Summer school in ETH Zurich, Switzerland, 30-08-04 to 3-09-04

Multiscale geometric data representation
Complexity, Analysis and Applications

Organizers : Albert Cohen and Chris Schwab.

Topics and Lecturers : Curvelets (Emmanuel Candes), Bandlets (Stéphane Mallat and Erwan Le Pennec), Contourlets and Directional Filterbanks (Minh Do and Martin Vetterli), Sparse Grids and Hyperbolic Wavelets (Chris Schwab), Anisotropic Triangulations (Nira Dyn), Edgelets and Wedgeprints (Richard Baraniuk), Edge-Adapted Multiresolution (Paco Arandiga and Albert Cohen).

Infos : http://www.sam.math.ethz.ch/news/conferences/zss04

Wavelets in Numerical Analysis

Albert Cohen

Laboratoire Jacques-Louis Lions
Université Pierre et Marie Curie
Paris

Collaborators : Wolfgang Dahmen, Ron DeVore,
Sidi-Mahmoud Kaber, Marie Postel, Sigfried Mueller

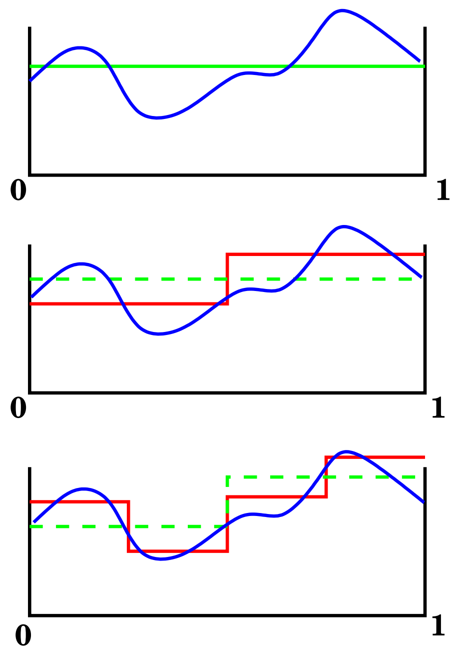
WAMA

Cargèse, July 2004

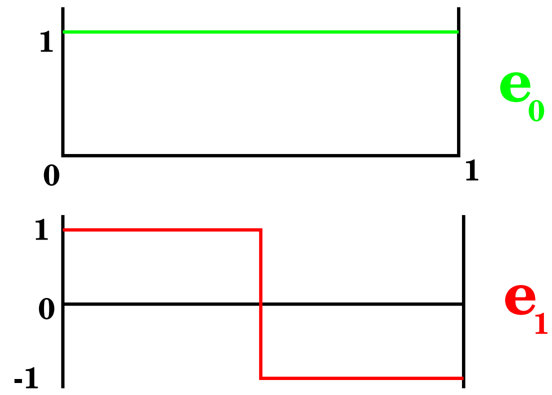
Agenda

1. Motivation: adaptive wavelet discretizations and PDE's
2. Adaptive space refinement for operator equations
3. Adaptive multiresolution processing for evolution equations

Basic example: the Haar system



$$f = \langle f, e_0 \rangle e_0$$



$$+ \langle f, e_1 \rangle e_1$$

$$+ \langle f, e_2 \rangle e_2 + \langle f, e_3 \rangle e_3$$

$$\cdots = \sum_\lambda f_\lambda \psi_\lambda \qquad f_\lambda := \langle f, \psi_\lambda \rangle$$

$\psi_\lambda := 2^{j/2}\psi(2^j \cdot -k)$, $\lambda = (j,k)$, $|\lambda| = j$.

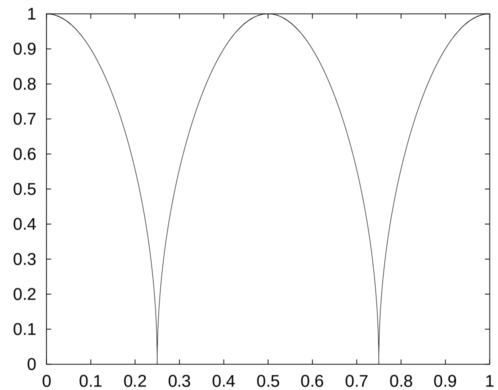
More general wavelets are constructed from similar multiscale approximation processes, using smoother functions such as splines or finite elements.

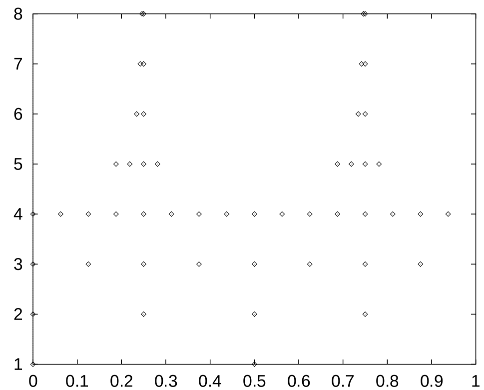
Approximating functions by wavelet bases

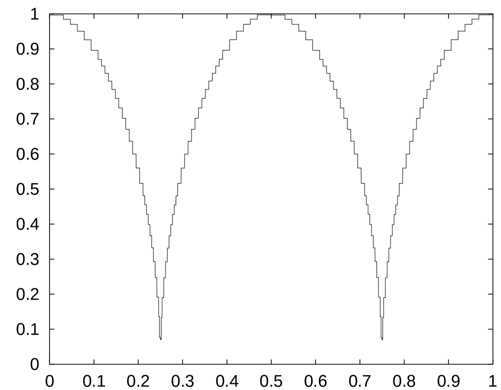
- Linear approximation at resolution level j by taking the truncated sum $f \mapsto P_j f := \sum_{|\lambda| \leq j} f_\lambda \psi_\lambda$.

