Lecture Notes in Mathematics

1825

J. H. Bramble A. Cohen
W. Dahmen

Multiscale Problems and Methods in Numerical Simulations

Martina Franca, Italy 2001

Editor: C. Canuto

Springer

J. H. Bramble A. Cohen W. Dahmen

Multiscale Problems and Methods in Numerical Simulations

Lectures given at the
C.I.M.E. Summer School
held in Martina Franca, Italy,
September 9-15, 2001

Editor: C. Canuto

Fondazione
C.I.M.E.

Springer

Editor and Authors

Claudio Canuto
Dipartimento di Matematica
Politecnico di Torino
Corso Duca degli Abruzzi 24
10129 Torino, Italy
e-mail: claudio.canuto@polito.it

James H. Bramble
Mathematics Department
Texas A&M University
College Station
Texas TX 77843-3368
USA
e-mail: james.bramble@math.tamu.edu

Albert Cohen
Laboratoire Jacques-Louis Lions
Université Pierre et Marie Curie
175 rue du Chevaleret
75013 Paris, France
e-mail: cohen@ann.jussieu.fr

Wolfgang Dahmen
Institut für Geometrie
und Praktische Mathematik
RWTH Aachen
Templergraben 55
52056 Aachen, Germany
e-mail: dahmen@igpm.rwth-aachen.de

Cataloging-in-Publication Data applied for
Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at http://dnb.ddb.de

Mathematics Subject Classification (2000): 82D37, 80A17, 65Z05

ISSN 0075-8434
ISBN 3-540-20099-1 Springer-Verlag Berlin Heidelberg New York

Springer-Verlag Berlin Heidelberg New York a member of BertelsmannSpringer
Science + Business Media GmbH

http://www.springer.de

Printed in Germany

Typesetting: Camera-ready TEX output by the authors

SPIN: 10953471 41/3142/du - 543210 - Printed on acid-free paper

These Lecture Notes are dedicated to the victims of the brutal attacks of September 11, 2001, including all who were affected. All of us who attended the C.I.M.E. course, Americans and non-Americans alike, were shocked and horrified by what took place.
We all hope for a saner world.

Preface

The C.I.M.E. course on "Multiscale Problems and Methods in Numerical Simulation" was held in Martina Franca (Italy) from September 9 to 15, 2001. The purpose of the course was to disseminate a number of new ideas that had emerged in the previous few years in the field of numerical simulation, bearing the common denominator of the "multiscale" or "multilevel" paradigm. This takes various forms, such as: the presence of multiple relevant "scales" in a physical phenomenon, with their natural mathematical and numerical counterparts; the detection and representation of "structures", localized in space or in frequency, in the unknown variables described by a model; the decomposition of the mathematical or numerical solution of a differential or integral problem into "details", which can be organized and accessed in decreasing order of importance; the iterative solution of large systems of linear algebraic equations by "multilevel" decompositions of finite-dimensional spaces.

Four world leading experts illustrated the multiscale approach to numerical simulation from different perspectives. Jim Bramble, from Texas A&M University, described modern multigrid methods for finite element discretizations, and the efficient multilevel realization of norms in Sobolev scales. Albert Cohen, from Université Pierre et Marie Curie in Paris, smoothly guided the audience towards the realm of "Nonlinear Approximation", which provides a mathematical ground for state-of-the-art signal and image processing, statistical estimation and adaptive numerical discretizations. Wolfgang Dahmen, from RWTH in Aachen, described the use of wavelet bases in the design of computationally optimal algorithms for the numerical treatment of operator equations. Tom Hughes, from Stanford University, presented a general approach to derive variational methods capable of representing multiscale phenomena, and detailed the application of the variational multiscale formulation to Large Eddy Simulation (LES) in fluid-dynamics, using the Fourier basis.

The "senior" lecturers were complemented by four "junior" speakers, who gave account of supplementary material, detailed examples or applications. Ken Jansen, from Rensselaer Polytechnic Institute in Troy, discussed variational multiscale methods for LES using a hierarchical basis and finite el-

ements. Joe Pasciak, from Texas A&M University, extended the multigrid and multilevel approach presented by Bramble to the relevant case of symmetric indefinite second order elliptic problems. Rob Stevenson, from Utrecht University, reported on the construction of finite element wavelets on general domains and manifolds, i.e., wavelet bases for standard finite element spaces. Karsten Urban, from RWTH in Aachen, illustrated the construction of orthogonal and biorthogonal wavelet bases in complex geometries by the domain decomposition and mapping approach.

