Multigrid: miscellaneous aspects

Harald van Brummelen

TU/e, Dept. of Mechanical Engineering, Multiscale Engineering Fluid Dynamics section

TU/e, Dept. of Mathematics and Computer Science, Centre for Analysis, Scientific computing and Applications (CASA)

Mathematical and Numerical Methods for Multiscale Problems
Multigrid Methods

Outline

Variational formulations

Consider:

1. An open bounded domain $\Omega \subset \mathbb{R}^d$
2. A Hilbert space $U \subset L^2(\Omega)$ with inner product $(\cdot,\cdot)_U$ (typically, Sobolev spaces)
3. The generic variation problem:

$$\boxed{\text{Find } u \in U: \qquad a(u,v) = b(v) \qquad \forall v \in U} \tag{P}$$

with $a : U \times U \to \mathbb{R}$ a (coercive and) bounded bilinear form and $b : V \to \mathbb{R}$ a bounded linear form

Variational formulations

Example: variational formulation of Poisson's problem

$$\begin{aligned} -\Delta u &= f \qquad \text{in } \Omega \\ u &= 0 \qquad \text{at } \partial\Omega \end{aligned}$$

$\Rightarrow$

1. $U = H_0^1(\Omega) = \left\{ u \in L^2(\Omega) : \nabla u \in L^2(\Omega, \mathbb{R}^d), u|_{\partial\Omega} = 0 \right\}$
2. $a(u,v) = \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x \qquad b(v) = \int_\Omega f v \, \mathrm{d}x$

associated with a and b are an operators $A : U \to U'$ and a functional b:

$$a(u,v) = \langle v, Au \rangle_{U,U'} \qquad b(v) = \langle v, b \rangle_{U,U'}$$

Galerkin methods

Conforming approximation spaces

Let $\{U_l\}$ denote a sequence of finite-dimensional asymptotically-dense nested subspaces of U:

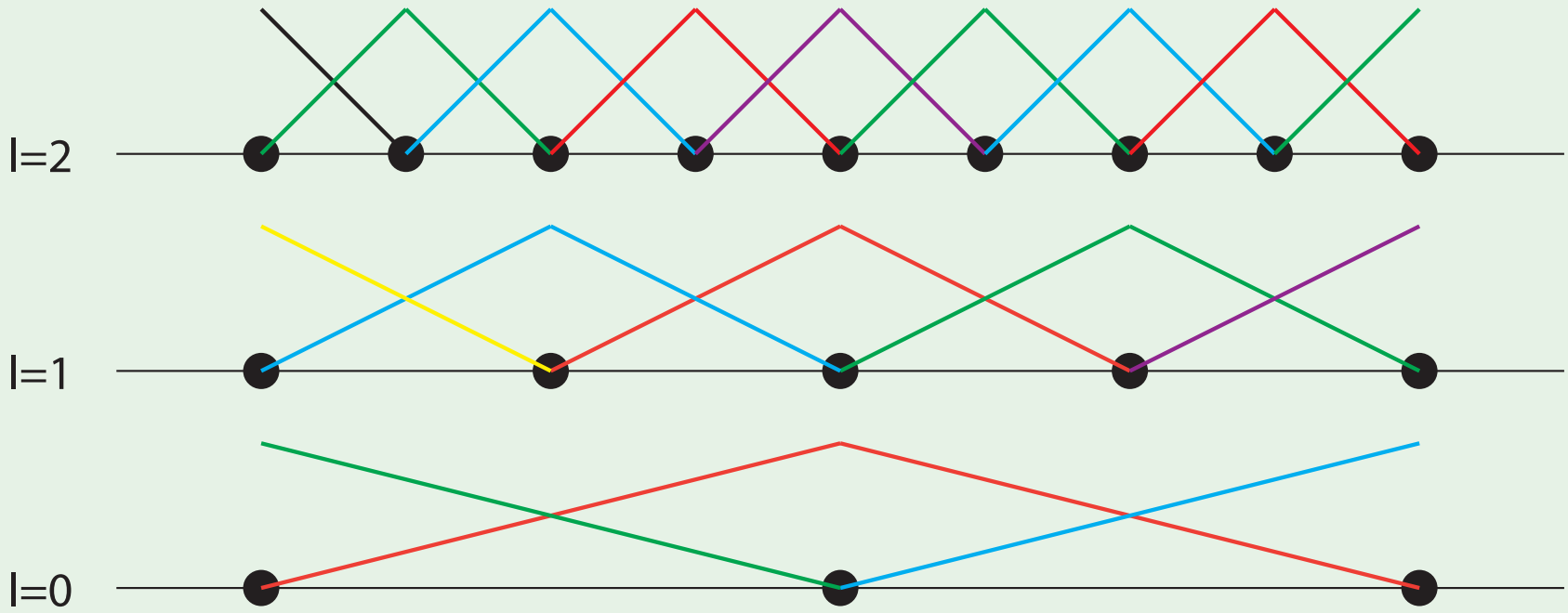
1. $U_0 \subset U_1 \subset \ldots \subseteq U$ (nesting)
2. $U_l \to U$ as $l \to \infty$ (asymptotic density)

for all $u \in U$ and all $\epsilon > 0$ exists $l := l_{\epsilon,u}$ s.t. $\inf_{w \in U_l} \|u - w\|_U < \epsilon$

Galerkin methods

Example

The standard hat-functions on a sequence of hierarchically refined meshes are nested and asymptotically dense in H^1.



Galerkin methods

Galerkin formulation

Consider $U_l \subset U$. The Galerkin approximation of (P) is:

$$\boxed{\text{Find } u \in U_l : \qquad a(u,v) = b(v) \qquad \forall v \in U_l} \qquad (\mathrm{P}_l)$$

Galerkin methods

Galerkin formulation

Consider $U_l \subset U$. The Galerkin approximation of (P) is:

$$\boxed{\text{Find } u \in U_l : \qquad a(u,v) = b(v) \qquad \forall v \in U_l} \qquad (\mathrm{P}_l)$$

Remarks

1. boundedness and coercivity transfer to subspaces: if (P) is well-posed, then so is (P_l)
2. Convergence by asymptotic density: $u_l \to u$ as $l \to \infty$

The VMS paradigm

Consider the additive decomposition $U = U_l \oplus U_l^{\perp}$ (or $U_{l+1} = U_l \oplus U_l^{\perp}$)
The variational problem can be decomposed into: Find $\bar{u} \in U_l$ (coarse-scale component) and $\hat{u} \in U_l^{\perp}$ (fine-scale component) s.t.*:

$$a(\bar{u}, \bar{v}) + a(\hat{u}, \bar{v}) = b(\bar{v}) \qquad \forall \bar{v} \in U_l$$
$$a(\bar{u}, \hat{v}) + a(\hat{u}, \hat{v}) = b(\hat{v}) \qquad \forall \hat{v} \in U_l^{\perp}$$

The coarse-scale problem can be recast into

$$\boxed{\text{Find } \bar{u} \in U_l : \qquad a(\bar{u}, \bar{v}) = b(\bar{v}) - a(\hat{u}, \bar{v}) \qquad \forall \bar{v} \in U_l} \qquad \text{(VMS)}$$

The term $-a(\hat{u}, \bar{v})$ represents the effect of the fine scales on the coarse scales.

*T.J.R. Hughes, *Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods*, Comput. Methods Appl. Mech. Engrg. **127** (1995), 387–401.

The VMS paradigm: interpretation

Orthogonal projections

The orthogonal projection P_X onto a subspace $X \subset U$ is implicitly defined via the inner product on U:

$$P_X u \in X : \qquad \big(P_X u, v\big)_U = (u, v)_U \qquad \forall v \in X$$

Consider an arbitrary $u \in U$, with additive decomposition $u = u_0 + u_1$, $u_0 \in U_l$ and $u_1 \in U_l^\perp$. By definition:

$$u_0 = P_{U_l} u \qquad u_1 = P_{U_l^\perp} u = \big(\mathrm{Id} - P_{U_l}\big) u$$

The VMS paradigm: interpretation

VMS interpretation

Implicit in the splitting $\bar{u} \in U_l$ and $\hat{u} \in U_l^{\perp}$ and the VMS equation:

$$\boxed{a(\bar{u}, \bar{v}) = b(\bar{v}) - a(\hat{u}, \bar{v}) \qquad \forall \bar{v} \in U_l} \qquad \text{(VMS)}$$

is $\bar{u} = P_{U_l} u$ with u the actual solution to (P)

$\Rightarrow$ the term $-a(\hat{u}, \bar{v})$ in the rhs ensures that the approximation $\bar{u}$ is the U-projection of the actual solution onto U_l.

