Nonlinear Wavelet Methods for the Solution of PDE's

tesi di dottorato di

Marco Verani

Università di Milano
Dottorato in Matematica Computazionale e Ricerca Operativa
XV Ciclo
Anno Accademico 2001-02

Università degli Studi di Milano

Dottorato di Ricerca in
Matematica Computazionale e Ricerca Operativa
(XV Ciclo)

[Consorzio fra Università degli Studi di Milano, Università di Genova,
Università di Pavia, Università di Torino e Politecnico di Milano]

Novembre 2002

Nonlinear Wavelet Methods for the Solution of PDE's

tesi di dottorato di

Marco Verani

Tutore: Dott.ssa Silvia Bertoluzza

Coordinatore del Dottorato in Matematica Computazionale e Ricerca Operativa: Prof. Stefano Paveri-Fontana

Dedicated to my mother and to my father

And though I have the gift of prophecy, and understand all mysteries, and all knowledge; and though I have all faith, so that I could remove mountains, and have no charity, I am nothing.

(I. Corinthians 13,2)

CONTENTS

NOTATIONS

Miscellaneous Symbols

$A \lesssim B$ $\quad A \leq CB$
$A \simeq B$ $\quad C_1 B \leq A \leq C_2 B$
$\mathbb{R}^d$ Euclidean space
span$\{\ldots\}$ linear space spanned by $\{\ldots\}$
$\#(E)$ cardinality of E
$\underline{u}$ finite or infinite array

Wavelets & Wavelet Packets

j scale index
k position index
$\lambda = (j, k)$
$|\lambda| = j$
$\varphi_{j,k}(x)$ scaling function
$\tilde{\varphi}_{j,k}(x)$ dual scaling function
$\psi_{j,k}(x)$ wavelet function
$\tilde{\psi}_{j,k}(x)$ dual wavelet function
$w_{p,\omega,s}(x)$ wavelet packet function

Nonlinear Approximation

$\mathbb{P}_N$ nonlinear projector
Σ_N continuous nonlinear space
σ_N discrete nonlinear space
ℓ^τ_w weak-ℓ^τ space

INTRODUCTION

Three Rings for the Elven-kings under the sky,
Seven for the Dwarf-lords in their halls of stone,
Nine for Mortal Men doomed to die,
One for the Dark Lord on his dark throne
In the Land of Mordor where the Shadows lie.
One Ring to rule them all, One Ring to find them,
One Ring to bring them all and in the darkness bind them
In the Land of Mordor where the Shadows lie.

(J.R.R. Tolkien, The Lord of the Rings)

Adaptive methods for differential equations

The goals of the design of any numerical computational method are

1. reliability
2. efficiency .

Reliability means that the computational error (i.e. the difference between the exact and the approximate solution measured in a suitable norm) is controlled on a given tolerance level. Efficiency means that the computational work to compute a solution within the given tolerance is essentially as small as possible.

It frequently happens in practical problems that due to the nature of the data a solution of a differential problem could exhibit some singularities. In this case to achieve the goals of reliability and efficiency, one would like to increase the accuracy of the approximate solution without using too many additional degrees of freedom. One way to do this to use a computational method which is adaptive with feedback from the computational process. Adaptive procedures for the numerical solution of partial differential equations started in the late 70's and are now standard tools in science and engineering: they consist of a discretization method together with an adaptive algorithm, whose main features are:

(a) a stopping criterium guaranteeing error control to a given tolerance level;

(b) a modification strategy in case the stopping criterium is not satisfied.

A posteriori error estimators [4], [6], [7], [83], [86], [87] are an essential ingredient of any modification strategy and hence of any adaptive procedure. They are computable quantities depending on the computed solution(s) and data that provide information about the quality of approximation and may thus be used to make efficient mesh modification.

The ultimate purpose of adaptive methods, such as adaptive finite element methods (FEM), is to construct a sequence of meshes that would eventually equidistribute the approximation errors and reduce, as a consequence, the computational effort. To this end, the a posteriori error estimators are split into local indicators which are then employed to make local mesh modifications by refinement and coarsening. This naturally leads to loops of the form

$$\text{Solve } \rightarrow \text{Estimate } \rightarrow \text{Refine/Coarsen.} \tag{0.0.1}$$

Although these methods have been show to be very efficient from a computational point of view, the theory describing the advantages of such methods over their non-adaptive counterparts is still not satisfactory. For results of this kind, at least in the case of the numerical solutions of elliptic equations by means of adaptive finite element methods, we refer to [49], [71], [21].

Recently a new class of numerical adaptive schemes has been developed, namely *adaptive wavelet methods*.

Adaptive wavelet schemes

Adaptive wavelet schemes for the numerical solution of both linear and nonlinear equations typically rely on the empirical idea that local error indicators are directly given by the size of the currently computed wavelet coefficients: a large coefficient indicates important fluctuations of the solution on the support of the corresponding wavelet, and suggests to refine the approximation by adding wavelets at finer scale in this region. This idea was first introduced in [67], for the discretization of initial value problem, then developed in [16], [65]. Recently it was also applied to stationary problems for which the possibility of computing a residual allows to derive more precise a posteriori error indicators from the computed wavelet coefficients. In the framework of linear elliptic PDE's, this approach leads to a more rigorous analysis of the wavelet adaptive strategy, as introduced in [10] and further developed in [36]. In this respect we also recall multilevel finite element approach [22], [5].

More recently new development on wavelets and nonlinear approximation [47], [48] have provided new tools for understanding and designing adaptive wavelet schemes for linear [17], [19], [28], [31] and nonlinear equations [30], [29], [79], [81], [82], requiring

1. the estimation-evaluation of the action of (linear or nonlinear) operators on functions expressed in terms of wavelet coefficients

2. the tracking of the significant coefficients as the iterative solution process progresses.

