcheva, Reali, Sesana, Speleers