The VMS paradigm: interpretation

VMS interpretation

Implicit in the splitting $\bar{u} \in U_l$ and $\hat{u} \in U_l^{\perp}$ and the VMS equation:

$$\boxed{a(\bar{u}, \bar{v}) = b(\bar{v}) - a(\hat{u}, \bar{v}) \qquad \forall \bar{v} \in U_l} \qquad \text{(VMS)}$$

is $\bar{u} = P_{U_l} u$ with u the actual solution to (P)

$\Rightarrow$ the term $-a(\hat{u}, \bar{v})$ in the rhs ensures that the approximation $\bar{u}$ is the U-projection of the actual solution onto U_l.

VMS challenge

Derive an (explicit/analytical) model $-a(\hat{u}(\bar{u}), \bar{v})$ for $-a(\hat{u}, \bar{v})$
('fine-scale Green's function')

The VMG paradigm

Smoother

Reconsider the splitting $U_{l+1} = U_l \oplus U_l^{\perp}$ (or $U = U_l \oplus U_l^{\perp}$). Let $\hat{u} \in U_l^{\perp}$ denote the fine scale component $\hat{u} = (\mathrm{Id} - P_{U_l})u$ of the actual solution. A smoother $S : U_{l+1} \to U_{l+1}$ is characterized by the property:

$$\left\| \left(\mathrm{Id} - P_{U_l}\right) S\check{u} - \hat{u} \right\|_U \leq C \left\| \left(\mathrm{Id} - P_{U_l}\right) \check{u} - \hat{u} \right\|_U \qquad \forall \check{u} \in U_{l+1}$$

with C a small constant.

$\Rightarrow$ The smoother provides an improved (very good ?) approximation of the fine-scale component!

The VMG paradigm

Variational Multi-Grid

Given an initial approximation $\check{u} \in U_{l+1}$,

1. Perform a smoothing step: $\check{u} \mapsto S\check{u}$
2. Insert $(\mathrm{Id} - P_{U_l})S\check{u}$ as approximation of the fine-scale component into the VMS equation for the coarse-scale component:

$$a(\bar{u}, \bar{v}) = b(\bar{v}) - a((\mathrm{Id} - P_{U_l})S\check{u}, \bar{v}) \qquad \forall \bar{v} \in U_l \qquad \text{(CG)}$$

3. Update $\check{u}$ according to $S\check{u} + (\bar{u} - P_{U_l}S\check{u})$
(replace coarse-scale component of $S\check{u}$ by $\bar{u}$)

The VMG paradigm

Remarks

1. Rearranging (CG) yields:

$$a(\bar{u} - P_{U_l} S\check{u}, \bar{v}) = b(\bar{v}) - a(S\check{u}, \bar{v}) =: \langle r(S\check{u}), \bar{v} \rangle \qquad \forall \bar{v} \in U_l \qquad \text{(CS)}$$

   with $r(\cdot) : U \to U'$ the *residual functional*
   $\Rightarrow$ (CG) in VMG is the variational form of the coarse-grid correction equation in the *correction scheme*

2. In the variational formulation, prolongation is intrinsically defined by injection and restriction coincides with U-projection. Restriction of equations in (CG) occurs via restriction of test functions.

3. Closeness of $(\mathrm{Id} - P_{U_l})S\check{u}$ to $\hat{u}$ for any $\check{u}$ relies on separation of scales. Generally, $P_{U_l}\check{u}$ must be close enough to the coarse-scale component of the actual solution and iteration is required.

VMS vs VMG (linear case)

Comparison

The structure of VMS and VMG is very similar. Main differences:

1. In VMS the effect of the fine scale is approximated explicitly/analytically. In VMG the smoother (=cheap numerical process) constructs an approximation of the fine scale.
2. VMS focuses only on the coarse-scale component, $\bar{u}$. VMG is concerned with the composition of the fine- and coarse-scale components, $\bar{u} + \hat{u}$.

VMG for nonlinear problems

The splitting $a(\bar{u}+\hat{u},\bar{v}) = a(\bar{u},\bar{v}) + a(\hat{u},\bar{v})$ is inadmissible for semi-linear functionals $a(\cdot;\cdot) : U \times U \to \mathbb{R}$.

VMS equations for NL problems

Consider the decomposition $U = U_l \oplus U_l^{\perp}$ and the nonlinear problem:

$$\boxed{\text{Find } u \in U : \qquad a(u;v) = b(v) \qquad \forall v \in U} \qquad \text{(P)}$$

Since a is linear in its second argument, (P) can be decomposed into

$$a(\bar{u}+\hat{u};\bar{v}) = b(\bar{v}) \qquad \forall \bar{v} \in U_l$$
$$a(\bar{u}+\hat{u};\hat{v}) = b(\hat{v}) \qquad \forall \hat{v} \in U_l^{\perp}$$

VMG for nonlinear problems

NL-VMS coarse-scale equation

The test-space decomposition yields the coarse-scale equation (assuming $\hat{u}$ is available):

$$\boxed{\text{Find } \bar{u} \in U_l : \qquad a(\bar{u} + \hat{u}; \bar{v}) = b(\bar{v}) \qquad \forall \bar{v} \in U_l} \qquad \text{(VMS}^*\text{)}$$

The functional $(\bar{u}, \bar{v}) \mapsto a(\bar{u} + \hat{u}; \bar{v})$ is too complicated to treat directly. So, instead, given an approximation $\check{\bar{u}}$ of the coarse-scale component, we consider the defect-correction approximation:

$$a(\bar{u}; \bar{v}) = a(\check{\bar{u}}; \bar{v}) + b(\bar{v}) - a(\check{\bar{u}} + \hat{v}; \bar{v}) \qquad \forall \bar{v} \in U_l \qquad \text{(VMS)}$$

with $(\bar{u}, \bar{v}) \mapsto a(\bar{u}, \bar{v})$ an approximation to $(\bar{u}, \bar{v}) \mapsto a(\bar{u} + \hat{u}, \bar{v})$.

VMG for nonlinear problems

NL-VMG

Given an initial approximation $\check{u} \in U_{l+1}$ with a 'not too bad' approximation of the coarse scales (e.g. injection of U_l solution):

1. Perform a smoothing step $\check{u} \mapsto S\check{u}$, where the smoother $S : U_{l+1} \to U_{l+1}$ improves the approximation of the fine scales
2. Solve the coarse-grid problem corresponding to $S\check{u}$:

$$\text{Find } \bar{u} \in U_l : \quad a(\bar{u};\bar{v}) = a(P_{U_l} S\check{u};\bar{v}) + b(\bar{v}) - a(S\check{u};\bar{v}) \quad \forall \bar{v} \in U_l \qquad \text{(FAS)}$$

3. Update $\check{u}$ according to $S\check{u} + (\bar{u} - P_{U_l} S\check{u})$

VMG for nonlinear problems

NL-VMG

Given an initial approximation $\check{u} \in U_{l+1}$ with a 'not too bad' approximation of the coarse scales (e.g. injection of U_l solution):

1. Perform a smoothing step $\check{u} \mapsto S\check{u}$, where the smoother $S : U_{l+1} \to U_{l+1}$ improves the approximation of the fine scales
2. Solve the coarse-grid problem corresponding to $S\check{u}$:

$$\text{Find } \bar{u} \in U_l : \quad \tilde{a}(\bar{u};\bar{v}) = \tilde{a}(P_{U_l} S\check{u};\bar{v}) + b(\bar{v}) - a(S\check{u};\bar{v}) \quad \forall \bar{v} \in U_l \qquad \text{(FAS)}$$

3. Update $\check{u}$ according to $S\check{u} + (\bar{u} - P_{U_l} S\check{u})$

The functional $(\bar{u},\bar{v}) \mapsto a(\bar{u},\bar{v})$ in (FAS) can be replaced by an approximation $(\bar{u},\bar{v}) \mapsto \tilde{a}(\bar{u},\bar{v})$ (e.g. linearization)

VMG for nonlinear problems

Remarks

1. Equation (FAS) is the variational form of the coarse-grid equation in the Full-Approximation Scheme
2. Does (FAS) have a counter part in variational multiscale methods?
3. Iteration required (inexact fine scale and inexact functional in coarse-grid equation!)
4. Because $b(v) - a(u, v) = 0$ for all $v \in U_{l+1} \supset U_l$, if $S\check{u} = u$ then $\bar{u} = P_{U_l} u \Rightarrow$ coarse-scale solution converges to U-projection of solution in U_{l+1}.

