Mathematisches Forschungsinstitut Oberwolfach

Report No. 34/2004

Wavelet and Multiscale Methods

Organised by
Albert Cohen (Paris)
Wolfgang Dahmen (Aachen)
Ronald A. DeVore (Columbia)
Angela Kunoth (Bonn)

July 11th–July 17th, 2004

Introduction by the Organisers

The Wavelet and Multiscale Methods workshop, organized by Albert Cohen (Paris), Wolfgang Dahmen (Aachen), Ronald A. DeVore (Columbia) and Angela Kunoth (Bonn) was held July 11–17, 2004. A central objective was to bring together leading researchers working in a variety of different areas where multiscale phenomena play an intrinsic role regarding the theoretical understanding of the problems as well as the numerical solution concepts. This meeting was well attended with over 45 participants with a broad geographic representation from all continents. With only very few exceptions, the most prominent representers of the respective fields participated in the workshop.

The main link being multiscale concepts, the overall audience was certainly less homogeneous than in most other workshops. In order to provide a solid ground for discussions and synergies, thirteen of the twenty six talks were longer survey–style lectures by leading experts who had been contacted by the organizers prior to the meeting. The following topics were addressed in the survey–type lectures:

- Adaptive and multiscale methods for integral equations and PDEs;
- problems of high spatial dimension;
- quasi-sparse representation of global operators and hierarchical matrices;
- nonlinear multiresolution techniques and new libraries for image analysis and compression;
- highly localized multiscale representations on the sphere;
- multiscale approaches for inverse problems;

- mathematical learning theory and statistics.

Several adaptive wavelet schemes for fluid flow problems were presented with varying emphasis on analysis and numerical realization. This included combinations of statistical and multiresolution concepts for understanding turbulence. Such approaches were complemented by a "two-scale" point of view where the concept of "residual free bubbles" allows one to model the influence of unresolved scales to the resolved ones, initiating vivid discussions about a possible combination of these approaches. From the perspective of nonlinear approximation, best N-term approximation in corresponding anisotropic function spaces was discussed, which is another important component in adaptive solution concepts in this context.

Another focal point was the quasi–sparse or data–sparse representation of global operators, for instance, via wavelet bases or through concepts like hierarchical matrices. On one hand, this was shown to lead to new shape–optimization techniques as well as to significant advancements of adaptive wavelet methods for boundary integral equations. On the other hand, hierarchical matrix concepts were seen to make extremely high dimensional problems of a certain separable structure numerically tractable. In fact, problems of high spatial dimension were treated in the context of quantum chemistry centering upon the solution of the Schrödinger equation for many particles as one of the grand challenges of the future. Here compression and sparse grid techniques were again complemented by theoretical investigations of the regularity of eigensolutions which provides the foundation of the potential success of such numerical methods.

Anisotropy is also a key issue in modern image compression techniques. The "wedgelet" concept was presented as a way of extending approximation libraries in combination with image encoding and model selection techniques.

Moreover, the meeting adequately reflected through several talks some very recent advances in the currently highly active area of Mathematical Learning Theory, where new synergies between approximation theory and statistics are just emerging. In particular, wavelet techniques have now led to new universal estimators in a model free regression context where the universality could be achieved through nonlinear approximation concepts.

Regarding the inhomogeneity of the participants' research fields, in order to maintain a high level of attendance and interest throughout the meeting, the topics of the lectures were deliberately mixed. This created a very open atmosphere and essentially all talks were followed by extensive discussions revealing a number of unforeseen connections between seemingly different subjects. The reaction of the participants during the meeting clearly indicated that the combination of different areas was very well perceived as a source of new ideas and perspectives. New collaborations were initiated and already existing cooperations partly across disciplines were pursued on a visibly intense level.

Workshop: Wavelet and Multiscale Methods

Table of Contents

Abstracts

Approximation and Compression of Piecewise Smooth Multidimensional Functions

Richard G. Baraniuk
(joint work with Michael Wakin, Justin Romberg, Hyeokho Choi, Venkat Chandrasekaran, and Dror Baron)

Efficient representations for discontinuities in data are important for many signal processing applications, but standard Fourier and wavelet representations fail to efficiently capture the structure of the discontinuities. In this talk, we discuss three new frameworks for approximating and compressing piecewise smooth multidimensional functions that aim to overcome the drawbacks of traditional representations.

First, for piecewise constant multidimensional functions we discuss *geometric tilings*. Unlike additive wavelet-based techniques, tilings use exactly one atom at each location in the function. In two dimensions (2D), *wedgelets* provide piecewise linear approximations to edges (and perform optimally for C^2 edges) [1]. In higher dimensions, we introduce *surflets*, which provide piecewise polynomial approximations to arbitrary C^K discontinuities [2]. We have developed multiscale statistical models for these tilings and compression algorithms with optimal rate-distortion performance.

Second, for piecewise smooth 2D images we discuss a wavelet-domain modeling framework motivated by the geometric tilings above. The work can be interpreted in two different ways: 1) as an extension to the "zerotree model" for wavelet coefficients that explicitly accounts for edge structure at fine scales, and 2) as a novel atomic representation that synthesizes images using a sparse combination of wavelets and *wedgeprints* — anisotropic atoms that are well-suited to edge singularities. Using this framework, we develop a prototype image coder that has near-optimal rate-distortion performance $D(R) \sim (\log R/R)^2$ for piecewise smooth images containing smooth C^2 regions separated by edges along smooth C^2 contours [3]. We also discuss generalizations of wedgeprints to higher dimensions and higher-order polynomials.

Third, we generalize the complex wavelet transform to higher dimensions using a multidimensional Hilbert transform. The resulting *hypercomplex wavelet transform* (HWT) is a flexible building block for constructing new classes of nearly shift-invariant wavelet frames that are oriented along lower-dimensional subspaces [4, 5]. The HWT can be computed efficiently using a 1D dual-tree complex wavelet transform along each signal axis. We demonstrate experimentally how the HWT can be used for fast detection of straight lines in 3-D.

References

[1] J. K. Romberg, M. B. Wakin, and R. G. Baraniuk, "Multiscale Wedgelet Image Analysis: Fast Decompositions and Modeling," *IEEE International Conference on Image Processing – ICIP'02*, Rochester, NY, September 2002. Available at www.dsp.rice.edu/publications.

[2] V. Chandrasekaran, M. Wakin, D. Baron, and R. G. Baraniuk, "Surflets: A Sparse Representation for Multidimensional Functions Containing Smooth Discontinuities," *IEEE International Symposium on Information Theory – ISIT*, Chicago, June 2004. Extended preprint available at www.dsp.rice.edu/publications.

[3] M. Wakin, J. K. Romberg, H. Choi, and R. G. Baraniuk, "Wavelet-domain Approximation and Compression of Piecewise Smooth Images," submitted to *IEEE Transactions on Signal Processing*, 2004. Available at www.dsp.rice.edu/publications.

[4] W. Chan, H. Choi, and R. G. Baraniuk, "Directional Hypercomplex Wavelets for Multidimensional Signal Analysis and Processing," *IEEE International Conference on Acoustics, Speech and Signal Processing – ICASSP'04*, Montreal, May 2004. Available at www.dsp.rice.edu/publications.

[5] W. Chan, H. Choi, R. G. Baraniuk, "Quaternion Wavelets for Image Analysis and Processing," *IEEE International Conference on Image Processing – ICIP-2004*, Singapore, October 2004. Available at www.dsp.rice.edu/publications.

Wavelet and Schur Complement Based Eigenvalue Elimination

Luise Blank

In many well posed problems properly scaled wavelet discretization can keep the condition numbers of the arising stiffness matrices bounded independently of the refinement level. In spite of the asymptotical assertion of convergence of the solution process, there may be still several outliers of eigenvalues which reduce drastically the convergence speed or even prohibit the application of an iterative solver. The purpose of this talk is to present a Schur complement method for the elimination of these eigenvalues for symmetric problems. In this context the Schur complement falls into the class of deflation methods. Deflation methods using invariant subspaces have been used in several ways and applications [4, 5, 6, 7, 8]. In general, the Schur complement does not employ an invariant subspace. However, it may use a subspace which contains an invariant subspace. Then, roughly speaking, the corresponding eigenvalues are eliminated [2]. Moreover, in practice an additional clustering of eigenvalues appears, much in favour for the application of iterative solvers.

The question arises, how to choose the relevant subspace which has to be very coarse to guarantee efficiency. This choice of the relevant subspace for the Schur complement is heavily based on wavelet discretization which connects the infinite dimensional problem with the arising sequence of discretized subproblems. Additionally, a priori knowledge of the eigenspace of the original problem is employed. Knowing the eigenfunctions corresponding to the smallest eigenvalues of the original problem, the compressed wavelet decomposition provides a coarse problem adapted subspace. One expects that this space also captures good approximations

of the eigenvectors corresponding to the smallest eigenvalues of the discretized systems, which are restricted to a finite dimensional subset of wavelets. Then, for the whole hierarchy of systems the Schur complement with respect to this coarse fixed subspace may give rise to well-conditioned systems. The way to choose a good subspace is, although theoretically motivated, still heuristic, and hence further analysis and investigations are necessary.

Despite that, the potential of this method is presented with some numerical results for an optimization problem arising in chemical engineering where on-line restriction leads to the application of a multiscale refinement approach. Using properly scaled wavelets the condition number can be bounded independently of the refinement level [1]. However, the in modulo smallest eigenvalues are still close to zero due to inherent features of the estimation problem, i.e., the original ill-posedness, the low measure of observability, and the large spectrum of the system matrices of the involved differential equations [3]. Guided by the eigenfunction corresponding to the smallest eigenvalues of the original problem for one state only [3], we choose the coarse subspace. In one example based on a simple model for a chemical process producing ethylene glycol, we were faced with condition numbers around 1e+20 for mesh size 2^{-7}. Scaling of the states and a choice of the regularization parameter in favour for the condition number leads roughly to 5.e+13 and wavelet preconditioning to 2e+9. However, the application of the above sketched Schur complement yields a condition number of 280 and provides additionally an eigenvalues distribution much in favour of the cg-method [2]. Such behaviour has been observed also in other numerical examples. Hence, for our application in state estimation, where we are typically faced with very large condition numbers, only the application of the Schur complement technique as a preconditioner enabled us to apply an iterative method in a successful way, providing high estimation quality with low computational cost.

References

[1] T. Binder, L. Blank, W. Dahmen, and W. Marquardt. Iterative algorithms for multiscale state estimation, part 1: Concept. *J. Optim. Theory Appl.*, 111(3):501–527, 2001.

[2] L. Blank. Preconditioning via a Schur complement method: an application in state estimation. *SIAM J. Sci. Comput.*, Vol. 25, No. 3, 942-960, 2003.

[3] L. Blank. State estimation without regularizing the initial data. Inverse Problems, volume 20, issue 5, pages 1357 - 1370, 2004.

[4] H.C. Elman, D.J. Silvester, and A.J. Wathen. Performance and analysis of saddle point preconditioners for the discrete steady–state Navier–Stokes equations. *Numer. Math.*, 90(4):665–688, 2002.

[5] J. Frank and C. Vuik. On the construction of deflation–based preconditioners. *SIAM J. Sci. Comput.*, 23(2):442–462, 2001.

[6] L. Mansfield. Damped Jacobi preconditioning and coarse grid deflation for conjugate gradient iteration on parallel computers. *SIAM J. Sci. Statist. Comput.*, 12:1314–1323, 1991.

[7] R.A. Nicolaides. Deflation of conjugate gradients with applications to boundary value problems. *SIAM J. Numer. Anal.*, 24:355–365, 1987.

[8] C. Vuik, A. Segal, and J.A. Meijerink. An efficient preconditioned CG method for the solution of a class of layered problems with extreme contrasts in the coefficients. *J. Comput. Phys.*, 152:1–19, 1999.

Bernstein Polynomials and Learning Theory

Dietrich Braess
(joint work with Thomas Sauer)

The point of departure is an encoding problem. We want to encode a (long) text such that the code length is minimized. The length of the codes for the symbols (letters) of the alphabet will be different; in particular, short codes are chosen for the letters that occur frequently. When the encoding is fixed, only a sample of n letters (from the beginning of the text) is known. It is the goal to determine the best code and the asymptotics of the loss that results from the sampling error.

The relative frequency of the symbols in the sample is given by the Binomial distribution. Therefore the Bernstein polynomials of the entropy function

$$f(x) := -x \log x - (1-x)\log(1-x) \tag{1}$$

enter into the expression for the expectation value of the loss. Unfortunately, the entropy function belongs to a class of functions for which the determination of the error function is considered a hard problem.

The encoding problem is as follows: The symbols $A_0, A_1, \dots, A_m$ of an alphabet with $m+1$ letters are to be encoded. It is possible to have a code with length $\log \frac{1}{q_i}$ for the letter A_i if the numbers q_i satisfy $\sum_{i=0}^m q_i = 1$. If the symbol A_i is found with the probability p_i, the expectation value of the code length is $\sum_{i=0}^m p_i \log \frac{1}{q_i}$. The minimum of this expression is attained if $q_i = p_i$ for all i. If the lengths q_i differ from the optimal values, there is a redundancy, i.e. a difference to the minimum of

$$\sum_{i=0}^{m} p_i \log \frac{p_i}{q_i}.$$

We restrict our attention to the special case $m = 1$. Here the loss function above may be rewritten as

$$L_{KL}(p,q) = p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}\,, \tag{2}$$

if we write $p_1 = p$, $p_0 = 1-p$, $q_1 = q$, and $q_0 = 1-q$. Usually L_{KL} is called the Kullback–Leibler distance.

The probability p is unknown, but we assume to have a sample with n letters. The encoding will be performed on the base of the information, how often the symbol A_1 is contained in the sample. Due to Bernoulli, the probability for finding it k times in the sample is $\binom{n}{k}(1-p)^{n-k}p^k = B_k^n(p)$ where B_k^n is a Bernstein polynomial. Now, an appropriate rule $k \mapsto Q(k)$, $0 \le k \le n$, is to be found for the encoding procedure. If the sample contains the symbol A_1 exactly k times, the encoding for the parameter $q_k = Q(k)$ will be chosen. The expectation value of the redundancy is

$$F_n(p) = \sum_{k=0}^{n} B_k^n(p) L_{KL}(p, q_k). \tag{3}$$

The following problem arises: *Find numbers* $q_k \in (0,1),\ k = 0, 1, \ldots, n$, *such that the worst case redundancy*

$$\sup_{0 \le p \le 1} F_n(p)$$

is minimized.

It was known that the minimax is asymptotically between $0.5/n$ and $0.50922/n$; cf. [BFSS, Kr]. We showed by an analysis of the Bernstein polynomials that the lower bound is the right one:

$$\inf_q \sup_{0 \le p \le 1} F_n(p) = \frac{1}{2n} + o(n^{-1}). \tag{4}$$

More generally, if the alphabet has $m+1$ symbols, then the functions are defined on an m-dimensional simplex and the asymptotics is given by $\frac{m}{2n}$; cf. [BS].

There are other loss functions for which the minimax problem has been often studied [Fe, Le], namely the quadratic loss L_{qu} and the standardized quadratic loss L_{sq}:

$$L_{qu}(p,q) := (p-q)^2, \qquad L_{sq}(p,q) := \frac{(p-q)^2}{p(1-p)}.$$

The expressions with Bernstein polynomials analogous to (3) are more easily treated. Typically the optimal estimators belong to the class of add-β rules

$$q_k = \frac{k+\beta}{n+2\beta}\,.$$

with appropriate β. This is also true for the Kullback-Leibler distance apart from fact that there are exceptional rules for the estimators for the boundary values. The add-β rules also show that the estimator q_k which can be understood as an estimator of the probability p does not always coincide with the observed relative frequency k/n.

The question whether a priori information on the probability p reduces the asymptotical behavior of the loss was recently studied [BD]. Assume that we know that $p \in [a,b]$ with $0 \le a < b \le 1$. It turned out that this does not diminish the asymptotic loss for L_{sq} and L_{KL}. The same is true for L_{qu}, however, only if $a \le 1/2 \le b$.

Now we turn to the Bernstein polynomial of the entropy function (1). It is of interest for determining the loss since the loss (3) is a polynomial minus the error of the approximation of f by its Bernstein polynomial. From Voronovskaja's famous theorem [Lo] it is known that

$$\lim_{n\to\infty} n(f - B_n[f])(x) = 1/2 \quad \text{for } 0 < x < 1,$$

but the convergence cannot be uniform since $(f - B_n[f])(0) = (f - B_n[f])(1) = 0$ for all n. We use the monotonicity of the mapping $f \mapsto B_n[f]$, the nonpositivity of $f^{(6)}$ and Taylor's polynomial of degree 5 to establish

$$f(x) - B_n[f](x) \ge \frac{1}{2n} + \frac{1}{20n^2x(1-x)} - \frac{1}{12n^2} \quad \text{for } 15n^{-1} \le x \le 1 - 15n^{-1}$$

and

$$f(x) - B_n[f](x) \leq \frac{1}{2n} + \frac{c(a,b)}{n^2} \quad \text{for } a \leq x \leq b,$$

if $0 < a < b < 1$; see [BS]. The region next to the boundary points is delt by the transformation $x = z/n$. The limit

$$\lim_{n\to\infty} n(f - B_n[f])(\frac{z}{n}) = -z\log z + ze^{-z}\sum_{k=1}^{\infty}\frac{z^k}{k!}k\log k$$

is determined numerically, and the numerical results are sufficient to complete the proof of (4). The limit above is the error of the approximation by the Szász operator. It can be understood from the connection of the Szász operators with the Poisson distribution.

References

[BD] D. Braess, H. Dette, The asymptotic minimax risk for the estimation of constrained binomial and multinomial probabilities. Sankhyā (to appear) – also available on http://homepage.rub.de/Dietrich.Braess/ftp.html

[BFSS] D. Braess, J. Forster, T. Sauer, and H. U. Simon, How to achieve minimax expected Kullback-Leibler Distance from an unknown finite distribution. In: N. Cesa-Bianchi, N. Numao, and R. Reischuk (eds.), "Algorithmic Learning Theory", Proceedings of the 13th International Conference ALT 2002", Springer-Verlag, Heidelberg, 2002, 380–394.