- Nonlinear (adaptive) approximation obtained by thresholding

$$f \mapsto \mathcal{T}_\Lambda f := \sum_{\lambda \in \Lambda} f_\lambda \psi_\lambda, \quad \Lambda = \Lambda(\eta) = \{\lambda \ \text{s.t.} \ |f_\lambda| \geq \eta\}.$$







Applications to image compression

digital picture | decomposition | 10^3 largest coefficients

Linear approximation results

- V_h: finite element space discretizing a domain $\Omega \subset \mathbb{R}^d$.
- $N := \dim(V_h) \sim \mathrm{vol}(\Omega) h^{-d}$
- $W^{s,p} := \{f \in L^p(\Omega) \text{ s.t. } D^\alpha f \in L^p(\Omega),\ |\alpha| \leq s\}$

Classical finite element approximation theory (Bramble-Hilbert, Ciarlet-Raviart, Strang-Fix): provides with the classical estimate

$$f \in W^{s+t,p} \Rightarrow \inf_{g \in V_h} \|f - g\|_{W^{s,p}} \leq C h^t \sim C N^{-t/d},$$

assuming that V_h has enough polynomial reproduction and is contained in W_p^s.

Nonlinear approximation results

N-terms approximations: $\Sigma_N := \{\sum_{\lambda\in\Lambda} d_\lambda\psi_\lambda \; ; \; \#(\Lambda) \leq N\}$.

Rate of decay governed by weaker smoothness conditions (DeVore): with $1/q = 1/p + t/d$

$$f \in W^{s+t,q} \Rightarrow \inf_{g\in\Sigma_N} \|f-g\|_{W^{s,p}} \leq CN^{-t/d},$$

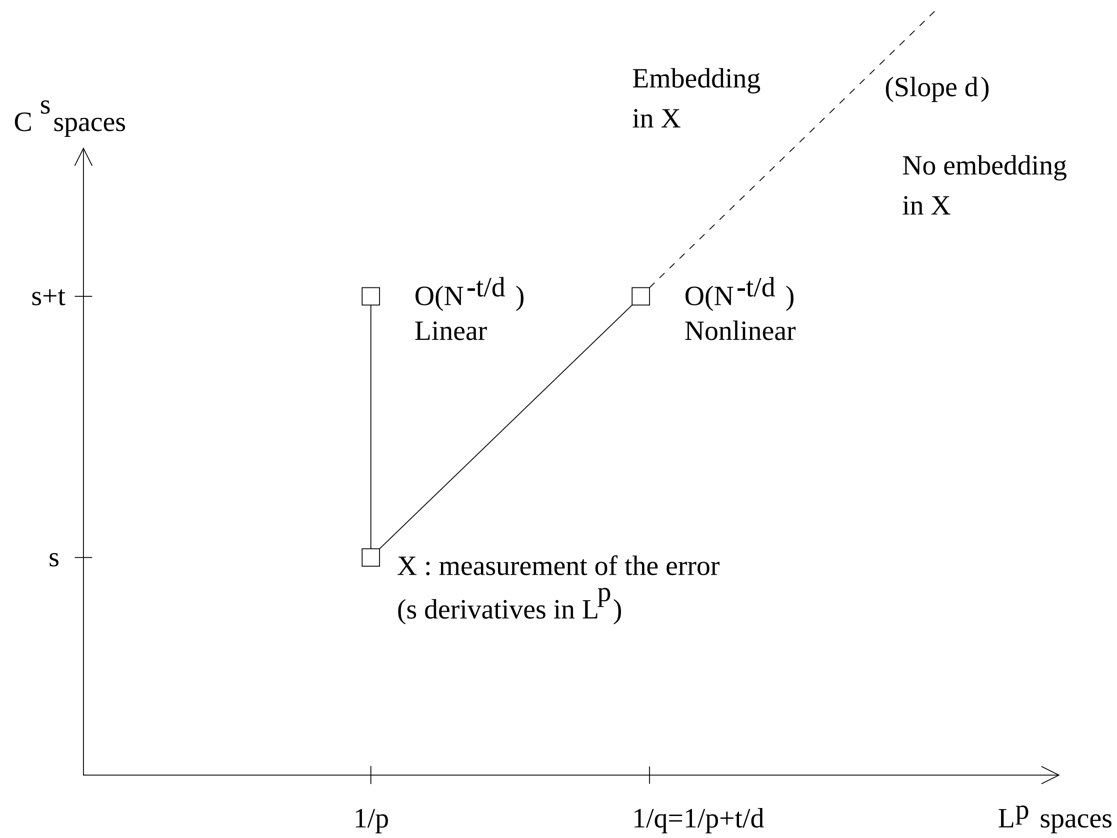
- For most error norm X (e.g. L^p, $W^{s,p}$, $B^s_{p,q}$), a near optimal approximation is obtained by thresholding : if $f = \sum_\lambda d_\lambda\psi_\lambda$, and $f_N := \sum_{N \text{ largest } \|d_\lambda\psi_\lambda\|_X} d_\lambda\psi_\lambda$, we then have

$$\|f - f_N\|_X \leq C \inf_{g\in\Sigma_N} \|f-g\|_X$$

with C independent of f and N.

- Remark: a similar theory for piecewise polynomial approximation on N adaptive triangles is still to be completed.

Pictorial interpretation of approximation results



General program for PDE's

- Theoretical: revisit regularity theory for PDE's. Solutions of certain PDE's might have substantially higher regularity in the scale governing nonlinear approximation than in the scale governing linear approximation. Examples : hyperbolic conservation laws (DeVore and Lucier 1987), elliptic problems on corner domains (Dahlke and DeVore, 1997).