Conclusion

- The structure of Variational Multi-Scale methods and Variational Multi-Grid methods is very similar. The main differences are:
  1. In VMS the effect of the fine scale is approximated explicitly/analytically. In VMG an inexpensive computational approximation (relaxation/smoother) is used to construct an approximation of the fine scale.
  2. VMS considers only the coarse-scale component. VMG considers the full approximation (composition of the fine- and coarse-scale components)
- The methods have several dissimilarities $\Rightarrow$ there are many opportunities for one methodology to borrow concepts/insights from the other (iteration, coarse-grid equations, etc.)

Further reading

- D. Braess and W. Hackbusch, *A new convergence proof for the multigrid method including the v-cycle*, SIAM Journal on Numerical Analysis **20** (1983), 967–975.
- T.J.R. Hughes, *Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods*, Comput. Methods Appl. Mech. Engrg. **127** (1995), 387–401.
- T.J.R. Hughes, G.R. Feijóo, L. Mazzei, and J.-B. Quincy, *The variational multiscale method – a paradigm for computational mechanics*, Comput. Methods Appl. Mech. Engrg. **166** (1998), 3–24.
- E.H. van Brummelen, K.G. van der Zee, and R. de Borst, *Space/time multigrid for a fluid-structure-interaction problem*, Applied Numerical Mathematics **58** (2008), 1951–1971.

Outline

Solution methods for FSI: categorization

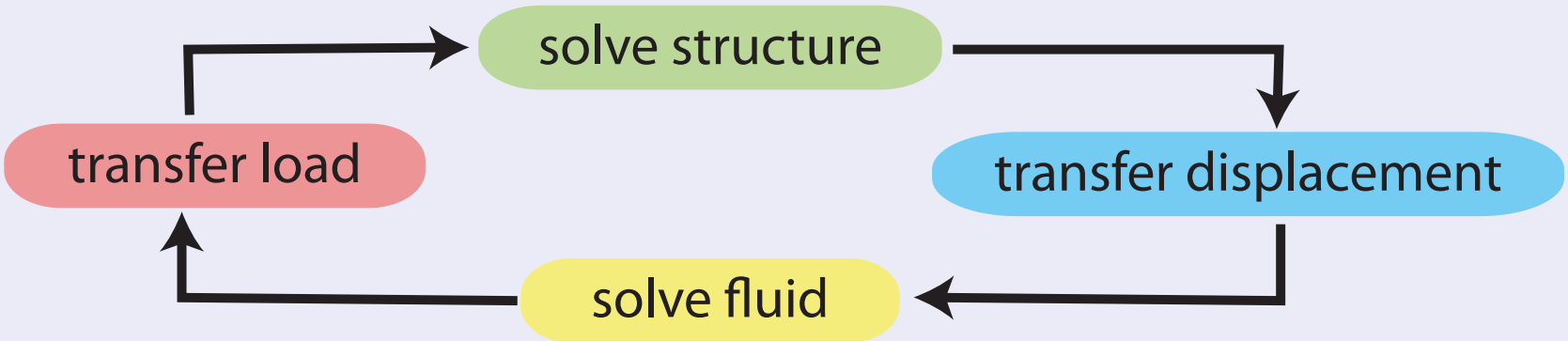
monolithic methods

Solve the fluid-structure (+pseudo-structure) system simultaneously

partitioned method (subiteration)

solve the fluid-structure system asynchronously by iteration:

1. solve fluid subject to structure displacement
2. solve structure subject to fluid load



Solution methods for FSI

Monolithic

- advantages: stable
- disadvantages: XL linear systems, inefficient, ill-conditioned matrices, non-modular, dense matrices, no standard preconditioners, ...

Partitioned

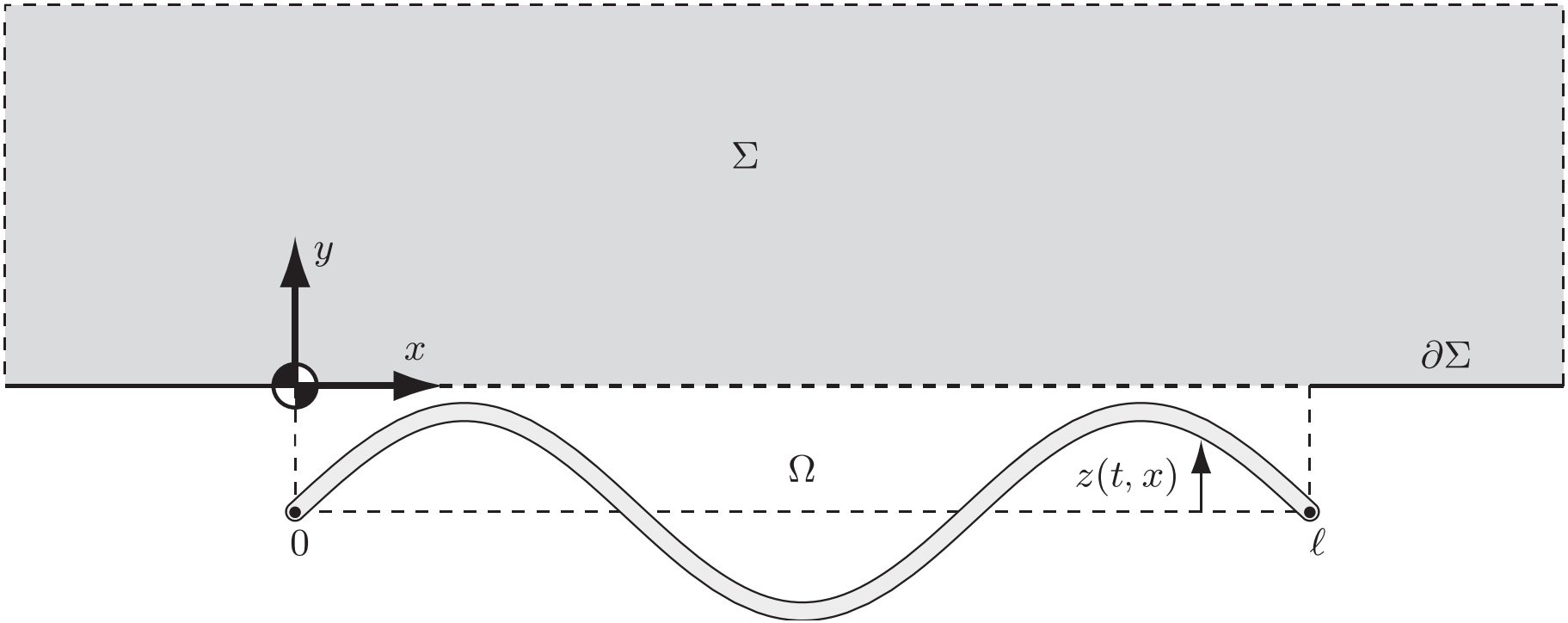
- advantages: standard linear systems corresponding to fluid and structure, modular
- disadvantages: potentially unstable, slow convergence

MG for FSI: subiteration as smoother

Objective

Show for a model problem that subiteration is an excellent smoother
$\Rightarrow$ enormous potential for multigrid!