[BS] D. Braess and T. Sauer. Bernstein polynomials and learning theory. J. Approximation Theory 128, 2004, 187–206.

[Fe] T.S. Ferguson. Mathematical Statistics. Academic Press, New York 1967.

[Kr] R.E. Krichevskiy. Laplace's law of succession and universal encoding. IEEE Transactions on Information Theory 44, 1998, 296–303.

[Le] E.L. Lehmann. Theory of Point Estimation. Wiley, N.Y. 1983.

[Lo] G. G. Lorentz. Bernstein Polynomials. University of Toronto Press, 1953.

A Total Curvature Diminishing Property for P_1 Finite Element Interpolation

Martin Campos Pinto

1. Introduction

We consider here a surface $f \in W^{2,1}(\Omega, \mathbb{R})$ and define its *total curvature* by the semi-norm

$$|f|_{W^{2,1}} := \|\partial^2_{xx} f\|_{L^1} + \|\partial^2_{xy} f\|_{L^1} + \|\partial^2_{yy} f\|_{L^1}. \tag{1}$$

We shall assume for convenience that Ω is a polygonal subset of $\mathbb{R}^2$, so that we may consider a triangulation $\mathcal{T}$ on it, and denote by $P_{\mathcal{T}}$ the associated P_1 finite element interpolation which maps f to $f_{\mathcal{T}}$. We know from the continuous embedding of $W^{2,1}$ into L^∞ that f is continuous, so that $P_{\mathcal{T}}$ is well defined on $W^{2,1}$.

What we are interested in is the alteration of the curvature. Of course, $f_{\mathcal{T}}$ is not in $W^{2,1}$ anymore: its first derivatives $Df_{\mathcal{T}} = (\partial_x, \partial_y)f_{\mathcal{T}}$ are a piecewise constant vector field, and its second derivatives $D^2 f_{\mathcal{T}} = (\partial^2_{xx}, \partial^2_{xx}, \partial^2_{xx})f_{\mathcal{T}}$ are

Radon (Dirac) measures, but $|f_{\mathcal{T}}|_{W^{2,1}}$ still makes sense if we consider the total mass $\int |D^2 f_{\mathcal{T}}|$ (in analogy with BV functions). It is not difficult, then, to see that the interpolation is stable, in the sense that there exists a constant $C(\mathcal{T})$ that depends on the triangulation, such that

$$|f_{\mathcal{T}}|_{W^{2,1}} \leq C(\mathcal{T})|f|_{W^{2,1}}. \tag{2}$$

Our aim is to prove that this holds with $C(\mathcal{T}) = 1$, which is equivalent to saying that $f_{\mathcal{T}}$ has the minimal curvature among all interpolants of f. The motivation for such a result lies in the study of semi-Lagrangian schemes for the Vlasov equation, where the solution at each time step is first advected along the caracteristic curves, and then interpolated on some finite element space (see for instance [1]). In order to save computational time and memory space, these schemes have recently been performed on adaptive meshes (see [2], [5], [6], and [3], [4] for general introductions to the use of adaptive methods in numerical analysis). Using classical estimates like

$$\|f - f_{\mathcal{T}}\|_{L^\infty} \leq C \sup_{T \in \mathcal{T}} |f|_{W^{2,1}(T)}, \tag{3}$$

we see that a possible "local indicator" for balancing the L^∞ error is precisely the local curvature, and that it is a crucial issue, for estimating the size of the adaptive mesh, to bound the time evolution of the total curvature of the solutions. From this point of view, we see that stability (2) is not sufficient, since the number of time steps (and hence the number of interpolations) has no reason to be bounded, even for a given time range $[0, t^*]$.

2. Discrete curvatures

It is not true, however, that the curvature defined by (1) is decreased by the interpolation, as we may see from the following example, but we can prove that a diminishing property holds for an equivalent semi-norm. Indeed, consider for $\varepsilon > 0$ the four points

$$a_\varepsilon = (-\varepsilon, 0), \quad b = (0,0), \quad c_\varepsilon = (\varepsilon, 1), \quad d = (0,1), \tag{4}$$

and let $\mathcal{T}_\varepsilon$ be made of the two triangles $A_\varepsilon = (a_\varepsilon, b_\varepsilon, d)$ and $B_\varepsilon = (b, c_\varepsilon, d)$ on $\Omega_\varepsilon = A_\varepsilon \cup B_\varepsilon$. Since the Dirac measure $D^2 f_{\mathcal{T}_\varepsilon}$ is concentrated on the edge $\gamma := [b, d]$, letting ε tend to zero does not change $|f_{\mathcal{T}_\varepsilon}|_{W^{2,1}(\Omega_\varepsilon)}$, while $|f|_{W^{2,1}(\Omega_\varepsilon)}$ goes to zero. In fact, this example shows that a curvature measure cannot be diminished by the interpolation unless it depends on the triangulation.

We therefore define the following quantity that "simulates" $|f_{\mathcal{T}}|_{W^{2,1}}$ at the discrete level:

$$|f_{\mathcal{T}}|_{\star,0} := \sum_{\gamma \in \mathcal{E}} |\gamma| \; \|[Df_{\mathcal{T}}]_\gamma\| \tag{5}$$

where $[\cdot]_\gamma$ and $\|\cdot\|$ denote respectively a jump over γ and the euclidian norm of $\mathbb{R}^2$. The idea now is to define a continuous semi-norm by iterative refinements, and establish a diminishing property between two consecutive levels. For any integer j, we let then $\mathcal{T}_j$ be the j-th structured refinement of $\mathcal{T}$, in the sense that at

each level, each triangle is divided into 4 similar triangles. Using $\mathcal{T}_j$ we define the following refined discrete curvature

$$|f|_{\star,j} := |f_{\mathcal{T}_j}|_{\star,0} \tag{6}$$

and this allows to define a *limit curvature*, as we have the

Lemma 1. (The limit curvature) *For any $f \in W^{2,1}$ and $\mathcal{T}$, the limit*

$$|f|_\star := \lim_{j\to\infty} |f|_{\star,j} \tag{7}$$

defines a semi-norm, and there are two constants $A(\mathcal{T})$ and $B(\mathcal{T})$ such that

$$A(\mathcal{T})\,|f_{\mathcal{T}}|_{W^{2,1}} \le |f|_\star \le B(\mathcal{T})\,|f_{\mathcal{T}}|_{W^{2,1}}. \tag{8}$$

The reason why this limit curvature is the right one is given by the next lemma:

Lemma 2. *For any continuous f, the refined curvatures do satisfy*

$$|f|_{\star,j} \le |f|_{\star,j+1} \tag{9}$$

for any $j \ge 0$.

These two lemmas are in fact very easy to prove, as we shall see below (in the case where $\mathcal{T}$ is simply derived from a cartesian mesh). They together lead to the

Theorem 1. *For any triangulation $\mathcal{T}$, there is a semi-norm $|\cdot|_\star$ which is equivalent to $|\cdot|_{W^{2,1}}$ (with constants that depend on $\mathcal{T}$), for which the piecewise affine interpolation $f_{\mathcal{T}}$ of any $f \in W^{2,1}$ satisfies*

$$|f_{\mathcal{T}}|_\star \le |f|_\star \tag{10}$$

and we have in addition

$$|f_{\mathcal{T}}|_\star = \sum_{\gamma\in\mathcal{E}} |\gamma|\ \|[Df_{\mathcal{T}}]_\gamma\|. \tag{11}$$

3. A sketch of proof

If the triangulation is made of half squares, there are only three types of edges. Considering one of them by setting $\varepsilon = 1$ in the previous example (4) (and dropping this index), we first compute

$$|\gamma|\ \|[Df_{\mathcal{T}}]_\gamma\| = |f(a) + f(c) - f(b) - f(d)|, \tag{12}$$

while on the second hand we have

$$f(c) - f(d) - f(b) + f(a) = \iint_{A\cup B} \partial^2_{xy} f + \partial^2_{xx} f. \tag{13}$$

By doing so for the two other edges, we first verify that

$$|f|_{\star,j} \to \|\partial^2_{xy} f + \partial^2_{xx} f\|_{L^1} + \|\partial^2_{xy} f + \partial^2_{yy} f\|_{L^1} + 2\|\partial^2_{xy} f\|_{L^1} \tag{14}$$

as j tends to infinity, therefore we have

$$2/3\,|f|_{W^{2,1}} \le |f|_\star \le 4|f|_{W^{2,1}} \tag{15}$$

which completes the proof of the first lemma. To prove the second lemma, we introduce the following midpoints:

$$k = \mathrm{m}(a,b), \quad l = \mathrm{m}(b,c), \quad m = \mathrm{m}(c,d), \quad n = \mathrm{m}(d,a), \quad o = \mathrm{m}(b,d), \tag{16}$$

and denote $\gamma_1 := [k,n]$, $\gamma_2 := [b,o]$, $\gamma_3 := [o,d]$ and $\gamma_4 := [l,m]$: from (12), we have then

$$\begin{aligned} |\gamma| \, \|[Df_{\mathcal{T}}]_\gamma\| &= |f(a)+f(c)-f(b)-f(d)| \\ &\le |f(a)+f(o)-f(k)-f(n)| + |f(k)+f(l)-f(b)-f(o)| \\ &\quad + |f(n)+f(m)-f(o)-f(d)| + |f(o)+f(c)-f(l)-f(m)| \\ &= \textstyle\sum_{i=1}^4 |\gamma_i| \, \|[Df_{\mathcal{T}}]_{\gamma_i}\|, \end{aligned}$$

and doing so for the other two types of edges and summing over all the pairs of triangles completes the proof of the second lemma.

References

[1] N. Besse, E. Sonnendrücker, *Semi-Lagrangian schemes for the Vlasov equation on an unstructured mesh of phase space*, to appear in Comput. Phys. Comm.

[2] M. Campos Pinto, M. Mehrenberger, *Adaptive numerical resolution of the Vlasov equation,* submitted to Numerical methods for hyperbolic and kinetic problems, Cemracs 2003 Proceedings.

[3] A. Cohen, *Numerical analysis of wavelet methods,* North-Holland, 2003.

[4] R. DeVore, *Nonlinear approximation,* Acta numerica, No. 7, Cambridge Univ. Press, Cambridge, 1998, 51-150.

[5] M. Gutnic, I. Paun, E. Sonnendrücker, *Vlasov simulations on an adaptive phase-space grid* to appear in Comput. Phys. Comm.

[6] E. Sonnendrücker, J. Roche, P. Bertrand, A. Ghizzo, *The Semi-Lagrangian method for the numerical resolution of Vlasov equations*, J. Comput. Phys. 149 (1999), 201-220.

Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information

Emmanuel J. Candes

(joint work with J. Romberg and T. Tao)

In many applications of practical interest, we often wish to reconstruct an object (a discrete signal, a discrete image, etc.) from incomplete Fourier samples. In a discrete setting, we may pose the problem as follows; let $\hat{f}$ be the Fourier transform of a discrete object $f(t)$, $t \in \mathbb{Z}_N^d := \{0,1,\ldots,N-1\}^d$,

$$\hat{f}(k) = \sum_{t \in \mathbb{Z}_N^d} f(t) e^{-i2\pi k \cdot t/N}.$$

The problem is then to recover f from partial frequency information, namely, from $\hat{f}(k)$, where $k = (k_1,\ldots,k_d)$ belongs to some set Ω of cardinality less than N^d—the size of the discrete object.

Is it possible to reconstruct f from the partial knowledge of its Fourier coefficients on the set Ω? Common wisdom says that this is actually impossible since one has far fewer equations than unknowns. This talk, however, will explain a

surprising phenomenon which says exact recovery is actually possible under certain sparsity conditions. The recovery consists of solving a convex optimization problem.

It might be best to illustrate the theory with an example. Suppose we wish to reconstruct a 2-dimensional image $f(t_1, t_2)$ from samples $\hat{f}|_\Omega$ of its discrete Fourier transform on a star-shaped domain Ω (this type of acquisition geometry arises frequently in medical applications such as Computed Tomography (CT) or MRI angiography). To recover f from partial Fourier samples, we find a solution $f^\sharp$ to the optimization problem

$$\min \|g\|_{BV} \qquad \text{subject to} \qquad \hat{g}(k) = \hat{f}(k) \text{ for all } k \in \Omega. \tag{1}$$

In a nutshell, given partial observation $\hat{f}_\Omega$, we seek a solution $f^\sharp$ with minimum complexity—here Total Variation (TV)—and whose 'visible' coefficients match those of the unknown object f.

In typical numerical experiments mimicking current or future medical devices, one gathers 512 samples along each of 22 radial lines (that is, we only acquire about 4% of all Fourier coefficients). When we use (1) for the recovery problem with the popular Logan-Shepp phantom as a test image, the results are surprising. The reconstruction is **exact**; that is $f^\sharp = f$! In fact, there is nothing special about the Logan-Shepp phantom; indeed, we performed a series of experiments of this type and obtained perfect reconstruction on many similar test phantoms.

This talk develops a quantitative understanding of this remarkable phenomenon and a typical result is as follows: suppose f is a one dimensional signal obeying

$$|T| = \#\{t,\ f(t-1) \neq f(t)\} = O(|\Omega|/\log N)$$

then for nearly all sets Ω (of cardinality $|\Omega|$), f can be reconstructed exactly as the solution to the ℓ_1 minimization problem

$$\|g\|_{BV} = \min_g \sum_{t=0}^{N-1} |g(t) - g(t-1)|, \quad \text{s.t. } \hat{g}(k) = \hat{f}(k) \text{ for all } k \in \Omega.$$

Except for the logarithmic factor, the condition on the size of the support is sharp.

The methodology extends to a variety of other setups and higher dimensions. For example, we show how one can reconstruct a sparse signal (here a sparse signal is understood as being made out of relatively few spikes in one or two or higher-dimensions) from incomplete frequency samples—provided that the number of spikes obeys the condition above—by minimizing other convex functionals such as the ℓ_1-norm of f

$$\|g\|_{\ell_1} \min_g \sum_{t=0}^{N-1} |g(t)|.$$

Finally, we show that similar exact reconstruction phenomena hold for other synthesis/measurement pairs. Suppose one is given a pair of of bases (B_1, B_2) and randomly selected coefficients of an object f in one basis, say B_2. Then, f can be recovered exactly provided that it may be synthesized as a sparse superposition of elements in B_1. The relationship between the number of nonzero terms in B_1

and the number of observed coefficients depends upon the *incoherence* between the two bases. The more incoherent, the fewer coefficients needed.

We conclude by pointing out that underlying our analysis are two results which might be of independent interest:

- First, we develop a new notion of uncertainty principle holding for almost any pair (T, Ω) which says that it is impossible to design a signal f with $T = \mathrm{supp}(f)$ and $\Omega = \mathrm{supp}(\hat{f})$ unless
$$|T| + |\Omega| \asymp O(N/\log N).$$
- And second, our results establish—with overwhelming probability—exact equivalence between a combinatorial and a convex optimization problem.

References

[1] E. J. Candès, J. Romberg, and T. Tao, Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information, Technical Report, California Institute of Technology. *Submitted for publication.*

[2] E. J. Candès, J. Romberg, and T. Tao, The Role of Sparsity and Incoherence for Exact Reconstruction from Incomplete Measurements. Manuscript.

[3] D.L. Donoho and X. Huo, Uncertainty principles and ideal atomic decomposition, *IEEE Transactions on Information Theory*, **47** (2001), 2845–2862.

Universal Algorithms for Learning Theory

Albert Cohen

(joint work with P. Binev, W. Dahmen, R. DeVore, and V. Temlyakov)

We shall present the construction and analysis of a universal estimator for the mathematical problem of supervised learning. In this problem, we observe the data $\mathbf{z} = (z_1, \ldots, z_m) \subset X \times Y$ of m independent random observations $z_i = (x_i, y_i)$, $i = 1, \ldots, m$, identically distributed according to a probability ρ on a product space $X \times Y$. We are interested in estimating the *regression function* $f_\rho(x)$ defined as the conditional expectation of the random variable y at x:

$$f_\rho(x) := \int_Y y d\rho(y|x)$$

with $\rho(y|x)$ the conditional probability measure with respect to x. The estimator $f_{\mathbf{z}}$ is assessed by the measure of the error $\|f_\rho - f_{\mathbf{z}}\|$ in the $L^2(X, \rho_X)$ metric with ρ_X the marginal probability which is unknown. This type of problem is refered to as *distribution-free*, see [3] for a general introduction.

Universal means that the estimator does not depend on any a priori assumptions about the regression function to be estimated. Our universal estimator, introduced in [1], consists of a least-square fitting procedure using piecewise constant

functions on a partition which depends adaptively on the data. The partition is generated by a splitting procedure which somehow differs from those used in CART algorithms [2] in the sense that it is based on thresholding empirical quantities which play the role of wavelet coefficients.

It is proven that this estimator performs at the optimal convergence rate for a wide class of priors on the regression function. Namely if the regression function is in a smoothness space of order not exceeding one (a limitation resulting because the estimator uses piecewise constants) then the estimator converges to the regression function (in the least squares sense) with an optimal rate of convergence in terms of the number of samples. The estimator is also numerically feasible and can be implemented on-line.

References

[1] P. Binev, A. Cohen, W. Dahmen, R. DeVore and V. Temlyakov, Universal algorithms for learning theory - Part I : piecewise constant functions, preprint, 2004

[2] D.L. Donoho, CART and best-ortho-basis : a connection, Ann. Stat. **25**, 1870-1911, 1997.

[3] L. Györfy, M. Kohler, A. Krzyzak, A. and H. Walk, A distribution-free theory of nonparametric regression, Springer, Berlin, 2002.