Both the senior and the junior lecturers contributed to the scientific success of the course, which was attended by 48 participants from 13 different countries. Not only the speakers presented their own material and perspective in the most effective manner, but they also made a valuable effort to dynamically establishing cross-references with other lecturers' topics, leading to a unitary picture of the course theme.

On Tuesday, September 11, we were about to head for the afternoon session, when we were hit by the terrible news coming from New York City. Incredulity, astonishment, horror, anger, worry (particularly for the families of our American friends) were the sentiments that alternated in our hearts. No space for Mathematics was left in our minds. But on the next day, we unanimously decided to resume the course with even more determination than before; we strongly believe, and we wanted to testify, that only rationality can defeat irrationality, that only the free circulation of ideas and the mutual exchange of experiences, as it occurs in science, can defeat darkness and terror.

The present volume collects the expanded version of the lecture notes by Jim Bramble, Albert Cohen and Wolfgang Dahmen. I am grateful to them for the timely production of such high quality scientific material.

As the scientific director of the course, I wish to thank the former Director of C.I.M.E., Arrigo Cellina, and the whole Scientific Board of the Centre, for inviting me to organize the event, and for providing us the nice facilities in Martina Franca as well as part of the financial support. Special thanks are due to the Secretary of C.I.M.E., Vincenzo Vespri. Generous funding for the course was provided by the I.N.D.A.M. Groups G.N.C.S. and G.N.A.M.P.A. Support also came from the Italian Research Project M.U.R.S.T. Cofin 2000 "Calcolo Scientifico: Modelli e Metodi Numerici Innovativi" and from the European Union T.M.R. Project "Wavelets in Numerical Simulation".

The organization and the realization of the school would have been by far less successful without the superb managing skills and the generous help of Anita Tabacco. A number of logistic problems were handled and solved by Stefano Berrone, as usual in the most efficient way. The help of Dino Ricchiuti, staff member of the Dipartimento di Matematica at the Politecnico di Torino, is gratefully acknowledged. Finally, I wish to thank Giuseppe Ghibò for his accurate job of processing the electronic version of the notes.

Torino, February 2003 *Claudio Canuto*

CIME's activity is supported by:

Ministero dell'Università Ricerca Scientifica e Tecnologica COFIN '99;
Ministero degli Affari Esteri - Direzione Generale per la Promozione e la Cooperazione - Ufficio V;
Consiglio Nazionale delle Ricerche;
E.U. under the Training and Mobility of Researchers Programme;
UNESCO - ROSTE, Venice Office.

Contents

Theoretical, Applied and Computational Aspects of Nonlinear Approximation
Albert Cohen 1
1 Introduction 1
2 A Simple Example 4
3 The Haar System and Thresholding 7
4 Linear Uniform Approximation 9
5 Nonlinear Adaptive Approximation 15
6 Data Compression 18
7 Statistical Estimation 21
8 Adaptive Numerical Simulation 24
9 The Curse of Dimensionality 26
References 28

Multiscale and Wavelet Methods for Operator Equations
Wolfgang Dahmen 31
1 Introduction 31
2 Examples, Motivation 32
2.1 Sparse Representations of Functions, an Example 32
2.2 (Quasi-) Sparse Representation of Operators 36
2.3 Preconditioning 37
2.4 Summary 39
3 Wavelet Bases – Main Features 39
3.1 The General Format 39
3.2 Notational Conventions 40
3.3 Main Features 40
4 Criteria for (NE) 45
4.1 What Could Additional Conditions Look Like? 45
4.2 Fourier- and Basis-free Criteria 46