Model problem



Model problem

simplifications

- (geometric) linearization
- incompressible flow with slip boundary condition on bottom boundary

procedure

- formal asymptotic expansion
- expand structural displacement in eigen functions (sinusoides)
- derive Fourier symbol of relation between displacement and pressure at the boundary of the fluid domain (Poincaré-Steklov operator)
- derive error-amplification for each eigenmode

Spectral decomposition

Structure model

$$\begin{aligned} \mu z'' + \lambda^2 D^4 z &= -p \quad \text{on } (0,\ell) =: \Omega \\ z = D^2 z &= 0 \quad \text{at } \{0,\ell\} \\ z'(0) = z(0) &= 0 \end{aligned}$$

Subiteration model

$$\begin{aligned} \mu z_n'' + \lambda^2 D^4 z_n &= -p(z_{n-1}) \quad \text{on } (0,\ell) \quad \text{(S)} \\ z_n = D^2 z_n &= 0 \quad \text{at } \{0,\ell\} \quad \text{(BC)} \\ z_n'(0) = z_n(0) &= 0 \quad \text{(IC)} \end{aligned}$$

Spectral decomposition

Structural eigenmodes

The eigenvalues and eigenmodes of the structure are

$$\sigma_k = \lambda^2 (k\pi/\ell)^4 \qquad \psi_k(x) = \sqrt{2/\ell}\sin(k\pi x/\ell) \qquad (k \in \mathbb{N})$$

Proposition

The eigenmodes are orthonormal in $L^2(\Omega)$ and form a countable orthogonal basis of the displacement space $H^2(\Omega) \cap H_0^1(\Omega)$

Corollary

There exist functions $\bar{z}_{n,k} : (0,T) \to \mathbb{R}$ such that $z_n(x,t) = \sum_k \bar{z}(t)\psi_k(x)$

Reduction

ODE for coefficients

L^2 projection of (S) onto ψ_k yields:

$$\mu \bar{z}''_{n,k}(t) + \sigma_k \bar{z}_{n,k}(t) = -\left(\psi_k, p\left(\sum_l \bar{z}_{n-1,l}\psi_l\right)(\cdot,t)\right)_{L^2(\Omega)} \qquad \text{(ODE)}$$

Green's function

By means of the Green's function for $\mu(\cdot)'' + \sigma(\cdot)$ we obtain:

$$\bar{z}_{n,k}(t) = -\int_0^t g(t,s)\left(\psi_k, p\left(\sum_l \bar{z}_{n-1,l}\psi_l\right)(\cdot,s)\right)_{L^2(\Omega)} \mathrm{d}s$$

where

$$g(t,s) = -(\mu\sigma)^{-1/2}\,\sin\left((\sigma/\mu)^{1/2}(s-t)\right) \qquad \text{(G)}$$

Reduction

A tedious derivation . . .

Solution of the linearized incompressible flow equations gives

$$\left(\psi_k, p\left(\sum_l \bar{z}_{n-1,l}\psi_l\right)(\cdot,t)\right)_{L^2(\Omega)}$$
$$= \rho\Big((k\pi/\ell)^{-1}\bar{z}''_{n-1,k}(t) + 2\nu(k\pi/\ell)\bar{z}'_{n-1,k}(t) - U^2(k\pi/\ell)\bar{z}_{n-1,k}(t)\Big)$$

- ρ: fluid density
- ν: fluid viscosity
- U: mean flow velocity

Amplification / smoothing

Error amplification

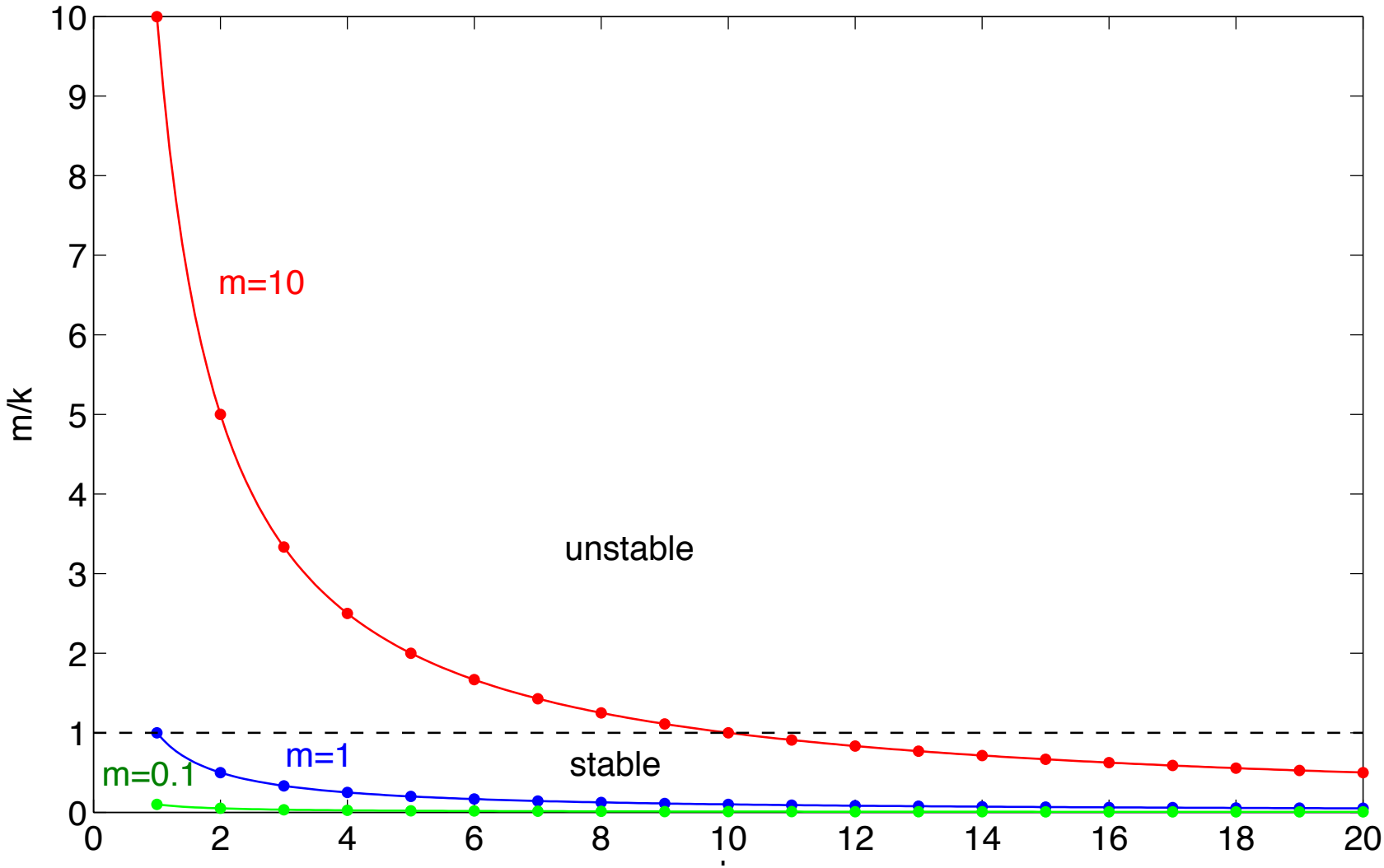
The subiteration process is characterized by the map $\bar{z}_{n-1,k} \mapsto \bar{z}_{n,k}$:

$$\begin{aligned}\bar{z}_{n,k}(t) &= -\frac{\rho\ell}{\mu\pi k}\bar{z}_{n-1,k}(t) - \int_0^t \beta_k(t,s)\,\bar{z}_{n-1,k}(s)\,\mathrm{d}s \\ &= -\frac{\rho\ell}{\mu\pi k}\bar{z}_{n-1,k}(t) + O(t\bar{z}_{n-1,k}) \qquad \text{as } t \to 0\end{aligned}$$

Conclusion

$1/k$ proportionality in the short-time-interval limit $t \to 0 \Rightarrow$ subiteration is a very effective smoother

Amplification / smoothing



Further reading

- E.H. van Brummelen, K.G. van der Zee, and R. de Borst, *Space/time multigrid for a fluid-structure-interaction problem*, Applied Numerical Mathematics **58** (2008), 1951–1971.
- E.H. van Brummelen, *Partitioned iterative solution methods for fluid-structure interaction*, Int. J. Numer. Meth. Fluids **65** (2011), 3–27.
- E.H. van Brummelen, *Added mass effects of compressible and incompressible flows in fluid-structure interaction*, Journal of Applied Mechanics **76** (2009), 021206–7.

Outline

Philosophies

Multigrid purism

For *any* problem, there exists a suitable relaxation scheme and coarse grid correction that allows you to solve that problem in only a few operations per grid point. If your scheme does not do that, then work harder!

Multigrid pragmatism

For difficult problems, the multigrid method eliminates most error components, but not all. The few remaining ones (possibly unstable!), can be effectively handled by a Krylov-subspace method $\Rightarrow$ multigrid as preconditioner (or Krylov acceleration of multigrid)

Krylov-subspace methods

Krylov subspace

Consider the linear problem $Au = b$ ($A \in \mathbb{R}^{N\times N}$, $b \in \mathbb{R}^N$) and an approximation $\check{u}$. Define $r = b - A\check{u}$. The Krylov subspace, $\mathcal{K}_m$, is built by recursion

$$\mathcal{K}_m = \text{span}\{r, Ar, A^2r, \ldots, A^{m-1}r\}$$

Interpretation

Let $e = \check{u} - \bar{u}$ denote the error. Then $Ae = r$ and

$$\mathcal{K}_m = \text{span}\{Ae, A^2e, A^3e, \ldots, A^me\}$$

$\Rightarrow$ Krylov space essentially contains error components corresponding to largest eigenvalues (=good approximation space!)