Convergence for AdaBoost-type Algorithms in Learning Theory

Ingrid Daubechies

(joint work with Cynthia Rudin and Rob Schapire)

Given a vast number of data points, a huge number of "hypotheses" (i.e. yes/no questions about the data points), and a partition of the data points into two classes, a standard "classification" problem in learning theory is to construct a convex combination of the hypotheses that best reproduces the partition. More precisely, given

$$\begin{aligned} \{x_1, \cdots, x_D\} &= X \\ h_1, \cdots, h_H &; \quad h_i : X \to \{-1, 1\} \\ y &: \quad X \to \{-1, 1\} \,, \end{aligned}$$

one seeks $\lambda \in \mathbb{R}_+^H$ with $\lambda_1 + \cdots + \lambda_H = 1$, so that λ maximizes either

- $\#\{i \ ; \ \sum_{j=1}^H \lambda_j h_j(x_i)$ and $y(x_i)$ have the same sign $\}$
  (in the *non-separable* case, where for each λ, there exist i_1 and i_2 so that $\sum_{j=1}^H \lambda_j h_j(x_{i_1}) y(x_{i_1})$ and $\sum_{j=1}^H \lambda_j h_j(x_{i_2}) y(x_{i_2})$ have different signs)
- $\min_{i=1,\cdots,D} \left[\left(\sum_{j=1}^H \lambda_j h_j(x_i) \right) y(x_i) \right]$
  (in the *separable* case, where there exists a λ such that for all i the expression $\sum_{j=1}^H \lambda_j h_j(x_i)$ has the same sign as $y(x_i)$).

Because H, D are extremely large, standard (low-dimensional) optimization techniques cannot be applied effectively, and other search techniques have to be used. Boosting algorithms are iterative methods that seek to improve successive "guesses" for λ_t by giving more weight, in the construction of λ_{t+1}, to the data points that were misclassified by the current λ_t, and down-playing the data points that were correctly classified by λ_t.

The most popular boosting algorithm is AdaBoost, first proposed by Freund and Schapire in [1]. In each iteration t, it singles out one particular j, and it "updates" the corresponding component $(\lambda_t)_j$ to construct λ_{t+1}. In order to select j, it finds the direction in which a "guiding function" $F(\lambda) := \sum_{i=1}^{D} \exp[-(M\lambda)_i]$ decreases fastest, where the $D \times H$-matrix M is defined by $M_{ij} = h_j(x_i)y(x_i)$. More precisely, the iteration step t consists in:

$$\begin{aligned} j_t &\in \operatorname{argmax}_j \left(\sum_{i=1}^{D} M_{ij} \exp[-(M\lambda)_i] \right) \\ r_t &:= \sum_{i=1}^{D} M_{ij_t} \exp[-(M\lambda)_i] \\ \alpha_t &:= \frac{1}{2} \ln \left(\frac{1+r_t}{1-r_t} \right) \\ (\lambda_{t+1})_{j_t} &= (\lambda_t)_{j_t} + \alpha_t \\ \text{and } (\lambda_{t+1})_j &= (\lambda_t)_j \text{ for } j \neq j_t \ . \end{aligned}$$

Note that the λ_t in this iterative scheme are not normalized to 1; in fact, it is their normalized versions $\widetilde{\lambda}_t = \lambda_t [\sum_{j=1}^{H} (\lambda_t)_j]^{-1}$ that provide the desired convex combinations of the h_j. In the non-separable case the properties of AdaBoost are well-known; in the separable case, however, it was until recently *not* known whether the $\widetilde{\lambda}_t$ produced by AdaBoost converge (they certainly seem to converge numerically in examples), and, if they converge, whether their limit is a true maximizer. In her Ph.D. thesis, co-advised by R. Schapire and the author, Cynthia Rudin showed the following results:

- AdaBoost can be recast, via duality, into an iteration on

$$S = \{(d_1, \cdots, d_D); d_I \geq 0 \text{ for all } i, \text{ and } \sum_{i=1}^{D} d_i = 1\};$$

- for many classifications y, this new iterative algorithm exhibits stable cycles;
- when the dual iteration map on S converges to a stable cycle, AdaBoost itself converges to a $\widetilde{\lambda}_{\text{limit}}$;
- however, there exist classifications y with a stable cycle for the dual map, for which the corresponding $\widetilde{\lambda}_{\text{limit}}$ does **not** achieve the desired maximization;
- One can replace AdaBoost by a smoothened version that is equally easy to implement, and that always converges, for all y, to a correct maximizing $\widetilde{\lambda}$. Like AdaBoost, this new algorithm selects, at every t, one index j_t for which λ_{j_t} is updated, while the other components of λ_t are not touched. In fact, the criterion for the selection of j_t is exactly the same as listed above for AdaBoost, but the update α_t for λ_{j_t} is smaller.

- The convergence of this smoothened version of AdaBoost holds even in the so-called "non-optimal" case, where the search for the index j_t is stopped as soon as a reasonable candidate for j has been identified, even if it is not necessarily the best.

These results are presented in the papers [2] and [3].

References

[1] Y. Freund and R. Schapire, A decision-theoretic generalization of online learning and an application to boosting. J. of Computer and System Sciences **55** pp. 119-139, 1997.
[2] C. Rudin, I. Daubechies and R. Schapire, The dynamics of AdaBoost: Cyclic behavior and convergence of margins, to appear in J. of Machine Learning Research, 2004.
[3] C. Rudin, R. Schapire and I. Daubechies, Boosting based on a smooth margin, preprint, PACM, Princeton University, 2004.

Regularizing Linear Inverse Problems by Sparsity Constraints

Christine De Mol
(joint work with Ingrid Daubechies and Michel Defrise)

Many linear or linearized inverse imaging and scattering problems amount to solving an operator equation of the form $Kf = g$, where the solution f and the data g belong to Hilbert spaces and K is a bounded linear operator, e.g. an integral operator with a kernel representing the response of the imaging device. Such problems are typically ill-posed (or ill-conditioned) and to define regularized solutions, i.e. solutions which are stable in the presence of noise, the problem can be reformulated as the minimization of a penalized least-squares functional. Classical regularization methods like Tikhonov's regularization use quadratic penalties, such as the norm of f in a Hilbert or quadratic Sobolev space.
Recently, in [1], we have considered the use of non-quadratic penalties to enforce sparsity in the solution of such problems. In many problems of practical relevance, indeed, the underlying ideal noiseless solution is expected to have a sparse expansion on an arbitrary pre–assigned orthonormal basis. As it has often been advocated, to enforce this type of prior knowledge, we should penalize the ℓ^1-norm of the sequence of coefficients of this expansion, instead of its ℓ^2-norm. We have proved that this type of penalty still provides a proper regularization method "à la Tikhonov", albeit nonlinear.
To compute the regularized solutions, i.e. to minimize the corresponding penalized least-squares functionals, we must then solve a complicated nonlinear optimization problem. To this aim, we have used an optimization transfer technique to derive an iterative algorithm in which the successive iterates are obtained as the minimizers of a sequence of surrogate functionals that are each easy to minimize. The resulting algorithm is simply a Landweber iteration with soft-thresholding applied at each iteration step. We have proved that this algorithm converges strongly in the Hilbert norm. The results described above extend to the case of weighted ℓ^p-penalties, with $1 \leq p \leq 2$. For wavelet expansions, this amounts to penalizing the

norm of the solution in a Besov space. The above results also apply to the statistical counterpart of this deterministic regularization setting, namely to penalized maximum likelihood with a Gaussian noise model and a (generalized) Laplacian prior for the distribution of the solution coefficients in the chosen basis.
In [2] we show how the framework can be generalized to treat the case of mixed smoothness and sparsity constraints, i.e. the case where the solution is assumed to be the sum of a smooth part, subject to a quadratic penalty, and of a sparse part, subject to a non-quadratic penalty. Several iterative algorithms are proposed to compute such solutions and practical applications to image restoration are discussed. The convergence of these schemes is established in [2] only for finite-dimensional problems, but strong convergence results and regularization properties can also be obtained in an infinite-dimensional setting similar to the one used in [1] (M. Defrise and C. De Mol, paper in progress). A similar approach is proposed in [3] for the decomposition of images in oscillating and cartoon components using wavelet bases.

References

[1] Daubechies, I., Defrise, M. and De Mol, C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, preprint June 2003 (available from http://www.arXiv.org, math.FA/0307152), to appear in *Comm. Pure Appl. Math.*

[2] De Mol, C. and Defrise, M. Inverse imaging with mixed penalties, *Proceedings URSI EMTS 2004* (Ed. PLUS Univ. Pisa), pp. 798–800.

[3] Daubechies, I. and Teschke, G. Variational image restoration by means of wavelets: simultaneous decomposition, deblurring and denoising, submitted to *Appl. Comp. Harm. Anal.*

Nonlinear MR transformations and Error Control

Rosa Donat
(joint work with F. Arandiga)

Multiscale transformations are being used in the first step of Transform Coding Algorithms (TCA). Ideally, a multiscale transformation allows for an *efficient representation* of the signal data, which is then processed using a non-reversible quantizer and passed on to the encoder. The latter produces the final set of compressed data which is ready to be stored or transmitted.

In such algorithms, the properties of the multiscale transformation are most important in the overall performance of the TCA. Compression is achieved because many small coefficients in the multiscale representation can be quantized or discarded with little loss of *real information* contents.

Recently, several authors have tried to build adaptive wavelet transforms that allow for more flexibility in signal/image representations and result in fewer large wavelet coefficients. Adaptivity is a hard issue in the classical, Fourier-based, wavelet framework, and other frameworks, which are based on a completely spatial construction of the multiscale transformation turn out to be more appropriate.

The framework of Ami Harten, with its intrinsic link to Approximation Theory, turns out to be specially adequate to design such nonlinear multiscale transformations. We shall revisit ENO-based nonlinear MR transformation and their performance in image compression. These can be interpreted as nonlinear filter banks, where the data determines a particular choice of a linear prediction filter among a finite pool. The central issue here becomes that of stability. We shall analyze the Error-Control mechanism designed by A. Harten in [Ha] and compare its performance and capabilities to other mechanisms considered in [CDSB] that ensure stability through synchronization within the nonlinear Lifting framework.

In addition, we shall also describe how Harten's strategies can be exploited to obtain near-lossless compression algorithms for which the maximum deviation from the original signal/image can be prescribed *a priori*.

References

[AD00] F. Arandiga, R. Donat, Nonlinear Multiscale Transformations: The approach of Ami Harten Numerical Algorithms 23, pp. 175-216

[AD03] F. Arandiga, R. Donat, Nonlinear Multiscale Transformations: From Synchronization to Error Control, Algorithms for Approximation IV, pp. 306-313 (2002)

[AD03b] F. Arandiga, R. Donat, Data Compression Algorithms with Maximum Error Control, IEEE Intl. Conf. on Image Processing, ICIP-03, Barcelona, Spain.

[CDSB] R. Claypole, G.M. Davis, W. Sweldens, R. Baraniuk. Nonlinear wavelet transforms for image coding via lifting. IEEE Transactions on Image Processing, 12: 1449-1459 (2003).

[Ha] A. Harten, Discrete Multiresolution Analysis and generalized wavelets. J. of Applied Num. Math. 12:153-193 (1993).

The Technique of Hierarchical Matrices

Wolfgang Hackbusch

1. Introduction

1.1. **Large and Dense Matrices.** The discretisation of partial differential equations (PDEs) lead to large systems of equations. The boundary element method (BEM; cf. [21]) produces fully populated (dense) matrices. Since the cost for their generation as well as the storage of dense $n \times n$-matrices costs at least $\mathcal{O}(n^2)$, several techniques were developed to reduce the cost. Examples are the panel clustering method [19], the multipole method [20] and compression techniques in the case of wavelets bases [4]. These methods support the storage of the matrix and the matrix-vector multiplication. Both is reduced to almost linear complexity, by which we mean $\mathcal{O}(n \log^q n)$ for some $q \geq 0$. One must emphasise that all methods replace the original matrix A by another matrix A'. The matrix-vector multiplication $A'x$ is only an approximation to Ax, but the error $\|A - A'\|$ can be controlled. The error bound $\|A - A'\| \leq \varepsilon$ requires a cost $\mathcal{O}(n \log \frac{1}{\varepsilon} \log^q n)$.

The hierarchical matrices generalise the panel clustering method in order to support *all* matrix operations, i.e., matrix addition, matrix multiplication and matrix inversion. This is also of interest for the sparse matrices from FEM, since

the dense inverse can be computed. The mentioned matrix operations are performed only approximately. As above, an accuracy of ε leads to the factor $\log \frac{1}{\varepsilon}$ in the costs.

1.2. **Construction of Hierarchical Matrices.** The main idea in the construction of hierarchical matrices is the combination of two steps.

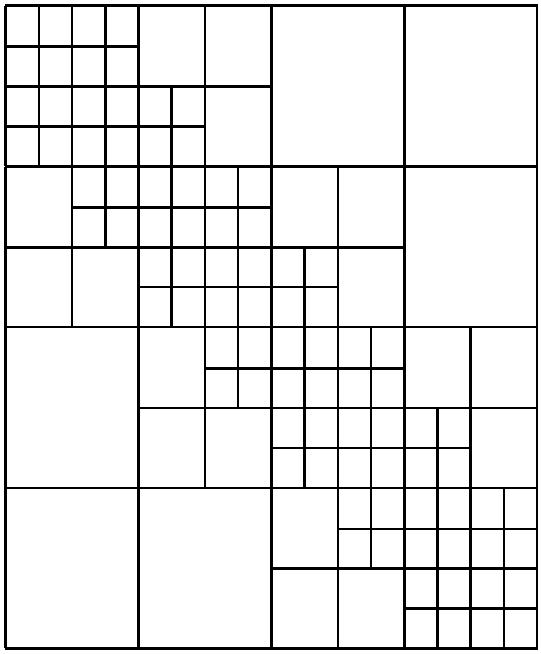

FIGURE 1. Example of a block partitioning

Step I: The matrix is partitioned into certain blocks defining the partition $\mathcal{P}$. Let I be the set of degrees of freedom. A block $b \subset I \times I$ is the product $b = \tau \times \sigma$ of subsets ("clusters") $\tau, \sigma \subset I$. More precisely, the sets of clusters is organised in the form of a tree (cluster tree), which yields the hierarchical structure. The size of the blocks is essential. Each degree of freedom is associated with a nodal point of the FEM/BEM or with the support of the corresponding basis function. This allows to define a diameter $\operatorname{diam}(\tau)$ and a distance $\operatorname{dist}(\tau, \sigma)$ in a natural way. We allow a block $b = \tau \times \sigma$ in the partitioning $\mathcal{P}$, if they are *admissible*, i.e.,

$$\min\{\operatorname{diam}(\tau), \operatorname{diam}(\sigma)\} \leq 2\eta \operatorname{dist}(\tau, \sigma)$$

for some fixed η, or in b is trivial in the sense that the size $\min\{\#\tau, \#\sigma\}$ is below some number $b_{\min}$. The admissibility condition leads to small blocks along the diagonal (where $\operatorname{dist}(\tau, \sigma)$ is small) and to large block far from the diagonal (see Figure 1).

Step II: We fill all blocks in the partitioning $\mathcal{P}$ from above by rank-k matrices, i.e., the matrix block $M|_b = (M_{ij})_{i \in \tau, j \in \sigma}$ (for $b = \tau \times \sigma \in \mathcal{P}$) must satisfy $rank(M|_b) \leq k$. Each of these matrix blocks is represented two matrices $A_b \in \mathbb{R}^{\#\tau \times k}$, $B_b \in \mathbb{R}^{\#\sigma \times k}$ so that

$$M|_b = A_b B_b^\top .$$

Note that the storage costs of A_b, B_b are $(\#\tau + \#\sigma)\, k$ instead of $\#\tau * \#\sigma$ for the naive method. Of course, the use of A_b, B_b is only advantageous if $k \ll \min\{\#\tau, \#\sigma\}$.

The resulting definition of hierarchical matrices of local rank k (more precisely $\leq k$) is

$$\mathcal{H}_k := \left\{ M \in \mathbb{R}^{I\times I} : rank(M|_b) \leq k \text{ for all } b \in \mathcal{P} \right\}.$$

The similar construction is possible for rectangular matrix from $\mathbb{R}^{I\times J}$.

Introductory papers on hierarchical matrices are [12], [2], [15], [8]. Details of the construction are explained in [10].

1.3. **Accuracy.** The stiffness matrices from the boundary element method as well as the inverse of the finite element stiffness matrix (for elliptic boundary value problems) have the following property: Let A be the exact matrix. The singular values of $A|_b$ (not of the whole matrix A) decay exponentially. This allows to approximate $A|_b$ up to an error ε with a rank-k matrix where $k = \mathcal{O}(\log \frac{1}{\varepsilon})$. Hence, A can be well approximated by some $A_{\mathcal{H}} \in \mathcal{H}_k$.

In the BEM case, this follows easily from the smoothness of the fundamental solution. Concerning the inverse FEM stiffness matrix, this result is proved in [1] and holds even for nonsmooth coefficients.

1.4. **Computational Costs.** The following asymptotics hold for $k \ll n = \#I$.

- The *storage* of hierarchical matrices from $\mathcal{H}_k$ requires $\mathcal{O}(nk \log n)$ units.
- The *matrix-vector multiplication* $(A \in \mathcal{H}_k, x \in \mathbb{R}^I \mapsto Ax)$ is exact and requires $\mathcal{O}(nk \log n)$ arithmetical operations.
- The *matrix-matrix addition* $(A, B \in \mathcal{H}_k \mapsto C := A \oplus B \approx A + B)$ computes the approximate sum in $\mathcal{O}(nk^2 \log n)$ operations.
- The *matrix-matrix multiplication* $(A, B \in \mathcal{H}_k \mapsto C := A \odot B \approx A * B)$ computes the approximate product in $\mathcal{O}(nk^2 \log^2 n)$ operations.
- The same cost estimate holds for the *matrix inversion* $(A \in \mathcal{H}_k \mapsto C \approx A^{-1})$.

The details, in particular characterisations of the constants involved in $\mathcal{O}(\cdot)$ are to be found in [10].

2. Application Fields

BEM matrices: The first goal in the boundary element method (BEM) is to generate a data-sparse approximation to the dense system matrix in order to reduce the storage requirements. The error should be comparable with the already existing discretisation error. This can be achieved with hierarchical matrices of the local rank $k = \mathcal{O}(\log n)$.

FEM preconditioning: Linear equations with sparse FEM matrices A are usually solved iteratively, where a good preconditioner (i.e., a fast iterative method) is needed. Sometimes it is hard to find a good one. Since any rough approximation B to A^{-1} is a good preconditioner, one can compute the inverse of A in $\mathcal{H}_k$ for rather small k. This approach is in particular of interest, if A is a Schur complement

matrix (as it appears in saddle point problems, mixed finite element formulations, etc.). Instead of an approximate inverse also approximate LU-factors are available.