5 Multiscale Decompositions – Construction and Analysis Principles ... 51
5.1 Multiresolution ... 51
5.2 Stability of Multiscale Transformations ... 52
5.3 Construction of Biorthogonal Bases – Stable Completions ... 53
5.4 Refinement Relations ... 53
5.5 Structure of Multiscale Transformations ... 55
5.6 Parametrization of Stable Completions ... 56
6 Scope of Problems ... 57
6.1 Problem Setting ... 57
6.2 Scalar 2nd Order Elliptic Boundary Value Problem ... 59
6.3 Global Operators – Boundary Integral Equations ... 59
6.4 Saddle Point Problems ... 61
7 An Equivalent ℓ_2-Problem ... 66
7.1 Connection with Preconditioning ... 67
7.2 There is always a Positive Definite Formulation – Least Squares . 68
8 Adaptive Wavelet Schemes ... 68
8.1 Introductory Comments ... 68
8.2 Adaptivity from Several Perspectives ... 70
8.3 The Basic Paradigm ... 70
8.4 (III) Convergent Iteration for the ∞-dimensional Problem ... 71
8.5 (IV) Adaptive Application of Operators ... 74
8.6 The Adaptive Algorithm ... 75
8.7 Ideal Bench Mark – Best N-Term Approximation ... 76
8.8 Compressible Matrices ... 76
8.9 Fast Approximate Matrix/Vector Multiplication ... 77
8.10 Application Through Uzawa Iteration ... 79
8.11 Main Result – Convergence/Complexity ... 79
8.12 Some Ingredients of the Proof of Theorem 8 ... 80
8.13 Approximation Properties and Regularity ... 85
9 Further Issues, Applications ... 88
9.1 Nonlinear Problems ... 88
9.2 Time Dependent Problems ... 90
10 Appendix: Some Useful Facts ... 90
10.1 Function Spaces ... 90
10.2 Local Polynomial Approximation ... 91
10.3 Condition Numbers ... 92
References ... 93

Multilevel Methods in Finite Elements
James H. Bramble ... 97
1 Introduction ... 97
1.1 Sobolev Spaces ... 97
1.2 A Model Problem ... 98
1.3 Finite Element Approximation of the Model Problem ... 100
1.4 The Stiffness Matrix and its Condition Number ... 101
1.5 A Two-Level Multigrid Method ... 102

2 Multigrid I 106
2.1 An Abstract V-cycle Algorithm 107
2.2 The Multilevel Framework 107
2.3 The Abstract V-cycle Algorithm, I 108
2.4 The Two-level Error Recurrence 109
2.5 The Braess-Hackbusch Theorem 110
3 Multigrid II: V-cycle with Less Than Full Elliptic Regularity 112
3.1 Introduction and Preliminaries 112
3.2 The Multiplicative Error Representation 116
3.3 Some Technical Lemmas 117
3.4 Uniform Estimates 119
4 Non-nested Multigrid 121
4.1 Non-nested Spaces and Varying Forms 121
4.2 General Multigrid Algorithms 122
4.3 Multigrid V-cycle as a Reducer 125
4.4 Multigrid W-cycle as a Reducer 127
4.5 Multigrid V-cycle as a Preconditioner 131
5 Computational Scales of Sobolev Norms 133
5.1 Introduction 133
5.2 A Norm Equivalence Theorem 135
5.3 Development of Preconditioners 138
5.4 Preconditioning Sums of Operators 138
5.5 A Simple Approximation Operator $\widetilde{Q}_k$ 139
Some Basic Approximation Properties 140
Approximation Properties: the Multilevel Case 141
The Coercivity Estimate 144
5.6 Applications 145
A Preconditioning Example 146
Two Examples Involving Sums of Operators 147
$H^1(\Omega)$ Bounded Extensions 148
References 150

List of Participants 153

Theoretical, Applied and Computational Aspects of Nonlinear Approximation

Albert Cohen

Laboratoire d'Analyse Numérique, Université Pierre et Marie Curie, Paris
cohen@ann.jussieu.fr

Summary. Nonlinear approximation has recently found computational applications such as data compression, statistical estimation or adaptive schemes for partial differential or integral equations, especially through the development of wavelet-based methods. The goal of this paper is to provide a short survey of nonlinear approximation in the perspective of these applications, as well as to stress some remaining open areas.

1 Introduction

Approximation theory is the branch of mathematics which studies the process of approximating general functions by simple functions such as polynomials, finite elements or Fourier series. It plays therefore a central role in the accuracy analysis of numerical methods. Numerous problems of approximation theory have in common the following general setting: we are given a family of subspaces $(S_N)_{N\geq 0}$ of a normed space X, and for $f \in X$, we consider the *best approximation error*

$$\sigma_N(f) := \inf_{g\in S_N} \|f-g\|_X. \tag{1}$$

Typically, N represents the number of parameters needed to describe an element in S_N, and in most cases of interest, $\sigma_N(f)$ goes to zero as this number tends to infinity.