Krylov-subspace methods

Example: GMRES

In the GMRES method*, we construct a new approximation $\tilde{u} \in \check{u} + \mathcal{K}_m$ such that

$$\tilde{u} = \underset{u \in \check{u} + \mathcal{K}_m}{\arg\inf} \|Au - b\| \qquad \text{(LSQ)}$$

Remarks

- (LSQ) corresponds to least-squares problem of dim m.
- many implementational details

*Y. Saad and M.H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput. **7** (1986), 856–869.

Convergence of GMRES

$$
\begin{aligned}
\|r_m\| :=& \|A\tilde{u} - b\| && \text{(Def)} \\
=& \inf_{u \in \check{u} + \mathcal{K}_m} \|Au - b\| && \text{(from (LSQ))} \\
=& \inf_{y \in \mathcal{K}_m} \|r_0 + Ay\| && (b - A\check{u} = r_0) \\
=& \inf_{a_1,\ldots,a_m} \|r_0 + a_1 A r_0 + \cdots + a_m A^m r_0\| && \text{(Def } \mathcal{K}_m) \\
=& \inf_{\psi \in \mathcal{P}_1^m} \|\psi(A) r_0\| \\
\leq& \inf_{\psi \in \mathcal{P}_1^m} \|\psi(A)\| \|r_0\| && \text{(Def)}
\end{aligned}
$$

where $\mathcal{P}_1^m$ denotes the the space of polynomials of degree m that evaluate to 1 at zero.

Convergence of GMRES (cont'd)

Spectral mapping Thm.

Consider a bounded linear operator $A : X \to X$ with spectrum $\sigma(A) \subset \mathbb{C}$. For any analytic function f it holds that

$$\sigma(f(A)) = f(\sigma(A))$$

(the spectrum of an analytic function of an operator is the function applied to the spectrum of the operator)

Convergence of GMRES (cont'd)

Spectral decomposition

Any (non-degenerate) matrix $A \in \mathbb{R}^{N \times N}$ can be decomposed as:

$$A = V \Sigma V^{-1}$$

with $\Sigma = \mathrm{diag}(\sigma_1(A), \ldots, \sigma_N(A))$ and V the matrix of eigenvectors

Convergence of GMRES (cont'd)

$$
\begin{aligned}
\frac{\|r_m\|}{\|r_0\|} &\leq \inf_{\psi \in \mathcal{P}_1^m} \|\psi(A)\| \\
&= \inf_{\psi \in \mathcal{P}_1^m} \|V \psi(\Sigma) V^{-1}\| && \text{(Sp. Map. Thm.)} \\
&\leq \|V\| \|V^{-1}\| \inf_{\psi \in \mathcal{P}_1^m} \|\psi(\Sigma)\| \\
&= \varkappa(V) \inf_{\psi \in \mathcal{P}_1^m} \|\psi(\Sigma)\| \\
&\leq \varkappa(V) \inf_{\psi \in \mathcal{P}_1^m} \sup_{z \in \sigma(A)} |\psi(z)|
\end{aligned}
$$

where $\varkappa(V)$ denotes the *condition number* of V.

Convergence of GMRES (cont'd)

Corollary

The residual reduction in the GMRES method is bounded by:

$$\frac{\|r_m\|}{\|r_0\|} \leq \varkappa \inf_{\psi \in \mathcal{P}_1^m} \sup_{z \in \sigma(A)} |\psi(z)|$$

$\Rightarrow$ Very good convergence if A has only 'a few' large eigen values

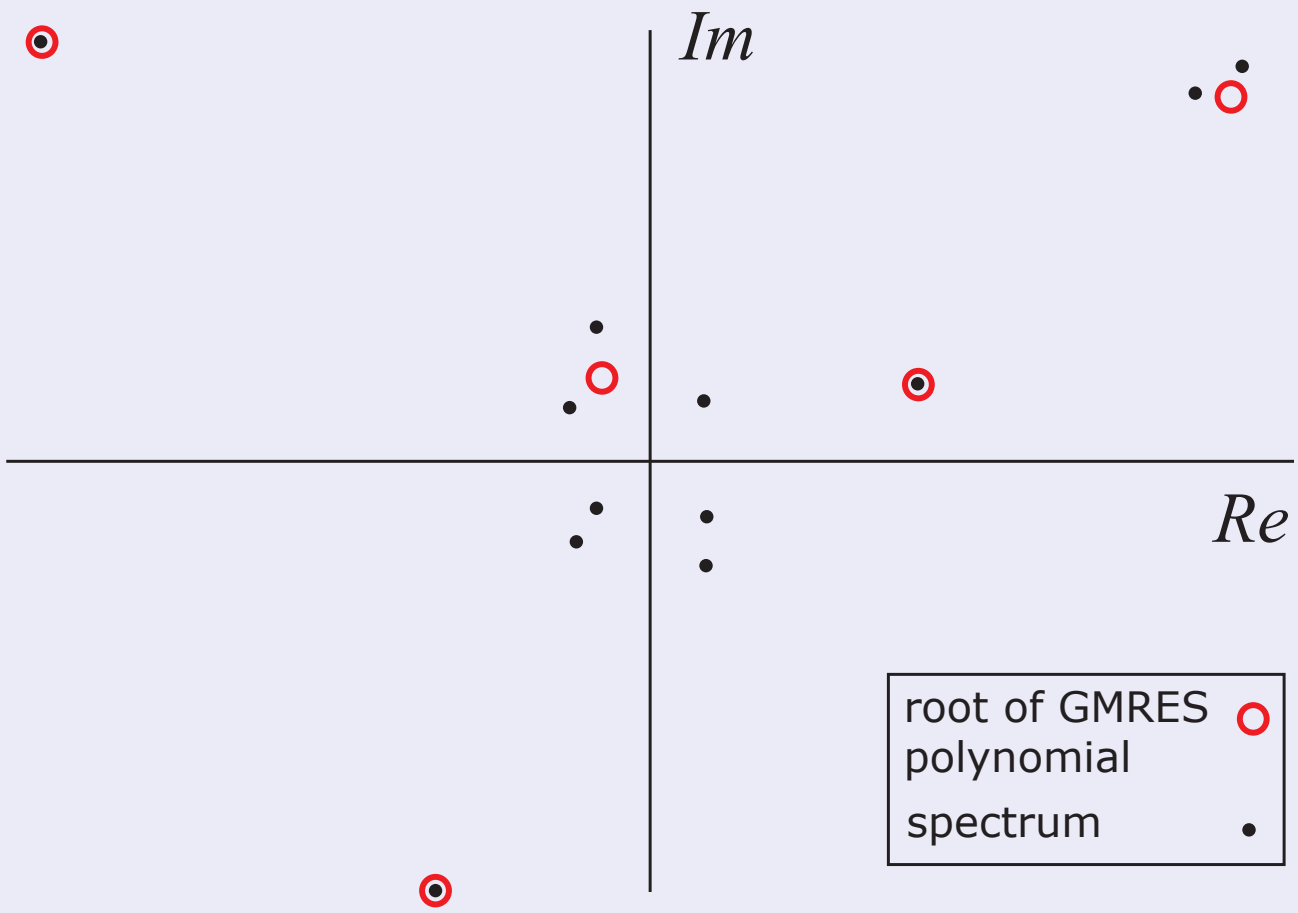
$\Rightarrow$ Very good convergence if eigenvalues are clustered

$\Rightarrow$ Bad if eigenvalues are dispersed in $\mathbb{C}$.

$\Rightarrow$ Monotone convergence

Convergence of GMRES (cont'd)

Illustration



Multigrid as preconditioner

One iteration of a multigrid method, $\check{u}_i \mapsto \check{u}_{i+1}$ can be regarded as an application of an approximate inverse, $\tilde{A}^{-1}$. From the sequence of approximations, we can construct a Krylov space:

$$\mathcal{K}_m(A\tilde{A}^{-1}, r_0) = \text{span}\{r_0, A\tilde{A}^{-1}r_0, \ldots, (A\tilde{A}^{-1})^{m-1}r_0\}$$

Convergence of Multigrid-preconditioned GMRES

$$\|r_m\| = \inf_{\psi \in \mathcal{P}_1^m} \left\|\psi\left(A\tilde{A}^{-1}\right)r_0\right\|$$

$\Rightarrow$ How are eigenvalues of $A\tilde{A}^{-1}$ distributed?