Domain decomposition: Eliminating the interior unknowns in a domain decomposition method with non-overlapping subdomains, one obtains a dense matrix (Schur complement) for the nodal points on the skeleton (union of interfaces). The hierarchical matrix technique allows the elimination as well as the treatment of the skeleton matrix (cf. [13], [17]).

Matrix equations: The following matrix equations arise in control theory. The Lyapunov equation $AX + XA = C$ or the nonlinear Riccati equation $A^\top X + XB - XFX + G = 0$ define a system of n^2 equations for the n^2 unknown entries of X. Therefore the best possible solve seems to need a work of $\mathcal{O}(n^2)$. If the coefficient matrix A arises from an elliptic operator (as in control problems with a state governed by an elliptic boundary value problem), it turns out that the solution X can be well approximated by an $X' \in \mathcal{H}_k$. The costs add up to $\mathcal{O}(nk^2 \log^3 n)$ even in the case of the nonlinear Riccati equation.

Matrix functions: For parabolic problems the matrix exponential function $\exp(-tA)$ is of interest, where A is the positive definite discretisation of an elliptic operator. The method proposed and analysed in [6] represents $\exp(-tA)$ by a Cauchy integral. Replacing the integral by a numerical quadrature, we are able to compute $\exp(-tA)$ with accuracy ε with a cost of order $\mathcal{O}(n \log^p \frac{1}{\varepsilon} \log^q n)$. Similarly, other matrix functions can be computed, e.g., $A^{-\alpha}$, $\cos(t\sqrt{A})A^{-k}$, $\left(\sinh \sqrt{A}\right)^{-1} \sinh(x\sqrt{A})$ (see [5], [7]). Another very interesting function is the sign-function $\operatorname{sign}(A)$ (see [11]).

Problems in high spatial dimensions: Related techniques be applied to problems in high spatial dimensions when Kronecker products of matrices can be used. An example is given in [9]: The discrete Laplace operator A in $[0,1]^d$ corresponding to $n = 1024$ nodal points in each direction and $d = 2048$ is an $N \times N$-matrix of size $N = 1024^{2048} \approx 1.2 \times 10^{6165}$. Nevertheless the inverse can be computed with high accuracy in 5 minutes (see [9] for older results).

3. $\mathcal{H}^2$-Matrices

Besides the hierarchy in the cluster tree, one can install a second hierarchy. Instead of the general rank-k matrices for the matrix blocks one requires the matrix blocks $M|_b$ ($b = \tau \times \sigma$) to belong to a fixed tensor space $V_\tau \times V_\sigma$ (i.e., $M|_b = A_b B_b^\top$ with $range(A_b) \subset V_\tau$ and $range(B_b) \subset V_\sigma$). In addition one need a compatibility condition for V_τ and $V_{\tau'}$ when τ' is the son of τ in the cluster tree. The arising $\mathcal{H}^2$-matrices lead to lower cost since instead of the matrices A_b, B_b per block, one has to store only a $k \times k$-matrix and transformation matrices of similar size. The saving can be seen in one or two factors $\log n$ less. The $\mathcal{H}^2$-technique has first been presented in [18]. Further papers are [3] and [14].

References

[1] M. Bebendorf and W. Hackbusch: Existence of $\mathcal{H}$-matrix approximants to the inverse FE-matrix of elliptic operators with L^∞-coefficients. *Numer. Math. 95* (2003), 1-28.

[2] S. Börm, L. Grasedyck, W. Hackbusch: Introduction to hierarchical matrices with applications. *Eng. Anal. Bound. Elem. 27* (2003), 405-422.

[3] S. Börm and W. Hackbusch: Data-sparse approximation by adaptive $\mathcal{H}^2$-matrices. *Computing 69* (2002), 1–35.

[4] W. Dahmen, S. Prössdorf, and R. Schneider, Wavelet approximation methods for pseudodifferential equations II: Matrix compression and fast solution, *Adv. Comput. Math. 1* (1993), 259-335.

[5] I.P. Gavrilyuk, W. Hackbusch, and B.N. Khoromskij: $\mathcal{H}$-Matrix approximation for elliptic solution operators in cylinder domains. *East-West J. of Numer. Math. 9* (2001). 25-59.

[6] I.P. Gavrilyuk, W. Hackbusch, and B.N. Khoromskij: $\mathcal{H}$-Matrix approximation for the operator exponential with applications. *Numer. Math. 92* (2002), 83-111.

[7] I.P. Gavrilyuk, W. Hackbusch, and B.N. Khoromskij: Data-sparse approximation to a hierarchy of operator-valued functions. In: Proceedings of the 18th GAMM-Seminar Leipzig 2002, pages 31-52. Electronic publication, ISBN 3-00-009258-7, http://www.mis.mpg.de/conferences/gamm/2002.

[8] L. Grasedyck: *Theorie und Anwendungen Hierarchischer Matrizen.* Dissertation, Universität zu Kiel, 2001.

[9] L. Grasedyck. Existence and computation of a low Kronecker-rank approximant to the solution of a tensor system with tensor right-hand side. *Computing*, 2004. To appear.

[10] L. Grasedyck and W. Hackbusch: Construction and arithmetics of $\mathcal{H}$-matrices. *Computing 70* (2003), 295-334.

[11] L. Grasedyck, W. Hackbusch, and B.N. Khoromskij: Solution of large scale algebraic matrix Riccati equations by use of hierarchical matrices. *Computing 70* (2003), 121-165.

[12] W. Hackbusch: A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part I: Introduction to $\mathcal{H}$-matrices. *Computing 62* (1999), 89-108.

[13] W. Hackbusch: Direct domain decomposition using the hierarchical matrix technique. In I. Herrera, D.E. Keyes, O.B. Widlund, and R. Yates, editors, *Domain decomposition methods in science and engineering. Fourteenth international conference on domain decomposition methods*, pages 39–50. National Autonomous University of Mexico, Mexico City, 2003.

[14] W. Hackbusch and S. Börm: $\mathcal{H}^2$-Matrix approximation of integral operators by interpolation. *Appl. Numer. Math. 43* (2002), 129-143.

[15] W. Hackbusch, L. Grasedyck, and S. Börm. An introduction to hierarchical matrices. *Math. Bohem. 127* (2002), 229–241.

[16] W. Hackbusch and B.N. Khoromskij: A sparse $\mathcal{H}$-matrix arithmetic. Part II: Application to multi-dimensional problems. *Computing 64* (2000), 21-47.

[17] W. Hackbusch, B.N. Khoromskij, and R. Kriemann: *Direct Schur complement method by domain decomposition based on $\mathcal{H}$-matrix approximation.* (DDM 15, to appear).

[18] W. Hackbusch, B.N. Khoromskij, and S.A. Sauter: On $\mathcal{H}^2$-matrices. In H.-J. Bungartz, R.H.W. Hoppe, and C. Zenger, editors, *Lectures on Applied Mathematics*, pages 9–29. Springer-Verlag, Berlin, 2000.

[19] W. Hackbusch and Z.P. Nowak: On the fast matrix multiplication in the boundary element method by panel clustering, *Numer. Math. 54* (1989), 463–491.

[20] V. Rokhlin: Rapid solution of integral equations of classical potential theory, *J. Comput. Phys. 60* (1985), 187–207.

[21] S. Sauter and C. Schwab : *Randelementmethoden*. Teubner, Stuttgart, 2004.

Shape Optimization using Wavelet BEM

Helmut Harbrecht
(joint work with Karsten Eppler)

This talk is concerned with the efficient numerical solution of shape optimization problems

$$J(\Omega) = \int_\Omega j(u, \mathbf{x}) d\mathbf{x} = \min, \tag{1}$$

which arise from the identification of obstacles or inclusions, the computation of free surfaces and the optimal shape design of electromagnets, wings, etc. The *state* u in (1) is assumed to solve the possibly exterior Dirichlet boundary value problem

$$\Delta u = f \text{ in } \Omega, \qquad u = g \text{ on } \Gamma := \partial\Omega. \tag{2}$$

Such problems are often subject to certain constraints

$$C_i(\Omega) = \int_\Omega h_i(\mathbf{x}) d\mathbf{x} = c_i, \quad i = 1, \dots, m,$$
$$C_i(\Gamma) = \int_\Gamma h_i(\mathbf{x}) d\sigma = c_i, \quad i = m+1, \dots, n.$$

For instance, domain integral constraints of practical interest are the volume of the domain while its perimeter defines a boundary integral constraint. Of course, inequality constraints can be considered as well. To ensure well–posedness the functions f, g, and $h_1, \dots, h_n$ are assumed to be sufficiently regular on a sufficiently large security set $D \subset \mathbb{R}^d$, $d = 2, 3$.

The directional derivative with respect to a smooth variation field $\mathbf{V}$ is a scalar quantity defined on the boundary Γ. If we restrict the shape functional (1) according to

$$J(\Omega) = \int_\Omega h(\mathbf{x}) u(\mathbf{x}) + h_0(\mathbf{x}) d\mathbf{x}, \tag{3}$$

the shape gradient involves only the Neumann data of the state and its adjoint

$$\nabla J(\Omega)[\mathbf{V}] = \int_\Gamma \langle \mathbf{V}, \mathbf{n} \rangle \left\{ hg + h_0 - \frac{\partial p}{\partial \mathbf{n}} \frac{\partial (g-u)}{\partial \mathbf{n}} \right\} d\sigma,$$

where the *adjoint state* p satisfies

$$\Delta p = h \text{ in } \Omega, \qquad p = 0 \text{ on } \Gamma. \tag{4}$$

Therefore, one has to compute only the Dirichlet–to–Neumann map to evaluate the shape gradient. Likewise, the shape Hessian admits also a representation as boundary integral. It depends on first and second order tangential and normal derivatives of the state and its adjoint, and the Neumann data of the associated local shape derivatives, cf. [2, 3] for the details.

Integration by parts shows that the class of functionals (3) includes also the Dirichlet energy

$$J(\Omega) = \int_\Omega \|\nabla u(\mathbf{x})\|^2 d\mathbf{x}$$

provided that u admits homogeneous Dirichlet data $g = 0$ in (2). In particular, in this case the adjoint state coincides with the state, that is $u = p$. In accordance with [5] the functional (3), its gradient as well as its Hessian are computable knowing only the boundary data associated with the state and its adjoint. Involving suitable newton potentials, these boundary data are derived via boundary integral equations.

The necessary condition $\nabla L_\alpha(\Omega, \boldsymbol{\lambda}) = 0$ of the Augmented Lagrangian for equality constrains

$$L_\alpha(\Omega, \boldsymbol{\lambda}) = J(\Omega) + \boldsymbol{\lambda}^T \begin{bmatrix} C_1(\Omega) - c_1 \\ \vdots \\ C_m(\Omega) - c_m \\ C_{m+1}(\Gamma) - c_{m+1} \\ \vdots \\ C_n(\Gamma) - c_n \end{bmatrix} + \frac{\alpha}{2} \left\| \begin{bmatrix} C_1(\Omega) - c_1 \\ \vdots \\ C_m(\Omega) - c_m \\ C_{m+1}(\Gamma) - c_{m+1} \\ \vdots \\ C_n(\Gamma) - c_n \end{bmatrix} \right\|^2, \qquad \alpha > 0,$$

is discretized by a Galerkin scheme employing spherical harmonics as ansatz and test functions. The boundary data of the state and its adjoint have to be computed very often with respect to different domains during the shape optimization algorithm. Hence, we propose to solve the boundary integral equations efficiently by a wavelet Galerkin scheme (cf. [1, 10, 12]). Particularly, in view of the calculation of the local shape derivatives, the arising very sparse system of equations can be solved very fast for different right hand sides.

Compared to first order optimization methods, the Newton method provides a higher performance since a line search becomes nearly obsolete. Combining the Newton method with a second order update rule for the Langrange multiplier (cf. [7, 11]) yields a very efficient and robust algorithm to solve the shape optimization problems under consideration. We demonstrate the efficiency of our method by numerical examples concerning a class of shape optimization problems from planar elasticity (cf. [4, 5]), electromagnetic shaping (cf. [6, 8]) and electrical impedance tomography (cf. [9]).

References

[1] W. Dahmen, H. Harbrecht and R. Schneider. Compression Techniques for Boundary Integral Equations – Optimal Complexity Estimates. *Preprint SFB393/02-06*, Faculty of Mathematics, Technical University of Chemnitz, 2002. submitted to SINUM.

[2] K. Eppler Optimal shape design for elliptic equations via BIE-methods. *J. of Applied Mathematics and Computer Science* 10:487–516, 2000.

[3] K. Eppler. Boundary integral representations of second derivatives in shape optimization. *Discussiones Mathematicae (Differential Inclusion Control and Optimization)*, 20:63–78, 2000.

[4] K. Eppler and H. Harbrecht. Numerical solution of elliptic shape optimization problems using wavelet–based BEM. *Optim. Methods Softw.*, 18:105–123, 2003.

[5] K. Eppler and H. Harbrecht. 2nd Order Shape Optimization using Wavelet BEM. *Preprint 06-2003*, Preprint Series of the Institute of Mathematics, Technical University of Berlin, 2003. submitted to Optim. Methods Softw.

[6] K. Eppler and H. Harbrecht. Exterior Electromagnetic Shaping using Wavelet BEM. *Technical Report 13-2003*, Preprint Series of the Institute of Mathematics, Technical University of Berlin, 2003. to appear in Math. Meth. Appl. Sci.

[7] K. Eppler and H. Harbrecht. Second order Lagrange multiplier approximation for constrained shape optimization problems. *Technical Report 35-2003*, Preprint Series of the Institute of Mathematics, Technical University Berlin, 2003. to appear in Proceedings of the 21st IFIP TC 7 Conference on System Modeling and Optimization.

[8] K. Eppler and H. Harbrecht. Fast wavelet BEM for 3d electromagnetic shaping. *Bericht 03-9*, Berichtsreihe des Mathematischen Seminars der Christian-Albrechts-Universität zu Kiel, 2003. to appear in Appl. Numer. Math.

[9] K. Eppler and H. Harbrecht. A regularized Newton method in electrical impedance tomography using shape Hessian information. *WIAS Preprint No. 943*, WIAS Berlin, 2004. submitted to Control Cybern.

[10] H. Harbrecht. Wavelet Galerkin schemes for the boundary element method in three dimensions. *PHD Thesis*, Technical University of Chemnitz, 2001.

[11] K. Mårtensson. A new approach to constrained function approximation. *J. of Optimization Theory and Application* 12:531–554, 1973.

[12] R. Schneider. *Multiskalen- und Wavelet-Matrixkompression: Analysisbasierte Methoden zur Lösung großer vollbesetzter Gleichungssysteme*. B.G. Teubner, Stuttgart, 1998.

Adaptive Wavelet Methods for Partial Differential Equations

Nicholas Kevlahan
(joint work with O.V. Vasilyev, D. Goldstein and J. Alam)

In this talk I will review three approaches we have been developing for solving nonlinear partial differential equations adaptively. Each method uses the adaptive wavelet collocation method (AWCM) based on bi-orthogonal lifted wavelets [Sw] to construct a computational grid adapted to the solution. Derivatives are calculated using high-order finite differences, and the wavelet transform is used to interpolate on the adapted grid. The wavelet decomposition naturally provides a set of nested multiscale grids adapted to the solution, and we take advantage of this property in developing our methods.

In the first method we implement a traditional time marching scheme for time-dependent partial differential equations, but use AWCM to adapt the computational grid to the solution at each time step [KeVa, VaKe02]. The second method uses the multiscale wavelet decomposition as the basis for an adaptive multilevel method for nonlinear elliptic equations [VaKe04]. Finally, we have begun to investigate a combination of the first two approaches to produce an adaptive simultaneous space–time method. In this case, both the space grid and time grid adapt locally to the solution, and the final solution is obtained simultaneously on the entire space–time domain of interest. In the following I briefly outline each method and give some examples.

1. AWCM with time marching

The first approach is to use a classical time integration scheme, but to dynamically adapt the computational grid at each time step using AWCM [KeVa, VaKe02]. This is done by filtering wavelet coefficients with magnitude less than a threshold ϵ (removing a wavelet coefficient removes a grid point), and then reconstructing the solution on the adapted grid in physical space. The method allows for the change in the solution over one time step by adding nearest neighbour wavelets (i.e. grid points) in both position and scale. The nearest neighbours in space correspond to a CFL criterion of one (adding the nearest two neighbours would correspond to CFL criterion of two, if necessary for the time scheme), while adding nearest neighbours in scale allows for the creation of a scale twice as small by a quadratic nonlinearity. Since all operations are $O(\mathcal{N})$ (where $\mathcal{N}$ is the number of wavelets retained), the grid adaptation procedure is efficient for large problems. Note that in this case the time step is uniform for all locations and length-scales (although it can be adjusted to maintain a given tolerance at each discrete time).

I will first illustrate the basic features of the method by applying it to the one-dimensional Burgers and moving shock equations. We have implemented this approach to solve the penalized two- and three-dimensional incompressible Navier–Stokes equations,

$$\begin{aligned} \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}+\mathbf{U})\cdot\nabla\mathbf{u} + \nabla P &= \nu\Delta\mathbf{u} - \frac{1}{\eta}\chi(\mathbf{x},t)(\mathbf{u}+\mathbf{U}-\mathbf{U}_o), \\ \nabla\cdot\mathbf{u} &= 0, \end{aligned} \tag{1}$$

where the last term on the rhs of (1) approximates the no-slip boundary conditions on the surface of an obstacle in the flow as $\eta \to 0$ ($\chi(x,t)$ defines the solid regions of the flow, and $\mathbf{u}_o$ is the obstacle's velocity).

The AWCM method achieves steady-state compression ratios of 270 times for fluid–structure interaction on large domains, and uses fewer computational elements than vortex methods for the same flow (although the cost per element is likely higher for AWCM).

2. AWCM multilevel method for elliptic problems

As mentioned in the introduction, the wavelet multiresolution analysis provides a natural framework for implementing an adaptive version of the multilevel method for elliptic equations [VaKe04]. The multilevel iterative algorithm is similar in spirit to multigrid methods [Br], but is different in the details of its implementation. First, in contrast to multi-grid methods, the lower level grid is not necessarily coarser at every region of the domain. Secondly, lower-order wavelet differentiation is used for the approximate solver (smoother). Thirdly, wavelet interpolation is used for both prolongation and restriction operators. Finally, as in the usual multi-grid methods, GMRES [Sa] is used as the exact solver.