For a given f, we can then study the *rate of approximation*, i.e., the range of $r \geq 0$ for which there exists $C > 0$ such that

$$\sigma_N(f) \leq CN^{-r}. \tag{2}$$

Note that in order to study such an asymptotic behaviour, we can use a sequence of *near-best approximation*, i.e., $f_N \in S_N$ such that

$$\|f-f_N\|_X \leq C\sigma_N(f), \tag{3}$$

with $C > 1$ independent of N. Such a sequence always exists even when the infimum is not attained in (1), and clearly (2) is equivalent to the same estimate with $\|f - f_N\|_X$ in place of $\sigma_N(f)$.

Linear approximation deals with the situation when the S_N are linear subspaces. Classical instances of linear approximation families are the following:
1) Polynomial approximation: $S_N := \Pi_N$, the space of algebraic polynomials of degree N.
2) Spline approximation with uniform knots: some integers $0 \leq k < m$ being fixed, S_N is the spline space on $[0,1]$, consisting of C^k piecewise polynomial functions of degree m on the intervals $[j/N, (j+1)/N]$, $j = 0, \cdots, N-1$.
3) Finite element approximation on fixed triangulations: S_N are finite element spaces associated with triangulations $\mathcal{T}_N$ where N is the number of triangles in $\mathcal{T}_N$.
4) Linear approximation in a basis: given a basis $(e_k)_{k\geq 0}$ in a Banach space, $S_N := \mathrm{Span}(e_0, \cdots, e_N)$.
In all these instances, N is typically the dimension of S_N, possibly up to some multiplicative constant.

Nonlinear approximation addresses in contrast the situation where the S_N are not linear spaces, but are still typically characterized by $\mathcal{O}(N)$ parameters. Instances of nonlinear approximation families are the following:
1) Rational approximation: $S_N := \{\frac{p}{q} \ ; \ p, q \in \Pi_N\}$, the set rational functions of degree N.
2) Free knot spline approximation: some integers $0 \leq k < m$ being fixed, S_N is the spline space on $[0,1]$ with N free knots, consisting of C^k piecewise polynomial functions of degree m on intervals $[x_j, x_{j+1}]$, for all partitions $0 = x_0 < x_1 \cdots < x_{N-1} < x_N = 1$.
3) Adaptive finite element approximation: S_N are the union of finite element spaces $V_{\mathcal{T}}$ of some fixed type associated to all triangulations $\mathcal{T}$ of cardinality less or equal to N.
4) N-term approximation in a basis: given a basis $(e_k)_{k\geq 0}$ in a Banach space, S_N is the set of all possible combinations $\sum_{k\in E} x_k e_k$ with $\#(E) \leq N$.
Note that these examples are in some sense nonlinear generalizations of the previous linear examples, since they include each of them as particular subsets. Also note that in all of these examples (except for the splines with uniform knots), we have the natural property $S_N \subset S_{N+1}$, which expresses that the approximation is "refined" as N grows.

On a theoretical level, a basic problem, both for linear and nonlinear approximation can be stated as follows:

Problem 1: *Given a nonlinear family $(S_N)_{N\geq 0}$, what are the analytic properties of a function f which ensure a prescribed rate $\sigma_N(f) \leq CN^{-r}$?*

By "analytic properties", we typically have in mind smoothness, since we know that in many contexts a prescribed rate r can be achieved provided that f belongs to some smoothness class $X_r \subset X$. Ideally, one might hope to identify the *maximal class* X_r such that the rate r is ensured, i.e., have a sharp result of the type

$$f \in X_r \Leftrightarrow \sigma_N(f) \leq CN^{-r}. \tag{4}$$

Another basic problem, perhaps on a slightly more applied level, is the effective construction of near-best approximants.

Problem 2: *Given a nonlinear family $(S_N)_{N\geq 0}$, find a simple implementable procedure $f \mapsto f_N \in S_N$ such that $\|f - f_N\|_X \leq C\sigma_N(f)$ for all $N \geq 0$.*

In the case of linear approximation, this question is usually solved if we can find a sequence of projectors $P_N : X \mapsto S_N$ such that $\|P_N\|_{X\to X} \leq K$ with K independent of N (in this case, simply take $f_N = P_N f$ and remark that $\|f - f_N\|_X \leq (1 + K)\sigma_N(f)$). It is in general a more difficult problem in the case of nonlinear method. Since the 1960's, research in approximation theory has evolved significantly toward nonlinear methods, in particular solving the two above problems for various spaces S_N.