Residual-convergence of Multigrid separately

For MG separately:

$$\|r_m\| = \|(I - A\tilde{A}^{-1})^m r_0\|$$

Multigrid as preconditioner

Residual-convergence of Multigrid separately (proof)

The multigrid method can be written in defect-correction form as:

$$\tilde{A}u_n = \tilde{A}u_{n-1} + (b - Au_{n-1})$$

Defining error as $e_n = u_n - \bar{u}$, it holds that:

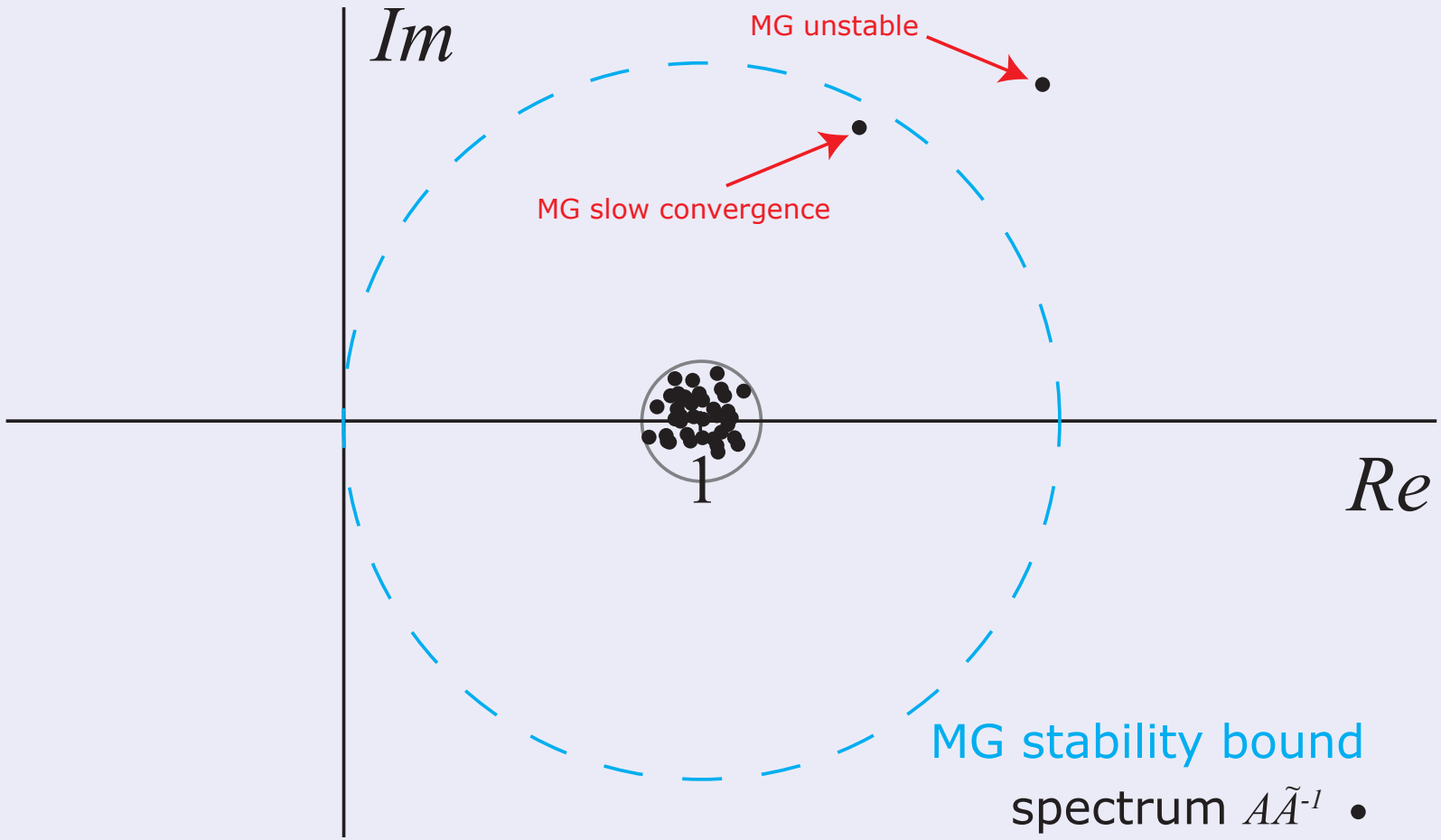
$$e_n = (I - \tilde{A}^{-1}A)e_{n-1} = (I - \tilde{A}^{-1}A)^n e_0$$

The residual and error are related by $r_n = Ae_n$:

$$\begin{aligned} r_m = Ae_m &= A(I - \tilde{A}^{-1}A)^m A^{-1} r_0 = A(I - \tilde{A}^{-1}A)^{m-1}(I - \tilde{A}^{-1}A)A^{-1} r_0 \\ &= A(I - \tilde{A}^{-1}A)^{m-1} A^{-1}(I - A\tilde{A}^{-1}) r_0 = \cdots = (I - A\tilde{A}^{-1})^m r_0 \end{aligned}$$

Multigrid as preconditioner

Illustration



Conclusion

Message

An imperfect multigrid method can be a perfect preconditioner for Krylov methods!

Further reading

- T. Washio and C.W. Oosterlee, *Krylov subspace acceleration for nonlinear multigrid schemes*, Electronic Transactions on Numerical Analysis **6** (1997), 271–290.
- C.W. Oosterlee and T. Washio, *An evaluation of parallel multigrid as a solver and a preconditioner for singularly perturbed problems*, SIAM J. on Sci. Comput. **19** (1998), 87–110.
- H.C. Elman, O.G. Ernst, and D.P. O'Leary, *A multigrid method enhanced by Krylov subspace iteration for discrete Helmholtz equations*, SIAM J. Sci. Comput. **23** (2002), 1291–1315.

Outline

Integral transforms and multi-summations

Integral transforms and multi-summations

$$u(x) = \int_{\Omega} K(|x-y|)\, v(y)\, \mathrm{d}\mu(y)$$

$$u_i = \sum_{j=1}^{N} K_{ij} v_j \qquad \text{for } i \in \{1, \ldots, N\}$$

Applications

- gravitational forces, coulomb forces, ...
- solid mechanics: interaction of dislocations, Herz' law of elastic deformation
- molecular dynamics: vd Waals forces (gen. potentials)
- ...

Integral transforms and multi-summations

Asymptotic smoothness

In many applications, the kernel K is asymptotically smooth (or *singularly smooth*):

$$D^n K(r) \to 0 \qquad \text{as } |r| \to \infty$$

Remarks:

- often monotonous decay
- generally $D^n K(r)/D^m K(r) \to 0$ as $|r| \to \infty$ for $n > m$
- generally $K(r)$ is singular as $r \to 0$.

Integral transforms and multi-summations

Examples

- Newton gravitation:

$$K_{ij} = -Gm^2 \frac{x_i - x_j}{|x_i - x_j|^3} \qquad (1/r^2)$$

- Coulomb forces:

$$K_{ij} = k_e q^2 \frac{x_i - x_j}{|x_i - x_j|^3} \qquad (1/r^2)$$

Integral transforms and multi-summations

Complexity

For each index $i \in \{1, \ldots, N\}$, all other indices $j \in \{1, \ldots, N\}$ have to be visited $\Rightarrow$ direct evaluation of a multi-summation amounts to $O(N^2)$ operations.

Integral transforms and multi-summations

Complexity

For each index $i \in \{1, \ldots, N\}$, all other indices $j \in \{1, \ldots, N\}$ have to be visited $\Rightarrow$ direct evaluation of a multi-summation amounts to $O(N^2)$ operations.

Multi-level Multi-summation

The main concept of multi-level multi-summation is to reduce the computational cost by exploiting the smoothness properties of the kernel: smooth functions can be accurately approximated by interpolation from a coarse grid.