- Numerical: develop for the unknown u of the PDE $\mathcal{F}(u) = 0$ appropriate adaptive resolution strategies which perform essentially as well as thresholding : produce $\tilde{u}_N$ with N terms such that $\|u - \tilde{u}_N\|$ has the same rate of decay N^{-s} as $\|u - u_N\|$ in some prescribed norm, if possible in $\mathcal{O}(N)$ computation.

Two approaches toward adaptive wavelet methods

- First approach (mostly applied to steady steate problems $\mathcal{F}(u) = 0$) : iterative space refinement techniques to access appropriate discretization sets $\{\psi_\lambda\}_{\lambda\in\Lambda_n}$, based on a variational formulation of the problem. (Bertoluzza, Perrier, Liandrat, Canuto, Dahlke, Hochmuth, Urban, Masson, Dahmen, DeVore, AC).

- Second approach (mostly applied to evolution problems $\partial_t u = \mathcal{E}(u)$) : multiresolution adaptive post-processing, i.e. start from a classical and reliable scheme on a uniform grid and use a discrete multiresolution decomposition in order to compress computational time and memory size, while preserving the accuracy of the initial scheme (Harten, Abgrall, Arandiga, Chiavassa, Donat, Dahmen, Mueller, Farge, Schneider, Kaber, Postel, AC)

General variational problems

$\mathcal{H}$ Hilbert space, $\mathcal{F} : \mathcal{H} \to \mathcal{H}'$ continuous mapping, u nonsingular solution of $\mathcal{F}(u) = 0$, i.e. $D\mathcal{F}(u)$ is an isomorphism from $\mathcal{H}$ to $\mathcal{H}'$.

Variational formulation : find $u \in \mathcal{H}$ such that

$$\langle \mathcal{F}(u), v \rangle = 0$$

for all $v \in \mathcal{H}$.

Simple linear examples: $\mathcal{F}(u) = \mathcal{A}u - f$

- Laplace: $\mathcal{A} := -\Delta$ and $\mathcal{H} := H_0^1$
- Stokes: $\mathcal{A}(u,p) := (-\Delta u + \nabla p, -\mathrm{Div}\, u)$ and $\mathcal{H} := (H_0^1)^3 \times L_0^2$.
- Single layer potential $\mathcal{A}u(x) := \int_\Gamma \frac{u(y)}{4\pi|x-y|} dy$ and $\mathcal{H} := H^{-1/2}$.

Standard (FEM) approach to discretisation

1. Well posed problem in infinite dimension $\mathcal{F}(u) = 0$.

2. Finite dimensional discretization $\mathcal{H} \to V_h$ by a Petrov-Galerkin type method ($\langle \mathcal{F}(u_h), v_h \rangle = 0$ for all $v_h \in W_h$).

Difficulties: not always well-posed (compatibility conditions, e.g. LBB for Stokes : $\inf_{p_h \in P_h} \sup_{u_h \in U_h} \frac{\int p_h \mathrm{Div}\, u_h}{\|p_h\|_{L^2} \|u_h\|_{H^1}} \geq \beta_h > 0$).

3. Iterative solver $u_h^0 \to u_h^1 \cdots \to u_h$.

Difficulties: ill-conditionning and dense matrices

4. Adaptivity: derive local error indicators by a-posteriori analysis of residual $\mathcal{F}(u_h)$, and apply local mesh refinement based on these indicators $V_h = V_r^0 \to V_r^1 \to \cdots$, $u_h = u_r^0 \to u_r^1 \to \cdots$

Difficulties: hanging nodes, convergence analysis of such refinement strategies (Dörfler 1996, Morin-Nocetto-Siebert 2000).

Wavelet adaptive discretizations: new paradigm

1. Well posed problem in infinite dimension $\mathcal{F}(u) = 0$.

2. Equivalent discrete problem in infinite dimension by wavelet-Galerkin: find $U = (u_\lambda)_{\lambda\in\nabla}$ such that

$$F(U) := (\langle \mathcal{F}(\sum u_\lambda \psi_\lambda), \psi_\mu \rangle)_{\mu\in\nabla} = 0.$$

Well-posed : $F : \ell^2 \to \ell^2$ if $(\psi_\lambda)_{\lambda\in\nabla}$ is a Riesz basis for $\mathcal{H}$ after renormalization, i.e. $\|u\|_{\mathcal{H}}^2 \sim \sum |u_\lambda|^2$ and $\|u\|_{\mathcal{H}'}^2 \sim \sum |\langle u, \psi_\lambda\rangle|^2$.

3. Converging iteration in infinite dimension $U^0 \to U^1 \to \cdots \to U$.

4. Adaptive approximation of this iteration up to prescribed tolerances in finite dimension: U^n supported by finite wavelet set $\Lambda_n \subset \nabla$.

$\Rightarrow$ allows to establish optimal accuracy and complexity results in the energy $\|u\|_{\mathcal{H}} \sim \|U\|$ norm.

The linear elliptic case

Assume $\mathcal{A}$ is an $\mathcal{H}$-elliptic operator. Equivalent problem :

$$AU = F$$

where A is ℓ^2-elliptic. For a suitable κ the iteration,

$$U^{n+1} = U^n + \kappa[F - AU^n],$$

converges with fixed error reduction rate $\rho < 1$.