More recently, nonlinear approximation became attractive on a more applied level, as a tool to understand and analyze the performance of *adaptive methods* in signal and image processing, statistics and numerical simulation. This is in part due to the emergence of *wavelet bases* for which simple N-term approximations (derived by thresholding the coefficients) yield in some sense optimal adaptive approximations. In such applications, the problems that arise are typically the following ones.

Problem 3 (data compression): *How can we exploit the reduction of parameters in the approximation of f by $f_N \in S_N$ in the perspective of optimally encoding f by a small number of bits ? This raises the question of a proper quantization of these parameters.*

Problem 4 (statistical estimation): *Can we use nonlinear approximation as a denoising scheme ? In this perspective, we need to understand the interplay between the approximation process and the presence of noise.*

Problem 5 (numerical simulation): *How can we compute a proper nonlinear approximation of a function u which is not given to us as a data but as the solution of some problem $F(u) = 0$? This is in particular the goal of adaptive refinement strategies in the numerical treatment of PDE's.*

The goal of the present paper is to briefly survey the subject of nonlinear approximation, with a particular focus on questions 1 to 5, and some emphasis on wavelet-based methods. We would like to point out that these questions are also addressed in the survey paper [15] which contains a more substantial

development on the theoretical aspects. We hope that our notes might be helpful to the non-expert reader who wants to get a first general and intuitive vision of the subject, from the point of view of its various applications, before perhaps going into a more detailed study.

The paper is organized as follows. As a starter, we discuss in §2 a simple example, based on piecewise constant functions, which illustrate the differences between linear and nonlinear approximation, and we discuss a first algorithm which produces nonlinear piecewise constant approximations. In §3, we show that such approximations can also be produced by thresholding the coefficients in the Haar wavelet system. In §4, we give the general results on linear uniform approximation of finite element or wavelet types. General results on nonlinear adaptive approximations by wavelet thresholding or adaptive partitions are given in §5. Applications to signal compression and estimation are discussed in §6 and §7. Applications to adaptive numerical simulation are shortly described in §8. Finally, we conclude in §9 by some remarks and open problems arising naturally in the multivariate setting.

2 A Simple Example

Let us consider the approximation of functions defined on the unit interval $I = [0,1]$ by piecewise constant functions. More precisely, given a disjoint partition of I into N subintervals $I_0, \cdots, I_{N-1}$ and a function f in $L^1(I)$, we shall approximate f on each I_k by its average $a_{I_k}(f) = |I_k|^{-1} \int_{I_k} f(t)dt$. The resulting approximant can thus be written as

$$f_N := \sum_{k=1}^{N} a_{I_k}(f)\chi_{I_k}. \tag{5}$$

If the I_k are fixed independently of f, then f_N is simply the orthogonal projection of f onto the space of piecewise constant functions on the partition I_k, i.e., a *linear approximation* of f. A natural choice is the uniform partition $I_k := [k/N, (k+1)/N]$. With such a choice, let us now consider the error between f and f_N, for example in the L^∞ metric. For this, we shall assume that f is in $C(I)$, the space of continuous functions on I. It is then clear that on each I_k we have

$$|f(t) - f_N(t)| = |f(t) - a_{I_k}(f)| \leq \sup_{t,u \in I_k} |f(t) - f(u)|. \tag{6}$$

We thus have the error estimate

$$\|f - f_N\|_{L^\infty} \leq \sup_k \sup_{t,u \in I_k} |f(t) - f(u)|. \tag{7}$$

This can be converted into an estimate in terms of N, under some additional smoothness assumptions on f. In particular, if f has a bounded first derivative, we have $\sup_{t,u \in I_k} |f(t) - f(u)| \leq |I_k| \|f'\|_{L^\infty} = N^{-1} \|f'\|_{L^\infty}$, and thus

$$\|f - f_N\|_{L^\infty} \leq N^{-1}\|f'\|_{L^\infty}. \tag{8}$$

Similarly, if f is in the Hölder space C^α for some $\alpha \in]0,1[$, i.e., if for all $x, y \in [0,1]$,