Multi-level multi-summation

Consider

- an interval $\Omega := (0, \ell)$
- a uniform partition $\{0, h, \ldots, Nh (:= \ell)\}$
- a coarse partition $\{0, H, \ldots, (N/2)H\}$ with $H = 2h$
- a fine-grid multi-summation $v^h_{(\cdot)} \mapsto u^h_{(\cdot)}$

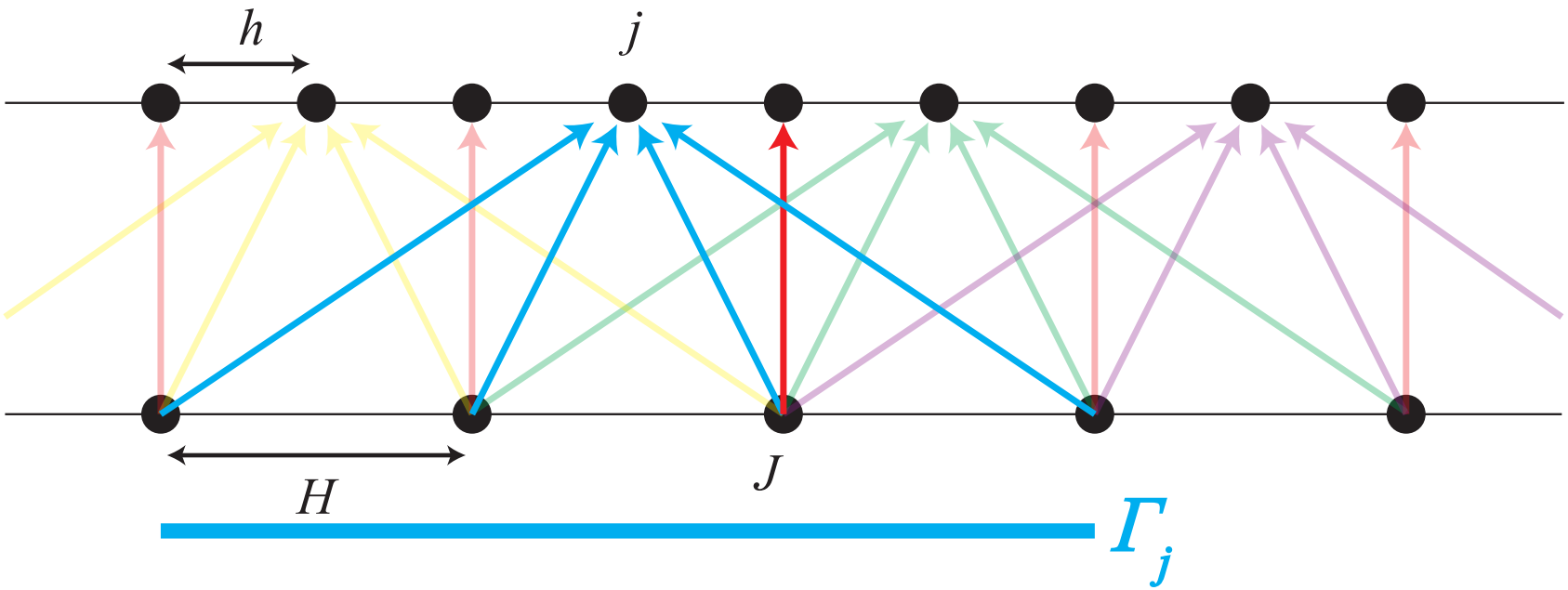
$$u^h_i = \sum_j K^{hh}_{ij} v^h_j$$

Multi-level multi-summation

Assume that K_{ij}^{hh} is 'sufficiently smooth' (for the moment, globally!), so that at an 'acceptable error' for each fixed i we can replace K_{ij}^{hh} by an interpolation of its values on the H-grid:

$$K_{ij}^{hh} \approx \sum_{J \in \Gamma_j} w_{jJ} K_{iJ}^{hH}$$

Γ_j: a neighborhood of j of H-grid points; w_{jJ}: interpolation weights.

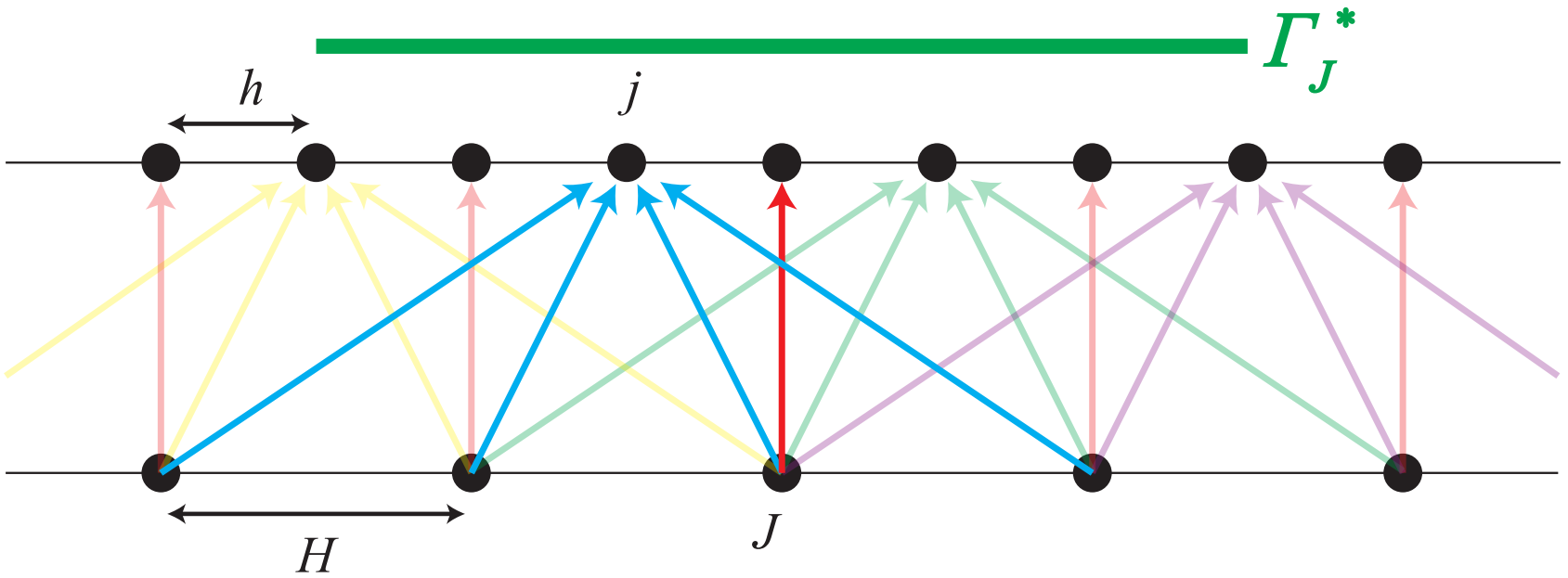


Multi-level multi-summation

Conversely, by reversing the order of summation:

$$\begin{aligned}\check{u}_i^h &= \sum_j \sum_{J \in \Gamma_j} w_{jJ} K_{iJ}^{hH} v_j^h \\ &= \sum_J \sum_{j \in \Gamma_J^*} w_{jJ} K_{iJ}^{hH} v_j^h\end{aligned}$$

where Γ_J^* is a dual neighborhood of h-grid point of point J.



Multi-level multi-summation

Conversely, by reversing the order of summation:

$$
\begin{aligned}
\check{u}_i^h &= \sum_j \sum_{J \in \Gamma_j} w_{jJ} K_{iJ}^{hH} v_j^h \\
&= \sum_J \sum_{j \in \Gamma_J^*} w_{jJ} K_{iJ}^{hH} v_j^h \\
&= \sum_J K_{iJ}^{hH} \sum_{j \in \Gamma_J^*} w_{jJ} v_j^h \\
&= \sum_J K_{iJ}^{hH} v_J^H \qquad \text{(coarse-grid summation)}
\end{aligned}
$$

Multi-level multi-summation

By the smoothness of K_{iJ}^{hH} with respect to the i index, we can replace K_{iJ}^{hH} by its values K_{IJ}^{HH} on the H-grid:

$$K_{iJ}^{hH} \approx \sum_{I \in \Gamma_i} w_{iI} K_{IJ}^{HH} \qquad (*)$$

$\Rightarrow$

$$\begin{aligned}
\check{u}_i^h &= \sum_J K_{iJ}^{hH} v_J^H \\
&= \sum_J \sum_{I \in \Gamma_i} w_{iI} K_{IJ}^{HH} v_J^H \qquad \text{(by } *\text{)} \\
&= \sum_{I \in \Gamma_i} w_{iI} \sum_J K_{IJ}^{HH} v_J^H \\
&= \sum_{I \in \Gamma_i} w_{iI} u_I^H
\end{aligned}$$