Approximate iteration with prescribed tolerance $\varepsilon > 0$,

$$U^{n+1} = U^n + \kappa[\text{APPROX}(F, \varepsilon) - \text{APPROX}(AU^n, \varepsilon)],$$

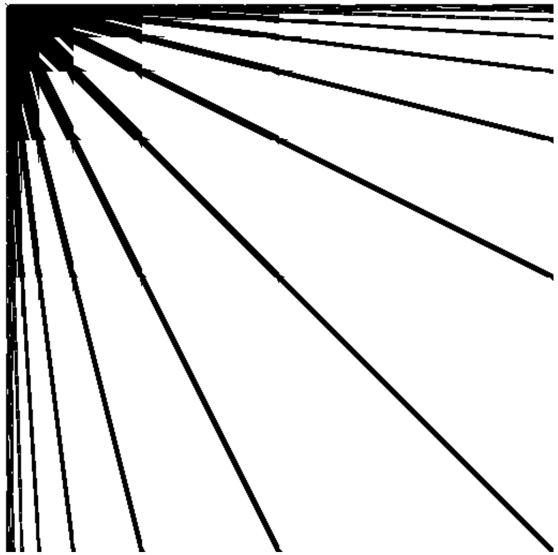
with $\|\text{APPROX}(AU^n, \varepsilon) - AU\| \leq \varepsilon$ and $\|\text{APPROX}(F, \varepsilon) - F\| \leq \varepsilon$. converges with reduction rate ρ until error is of order ε.

The procedure $\text{APPROX}(F, \varepsilon)$ amounts in thresolding F in ℓ^2, or equivalently the data f in the $\mathcal{H}'$ norm.

Matrix-vector approximation

The procedure APPROX(AU^n, ε) is made possible by matrix compression: one can build A_N with N coefficients per rows and colums such that $\|A - A_N\| \le CN^{-r}$

(V_1,V_1)	(W_1,V_1)	(W_2,V_1)	(W_3,V_1)
(V_1,W_1)	(W_1,W_1)	(W_2,W_1)	(W_3,W_1)
(V_1,W_2)	(W_1,W_2)	(W_2,W_2)	(W_3,W_2)
(V_1,W_1)	(W_1,W_3)	(W_2,W_3)	(W_3,W_3)



Analysis : based on the Schur lemma, using esimates of the type

$$|\langle \mathcal{A}\psi_\lambda, \psi_\mu \rangle| \le C[1 + \mathrm{dist}(\lambda,\mu)]^{-\beta} 2^{-\gamma||\lambda|-|\mu||},$$

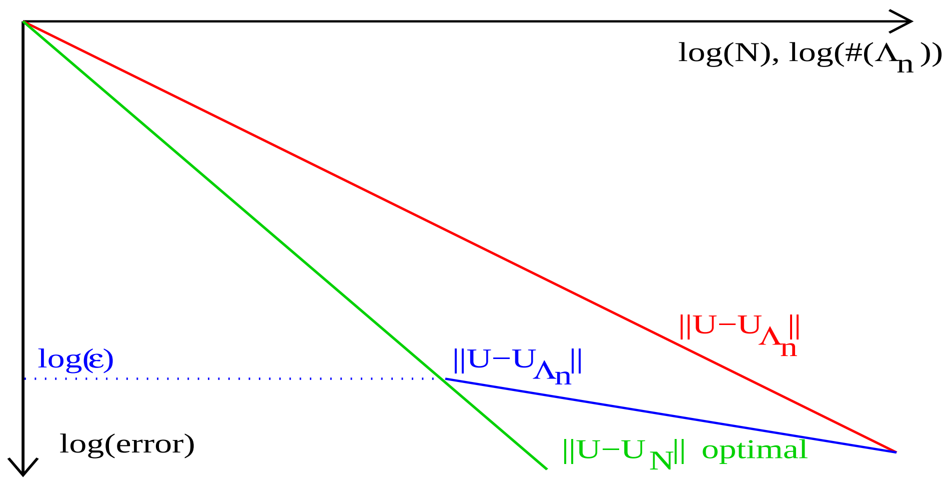
derived from the smoothness and vanishing moments of the ψ_λ.

The role of thresholding

Lemma : if U is such that $\|U - U_N\| \leq CN^{-s}$ and V is such that $\|V - U\| \leq \varepsilon$, then with $a > 1$ fixed and W the smallest subvector of V such that $\|V - W\| \leq a\varepsilon$, one has

$$\|U - W\| \leq (1+a)\varepsilon \text{ and } \#(W) \leq C\varepsilon^{-1/s}, \text{ i.e. } \|U - W\| \leq C[\#(W)]^{-s}$$

Thresholding ensures optimality



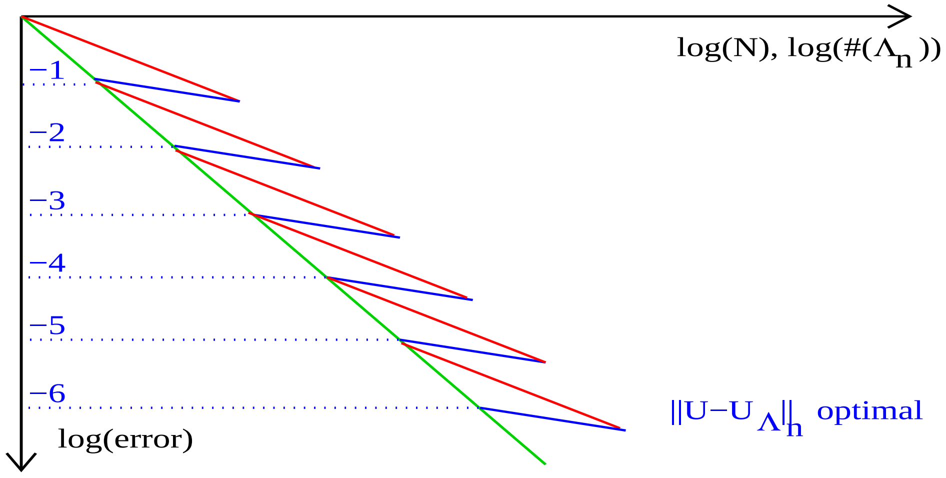
Problem : intermediate memory size and computational time should also be optimal, i.e. $\mathcal{O}(\varepsilon^{-1/s})$.