As an example we consider the three-dimensional Poisson equation $\Delta\mathbf{u} = \mathbf{f}$ with a localized source $\mathbf{f}$. The grid adapts to the three-dimensional solution at

the same time as the multilevel methods converges. We find that the L_∞ error converges linearly with the number of V-cycle iterations.

3. Simultaneous space–time solution using AWCM

The main drawback of our time-marching approach to evolution equations described in §1 is that the time step is not truly adaptive. In intermittent problems such as turbulence the time step (as well as the spatial grid) should depend on position and scale. The most straightforward way to achieve this is to apply the multilevel method described in §2 to a time evolution problem on a space–time domain. There is no obvious choice for the boundary condition at $t = t_{max}$, but we have found that a dynamical boundary condition (i.e. simply solving the equation at that boundary) works well.

We consider the Burgers equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = \nu\frac{\partial^2 u}{\partial x^2},$$

on the space–time domain $[0, 1] \times [0, 0.2]$, with periodic spatial boundary conditions, initial condition $u(x, 0) = \sin(2\pi x)$, and viscosity $\nu = 10^{-2}$. The space–time approach gives accurate results, and we obtain a truly adapted grid, in both space and time. Note that small time steps are necessary only in the vicinity of the shock. Further investigation is required to determine whether this approach is more efficient than the classical time marching scheme. The correct size of the adjacent zone is an outstanding question: unlike the case of time-marching methods, there are no theoretical results on the appropriate number of neighbouring points to retain for simultaneous space–time solution.

References

[Br] J.H. Bramble, Multigrid Methods, 1993, CRC Press.

[KeVa] N.K.-R. Kevlahan and O.V. Vasilyev An adaptive wavelet collocation method for fluid–structure interaction at high Reynolds numbers. 2004, preprint, available at `http://www.math.mcmaster.ca/kevla/publications.html`.

[Sa] Y. Saad and M. H. Schultz, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, SIAM J. Sci. Stat. Comput., 7(3), 1986, 856–869.

[Sw] W. Sweldens, The lifting scheme: A custom-design construction of biorthogonal wavelets, Appl. Comput. Harmon. Anal., 3 (2), 1996, 186–200.

[VaKe04] O.V. Vasilyev and N.K.-R. Kevlahan, An adaptive multilevel wavelet collocation method for elliptic problems. 2004 preprint, available at `http://www.math.mcmaster.ca/kevla/publications.html`.

[VaKe02] O.V. Vasilyev and N.K.-R Kevlahan, Hybrid wavelet collocation-Brinkman penalization method for complex geometry flows, Int. J. Num. Meth. Fluids 30, 2002, 531–538.

B–Spline–Based Monotone Multigrid Methods

Angela Kunoth
(joint work with Markus Holtz)

We consider an elliptic variational inequality

$$\text{find } u \in \mathcal{K}: \;\; a(u, v-u) \geq f(v-u) \;\; \text{for all } v \in \mathcal{K},$$

employing a continuous, symmetric and H_0^1– elliptic bilinear form $a(\cdot,\cdot)$, which is to be solved on a closed convex subspace

$$\mathcal{K} := \{v \in H_0^1(\Omega): \;\; v(x) \leq g(x) \;\; \text{for all } x \in \Omega\} \subset H_0^1(\Omega),$$

where $g \in H_0^1(\Omega)$ represents an upper obstacle.

For the efficient numerical solution of such variational inequalities, monotone multigrid (MMG) methods based on piecewise linear finite elements have been investigated over the past decades, see [Ko1, Ko2] and the references therein. Essential for the success of these methods is the appropriate approximation of the obstacle on coarser grids. Since piecewise linear approximations work with geometric considerations employing point values, the extension of the MMG method to higher order basis functions has appeared to be difficult. On the other hand, there are a number of problems which profit from higher order approximations. Among these is the problem of prizing American options, formulated as a parabolic boundary value problem involving Black–Scholes' equation with a free boundary which indicates when the option is to be exerted. In addition to computing this free boundary, of particular importance are pointwise derivatives of the solution, the value of the stock option, up to order two, the so–called Greek letters, to high precision.

In this talk based on [HK], a monotone multigrid method has been presented for discretizations in terms of B–splines of arbitrary order to solve variational inequalities of the above form. In order to maintain monotonicity (upper bound) and quasi–optimality (lower bound) of the coarse grid corrections for the equivalent linear complementary problem, we have proposed an optimized coarse grid correction (OCGC) algorithm which is based on B–spline evaluation coefficients. The OCGC scheme has been formulated by solving a linear constrained optimization problem. For the solution process, we have exploited essential properties of B–Splines, namely, positivity of B–Splines and total positivity of their refinement matrices. We have proved that the OCGC algorithm is of optimal complexity of the degrees of freedom of the coarse grid. Moreover, the resulting monotone multigrid method has been shown to be of optimal multigrid complexity and converges with an optimal rate independent of the discretization.

Finally, the method has been applied to the valuation of American options. It has been shown that a discretization based on B–Splines of order four in particular meets the requirement of computing the pointwise error of the second derivative of the value of the stock option up to high precision. Further results for the prizing of options applying monotone multigrid methods can be found in [H].

References

[H] M. Holtz, *On the computation of American option price sensitivities via multigrid schemes*, Manuscript, August 2004.

[HK] M. Holtz, A. Kunoth, *Construction of B–spline based higher order monotone multigrid methods*, Manuscript, August 2004.

[Ko1] R. Kornhuber, *Monotone multigrid methods for elliptic variational inequalities I*, Numer. Math. 69, 167–184, 1994.

[Ko2] R. Kornhuber, *Adaptive Monotone Multigrid Methods for Nonlinear Variational Problems*, Teubner, 1997.

Adaptive Multi-scale Methods for Inverse Problems

Peter Maass

(joint work with Stephan Dahlke)

We are analyzing adaptive methods for solving operator equations $Af = g$ with noisy data $||g^\delta - g|| \leq \delta$, where $A : X \longrightarrow Y$ is supposed to be a linear or non-linear mapping between Hilbert spaces X and Y.

Classical adaptive methods use local refinements of the approximate solution based on some a–posteriori error estimators. Those error estimators should be local to allow suitable refinement strategies, and they usually require a boundedly invertable operator in order to pull back a residual estimate in the image space to the function space where the desired solution is living.

Typical inverse problems like applications in tomography or parameter identification problems for partial differential equations share two properties which make them - at least at a first glance - particular unsuitable for adaptive methods. First of all, they lead to ill-posed operator equations, i.e., the inverse operator is not continuous and it is not defined on all of Y, see [8, 7]. This is taken care of by applying regularization schemes like Tikhonov regularization with $T_\alpha = (A^*A + \alpha I)^{-1}A^*$. Moreover, local refinements of f lead to global changes of the data, i.e., no local error estimates are possible.

¿From a philosophical point of view, there are two players in this game: the operator A and the solution f. From this point of view, Tikhonov regularization only pays attention to the operator, no information about f enters the construction of T_α. As a consequence, the missing information has to be introduced via a rather expensive iteration scheme (e.g., α-descent methods for satisfying Morozov's discrepancy principle). This provides the playground for adaptive methods, which use information about the solution f at an earlier stage.

Probably the first systematic approach to adaptive methods for inverse problems were proposed in [10, 9]. These papers introduce mollifier methods, which do not aim at reconstructing f but rather functionals $\langle f, \varphi_x \rangle$, which can be tuned to give pseudo-local approximations.

More recently, different approaches concerning adaptive multi–scale methods have been published. In this context, adaptive methods can be used in different ways:

(1) presmoothing the noisy data g^δ prior to reconstruction;
(2) adaptive forward solvers with prescribed accuracy $||u_N - u|| \leq \epsilon$;
(3) regularization by adaptive discretization.

Presmoothing the data has been analyzed in various ways, e.g., by inverse Sobolev embedding [1, 11] or wavelet shrinkage [4]. The second approach is based on the analysis in [2, 3, 5], its application to inverse problems has been outlined in [6]. Obviously, the last problem is the most demanding one and the topic of future research.

In this talk, we will extend the first approach of Cohen et al. by analyzing two stage regularization schemes $T_{\alpha,\lambda} = R_\alpha S_\lambda$, where $S_\lambda : Y \longrightarrow Y$ is a pre-smoothing step (wavelet shringkage) followed by a regularization scheme R_α.

We can show, that optimal convergence rates can also be achieved by an incomplete pre-smoothing, i.e., we only smooth the data up to level below $range(A)$. Hence, we require a subsequent regularization scheme. The advantage of such a method as compared with classical methods without pre-smoothing is as follows: optimal convergence rates can be obtained with a substantially smaller regularization parameter, which avoids the typical oversmoothing of, e.g., Tikhonov regularization.

Test calculations with simulated SPECT data confirm the theoretical convergence rates.

References

[1] G. Bruckner, *On the stabilization of trigonometric collocation methods for a class of ill–posed first kind equations*, J. Inverse Ill-posed Problems **5** (1997), 117–127.

[2] A. Cohen, W. Dahmen, and R. DeVore, *Adaptive wavelet methods for elliptic operator equations – Convergence rates*, Math. Comp. **70** (2001), 22–75.

[3] A. Cohen, W. Dahmen, and R. DeVore, *Adaptive wavelet methods II: Beyond the elliptic case*, Found. Comput. Math. **2(3)** (2002), 1230–1262.

[4] A. Cohen, M. Hoffmann, and M. Reiss, *Adaptive wavelet Galerkin methods for linear inverse problems*, to appear in: SIAM J. Numer. Anal.

[5] S. Dahlke, W. Dahmen, R. Hochmuth, and R. Schneider, *Stable multiscale bases and local error estimation for elliptic problems,* Appl. Numer. Math. **23(1)** (1997), 21–48.

[6] S. Dahlke and P. Maaß, *An outline of adaptive wavelet Galerkin methods for Tikhonov regularization of inverse parabolic problems*, in: Recent Development in Theory and Numerics, Proceedings of the International Conference on Inverse Problems, World Scientific Publisher (2002), 56–66.

[7] H. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer Acad. Publ., Dordrecht, (1996).

[8] A.K. Louis, *Inverse und schlecht gestellte Probleme*, Teubner, Stuttgart, (1989).

[9] A.K. Louis, *A unified approach to regularization methods for linear ill-posed problems*, Inverse Prob. **15** (1999), 489–498.

[10] A.K. Louis and P. Maass, *A mollifier method for linear operator equations of the first kind,* Inverse Probl. **6** (1990), 427–440.

[11] R. Ramlau and G. Teschke, *Regularization of Sobolev embedding operators and applications part I: Fourier and wavelet based methods*, to appear in: Sampl. Theory Signal Image Process. **3(2)** (2004).

Multiresolution Finite Volume Schemes for Conservation Laws

Siegfried Müller

The solution of hyperbolic conservation laws typically exhibits locally steep gradients and large regions where it is smooth. To account for the highly nonuniform spatial behavior, we need numerical schemes that adequately resolve the different scales, i.e., use a high resolution only near sharp transition regions and singularities but a moderate resolution in regions with smooth, slowly varying behavior of the solution.

For this purpose, numerical schemes have been discussed or are under current investigation that aim at adapting the *spatial* grid to the local behavior of the flow field. A standard strategy is based on *local indicators* which are typically related to gradients in the flow field or local residuals. Although these concepts turn out to be very efficient in practice they offer no reliable error control. For this purpose, *a posteriori error estimates* have been derived which aim at equilibrating local errors. So far, this type of error estimators is only available for scalar problems. Another approach employs (dual-)weighted residual error indicators for adaptive grid refinement. This leads to meshes that are tailored to the cost-efficient computation of the quantity of interest, e.g. drag coefficient or lift coefficient. The approach gives very sparse grids but currently is not efficient for 3D instationary problems.

In the early 90's Harten [Ha] proposed to use *multiresolution techniques* in the context of finite volume schemes applied to hyperbolic conservation laws. He employed these techniques to transform the arrays of cell averages associated with any given finite volume discretization into a different format that reveals insight into the local behavior of the solution. The cell averages on a given highest level of resolution (reference mesh) are represented as cell averages on some coarse level where the fine scale information is encoded in arrays of *detail coefficients* of ascending resolution. By means of the multiresolution analysis the flux evaluation is controlled, i.e., cheap finite differences are employed in regions where the solution is smooth. By this strategy the computation is accelerated and the solution remains within the same accuracy as the *reference scheme*, i.e., the scheme on the finest computational mesh that uses the expensive flux evaluation throughout the entire domain. However, since one works still on a uniform mesh the computational complexity stays proportional to the number of cells on the finest grid. So far, Harten's concept has been successfully implemented for two-dimensional Cartesian meshes [BH1, BH2, CD, CDM, RCD], curvilinear meshes [DGM] and unstructured meshes [A, BOLR, CDKP].

Parallel to Harten's original idea a modified approach has been developed by Müller et al. [GM, CKMP, M] that is aiming at reducing the computational costs with regard to both computational time *and* memory requirements but still maintaining the accuracy of the reference scheme. In contrast to this, the *detail coefficients* will be used here to create *locally refined meshes* on which the discretization is performed. Of course, the crux in this context is to arrange this procedure in

such a way that at no stage of the computation there is ever made use of the fully refined uniform mesh. A central mathematical problem is then to show that the solution on the adapted mesh is of the same accuracy as the solution on the reference mesh. By now the fully adaptive multiresolution concept has been applied by several groups with great success to different real world applications, e.g., 2D/3D–steady state computations of compressible fluid flow around air wings modeled by the Euler and Navier–Stokes equations, respectively, as well as fluid-structure interactions on block–structured curvilinear grid patches [BGH+, BLM], non–stationary wave interactions in two-phase fluids on 2D Cartesian grids for Euler equations [M, ABMV, ABM, DMV], backward–facing step on 2D triangulations [CDM] and simulation of a flame ball modeled by reaction–diffusion equations on 3D Cartesian grids [RSTB].

So far a short-coming of this approach has been the lack of *temporal adaptivity*, i.e., all cell averages are evolved in time by the same time step size τ. For reasons of stability we are therefore obliged to choose τ such that the CFL condition for the cells on the *finest* mesh is satisfied. However, for cells corresponding to a coarser discretization we may use a larger time step to meet the local CFL condition. Therefore it is natural to use locally varying time stepping.

First results on *local time stepping* have been published by Osher and Sanders [OS] for one-dimensional scalar conservation laws. Here the space discretization is fixed but non-uniform. Each element is evolved in time either by an entire time step or a fixed number of smaller time steps. About the same time, Berger and Oliger [BO] proposed the by now classical Adaptive Mesh Refinement (AMR) technique. Here refined grids are laid over regions of the coarse mesh. In particular, the grids need not to be nested but can have a different orientation than the coarse grid. This allows for a local alignment of the grid with anisotropic effects such as shocks. Each refinement level is propagated with its own time step. Information is passed between the grids using injection and interpolation techniques.

Recently, a local time stepping strategy has been incorporated to the concept of fully adaptive multiresolution schemes, cf. [MS]. Here ideas similar to the predictor-corrector scheme [OS] are used. This has to be adjusted to the requirement that the resulting scheme provides a spatial accuracy that is comparable to the spatial accuracy of the reference mesh. Therefore we have to address properly the issues of (i) a conservation-preserving flux evaluation at interface points, (ii) the computation of appropriate prediction values on coarser levels, (iii) the synchronization of the time evolution and (iv) the local grid refinement on the intermediate time levels to track appropriately the movement of discontinuities. Note that the resulting concept can be applied to multidimensional conservation laws discretized by an explicit or implicit reference finite volume scheme,

The outline of the talk will be as follows. We start with a summary of the standard fully adaptive multiresolution concept recalling its core ingredients, namely, the multiresolution analysis and the local grid adaptation. Then we outline the concept for incorporating locally varying time stepping. Here we first consider an

explicit time integration. These ideas are then extended to an implicit time integration. Numerical results verify the efficiency and the accuracy of our method. We conclude with some remarks on open questions and future work.

References

[A] R. Abgrall, Multiresolution analysis on unstructured meshes: Applications to CFD, in: Chetverushkin et al. (eds.), Experimentation, modeling and computation in flow, turbulence and combustion, John Wiley & Sons, 1997.

[ABMV] S. Andreae, J. Ballmann, S. Müller, A. Voß, Dynamics of collapsing bubbles near walls, in: T.Y. Hou, E. Tadmor (eds.), Hyperbolic Problems: Theory, Numerics, Applications, Springer Verlag, 2003, 265–272.

[ABM] S. Andreae, J. Ballmann, S. Müller, Wave processes at interfaces, in: Lecture Notes on Computational Science and Engineering, Springer Verlag, 2004, to appear.

[BOLR] B.L. Bihari, D.K. Ota, Z. Liu, S.V. Ramakrishnan, The multiresolution method on general unstructured meshes, AIAA paper, AIAA 2001-2553, 2001.

[B1] M.J. Berger, Stability of interfaces with mesh refinement, Math. Comp., 45, 1985, 301-318.

[B2] M.J. Berger, On conservation at grid interfaces, SIAM J. Numer. Anal., 24, 1987, 967-984.

[BO] M.J. Berger, J. Oliger, Adaptive mesh refinement for hyperbolic partial differential equations, J. Comp. Physics, 53, 1984, 484-512.

[BH1] B. Bihari, A. Harten, Application of generalized wavelets: An adaptive multiresolution scheme, J. Comp. Appl. Math, 61, 1995, 275-321.

[BH2] B. Bihari, A. Harten, Multiresolution schemes for the numerical solution of 2–D conservation laws I, SIAM J. Sci. Comput., 18 (2), 1997, 315-354.

[BGH+] F. Bramkamp, B. Gottschlich–Müller, M. Hesse, Ph. Lamby, S. Müller, J. Ballmann, K.-H. Brakhage, W. Dahmen, *H*-adaptive Multiscale Schemes for the Compressible Navier–Stokes Equations — Polyhedral Discretization, Data Compression and Mesh Generation, in: J. Ballmann (ed.), Flow Modulation and Fluid-Structure-Interaction at Airplane Wings, Numerical Notes on Fluid Mechanics, Springer, 2003, 125-204.