Multi-level multi-summation

Algorithm: At the expense of an interpolation error, we can replace the fine-grid multi-summation by

1. restriction of v_j^h to the coarse-grid:

$$v_J^H = \sum_{j \in \Gamma_J^*} w_{jJ} v_j^h$$

2. coarse-grid multi-summation:

$$u_I^H = \sum_J K_{IJ}^{HH} v_J^H$$

3. interpolation:

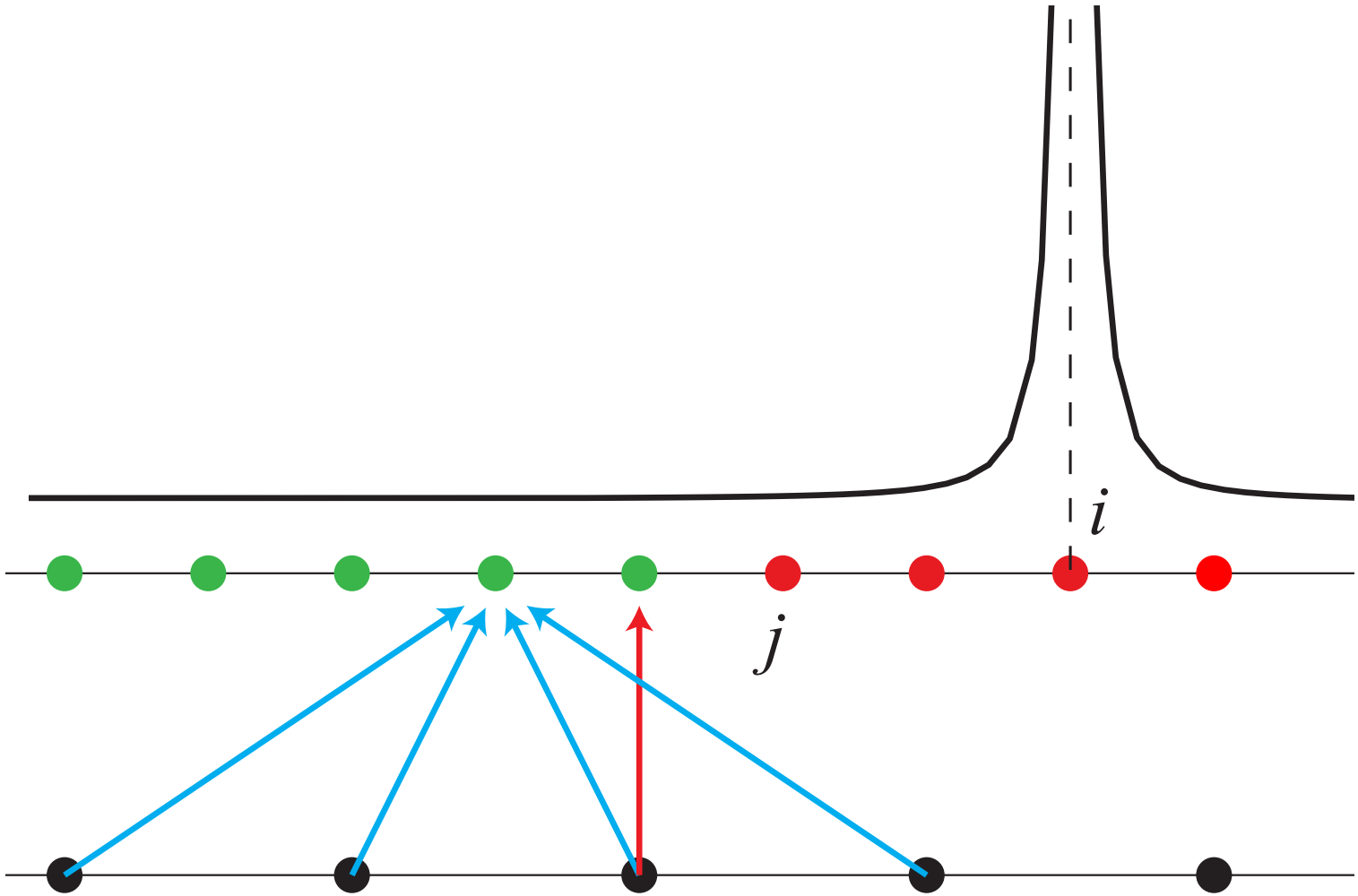
$$u_i^h = \sum_{I \in \Gamma_i} w_{iI} u_I^H$$

Multi-level multi-summation: singularly smooth kernels

For singularly-smooth kernels, the basic algorithm will not work on account of the excessive interpolation error near the singularity. Instead, we perform a local correction. Note that

$$
\begin{aligned}
u_i^h &= \check{u}_i^h + (u_i^h - \check{u}_i^h) \\
&= \check{u}_i^h + \sum_j \left(K_{ij}^{hh} - \sum_{I \in \Gamma_i} \sum_{J \in \Gamma_j} w_{iI} w_{jJ} K_{IJ}^{HH} \right) v_j^h
\end{aligned}
$$

Multi-level multi-summation: singularly smooth kernels



Multi-level multi-summation: singularly smooth kernels

For singularly-smooth kernels, the basic algorithm will not work on account of the excessive interpolation error near the singularity. Instead, we perform a local correction. Note that

$$
\begin{aligned}
u_i^h &= \check{u}_i^h + (u_i^h - \check{u}_i^h) \\
&= \check{u}_i^h + \sum_j \left(K_{ij}^{hh} - \sum_{I \in \Gamma_i} \sum_{J \in \Gamma_j} w_{iI} w_{jJ} K_{IJ}^{HH} \right) v_j^h \\
&= \check{u}_i^h + \sum_{j \in \Gamma_i^\epsilon} \left(K_{ij}^{hh} - \sum_{I \in \Gamma_i} \sum_{J \in \Gamma_j} w_{iI} w_{jJ} K_{IJ}^{HH} \right) v_j^h
\end{aligned}
$$

where Γ_i^ϵ is a neighborhood of i, depending on the admissible error ϵ.

Multi-level multi-summation: singularly smooth kernels

Algorithm: At the expense of a controlable interpolation error, we can replace the fine-grid multi-summation by

1. restriction of v_j^h to the coarse-grid:

$$v_J^H = \sum_{j \in \Gamma_J^*} w_{jJ} v_j^h$$

2. coarse-grid multi-summation:

$$u_I^H = \sum_J K_{IJ}^{HH} v_J^H$$

3. interpolation:

$$\tilde{u}_i^h = \sum_{I \in \Gamma_i} w_{iI} u_I^H$$

4. local correction:

$$u_i^h = \tilde{u}_i^h + \sum_{j \in \Gamma_i^\epsilon} \left(K_{ij}^{hh} - \sum_{I \in \Gamma_i} \sum_{J \in \Gamma_j} w_{iI} w_{jJ} K_{IJ}^{HH} \right) v_j^h$$

Multi-level multi-summation: singularly smooth kernels

Remarks

1. recursion: the coarse-grid multi-summation can again be evaluated by the same algorithm
2. the actual multi-summation can be performed in $O(N)$ operations on a grid with $O(\sqrt{N})$ points; in practice, one uses a grid with $O(1)$ points
3. on all but the coarsest grid, *only local operations are performed* (restriction, interpolation, correction) $\Rightarrow$ for fixed ϵ, multi-level multi-summation requires $O(N)$ operations

Further reading

- A. Brandt, *Multigrid Techniques: 1984 Guide with Applications to Fluid Dynamics*, Tech. Report GMD **85**, 1984.
- C.H. Venner and A.A. Lubrecht, *Multilevel methods in lubrication*, Tribology Series, vol. 73, Elsevier, Amsterdam, 2000.
- A. Brandt and C.H. Venner, *Fast evaluation of integral transforms on adaptive grids*, Multigrid Methods V (Proc. Stuttgart, 1996) (Berlin) (W. Hackbusch and G. Wittum, eds.), Lecture Notes in Computational Science and Engineering, vol. 3, Springer Verlag, 1998, also appeared as Internal Report Carl F. Gauss Minerva center for scientific computation, WI/CG-5, pp. 20–44.
- A. Brandt and A. A. Lubrecht, *Multilevel matrix multiplication and the fast solution of integral equations*, J. Comput. Phys. **90** (1990), 348–370.
- A. Brandt and C.H. Venner, *Multilevel evaluation of integral transforms with asymptotically smooth kernels*, SIAM J. Sci. Stat. Comput. 19 (1998), 468–492