Geometric tolerances

Idea: decrease

tolerances $\varepsilon_0 = 1$,

$\varepsilon_1 = \frac{1}{2}, \cdots, \varepsilon_j = 2^{-j}$



Fixed number of iteration at each step $j \to j+1$ involving sparse matrix-vector product: $W = \text{APPROX}(AV, \varepsilon)$ obtained by decomposing $V = V_1 + [V_2 - V_1] + [V_4 - V_2] + \cdots$, and taking

$$W := A_{2^J} V_1 + A_{2^{J-1}}[V_2 - V_1] + \cdots + A_1[V_{2^J} - V_{2^{J-1}}]$$

with J large enough such that
$\|W - AV\| \le \|A\| \|V - V_{2^J}\| + \sum_{j=1}^{J} \|A - A_{2^{J-j}}\| \|V_{2^j} - V_{2^{j-1}}\| \le \varepsilon$.

Results

Theorem (Dahmen, DeVore, AC - Math. Comp. 2000) : if V is such that $\|V - V_N\| \leq CN^{-s}$, and if $\|A - A_N\| \leq CN^{-r}$ with $r > s$, then $|\text{Supp}(W)| \leq C\varepsilon^{-1/s}$ and therefore $\|W - AU\| \leq C|\text{Supp}(W)|^{-s}$.

Theorem (Dahmen, DeVore, AC - FoCM 2002) : The general strategy for linear operator equations based on the above ingredients (thresholding, adaptive matrix vector multiplication) achieves the ultimate goal, namely production of U^n and $\Lambda_n = \text{Supp}(U^n)$, such that if $\|U - U_N\| \leq CN^{-s}$, then

$$\|U - U^n\| \leq C\#(\Lambda_n)^{-s},$$

with $\mathcal{O}(\#(\Lambda_n))$ computational cost.

Remarks on practical aspects

All wavelet properties are exploited : Sobolev norm equivalences, smoothness (not always available) and vanishing moments.

Coarsening is not needed in all practical cases studied so far, yet seems necessary in the proof of the optimality theorem ! Similar optimality results recently obtained for adaptive FEM by Binev, Dahmen and DeVore, using the Morin-Nocetto-Siebert algorithm combined with coarsening.

Complexity is dominated by assembling matrix elements, numerical quadratures, addressing the indices in Λ_n (key role of efficient data structures). Practical comparison between adaptive FEM and wavelets based on the same FE spaces : for a given error, wavelets may win for $N_{\mathrm{d.o.f.}}$ but lose (by a factor > 4) for computational cost.

Extension to more general problems

Saddle point problems $AU + B^T P = F$ and $BU = G$, e.g. based on adaptive approximation of the Uzawa iteration (Dahlke, Hochmuth and Urban 1999) :

$$AU^n = F - B^T P^{n-1} \text{ and } P^n = P^{n-1} + \kappa(BU^n - G)$$

No LBB is needed here, adaptivity stabilizes Similar result for adaptive FEM algorithm : Nocetto 2002. Question : do the same concepts apply to convection dominated problems, such as $-\varepsilon\Delta u + a.\nabla u = 0$ with convergence rate independent of ε ?

Extension to nonlinear problems : DeVore, Dahmen, A.C. 2002 (need specific adaptation of fast evaluation of $F(U)$), no available numerical results yet.

Problem dependent tuning seems unavoidable in order to optimize this type of algorithms.

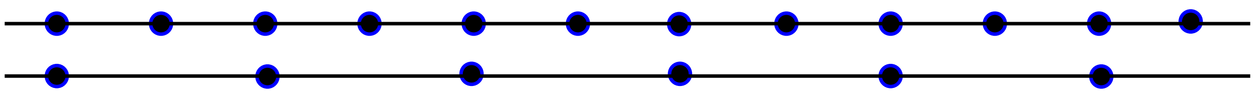
A discrete multiresolution framework

- Γ_j, $j = 0, \cdots, J$: sequence of discretisations at scales 2^{-j}.

- $U_j = (U_j(\gamma))_{\gamma \in \Gamma_j}$ discretisation of a fonction u on Γ_j, i.e. vector of $\mathcal{V}_j := \mathbb{R}^{\Gamma_j}$.

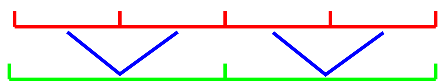
- Restriction operator P^j_{j-1} from $\mathcal{V}_j$ onto $\mathcal{V}_{j-1}$: computes coarser discretization $U_{j-1} = P^j_{j-1} U_j$ from the next finer.

Basic example 1: point values on nested grids $\Gamma_{j-1} \subset \Gamma_j$, i.e. $U_{j-1}(\gamma) = U_j(\gamma)$ for $\gamma \in \Gamma_{j-1}$.



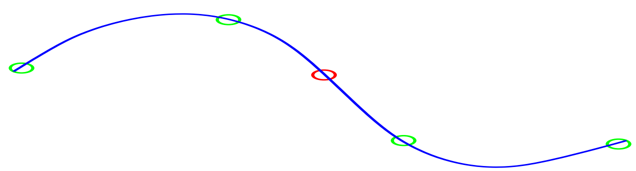
Basic example 2: cell averages on nested partitions, i.e.

$$U_{j-1}(\gamma) = \mathrm{vol}(\gamma)^{-1} \sum_{\mu \in \Gamma_j, \mu \subset \gamma} \mathrm{vol}(\mu) U_j(\mu)$$