[BLM] F. Bramkamp, Ph. Lamby, S. Müller, An adaptive multiscale finite volume solver for unsteady and steady flow computations, J. Comp. Physics, 2004, to appear.

[CD] G. Chiavassa, R. Donat, Point value multiresolution for 2D compressible flows, SIAM J. Sci. Comput., 23 (3), 2001, 805-823.

[CDM] G. Chiavassa, R. Donat, A. Marquina, Fine–Mesh Numerical Simulations for 2D Riemann Problems with a Multilevel Scheme, in: G. Warnecke, H. Freistühler (eds.), Hyperbolic Problems: Theory, Numerics, Applications, Birkhäuser, 2002, 247-256.

[CDKP] A. Cohen, N. Dyn, S.M. Kaber, M. Postel, Multiresolution finite volume schemes on triangles, J. Comp. Physics, 161, 2000, 264-286.

[CKMP] A. Cohen, S.M. Kaber, S. Müller, M. Postel, Fully adaptive multiresolution finite volume schemes for conservation laws, Math. Comp, 72 (241), 2003, 183-225.

[CDM] A. Cohen, S.M. Kaber, M. Postel, Multiresolution Analysis on Triangles: Application to Gas Dynamics, in: G. Warnecke, H. Freistühler (eds.), Hyperbolic Problems: Theory, Numerics, Applications, Birkhäuser, 2002, 257-266.

[DGM] W. Dahmen, B. Gottschlich–Müller, S. Müller, Multiresolution schemes for conservation laws, Num. Math, 88 (3), 2000, 399-443.

[DMV] W. Dahmen, S. Müller, A. Voß, Riemann Problem for the Euler equations with nonconvex equation of state including phase transitions, in: Lecture Notes on Computational Science and Engineering, Springer Verlag, 2004, to appear.

[GM] B. Gottschlich–Müller, S. Müller, Adaptive Finite Volume Schemes for Conservation Laws based on Local Multiresolution Techniques, in: M. Fey, R. Jeltsch (eds.), Hyperbolic Problems: Theory, Numerics, Applications, Birkhäuser, 1999, 385-394.

[Ha] A. Harten, Multiresolution algorithms for the numerical solution of hyperbolic conservation laws, Comm. Pure Appl. Math., 48 (12), 1995, 1305-1342.
[M] S. Müller, Adaptive Multiscale Schemes for Conservation Laws, in: Lecture Notes on Computational Science and Engineering, vol. 27, Springer–Verlag, Heidelberg, 2001.
[MS] S. Müller, Y. Stiriba, Fully adaptive multiscale schemes for conservation laws employing locally varying time stepping, IGPM-Report 238, RWTH Aachen, 2004.
[OS] S. Osher, R. Sanders, Numerical approximations to nonlinear conservation laws with locally varying time and space grids, Math. Comp, 41, 1983, 321-336.
[RCD] A. Rault, G. Chiavassa, R. Donat, Shock-vortex interactions at high Mach numbers, J. Scientific Computing, 19, 2003, 347-371.
[RSTB] O. Roussel K. Schneider, A. Tsigulin, H. Bockhorn, A conservative fully adaptive multiresolution algorithm for parabolic PDEs, J. Comp. Physics, 188 (2), 2003, 493-523.

Smoothness Spaces and Nonlinear Approximation on the Sphere

Pencho Petrushev

A new characterization of the Besov and Triebel-Lizorkin spaces on the n-dimensional sphere is given. Extremely well localized semi-orthogonal elements (called "needlets") are constructed on the sphere. It is shown that the Besov and Triebel-Lizorkin on the sphere can be decomposed via needlets for the full range of indices. The emphasis is placed on the Besov spaces involved in the theory of nonlinear approximation on the sphere. Some basic results of the theory of nonlinear n-term approximation from needlets and its applications are presented as well.

Atomic Decomposition on Warped Wavelets and Muckenhoupt Weights

Dominique Picard
(joint work with Gerard Kerkyacharian)

We consider the problem of replacing a standard (compactly supported) wavelet expansion of a function f :

$$f(x) = \sum_{I=(j,k)} \beta_I \psi_I(x), \tag{1}$$

by a decomposition on atoms of the form $\{\psi_{jk}(G), j,\ k\}$ where the standard wavelet basis has been warped by a function G.

$$f(x) = \sum_{I=(j,k)} \beta_I \psi_I(G(x)). \tag{2}$$

This kind of atoms naturally appear when the function is observed in a non regular design (in regression or denoising, see Kerkyacharian and Picard [KP04]), or to catch some geometric features (see LePennec and Mallat ([LM03]), or to handle local stationarity (see Clerc and Mallat ([CM03]).

The aim of this talk is to investigate the properties of this new basis, and especially to prove that if the warping function has a property of Muckenhoupt type, this new basis has a behavior quite similar to a regular wavelet basis.

A closely connected problem is the following : Generally spaces of regularity (Sobolev or Besov spaces) are defined with respect to the Lebesgue measure. However in some cases, it can be more natural to consider other measures and especially measures of the form $\omega(x)dx$ where ω is a weighting of the space.

In this case, we are interested in considering the standard expansion (1), but eventually to measure its approximation performances in the spaces $\mathbb{L}_p(\omega(x)dx)$. Using a simple change of variables (valid under mild conditions on G), it is obvious that the approximation properties of f in terms of the warped atoms correspond to the approximation properties of $f(G^{-1})$ in terms of the standard wavelet bases, but measured with the weight $\omega(x) = \frac{1}{g(G^{-1}(x))}$.

1. What properties for atoms ?

Many properties of the possible atoms that are shared by wavelet bases can be explored. However, we shall here concentrate on two special properties which are especially important in the treatment of most statistical applications.

1.1. **Shrinkage (or unconditional) property.** There exists an absolute constant K such that if $|\theta_i| \leq |\theta'_i|$ for all i, then

$$\| \sum_i \theta_i e_i \|_p \leq K \| \sum_i \theta'_i e_i \|_p. \tag{3}$$

This property means in particular, that by thresholding or shrinking the coefficients we do not risk exploding the norm, and has many more important properties.

1.2. **p-democratie property or Temlyakov-property.** There exist c_p and C_p such that for any finite set of integers F we have :

$$c_p \int \sum_{i \in F} |e_i|^p \leq \int (\sum_{i \in F} |e_i|^2)^{\frac{p}{2}} \leq C_p \int \sum_{i \in F} |e_i|^p. \tag{4}$$

This property is fundamental to express in a simple way the spaces of approximation by wavelet thresholding.

Both of the properties stated above are true for compactly supported wavelets. What can be said more generally for warped bases, if we assume good properties on G ?

2. Muckhenhoupt weights

The concept of Muckenhoupt weight has been introduced in [M72] (see also [GR85] and[CF74]) and widely used afterwards in the context of Calderon-Zygmund theory.

Definition. (Muckenhoupt weights) *For $1 < p < \infty,\ 1/p + 1/q = 1$, a measurable function $\omega \geq 0$ belongs to the Muckenhoupt class A_p if there exists $0 < C < \infty$ such that for any interval I included in $\mathbb{R}$,*

$$\left(\frac{1}{|I|}\int_I \omega(x)dx\right)^{1/p}\left(\frac{1}{|I|}\int_I \omega(x)^{-\frac{q}{p}}dx\right)^{1/q} \leq C$$

For $p = 1$, $\omega \geq 0$ belongs to the Muckenhoupt A_1 class if there exists $0 < C < \infty$ such that,

$$\omega^*(x) \leq C\omega(x)\ a.e.$$

where $\omega^*(x)$ is the Hardy-Littlewood maximal function.
For $p = \infty$, we set

$$A_\infty = \cup_{p\geq 1} A_p.$$

3. Results

3.1. **Warped bases. Theorem.** *Let $1 < p < \infty,\ \omega \in A_p,$ and $\psi_{j,k}$ be a compactly supported wavelet. Let T and S be two real measurable functions defined on $\mathbb{R}$ such that*

$$S(T(x)) = x,\ a.e.;\ T(S(x)) = x,\ a.e.$$

$$\forall h \geq 0,\ \text{measurable function},\ \int_{\mathbb{R}} h(T(x))dx = \int_{\mathbb{R}} h(y)\omega(y)dy \tag{5}$$

the family $\{\psi_{jk}(T(.)),\ j \geq -1,\ k \in \mathbb{Z}\}$ satisfies the properties of shrinkage and p-democratie.

3.2. **Weighted Besov spaces.** One of the major advantages of regular Besov spaces is that they can be expressed in terms of wavelet coefficients :

Under standard oscillating conditions on the wavelet ψ, we have,

$$f \in B_{s,p,q} \iff [\sum_{j\geq -1}\{2^{js}2^{j/2}(\sum_{k\in\mathbb{Z}}|\beta_{j,k}|^p 2^{-j})^{1/p}\}^q]^{1/q} < \infty.$$

If $B_{s,p,q}(\omega)$ denotes the Besov space where $\mathbb{L}_p(dx)$ is replaced by $\mathbb{L}_p(\omega(x)dx)$, we show in the following proposition that under conditions on ω, the direct sense of the implication above is still true, if in the sum in k, 2^{-j} is replaced by $\omega([\frac{k}{2^j}, \frac{k+1}{2^j}])$.

Proposition. *For $1 \leq p \leq \infty$, let us suppose that ω is in A_p, and let us put for every interval $I \subset \mathbb{R}$*

$$\omega(I) = \int_I \omega(x)dx$$

Then, if ψ is a compactly supported wavelet, such that

$$\int \psi(x)x^k dx = 0,\ k = 0, \ldots, N-1$$

then for

$$f = \sum_{j,k}\beta_{jk}\psi_{j,k}, \quad I_{j,k} = [\frac{k}{2^j}, \frac{k+1}{2^j}]$$

$$f \in B_{s,p,q}(\omega) \Longrightarrow [\sum_j \{2^{js} 2^{j/2} (\sum_{k\in\mathbb{Z}} |\beta_{j,k}|^p \omega(I_{j,k}))^{1/p}\}^q]^{1/q} < \infty$$

with the usual modification if $q = \infty$.

References

[CM03] Clerc, M., Mallat, S. (2001) Estimating deformations of stationary processes, Annals of Statistics (2003) vol 31, N. 6, P1772-1822.

[KP03] G. Kerkyacharian, D. Picard, Non linear approximation and Muckhenhoupt weights, submitted for publication. www.proba.jussieu.fr/mathdoc/preprints

[KP04] G. Kerkyacharian, D. Picard, Regression in random design and warped wavelets, submitted for publication, 2002. www.proba.jussieu.fr/mathdoc/preprints/

[LM03] Le Pennec, E., Mallat, S. (2003) Sparse Geometric Image Representation with Bandelets. http://www.cmap.polytechnique.fr/ mallat/biblio.html

[M72] Muckenhoupt, B. (1972) Weighted norm inequalitiesfor the hardy maximal function, Trans. Amer. Math. Hencec. 165, 207-226.

[GR85] Garcia-Cuerva, J. and Rubio de Francia J.L. (1985) *Weighted norm inequalities and related topics* North-Holland math. stud. 116.

[CF74] Coifman, R., and Fefferman, C. (1974) Weighted norm inequality for maximal functions and singular integrals. Studia Math. 109 (3) pp 241-250.

Combining Multiresolution Technique with a Semi–implicit Scheme for Two–phase Flows

Marie Postel

(joint work with F. Coquel, Q. H. Tran and N. Poussineau)

In the context of multiphase flows in oil pipelines, we study the following one dimensional model problem

$$\left\{\begin{array}{lcll} \partial_t(\rho) & + & \partial_x(\rho v) & = 0, \\ \partial_t(\rho Y) & + & \partial_x(\rho Y v) & = 0, \\ \partial_t(\rho v) & + & \partial_x(\rho v^2 + P) & = S(\rho, \rho v), \end{array}\right.$$

where the unknowns are the mean values of the physical unknowns over the cross section of a pipeline. We are solving this system for the density ρ, the mean velocity v, which is here assumed to be the same for both oil and gas, and the gas mass fraction Y. The pressure P is given by a thermodynamical equilibrium law as a function of the density and the gas fraction.

We don't solve this system numerically but consider instead a relaxed version

$$\left\{\begin{array}{lcll} \partial_t(\rho) & + & \partial_x(\rho v) & = 0, \\ \partial_t(\rho Y) & + & \partial_x(\rho Y v) & = 0, \\ \partial_t(\rho v) & + & \partial_x(\rho v^2 + \Pi) & = S(\rho, \rho v), \\ \partial_t(\rho \Pi) & + & \partial_x(\rho \Pi v + a^2 v) & = k\rho(P(\rho, \rho Y) - \Pi), \end{array}\right.$$

where k denotes the relaxation parameter. Here, $a > 0$ stands for a frozen (Lagrangian) sound speed to be prescribed according to some sub-characteristic conditions (see [1]). At each time step, we solve the system with $k = 0$ and then force the pressure Π to the equilibrium using the pressure law $\Pi = P(\rho, , \rho Y)$.

The relaxation system is hyperbolic, all the fields being linearly degenerate. It comes with two large eigenvalues corresponding to acoustic waves and a much smaller double eigenvalue, v, which is actually the transport speed. The wave of interest is the transport one. The other fast acoustic waves are not interesting but impose a very restrictive CFL condition if an explicit scheme is used. In a semi implicit finite volume scheme like the one developped in [3, 1], the fast waves are handled with a linearized implicit formulation and the slow wave remains explicitely solved. The CFL condition, governed by the explicit wave speed is then more reasonable. On the other hand, a tridiagonal block matrix whose terms involve the solution of the local Riemann problems, has to be constructed and inverted at every time step. The semi-implicit scheme is therefore still quite costly, and adaptive mesh refinement seems indicated to improve its performances since the interesting quantity, the gas mass fraction in the mixture for instance, presents localized singularities that are carried by the transport wave.

We combine this method with a multiscale analysis of the vector solution U in its conservative variables similar to the one developped in [2] for explicit schemes, which we will now present. The fine uniform grid of the underlying finite volume scheme is taken as the finest level of a hierarchy of nested dyadic grids. The finite volume solution can be represented by its mean values $U_k = (u_{j,k})$ on any level k, with 0 beeing the coarsest level. Starting from the representation on the finest level K, we can iterate a standard projection operator P_k^{k+1} to go from fine level $k+1$ to the immediately coarser one:

$$U_k = P_k^{k+1} U_{k+1}, \quad u_{j,k-1} = \frac{1}{2}(u_{2j,k} + u_{2j+1,k}),$$

and a centered reconstruction operator P_k^{k-1} based on a mean values preserving quadratic polynomial to recover mean values on the level k from the mean values on the immediately coarser level:

$$\hat{U}_k = P_k^{k-1} U_k, \quad \left\{ \begin{array}{lcl} \hat{u}_{2j,k} & = & u_{j,k-1} - (u_{j+1,k-1} - u_{j-1,k-1})/8, \\ \hat{u}_{2j+1,k} & = & u_{j,k-1} + (u_{j+1,k-1} - u_{j-1,k-1})/8. \end{array} \right.$$

There is a one to one mapping between the representation of the solution by its mean values on the finest level and its mean values on the coarsest level plus all the details defined at level k by

$$D_k, \quad d_{j,k} = u_{2j+1,k} - \hat{u}_{2j+1,k}.$$

At each time step n , the solution is encoded, its details are tested against a level dependent threshold $\varepsilon_k = 2^k \varepsilon$ and a tree of indices T_n^ε corresponding to significant details is computed. This tree is then enlarged in order to ensure that it captures all significant details also at next time step. The solution is then locally decoded according to this tree, meaning that the finest grid representation will be used only in the vicinity of singularities, while mean values on coarser cells will be sufficient in regular regions. This provides the representation of the solution by its mean value on the time dependent adaptive grid.

The semi-implicit scheme is then applied to this non-uniformly sampled solution. At this stage, the accuracy of the uniform scheme on the finest level is maintained because although the data can be sampled on much coarser cells where the solution is regular, the coefficients of the matrix and the residual are computed using data on the finest grid. This can be obtained easily in the one dimensional case by local reconstruction near the edges of the adaptive grid cells.

An important issue is the thresholding and prediction strategies used to design the adaptive grid, if we want to retain a good behavior of the adaptive scheme solution with respect to the uniform scheme one. First of all we checked numerically that although we are interested only in the gas mass fraction behaviour it is important to track the singularities on all the solution components.

Concerning the prediction strategy, we use here the heuristic rules introduced by Harten [4], which rely on the hyperbolicity of the equation: for an explicit scheme with a given CFL condition lower than 1, the singularities will not move more than one cell away in one time step. Therefore, knowing the tree T_n^ε of relevant details at time n, the predicted tree at time $\hat{T}_{n+1}^\varepsilon$ can be designed by adding to T_n^ε all its immediate neighbours, plus the subdivisions of those details that are greater than a given multiple of the threshold parameter. In our case, this prediction is only valid for the details that move along with the transport wave, because those are treated explicitely with a CFL less than 1. It might become insufficient for the details moving at the speed of the acoustics wave for which the CFL condition can be as high as 20. Fortunately these details are treated implicitely and are therefore smoothed out by the diffusion inherent to the implicit scheme. This, plus the gradedness that we impose to the tree in order to optimize the complexity of the decoding-encoding algorithm, ensures that the singularities that are carried away by the acoustic waves are still well captured by the tree.

We experiment on test cases including boundary conditions and source term treatment that the efficiency in terms of computing time and memory requirements is improved by a factor varying between 3 and 5 while retaining the accuracy performances. In particular the L_1 error between the adaptive scheme and the uniform scheme solutions depends linearly on the thresholding parameters, which is the theoretical behavior obtained for a non linear scalar equation and an explicit scheme in [2].

Further developments will include analysis of the prediction strategy for the semi-implicit scheme, better treatment of the boundary conditions and source terms and implementation of local time stepping.

References

[1] M. Baudin, F. Coquel, and Q.-H. Tran. A semi-implicit relaxation scheme for modeling two-phase flow in a pipeline. *(Submitted)*, 2004.

[2] A. Cohen, S.M. Kaber, S. Müller, and M. Postel. Fully adaptive multiresolution finite volume schemes for conservation laws. *Math. Comp.*, 72(241):183–225 (electronic), 2003.

[3] I. Faille and E. Heintzé. A rough finite volume scheme for modeling two phase flow in a pipeline. *Computers and Fluids*, 28:213–241, 1999.