$$|f(x) - f(y)| \leq C|x-y|^\alpha, \tag{9}$$

we obtain the estimate

$$\|f - f_N\|_{L^\infty} \leq CN^{-\alpha}. \tag{10}$$

By considering simple examples such as $f(x) = x^\alpha$ for $0 < \alpha \leq 1$, one can easily check that this rate is actually sharp. In fact it is an easy exercise to check that a converse result holds : if a function $f \in C([0,1])$ satisfies (10) for some $\alpha \in]0,1[$ then necessarily f is in C^α, and f' is in L^∞ in the case where $\alpha = 1$. Finally note that we cannot hope for a better rate than N^{-1}: this reflects the fact that piecewise constant functions are only first order accurate.

If we now consider an *adaptive partition* where the I_k depend on the function f itself, we enter the topic of *nonlinear approximation*. In order to understand the potential gain in switching from uniform to adaptive partitions, let us consider a function f such that f' is integrable, i.e., f is in the space $W^{1,1}$. Since we have $\sup_{t,u\in I_k} |f(t) - f(u)| \leq \int_{I_k} |f'(t)|dt$, we see that a natural choice of the I_k can be made by equalizing the quantities $\int_{I_k} |f'(t)|dt = N^{-1}\int_0^1 |f'(t)|dt$, so that, in view of the basic estimate (7), we obtain the error estimate

$$\|f - f_N\|_{L^\infty} \leq N^{-1}\|f'\|_{L^1}. \tag{11}$$

In comparison with the uniform/linear situation, we thus have obtained the same rate as in (8) for a larger class of functions, since f' is not assumed to be bounded but only integrable. On a slightly different angle, the nonlinear approximation rate might be significantly better than the linear rate for a fixed function f. For instance, the function $f(x) = x^\alpha$, $0 < \alpha \leq 1$, has the linear rate $N^{-\alpha}$ and the nonlinear rate N^{-1} since $f'(x) = \alpha x^{\alpha-1}$ is in $L^1(I)$. Similarly to the linear case, it can be checked that a converse result holds : if $f \in C([0,1])$ is such that

$$\sigma_N(f) \leq CN^{-1}, \tag{12}$$

where $\sigma_N(f)$ is the L^∞ error of best approximation by adaptive piecewise constant functions on N intervals, then f is necessarily in $W^{1,1}$.

The above construction of an adaptive partition based on balancing the L^1 norm of f' is somehow theoretical, in the sense that it pre-assumes a certain amount of smoothness for f. A more realistic adaptive approximation algorithm should also operate on functions which are not in $W^{1,1}$. We shall describe two natural algorithms for building an adaptive partition. The first algorithm is sometimes known as *adaptive splitting* and was studied e.g. in [17]. In this algorithm, the partition is determined by a prescribed tolerance $\varepsilon > 0$

which represents the accuracy that one wishes to achieve. Given a partition of $[0,1]$, and any interval I_k of this partition, we split I_k into two sub-intervals of equal size if $\|f - a_{I_k}(f)\|_{L^\infty(I_k)} \geq \varepsilon$ or leave it as such otherwise. Starting this procedure on the single $I = [0,1]$ and using a fixed tolerance $\varepsilon > 0$ at each step, we end with an adaptive partition $(I_1, \cdots, I_N)$ with $N(\varepsilon)$ and a corresponding piecewise constant approximation f_N with $N = N(\varepsilon)$ pieces such that $\|f - f_N\|_{L^\infty} \leq \varepsilon$. Note that we now have the restriction that the I_k are *dyadic intervals*, i.e., intervals of the type $2^{-j}[n, n+1]$.

We now want to understand how the adaptive splitting algorithm behaves in comparison to the optimal partition. In particular, do we also have that $\|f - f_N\|_{L^\infty} \leq CN^{-1}$ when $f' \in L^1$? The answer to this question turns out to be negative, but a slight strengthening of the smoothness assumption will be sufficient to ensure this convergence rate : we shall instead assume that the *maximal function* of f' is in L^1. We recall that the maximal function of a locally integrable function g is defined by

$$Mg(x) := \sup_{r>0} [\mathrm{vol}(B(x,r)]^{-1} \int_{B(x,r)} |g(t)|dt. \tag{13}$$

It is known that $Mg \in L^p$ if and only if $g \in L^p$ for $1 < p < \infty$ and that $Mg \in L^1$ if and only if $g \in L \log L$, i.e., $\int |g| + \int |g \log |g|| < \infty$. Therefore, the assumption that Mf' is integrable is only slightly stronger than $f \in W^{1,1}$.