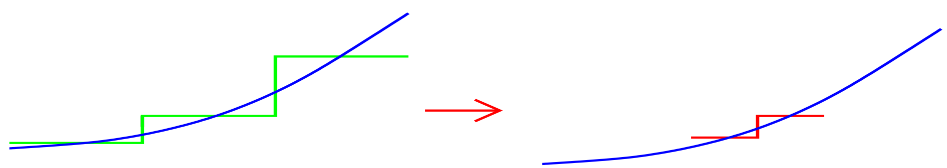


- Prediction operator P_j^{j-1} from $\mathcal{V}_{j-1}$ into $\mathcal{V}_j$: reconstructs an approximation $\hat{U}_j = P_j^{j-1} U_{j-1}$ of U_j.
- Consistancy assumption: $P_{j-1}^{j} P_j^{j-1} = I$

Point value example: $\hat{U}_j(\gamma) = U_{j-1}(\gamma)$ for $\gamma \in \Gamma_{j-1}$, and $\hat{U}_j(\gamma)$ obtained par local interpolation for $\gamma \in \Gamma_j \setminus \Gamma_{j-1}$.



Cell avergage example: $\hat{U}_j(\gamma)$ obtained by "interpolating" the averages in a consistant way, e.g. via polynomial reconstruction.

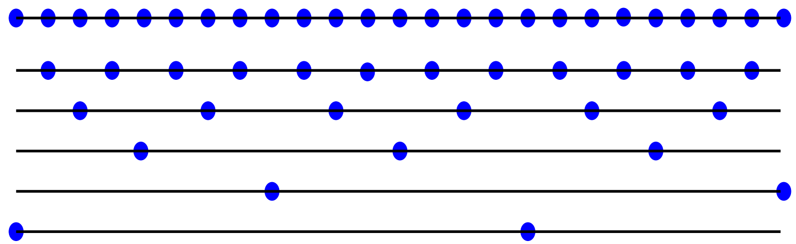


Multiscale decomposition

Prediction error $E_j := U_j - \hat{U}_j \in \mathcal{W}_{j-1} = \mathrm{Ker}(P^j_{j-1})$. Detail vector D_{j-1}: coordinates of E_j in a basis of $\mathcal{W}_{j-1}$.

Point value example: $D_{j-1}(\lambda) = E_j(\lambda)$, $\gamma \in \Gamma_j \setminus \Gamma_{j-1}$ interpolation error at intermediate point. Cell average example: on each coarse cell of Γ_{j-1} the prediction error E_j has null average $\Rightarrow$ define D_j by removing for each coarse cell γ one fine cell $\mu \subset \gamma$.

$$U_J \Leftrightarrow (U_{J-1}, D_{J-1}) \Leftrightarrow (U_{J-2}, D_{J-2}, D_{J-1}) \Leftrightarrow \cdots$$
$$\Leftrightarrow (U_0, D_0, \cdots, D_{J-1}) = \mathcal{M}U_J = (d_\lambda)_{\lambda \in \nabla_J}$$



Physical grid Γ_J

Multiscale grid (point values) ∇_J

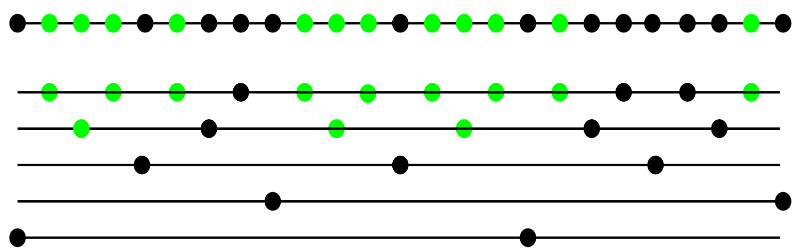
Complexity of $\mathcal{M}$ and $\mathcal{M}^{-1}$: $\mathcal{O}(\mathrm{Card}(\Gamma_J))$

Compression

Thresholding: given a level dependent threshold $\eta = (\eta_0, \cdots, \eta_{J-1})$ set to zero all coefficients $|d_\lambda| \leq \eta_{|\lambda|} \Leftrightarrow$ approximation of U_J by

$$\mathcal{T}_\eta U_J = \mathcal{T}_\Lambda U_J = \mathcal{M}^{-1}\mathcal{R}_\Lambda \mathcal{M} U_J,$$

$\mathcal{R}_\Lambda$: restriction of ∇_J to $\Lambda = \Lambda(\eta) = \{\lambda \in \nabla_J \ \text{t.q.} \ |d_\lambda| \geq \eta_{|\lambda|}\}$.



Adaptive mesh $\Gamma(\Lambda)$

Adaptive set Λ

Compressed representation $(d_\lambda)_{\lambda\in\Lambda}$ to the data of point values or cell averages $U_J(\gamma)$ on an adaptive physical mesh $\Gamma(\Lambda)$ associated to Λ, (need to impose that Λ has a tree structure up to enlarging it).

Complexity of adaptive decomposition and reconstruction algorithms: $\mathcal{O}(\text{Card}(\Lambda))$

Adaptive multiresolution processing

Reference scheme on Γ_J: approximation of $u(x, n\Delta t)$ by $U_J^n = (U_J^n(\gamma))_{\gamma\in\Gamma_J}$ with $U_J^{n+1} = E_J U_J^n$

$$U_J^{n+1}(\gamma) = U_J^n(\gamma) + F(U_J^n(\mu)\ ;\ \mu \in S(\gamma)).$$

$S(\gamma)$: local stencil (excludes implicit schemes).

In the case of FV conservative schemes, F has the form of a balance over the edges surrounding the cell γ

$$U_J^{n+1}(\gamma) = U_J^n(\gamma) + \sum_{\mu \text{ s.t. } |\Gamma_{\gamma,\mu}|\neq 0} F_{\gamma,\mu}^n$$

where $F_{\gamma,\mu}^n = -F_{\mu,\gamma}^n$ is a function of the $U_J^n(\nu)$ for ν in a local stencil surrounding γ and μ.