[4] A. Harten. Multiresolution algorithms for the numerical solutions of hyperbolic conservation laws. *Comm. on Pure and Appl. Math.*, 48:1305–1342, 1995.

Fast Iterative Solvers for Discrete Stokes Equations

Arnold Reusken
(joint work with Jörg Peters and Volker Reichelt)

We consider a class of Stokes equations on a bounded connected polyhedral Lipschitz domain Ω in d-dimensional Euclidean space. We use the notation $\mathbf{V} := H_0^1(\Omega)^d$ for the velocity space and $M = L_0^2(\Omega) := \{\, p \in L^2(\Omega) \mid \int_\Omega p(x)\,dx = 0 \,\}$ for the pressure space. The variational problem is as follows: given $\mathbf{f} \in L^2(\Omega)^d$ find $\{\mathbf{u}, p\} \in \mathbf{V} \times M$ such that

$$\left\{ \begin{array}{lll} (\nabla \mathbf{u}, \nabla \mathbf{v}) + \xi(\mathbf{u}, \mathbf{v}) - (\operatorname{div} \mathbf{v}, p) & = (\mathbf{f}, \mathbf{v}) & \text{for } \mathbf{v} \in \mathbf{V}, \\ (\operatorname{div} \mathbf{u}, q) & = 0 & \text{for } q \in M\ . \end{array} \right. \tag{1}$$

with a constant $\xi \geq 0$. The L^2 scalar product and associated norm are denoted by $(\cdot,\cdot)$, $\|\cdot\|$, respectively. The zero order term $\xi(\mathbf{u}, \mathbf{v})$ is included in view of implicit time integration methods applied to instationary Stokes equations.

For the discretization of this problem we use a pair of conforming LBB stable finite element spaces. This results in a saddle point problem of the form

$$\begin{pmatrix} \mathrm{A} & \mathrm{B}^T \\ \mathrm{B} & 0 \end{pmatrix} \begin{pmatrix} \mathrm{x} \\ \mathrm{y} \end{pmatrix} = \begin{pmatrix} \mathrm{f} \\ 0 \end{pmatrix} \tag{2}$$

Many different iterative methods for solving this discrete problem are known. One possible approach is to apply multigrid techniques to the whole coupled system in (2). Most other approaches are based on the prominent classical Uzawa method. This Uzawa method requires that $\mathrm{A}^{-1}\mathrm{x}$ can be computed exactly. In many variants of this method, which are often called *inexact* Uzawa methods, one tries to avoid the exact inversion by using an inner iterative method, for example, a Jacobi-like iteration or a multigrid method. A third possible strategy is to use (variants of) the preconditioned MINRES method. Finally we mention the approach presented in [2]. There the indefinite problem (2) is reformulated as a symmetric positive definite problem. For most of these different approaches theoretical convergence analyses are known. In [4] the performance of a few of these methods is compared by means of systematic numerical experiments for a stationary 2D Stokes problem.

In this talk we consider three representative methods from the large class of iterative Stokes solvers, namely the preconditioned CG method from [2] (denoted by BPCG), the preconditioned MINRES method from [6, 7, 8] (denoted by PMINRES) and the method from [1] (denoted by MGUZAWA). The topics discussed in this talk are the following:

- For these three methods we discuss costs per iteration, known theoretical convergence results and some implementation issues. This makes it possible to make a fair comparison of these methods.

- For the MGUZAWA method we present a convergence analysis. This analysis is much simpler than the analyses presented in [1, 9]. The result that we obtain is different from the ones in [1, 9] and gives a better explanation of the observation that if one uses a very good preconditioner for A (like multigrid) then even with a very low accuracy in the inner iteration the MGUZAWA method converges.
- We present a comparative study of the performance of the three methods. For this we consider a Stokes problem as in (1) in 3D. We treat both the stationary ($\xi = 0$) and instationary ($\xi > 0$) case. For the discretization we apply the popular Hood-Taylor $P_2 - P_1$ finite element pair. As a preconditioner for A a standard multigrid method is used. For the Schur complement preconditioner we use the mass matrix for the stationary case and a more sophisticated preconditioner analyzed in [3] for the instationary case.

The main results are a detailed comparative study of the three fast iterative solvers for a 3D Stokes problem and a new convergence analysis for the MGUZAWA method. A more detailed treatment can be found in [5].

References

[1] Bank, R.E., Welfert, B.D., and Yserentant, H.: A class of iterative methods for solving saddle point problems, Numer. Math., **56**, 1990, 645–666.

[2] Bramble, J.H., Pasciak, J.E.: A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems. Math. Comp., **50**, 1988, 1–17.

[3] Bramble, J.H., and Pasciak, J.E.: Iterative techniques for time dependent Stokes problems, Comput. Math. Appl. **33**, 1997, 13–30

[4] Elman, H.C.: Multigrid and Krylov subspace methods for the discrete Stokes equations. Int. J. Numer. Meth. Fluids, **227**, 1996, 755–770.

[5] Peters, J., Reichelt, V. and Reusken, A.: Fast iterative solvers for discrete Stokes equations. IGPM report **241**, RWTH-Achen, 2004.

[6] Rusten, T. and Winther, R.: A preconditioned iterative method for saddlepoint problems. SIAM J. Matrix Anal. Appl., **13**, 1992, 887–904.

[7] Silvester, D. Wathen, A.: Fast iterative solution of stabilised stokes systems. Part II: using general block preconditioners. SIAM J. Numer. Anal., **31**, 1994, 1352–1367.

[8] Wathen, A., Silvester, D.: Fast iterative solution of stabilised stokes systems. Part I: using simple diagonal preconditioners. SIAM J. Numer. Anal., **30**, 1993, 630–649.

[9] Zulehner, W.: Analysis of iterative methods for saddle point problems: A unified approach, Math. Comp., **71**, 2001, 479–505.

An Adaptive Multiresolution Method for Reaction–Diffusion Equations: Applications to Flame Instabilities

Kai Schneider

(joint work with Olivier Roussel)

We present an adaptive multiresolution method for solving parabolic PDEs in Cartesian geometry [2]. The numerical scheme is based on a finite volume discretization coupled with a discrete multiresolution analysis to adapt the grid in

physical space dynamically to track the evolution of the solution in scale and space. The time discretization is done by an explicit third order Runge–Kutta scheme, in space we use second order schemes, either of ENO type for the convective terms or centered schemes. The fluxes are evaluated on the adaptive grid where conservativity is ensured. The implementation is based on a graded tree data structure, which improves both CPU and memory performances, as no fine gridding is required in regions where the solution is smooth. A recursive procedure is used to address each element of the tree. Although this concept is slightly more complex, i.e. an $O(N \log N)$ complexity instead of $O(N)$ (where N denotes the number of active grid points), this choice enables us to avoid hash–tables, which require very large arrays and therefore much memory which may be prohibitive for large scale 3D computations.

To illustrate the features and the efficiency of the method, we compute different test problemens in one, two and three space dimensions, like convection–diffusion, viscous Burgers and reaction–diffusion equations [2].

Then we present several applications to compute thermo–diffusive flame instabilities. First, we show that pulsating planar flames can be efficiently computed even for very large activation energies [5]. Depending on the Lewis and Zeldovich numbers we observe stable or pulsating flames, the latter being either damped, periodic or aperiodic in time. A bifurcation diagram in the Lewis–Zeldovich plane is computed and the results are compared with previous FEM computations and theoretical predictions. For Lewis numbers larger than 6 we find that the stability limit is again increasing towards larger Zeldovich numbers and not monotonically decreasing as predicted by the asymptotic theory. A study of the flame velocities for different Zeldovich numbers shows that the amplitude of the pulsations strongly varies with the Lewis number. A Fourier analysis yields information on their frequency.

Secondly, we study the interaction of spherical flame structures with adiabatic walls [4]. We show that the Lewis number determines the behaviour of the flame-wall interaction. When the flame is approaching the wall we observe for Lewis numbers smaller than unity that the reaction rate is decreased, for unitary Lewis number the reaction rate neither increases nor decreases and for Lewis numbers larger than unity the reaction rate increases. Due to tangential diffusion the flame front curvature is also modified, i.e. for small Lewis number the spherical flame contracts, for large Lewis number it spreads out, while for unitary Lewis number the flame front remains perpendicular to the wall. The observed phenomena present similarities with capillarity effects in fluid mechanics when a droplet hits a wall.

Finally, we compute several flame ball-vortex interactions and study the role played by the fluid flow on the evolution of the flame ball [3]. We observe the roll-up of the flame ball around the vortex into a snail-like structure. We also put into evidence the flammability limit of the flame ball in function of both vortex and radiation intensities.

Current work is dealing with the parallel implementation of the algorithm on a PC cluster to perform large scale 3D computations. To reach this goal, the data structure is organized into a "forest", i.e. an ensemble of trees, each one working on a different processor. Future work will focus on the extension of the scheme to reactive Navier–Stokes equations, in order to take into account hydrodynamic effects in combustion problems and to use the CVS (Coherent Vortex Simulation) approach [1] to model and to simulate turbulent reactive flows on adaptive grids. Time–adaptivity will also be investigated, since the currently used explcit time discretization limits the time step to the smallest time scale to be resolved in order to guarantee numerical stability.

References

[1] M. Farge, K. Schneider. Coherent Vortex Simulation (CVS), a semi–deterministic turbulence model using wavelets. *Flow, Turbulence and Combustion*, **66**(4), 393–426, 2001.

[2] O. Roussel, K. Schneider, A. Tsigulin, H. Bockhorn. A conservative fully adaptive multiresolution algorithm for parabolic PDEs. *J. Comput. Phys.*, **188**, 493–523, 2003.

[3] O. Roussel, K. Schneider. An adaptive multiresolution method for combustion problems: application to flame ball–vortex interaction. *Comp. Fluids*, in press 2004.

[4] O. Roussel, K. Schneider. Numerical study of spherical flame structures interacting with adiabatic walls using an adaptive multiresolution scheme. *Combust. Theory Modelling*, submitted, 2004.

[5] O. Roussel, K. Schneider. Adaptive numerical simulation of pulsating planar flames for large Lewis and Zeldovich ranges. *Comm. Nonlinear Sci. Numer. Simul.*, submitted, 2003.

Wavelets for Linear Scaling Computation in Electronic Structure Calculation

Reinhold Schneider

(joint work with H.J. Flad, W. Hackbusch and B. Khoromskij)

Numerical simulation of electronic structures derived from the multi-particle Schrödinger's equation play a prominent role in molecular physics, chemistry, solid physics and material science. In the present talk we consider mean field equations like Kohn-Sham equations based on density functional theory as well as Hartree-Fock models for molecular systems with N electrons. At least the Hartree-Fock model contains a nonlocal operator. Instead of computing the orbitals we compute the so called *density matrix* which represents the kernel function in $\mathbb{R}^6$ of the spectral projection operator. We consider systematic basis functions subordinated to different scales e.g. wavelets for the discretisation. We exploit the potential of wavelets for hyperbolic cross approximation in high dimensional spaces together with their ability for sparse representation of nonlocal operators to achieve *linear scaling* with respect to the number of particles and basis functions (in a suitable setting). Optionally a new concept of Kronecker-product approximation of operators introduced by Beylkin et al. [HFYGB] and Tyrtyshnikov [T] will be presented.

References

[FHKS] H.J. Flad, W. Hackbusch, B. Khoromskij, R. Schneider, Kronecker product approximation for electronic structure calculation, in preparation.

[G] S. Goedecker, $O(N)$ methods for electronic structure calculations, Rev. of Mod. Phys. **71**, 1085–1125 (1999).

[HFYGB] R.J. Harrison, G.I. Fann, T. Yanai, Z. Gan, G. Beylkin, Multiresolution quantum chemistry: basic theory and initial applications Preprint.

[LB] C. LeBris (eds), Handbook of Numerical Analysis, Vol.X: Computational Chemnistry, North Holland, 2003.

[T] E. Tyrtyshnikov, Kronecker-product approximations for some function-related matrices, Lin. Alg. Appl. 379, 423-437 (2004).

Adaptive Sparse Approximation of Elliptic Singularities

Christoph Schwab

(joint work with Pal Andrej Nitsche)

We consider the best N-term approximation of solutions of elliptic BVPs

$$Lu = f \ \text{ in } \ \Omega, \quad Bu = g \ \text{ on } \ \partial\Omega$$

where $\Omega \subset \mathbb{R}^d$, d=2,3 is a polyhedron, L is a second order elliptic differential operator in divergence form and f,g are piecewise smooth boundary data by piecewise polynomial functions of degree $p \geq 1$.

In the case where $\Omega = (0,1)^d$, we show that for differential operators L which admit compressible matrices of order p in the sense of Cohen, Dahmen, DeVore, u can be adaptively approximated using their approximation algorithm [CDD2000] on anisotropic tensor products of univariate spline wavelets of degree p at rate $N^{-p'}$ where $p' < p$, i.e. the curse of dimension can be overcome in these cases.

This result corresponds to numerical results obtained by Ch. Zenger and M. Griebel and Bungartz and their students using adaptive sparse grids in hierarchical FE-Bases.

The results are part of the doctoral dissertation of P.A. Nitsche.

References

[CDD2000] A. Cohen, W. Dahmen and R. DeVore: Adaptive Wavelet Methods for elliptic operator equations: Convergence Rates Math. Comp. **70**(2001), 27-75.

On the Compressibility of Operators in Wavelet Coordinates

Rob Stevenson

(joint work with Tsogtgerel Gantumur)

In [CDD02], Cohen, Dahmen, and DeVore proposed an adaptive wavelet algorithm for solving operator equations. Assuming that the operator defines a boundedly invertible mapping between a Hilbert space and its dual, and that a Riesz basis of wavelet type for this Hilbert space is available, the operator equation can

be transformed into an equivalent well-posed infinite matrix-vector system. This system is solved by an iterative method, where each application of the infinite stiffness matrix is replaced by an adaptive approximation. For a certain range of $s > 0$, determined by the compressibility of the stiffness matrix, i.e., by how well it can be approximated by computable sparse matrices, it was proven that if the errors of best linear combinations from the wavelet bases with N terms are $\mathcal{O}(N^{-s})$, then approximations yielded by the adaptive method with N terms also have errors of $\mathcal{O}(N^{-s})$, where their computation takes only $\mathcal{O}(N)$ operations. With the available estimates for both differential and singular integral operators, the compressibility of the stiffness matrix appears to limit the rate of convergence of the adaptive method, in the sense that for solutions that have a sufficiently high (Besov) regularity, these best N-term approximations converge with a better rate than can be shown for the approximations produced by the adaptive method.

In this talk, using modified sparse matrix approximations and suitable quadrature rules to approximate their entries, we present improved results concerning compressibility. From these results it will follow that for suitable wavelets as biorthogonal spline wavelets with sufficiently many vanishing moments, for the full range of s for which, under appropriate smoothness conditions, convergence of the best N-term approximations of $\mathcal{O}(N^{-s})$ can be shown, the adaptive method converges with that rate.

References

[CDD02] A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet methods II - Beyond the elliptic case. *Found. Comput. Math.*, 2(3):203–245, 2002.

[Gan04] T. Gantumur. Computation of integral operators in wavelet coordinates. Technical report, Utrecht University, 2004. To appear.

[GS04] T. Gantumur and R.P Stevenson. Computation of differential operators in wavelet coordinates. Technical report, Utrecht University, 2004. To appear.

[Ste04] R.P. Stevenson. On the compressibility of operators in wavelet coordinates. *SIAM J. Math. Anal.*, 35(5):1110–1132, 2004.

Multiscale Finite Element Methods and Residual–Free Bubbles

Endre Süli

(joint work with Franco Brezzi, Andrea Cangiani, and Donatella Marini)

Classical computational algorithms for the numerical simulation of physical phenomena have been designed to operate at a certain preselected scale fixed by the choice of a discretisation parameter. The task of numerically computing or even representing all of the physically relevant scales present in a multiscale problem results in excessive algorithmic complexity; the increased complexity may take the form of an extremely large set of unknowns in the discretisation and/or in a demand for a large amount of computer memory and CPU time. Limitations of these kinds have lead in recent years to the development of multiscale finite element

methods for partial differential equations; see, for example, the survey paper of Brezzi [2], the work of Hou, Wu and Cai [4] or that of Schwab and Matache [5].

One such class of algorithms are residual-free bubble finite element methods which are parameter-free, stable numerical techniques that have been successfully used for the approximate solution of a wide range of boundary-value problems exhibiting multiple-scale behaviour.

The first half of the lecture was devoted to surveying the results of the article [1]. The paper developed the general a priori error analysis of residual-free bubble (RFB) finite element approximations to non-self-adjoint elliptic problems of the form $(\varepsilon A+C)u = f$ subject to a homogeneous Dirichlet boundary condition, where A is a linear symmetric second-order elliptic operator, C is a skew-symmetric first-order linear differential operator, and ε is a positive parameter. The computational domain $\Omega \subset \mathbb{R}^d$ was subdivided into a shape-regular set $\{\mathcal{T}_h\}_h$ of d-dimensional simplices, where h is a positive discretisation parameter and continuous piecewise polynomial finite elements of degree k were used in conjunction with $\mathrm{H}^1_0(T)$-bubbles on each element $T \in \mathcal{T}_h$.

Using a subtle result from function-space interpolation theory due to Luc Tartar [6] which states that

$$[\mathrm{L}^2(T), \mathrm{H}^1_0(T)]_{1/2,\infty} \supset [\mathrm{L}^2(T), \mathrm{H}^1(T)]_{1/2,1},$$

with continuous embedding, it was shown that the error between the analytical solution u to the boundary value problem and its RFB finite element approximation u_{RFB} satisfies the error bound

$$\varepsilon^{1/2}|u - u_{\mathrm{RFB}}|_{\mathrm{H}^1(\Omega)} + h^{-1/2}\|C(u - u_{\mathrm{RFB}})\|_{\mathrm{H}^{-1}(\Omega)} \leq M\left(\varepsilon^{1/2}h^k + h^{k+1/2}\right)|u|_{\mathrm{H}^k(\Omega)}.$$

Here, it is assumed that the analytical solution u belongs to $\mathrm{H}^k(\Omega) \cap \mathrm{H}^1_0(\Omega)$ and M is a positive constant which only depends on the shape-regularity constant of the mesh and on $\|C\|_{\mathrm{H}^1(\Omega)\to\mathrm{L}^2(\Omega)}$, but not on u or ε.