If $(I_1, \cdots, I_N)$ is the final partition, consider for each k the interval J_k which is the *parent* of I_k in the splitting process, i.e., such that $I_k \subset J_k$ and $|J_k| = 2|I_K|$. We therefore have

$$\varepsilon \leq \|f - a_{J_k}(f)\|_{L^\infty} \leq \int_{J_k} |f'(t)|dt. \tag{14}$$

For all $x \in I_k$, the ball $B(x, 2|I_k|)$ contains J_k and it follows therefore that

$$Mf'(x) \geq [\mathrm{vol}(B(x, 2|I_k|))]^{-1} \int_{B(x,2|I_k|)} |f'(t)|dt \geq [4|I_k|]^{-1}\varepsilon, \tag{15}$$

which implies in turn

$$\int_{I_k} Mf'(t)dt \geq \varepsilon/4. \tag{16}$$

If Mf' is integrable, this yields the estimate

$$N(\varepsilon) \leq 4\varepsilon^{-1} \int_0^1 Mf'(t)dt. \tag{17}$$

It follows that

$$\|f - f_N\|_{L^\infty} \leq CN^{-1} \tag{18}$$

with $C = 4\int_0^1 Mf'$. Note that in this case this is only a sufficient condition for the rate N^{-1} (a simple smoothness condition which characterizes this rate is still unknown).

3 The Haar System and Thresholding

The second algorithm is based on *thresholding* the decomposition of f in the simplest wavelet basis, namely the Haar system. The decomposition of a function f defined on $[0,1]$ into the Haar system is illustrated on Figure 1. The first component in this decomposition is the average of f, i.e., the projection onto the constant function $\varphi = \chi_{[0,1]}$, i.e.,

$$P_0 f = \langle f, \varphi \rangle \varphi. \tag{19}$$

The approximation is then recursively refined into

$$P_j f = \sum_{k=0}^{2^j-1} \langle f, \varphi_{j,k} \rangle \varphi_{j,k}, \tag{20}$$

where $\varphi_{j,k} = 2^{j/2}\varphi(2^j \cdot -k)$, i.e., averages of f on the intervals $I_{j,k} = [2^{-j}k, 2^{-j}(k+1)[$, $k = 0, \cdots, 2^j - 1$. Clearly $P_j f$ is the L^2-orthogonal projection of f onto the space V_j of piecewise constant functions on the intervals $I_{j,k}$, $k = 0, \cdots, 2^j - 1$. The orthogonal complement $Q_j f = P_{j+1} f - P_j f$ is spanned by the basis functions

$$\psi_{j,k} = 2^{j/2}\psi(2^j \cdot -k), \quad k = 0, \cdots, 2^j - 1, \tag{21}$$

where ψ is 1 on $[0, 1/2[$, -1 on $[1/2, 1[$ and 0 elsewhere. By letting j go to $+\infty$, we therefore obtain the expansion of f into an orthonormal system of $L^2([0,1])$

$$f = \langle f, \varphi \rangle \varphi + \sum_{j \geq 0} \sum_{k=0}^{2^j-1} \langle f, \psi_{j,k} \rangle \psi_{j,k} = \sum_{\lambda} d_\lambda \psi_\lambda. \tag{22}$$

Here we use the notation ψ_λ and $d_\lambda = \langle f, \psi_\lambda \rangle$ in order to concatenate the scale and space parameters j and k into one index $\lambda = (j,k)$, which varies in a suitable set ∇, and to include the very first function φ into the same notation. We shall keep track of the scale by using the notation

$$|\lambda| = j \tag{23}$$

whenever the basis function ψ_λ has resolution 2^{-j}. This simple example is known as the *Haar system* since its introduction by Haar in 1909. Its main limitation is that it is based on piecewise constant functions which are discontinuous and only allow for approximation of low order accuracy. We shall remedy to this defect by using smoother wavelet bases in the next sections.

We can use wavelets in a rather trivial way to build linear approximations of a function f since the projections of f onto V_j are given by

$$P_j f = \sum_{|\lambda| < j} \sum_{\lambda} d_\lambda \psi_\lambda. \tag{24}$$