Adaptive algorithm

Goal: compute approximations of $u(x, n\Delta t)$ by (V_J^n, Λ_η^n), where $V_J^n = (V_J^n(\gamma))_{\gamma\in\Gamma_J}$ is represented by its coefficients $(d_\lambda^n)_{\lambda\in\Lambda_\eta^n}$ or its physical values (point values or cell averages) on the adaptive mesh $(V_J^n(\gamma))_{\gamma\in\Gamma(\Lambda_\eta^n)}$ (we always impose the graded tree structure on Λ_η^n).

Benchmark: an ideal choice would be Λ_η^n the smallest graded tree containing $\{\lambda,\ \ |d_\lambda(U_J^n)| \geq \eta_{|\lambda|}\}$ but it is not be accessible. The adaptive solution V_J^n should still be comparable to $\mathcal{T}_\eta U_J^n$, i.e. the corresponding thresolding operator applied to the exact reference solution.

Initialization: define Λ_η^0 the smallest graded tree containing $\{\lambda,\ \ |d_\lambda(U_J^0)| \geq \eta_{|\lambda|}\}$ and set $V_J^0 := \mathcal{T}_\eta U_J^0$,

Derivation of $(V_J^{n+1}, \Lambda_{n+1})$ from (V_J^n, Λ_n)

Three basic steps:

- Refinement: predict a superset $\Lambda_\eta^n \subset \tilde{\Lambda}_\eta^{n+1}$ adapted to describe the solution at time $n+1$ (ideally such that $|d_\lambda(E_J V_J^n)| < \eta_{|\lambda|}$ if $\lambda \notin \tilde{\Lambda}_\eta^{n+1}$) and extend by $d_\lambda^n = 0$ for $\lambda \in \tilde{\Lambda}_\eta^{n+1} \setminus \Lambda_\eta^n$.

- Evolution: compute the new value $V_J^{n+1}(\gamma)$, for $\gamma \in \Gamma(\tilde{\Lambda}_\eta^{n+1})$ (ideally $V_J^{n+1} = \mathcal{T}_{\tilde{\Lambda}_\eta^{n+1}} E_J V_J^n$).

- Coarsening: apply level dependent thresholding operator to the computed vector $(d_\lambda^n)_{\lambda \in \tilde{\Lambda}_\eta^{n+1}} \Rightarrow$ new set $\Lambda_\eta^{n+1} \subset \tilde{\Lambda}_\eta^{n+1}$ and V_J^{n+1}.

The loss of accuracy with respect to the reference scheme depends on each of the three steps.

- Coarsening: accuracy is controled by the level of the threshold η and the stability properties of the multiscale reconstruction (existence of underlying continuous wavelet systems).

- Refinement: accuracy is controled by analyzing the action of the discrete evolution operator E_J on the size of the coefficients in the multiscale decomposition.

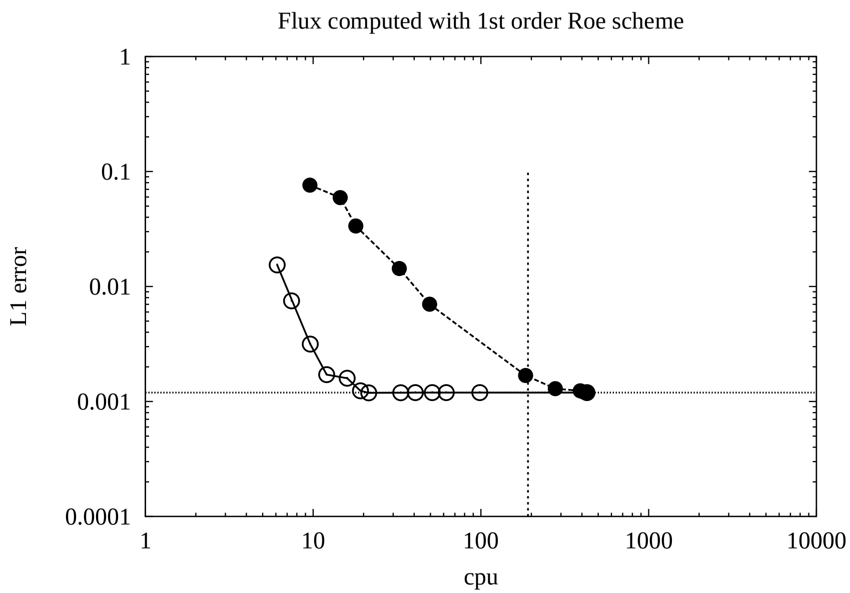
- Evolution: need an accurate evolution step in the compressed form. Two possible approaches:

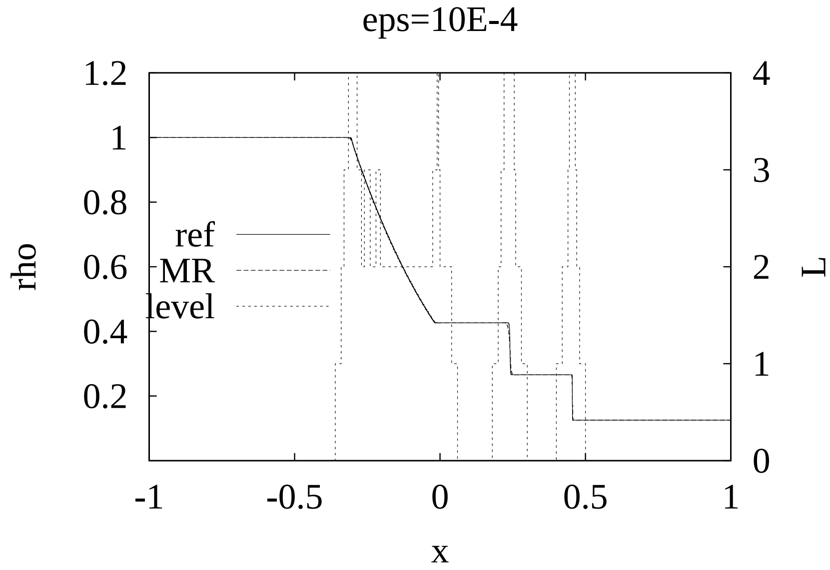
(i) direct application of the numerical scheme on the adaptive grid $\Gamma(\tilde{\Lambda}_\eta^{n+1})$: loss of accuracy for low order schemes.