If some local features of the solution u are known a priori, the approximation properties of the RFB finite element space can be improved through enrichment on selected edges of the partition by *edge-bubbles* that are supported on pairs of neighbouring elements. Based on this idea, we introduced and analysed the *enhanced residual free bubble* (RFBe) method for the numerical

approximation of convection-dominated diffusion equations. In two space dimensions we explored, both analytically and numerically, the accuracy of the RFB and RFBe approximations focusing on the practically relevant preasymptotic regime where the parameter ε is substantially smaller than the typical meshsize. In particular, in the case when the analytical solution to the problem exhibits an exponential boundary layer of thickness $\kappa = \varepsilon|\log\varepsilon|$ at the hyperbolic outflow boundary, the following result holds [3].

Theorem. *Suppose that $u \in \mathrm{H}^1_0(\Omega)$ is the solution to the boundary value problem with $f \in \mathrm{W}^2_\infty(\Omega)$ and with a sufficiently smooth convective velocity field. Assume, further, that $\mathcal{T}_h$ is an axiparallel shape-regular rectangular mesh on Ω of spacing $h \gg \kappa$ and that $\varepsilon \leq 1/\mathrm{e}$. Then, the RFB approximation u_{RFB} to u*

satisfies

$$\begin{aligned}\varepsilon^{1/2}|u-u_{RFB}|_{1,\Omega} \quad &+ \quad h^{-1/2}\|C(u-u_{\mathrm{RFB}})\|_{\mathrm{H}^{-1}(\Omega)} \\ &\leq \; C_1\left((\varepsilon/h)^{1/2}+\min\left(\varepsilon^{1/4},h^{1/2}\right)\right)+C_2,\end{aligned}$$

while the RFBe approximation u_{RFBe} *to* u *satisfies*

$$\begin{aligned}\varepsilon^{1/2}|u-u_{\mathrm{RFBe}}|_{1,\Omega} \quad &+ \quad h^{-1/2}\|C(u-u_{\mathrm{RFBe}})\|_{\mathrm{H}^{-1}(\Omega)} \\ &\leq \; C_1\left((\varepsilon/h)^{1/2}+\min\left(\varepsilon^{1/4},h^{1/2}\right)\right)+C_3h.\end{aligned}$$

Here C_1, C_2 and C_3 are positive constants, independent of ε.

This result, which illuminates the improved accuracy of RFBe over RFB, was substantiated by numerical experiments. For further details, we refer to [1] and [3].

References

[1] Brezzi, F., Marini, L.D., and E. Süli, Residual-free bubbles for advection-diffusion problems: the general error analysis. *Numer. Math.* 85, 1 (2000), 31-47.

[2] Brezzi, F., Recent results in the treatment of subgrid scales. In Canum 2000, *Actes du 32ème Congrès d'Analyse Numérique*, vol. 11., A. Blouza, I. Danaila, P. Joly, S.M. Kaber, B. Lucquin, F. Murat & M. Postel, Éditeurs; ESAIM: Proceedings, 2002, pp. 61–84.

[3] Cangiani, A. and Süli, E., Enhanced RFB method, *Oxford University Computing Laboratory Technical Report*, 2003/17. (Submitted for publication, 2004).

[4] Hou T.Y., Wu X.-H., and Cai, Z., Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. *Math. Comp.* 68, 227 (1999), 913–943.

[5] Matache, A.-M. and Schwab, C., Two-scale FEM for homogenization problems. *M2AN Math. Model. Numer. Anal.* 36 (2002), no. 4, 537–572.

[6] Tartar, L., Remarks on some interpolation spaces. In: Lions, J.L., Baiocchi, C. (eds.), *Boundary Value Problems for Partial Differential Equations and Applications.* Masson, pp. 229-252, 1993.

Some Error Estimates in Learning Theory

Vladimir Temlyakov

(joint work with R. DeVore, G. Kerkyacharian, S. Konyagin, D. Picard)

We continue investigation of some problems in learning theory in the setting formulated by F. Cucker and S. Smale [CS]. The goal is to find an estimator f_z on the base of given data $z := ((x_1,y_1),\dots,(x_m,y_m))$ that approximates well the regression function f_ρ of an unknown Borel probability measure ρ defined on $Z = X \times Y$. Following [CS] we consider a problem of approximate recovery of a projection f_W of an unknown regression function f_ρ onto a given class of functions W. It is known from [CS], [DKPT], [KT1], and [KT2] that the behavior of the entropy numbers $\epsilon_n(W)$ of W in the uniform norm plays an important role in the above problem. We obtain sharp (in the sense of order) estimates for the error between f_W and f_z for the classes W satisfying

$$\epsilon_n(W) \leq Dn^{-r}, \quad n = 1,2,\dots, \quad |f| \leq D, \quad f \in W. \tag{1}$$

We observe that the error estimates exhibit a saturation phenomenon for the range $r > 1/2$. We improve the error estimates by imposing one additional assumption on the relation between f_ρ and W, namely, we assume $f_\rho \in W$.

Let $X \subset R^d$, $Y \subset R$ be Borel sets, ρ be a Borel probability measure on $Z = X \times Y$. For $f : X \to Y$ define *the error*

$$\mathcal{E}(f) := \mathcal{E}_\rho(f) := \int_Z (f(x) - y)^2 d\rho.$$

Consider $\rho(y|x)$ - conditional (with respect to x) probability measure on Y and ρ_X - the marginal probability measure on X (for $S \subset X$, $\rho_X(S) = \rho(S \times Y)$). Define

$$f_\rho(x) := \int_Y y d\rho(y|x).$$

The function f_ρ is known in statistics as the *regression function* of ρ.

The next question is how to build $f_z \in \mathcal{H}$. We discuss a standard in statistics method of *empirical risk minimization* that takes

$$f_{z,\mathcal{H}} = \arg\min_{f \in \mathcal{H}} \mathcal{E}_z(f),$$

where

$$\mathcal{E}_z(f) := \frac{1}{m} \sum_{i=1}^m (f(x_i) - y_i)^2$$

is the *empirical error (risk)* of f. This $f_{z,\mathcal{H}}$ is called the *empirical optimum.*

We will assume that ρ and W satisfy the following condition.

(2) For all $f \in W$, $f : X \to Y$ is such that $|f(x) - y| \le M$ a.e.

One of important questions is to estimate the *defect function* $L_z(f) := \mathcal{E}(f) - \mathcal{E}_z(f)$ of $f \in W$. If ξ is a random variable (a real valued function on a probability space Z) then denote

$$E(\xi) := \int_Z \xi d\rho; \quad \sigma^2(\xi) := \int_Z (\xi - E(\xi))^2 d\rho.$$

The following useful inequality has been obtained in [CS].

Theorem 1. *Let W be a compact subset of $\mathcal{C}(X)$. Assume that ρ, W satisfy (2). Then, for all $\epsilon > 0$*

$$(3) \qquad Prob_{z \in Z^m}\{\sup_{f \in W} |L_z(f)| \ge \epsilon\} \le N(W, \epsilon/(8M)) 2e^{-\frac{m\epsilon^2}{2(\sigma^2 + M^2\epsilon/3)}}.$$

Here $\sigma^2 := \sigma^2(W) := \sup_{f \in W} \sigma^2((f(x) - y)^2)$. This theorem contains a factor $N(W, \epsilon/(8M))$ that may grow exponentially for classes W satisfying (1): $N(W, \epsilon) \le 2^{(D/\epsilon)^{1/r}+1}$. We prove a stronger (in a certain sense) estimate than (3) under assumption that W satisfies (1). For instance, in the case $r > 1/2$ we replace $N(W, \epsilon/(8M))$ in an analogue of (3) by a constant $C(M, D, r)$ independent of ϵ. This strengthening of Theorem 1 pays off in improved estimates for $\mathcal{E}(f_{z,W}) - \mathcal{E}(f_W)$ when we do not assume that $f_\rho \in W$. The following result is

essentially due to [CS] (see [DKPT]). Let W and ρ satisfy (1) and (2) then for $A \geq A_0(M, D, r)$

$$Prob_{z\in Z^m}\{\mathcal{E}(f_{z,W}) - \mathcal{E}(f_W) \leq Am^{-\frac{r}{1+2r}}\} \geq 1 - e^{-c(M)A^2 m^{\frac{1}{1+2r}}}.$$

We prove in [KT1], for instance, for $r > 1/2$ that for W, ρ satisfying (1) and (2) we have for $A \geq A_0(M, D, r)$

$$Prob_{z\in Z^m}\{\mathcal{E}(f_{z,W}) - \mathcal{E}(f_W) \leq Am^{-1/2}\} \geq 1 - e^{-c(M)A^2}.$$

We also prove that one cannot improve the error estimate of order $m^{-1/2}$ in the setting with no assumptions on f_ρ.

It turns out that if we assume that $f_\rho \in W$ we obtain significantly better estimates. We prove in [KT1] the following estimate.

Theorem 2. *Let $f_\rho \in W$ and let ρ and W satisfy (1) and (2). Then there exists an estimator f_z such that for $A \geq A_0(M, D, r)$*

$$Prob_{z\in Z^m}\{\mathcal{E}(f_z) - \mathcal{E}(f_\rho) \leq Am^{-\frac{2r}{1+2r}}\} \geq 1 - e^{-c(M)Am^{\frac{1}{1+2r}}}.$$

References

[CS] F. Cucker and S. Smale, On the mathematical foundations of learning, Bulletin of AMS, 39(2001),1–49.

[DKPT] R. DeVore, G. Kerkyacharian, D. Picard, V. Temlyakov, On Mathematical Methods of Learning, IMI Preprints, 2004:10, 1–23.

[KT1] S. Konyagin and V. Temlyakov, Some error estimates in Learning Theory, IMI Preprints, 2004:05, 1–18.

[KT2] S. Konyagin and V. Temlyakov, The entropy in the Learning Theory. Error estimates, IMI Preprints, 2004:09, 1–25.

The Electronic Schrödinger Equation — A Grand Challenge for Numerical Mathematics

Harry Yserentant

The electronic Schrödinger equation describes the motion of electrons under Coulomb interaction forces in the field of clamped nuclei and forms the basis of quantum chemistry. The problem is that its solutions, the electronic wavefunctions, depend on $3N$ variables in the case of an atom or molecule with N electrons. The direct numerical solution of the electronic Schrödinger equation therefore represents a grand challenge. The talk is first concerned with the regularity properties of the electronic wavefunctions that are compatible with the Pauli principle, another basic principle of quantum mechanics. It is shown that these solutions possess certain square integrable mixed weak derivatives of order up to $N + 1$ and in this sense become smoother and smoother for a growing number of electrons. This property is a possible point of attack for the development of numerical methods. A construction of this type is presented.

References

[Y1] H. Yserentant, On the regularity of the electronic Schrödinger equation in Hilbert spaces of mixed derivatives, Numer. Math, in print.

[Y2] H. Yserentant, Sparse grid spaces for the numerical solution of the electronic Schrödinger equation, submitted for publication.

Reporter: Angela Kunoth

Participants

Prof. Dr. Hans Wilhelm Alt

Institut für Angewandte Mathematik
Universität Bonn
Wegelerstr. 6
53115 Bonn

Prof. Dr. Francesc Arandiga

Departament de Matematica Aplicada
Universitat de Valencia
Dr. Moliner 50
E-46100 Burjassot

Prof. Dr. Richard Baraniuk

Dept. of Electrical and Computer Engineering
Rice University
6100 South Main Street
Houston, TX 77251-1892
USA

Prof. Dr. Peter G. Binev

Dept. of Mathematics
University of South Carolina
Columbia, SC 29208
USA

Dr. Luise Blank

Naturwissenschaftliche Fakultät I
Mathematik
Universität Regensburg
93040 Regensburg

Prof. Dr. Dietrich Braess

Fakultät für Mathematik
Ruhr-Universität Bochum
44780 Bochum

Martin Campos Pinto

Laboratoire Jacques-Louis Lions
Universite Pierre et Marie Curie
175 rue du Chevaleret
F-75013 Paris

Prof. Dr. Emanuel Candes

Applied Mathematics 217-50
California Institute of Technology
Pasadena, CA 91125
USA

Prof. Dr. Claudio Canuto

Dipartimento di Matematica
Politecnico di Torino
Corso Duca degli Abruzzi, 24
I-10129 Torino

Dr. Konstantinos Chrysafinos

Institut für Numerische Simulation
Universität Bonn
Wegelerstr. 6
53115 Bonn

Prof. Dr. Albert Cohen

Laboratoire Jacques-Louis Lions
Universite Pierre et Marie Curie
175 rue du Chevaleret
F-75013 Paris

Prof. Dr. Stephan Dahlke

FB Mathematik und Informatik
Universität Marburg
Hans-Meerwein-Strasse (Lahnbg)
35032 Marburg

Prof. Dr. Wolfgang Dahmen

Institut für Geometrie und
Praktische Mathematik
RWTH Aachen
Templergraben 55
52056 Aachen

Prof. Dr. Ingrid Daubechies

Department of Mathematics
Princeton University
Fine Hall
Washington Road
Princeton, NJ 08544-1000
USA

Prof. Dr. Christine De Mol

Department of Mathematics
Universite Libre de Bruxelles
C.P. 217
Bd. du Triomphe
B-1050 Bruxelles

Prof. Dr. Ronald A. DeVore

Dept. of Mathematics
University of South Carolina
Columbia, SC 29208
USA

Prof. Dr. Rosa Donat

Departament de Matematica Aplicada
Universitat de Valencia
Dr. Moliner 50
E-46100 Burjassot

Prof. Dr. Nira Dyn

School of Mathematical Sciences
Tel Aviv University
Ramat Aviv
Tel Aviv 69978
ISRAEL

Prof. Dr. Wolfgang Hackbusch

Max-Planck-Institut für Mathematik
in den Naturwissenschaften
Inselstr. 22 - 26
04103 Leipzig

Dr. Helmut Harbrecht

Institut für Informatik und
Praktische Mathematik
Christian-Albrechts-Universität
Olshausenstr. 40
24098 Kiel

HD Dr. Reinhard Hochmuth

Fakultät für Mathematik und
Informatik; Technische Universität
Bergakademie Freiberg
Agricolastr. 1
09599 Freiberg

Prof. Dr. Gerard Kerkyacharian

U. E. R. de Sciences Economiques
Universite de Paris X - Nanterre
200, Avenue de la Republique
F-92001 Nanterre Cedex

Prof. Dr. Nicholas Kevlahan

Department of Mathematics and
Statistics
Mc Master University
1280 Main Street West
Hamilton, Ont. L8S 4K1
CANADA

Prof. Dr. Angela Kunoth

Institut für Angewandte Mathematik
Universität Bonn
Wegelerstr. 6
53115 Bonn

Prof. Dr. Jacques Liandrat

ESM2/LATP
Technopole de Chateau - Gombert
38, rue F. Joliot Curie
F-13451 Marseille Cedex 20

Prof. Dr. Peter Maaß

Fachbereich 3
Mathematik und Informatik
Universität Bremen
Postfach 330440
28334 Bremen

Dr. Ralf Massjung

Institut für Geometrie und
Praktische Mathematik
RWTH Aachen
Templergraben 55
52062 Aachen

Priv.-Doz. Dr. Siegfried Müller

Institut für Geometrie und
Praktische Mathematik
RWTH Aachen
Templergraben 55
52062 Aachen

Prof. Dr. Peter Oswald

International University Bremen
School of Engineering and Science
Postfach 750561
28725 Bremen

Prof. Dr. Guergana Petrova

Department of Mathematics
Texas A & M University
College Station, TX 77843-3368
USA

Prof. Dr. Pencho P. Petrushev

Dept. of Mathematics
University of South Carolina
Columbia, SC 29208
USA

Prof. Dr. Dominique Picard

Laboratoire de Probabilites et
Modeles aleatoires
Univeriste Paris VII
4, Place Jussieu
F-75252 Paris Cedex 05

Prof. Dr. Bojan Popov

Department of Mathematics
Texas A & M University
College Station, TX 77843-3368
USA

Dr. Marie Postel

Laboratoire Jacques-Louis Lions
Universite Pierre et Marie Curie
175 rue du Chevaleret
F-75013 Paris

Prof. Dr. Arnold Reusken

Institut für Geometrie und
Praktische Mathematik
RWTH Aachen
Templergraben 55
52062 Aachen

Prof. Dr. Kai Schneider

Centre de Mathematiques et
d'Informatique
Universite de Provence
39, Rue Joliot-Curie
F-13453 Marseille Cedex 13

Prof. Dr. Reinhold Schneider

Lehrstuhl für Scientific Computing
Universitäts-Hochhaus 9 Stock
Christian-Albrechts-Platz 4
24098 Kiel

Prof. Dr. Christoph Schwab

Seminar für Angewandte Mathematik
ETH-Zentrum
Rämistr. 101
CH-8092 Zürich

Dr. Rob Stevenson

Department of Mathematics
Utrecht University
P.O.Box 80.010
NL-3508 TA Utrecht

Prof. Dr. Endre Süli

Computing Laboratory
Oxford University
Wolfson Building
Parks Road
GB-Oxford OX1 3QD

Prof. Dr. Anita Tabacco

Dipartimento di Matematica
Politecnico di Torino
Corso Duca degli Abruzzi, 24
I-10129 Torino

Prof. Dr. Vladimir N. Temlyakov

Dept. of Mathematics
University of South Carolina
Columbia, SC 29208
USA

Prof. Dr. Karsten Urban

Abteilung Numerik
Universität Ulm
Helmholtzstr. 18
89069 Ulm

Ralf Wildenhues

Institut für Numerische Simulation
Universität Bonn
Wegelerstr. 6
53115 Bonn

Prof. Dr. Gabriel Wittum

IWR Technische Simulation
Universität Heidelberg
Im Neuenheimer Feld 368
69120 Heidelberg

Prof. Dr. Przemek Wojtaszczyk

Instytut Matematyki
Stosowanej i Mechaniki
Uniwersytet Warszawski
ul. Banacha 2
02-097 Warszawa
POLAND

Prof. Dr. Harry Yserentant

Institut für Mathematik
Technische Universität Berlin
Straße des 17. Juni 136
10623 Berlin