(ii) exact computation of $\mathcal{T}_{\tilde{\Lambda}_\eta^{n+1}} E_J V_J^n$ by local adaptive reconstruction on fine grid : more accurate but more costly.

Numerical illustration

In 1D: comparison of AMR and local reconstruction on Sod tube test.





For a low order reference scheme, only local reconstruction preserves the accuracy with a substantial reduction of CPU time and memory space (1/20 at best).

Error Analysis

Remark: adaptive evolution with local reconstruction is given by

$$V_J^{n+1} = \mathcal{T}_{\Lambda_\eta^{n+1}} \mathcal{T}_{\tilde{\Lambda}_\eta^{n+1}} E_J V_J^n.$$

Compare $U_J^{n+1} = E_J U_J^n$ with $V_J^{n+1} = \mathcal{T}_{\Lambda_{n+1}} \mathcal{T}_{\tilde{\Lambda}_{n+1}} E_J V_J^n$.

Cumulative error analysis between both solutions:

$$\|U_J^{n+1} - V_J^{n+1}\| \leq \|E_J U_J^n - E_J V_J^n\| + d_n,$$

with $d_n = \|V_J^{n+1} - E_J V_J^n\| \leq t_n + c_n$ where

$t_n := \|\mathcal{T}_{\Lambda_{n+1}} \mathcal{T}_{\tilde{\Lambda}_{n+1}} E_J V_J^n - \mathcal{T}_{\tilde{\Lambda}_{n+1}} E_J V_J^n\|$, $c_n := \|\mathcal{T}_{\tilde{\Lambda}_{n+1}} E_J V_J^n - E_J V_J^n\|$,

denote the thresholding and refinement errors. The analysis of refinement and thresholding strategies should allow to control both terms with a prescribed precision ε.

Controling the thresholding error

Analysis based on underlying continuous wavelet system (ψ_λ):

$$\|U_J - \mathcal{T}_\Lambda U_J\| \leq \sum_{\lambda \notin \Lambda} \|d_\lambda \psi_\lambda\|.$$

For the L^1 norm, this gives $\|U_J - \mathcal{T}_\Lambda U_J\| \leq C \sum_{\lambda \notin \Lambda} 2^{-d|\lambda|}|d_\lambda|$, and therefore with $\eta_j = 2^{dj}\eta_0$,

$$\|U_J - \mathcal{T}_\eta U_J\| \leq C \sum_{2^{-d|\lambda|}|d_\lambda| < \eta_0} 2^{-d|\lambda|}|d_\lambda|$$

- Crudest estimate: $\eta_0 \#(\nabla_J) \sim \eta_0 2^{dJ} \Rightarrow$ take $\eta_0 = \varepsilon 2^{-dJ}$.
- Better estimate: $\eta_0 \#(\tilde{\Lambda}^{n+1}) \Rightarrow$ take $\eta_0 = \varepsilon / \#(\tilde{\Lambda}^{n+1})$.
- Even better: take largest η_0 s.t. $\sum_{2^{-d|\lambda|}|d_\lambda| < \eta_0} 2^{-d|\lambda|}|d_\lambda| \leq \varepsilon$.

Controling the refinement error

Harten's refinement rule for hyperbolic equations (assuming CFL condition for the reference scheme $\Delta t \leq C 2^{-J}$):

- If $|d_\lambda| > \eta_{|\lambda|}$ include in $\tilde{\Lambda}_\eta^{n+1}$ the neighbors of λ at the same level.

- If $|d_\lambda| > 2^{r-1}\eta_{|\lambda|}$ also include the childrens of λ at the finer level.

Here r represents the order of accuracy of the prediction operator. Not sufficient to prove that $|d_\lambda(E_J V_J^n)| < \eta_{|\lambda|}$ if $\lambda \notin \tilde{\Lambda}_\eta^{n+1}$.

This can be proved by a more severe refinement rule: refine of n level if $2^{n(s-1)}\eta_{|\lambda|} \leq |d_\lambda| < 2^{(n+1)(s-1)}\eta_{|\lambda|}$, with s the Hölder smoothness of the underlying wavelet system.

In practice, however, we observe that Harten's rule is sufficient and that the thresolding error dominates the refinement error.

A crude error estimate

Assuming stability in the prescribed norm $\|\cdot\|$ for the reference scheme in the sense that $\|E_J U - E_J V\| \leq (1 + c\Delta t)\|U - V\|$, this yields the cumulative (too pessimistic) estimate

$$\|U_J^{n+1} - V_J^{n+1}\| \leq (1 + c\Delta t)\|U_J^n - V_J^n\| + \varepsilon \leq \cdots \leq C(T)n\varepsilon \sim \frac{\varepsilon}{\Delta x}$$

Main defects of this analysis :

- In most practical cases, we observe that thresholding and refinement error does not accumulate linearly.

- No error bound available in terms of the number $N.$ of wavelet coefficients which is used to represent the adaptive solution (we are not ensured that $N << \#(\Gamma_J)$). Such a bound would require that the relevant Besov-Sobolev smoothness which governs the nonlinear approximation error in $\|\cdot\|$ is preserved under the action of evolution and thresholding.