Such class of adaptive wavelet schemes strongly rely on the sparsity of the wavelet representation of the solution and of the involved operators allowing for data compression, as well as the ability to perform accurate numerical computations in the compressed representation.

In this thesis we present and study adaptive wavelet schemes obtained by coupling iterative algorithms for the solution of linear and nonlinear problems and the techniques of nonlinear approximation. The approach we follow relies on a new paradigm which has been put forward recently for a class of linear problems [80], [31]. This new paradigm is based upon a convergent iterative scheme written for an equivalent infinite dimensional problem formulated in the wavelet coordinate domain and, as the iteration progresses, the adaptive evaluation of the involved linear and nonlinear infinite dimensional operators.

Classical approach & New approach

The new approach to the adaptive solution of well-posed PDE's has been very clearly presented in [29]. Let us point out, following such a paper, the differences between the classical paradigm and the new paradigm to numerically solving (linear and nonlinear) equations.

The classical approach is concerned with the following issues:

(c.1) variational formulation of the continuous problem;

(c.2) (adaptive) discretization of the infinite dimensional problem so as to obtain a finite system of algebraic equations;

(c.3) numerical solution of the finite system of equations, by means of a convergent iterative scheme;

(c.4) if the solution is not satisfactory, then perform a new (adaptive) discretization.

It is important to remark that performing the above approach yields a sequence of finite dimensional problems depending on the (adaptively) chosen discretizations. As a consequence the convergence of the iterative scheme depends itself on the chosen discretizations.

The new approach is performed by essentially using the same ingredients as in the classical approach, but "re-ordered" in a new meaningful way. The basic steps there read as follows:

(n.1) variational formulation of the continuous problem;

(n.2) transformation of the initial problem into an equivalent infinite dimensional problem in ℓ^2;

(n.3) derivation of a convergent iterative scheme for the infinite dimensional ℓ^2-problem;

(n.4) numerical realization of the iterative scheme by an approximate (possibly adaptive) application of the involved infinite dimensional operators within some strategy of dynamically updated accuracy tolerances.

The main difference between the two approaches is the discretization step. In the classical approach it is performed (step c.2) at the beginning of the procedure, by adaptively choosing a discrete space and then solving the resulting finite dimensional problem, by using an iterative scheme. On the contrary in the new approach a convergent iterative scheme is written directly for the ∞-dimensional problem and no discrete spaces are fixed in advance. The (adaptive) discretization is then performed at the very end of the procedure (step n.4), by an (adaptive) approximate application of the involved infinite dimensional operators, as the iterative scheme progresses.

It is important to remark that the convergence of the ∞-dimensional iterative scheme does not depend on the discretization, as it happens in the classical approach, but it is based on the "wavelet preconditioning" of the initial continuous problem. Wavelet preconditioning relies on the characterization of the involved functional spaces (typically Hilbert spaces) in terms of the decay of the wavelet coefficients, i.e. the norm of a function is equivalent to a weighted ℓ^2-norm of its wavelet coefficients.

In carrying out the new paradigm, for both linear and nonlinear equations, one has to face different issues, namely:

(a) the design of stable convergent iterative schemes for the ∞-dimensional discrete problem

(b) the design of economic approximate application schemes for the involved linear and nonlinear infinite dimensional operators.

(c) the choice of tolerances in (n.4) to ensure that the perturbed iteration converges to the correct solution;

(d) the estimate of the complexity of the scheme.

In this thesis we deal with issue (a) (only in the linear case) and issue (b) (both in the linear and nonlinear case).

It could happen (see e.g. nonconforming domain decomposition and Schur complement, Chapter 3) that the approximate application of an operator is equivalent to the approximate solution of an auxiliary problem. In this case issue (a) reduces to solve the auxiliary problem by using any strategy (e.g. finite elements) able to give an approximate solution within a prescribed tolerance.

The basic scheme

We consider the numerical solution of the problem:

$$R(u) = 0 \tag{0.0.2}$$

where $R : V \to W$ is a (linear or nonlinear) mapping between two Hilbert spaces V, W.
Chosen a wavelet basis $\{\psi_\lambda\}_\lambda$, it is possible to transform (0.0.2) into wavelet coordinates, obtaining an equivalent ∞-dimensional problem: find $\underline{u} \in \ell^2$ such that

$$\mathcal{R}(\underline{u}) = \underline{0}, \tag{0.0.3}$$

where $\mathcal{R} : \ell^2 \to \ell^2$ maps the sequence of the wavelet coefficients of v into the sequence of the wavelet coefficients of $R(v)$, while $\underline{u}$ is the unknown infinite array containing the wavelet coefficients of the unknown solution u to the initial problem (0.0.2).

Remark that problem (0.0.3) is indeed equivalent to problem (0.0.2). No approximation is performed at this point. According to (n.3) we wish to devise an iterative convergent scheme for the problem (0.0.3). The schemes we shall consider will have this form

$$\underline{u}^{n+1} = \underline{u}^n - \mathcal{B}_n \mathcal{R}(\underline{u}^n) \tag{0.0.4}$$

where the (infinite, possibly iteration dependent) matrix $\mathcal{B}_n$ is yet to be chosen in order to guarantee the convergence:

(1) for $\mathcal{B}_n = \theta\mathcal{I}$ we will obtain an ∞**-dimensional Richardson** scheme,

(2) for $\mathcal{B}_n = [\mathcal{R}'(\underline{u}^n)]^{-1}$, where $\mathcal{R}'(\underline{u})$ is the Fréchet derivative of $\mathcal{R}$ at $\underline{u}$, we will obtain an ∞ **dimensional Newton** scheme.

In order to arrive at computable versions of the schemes (0.0.5), we will couple such iterative algorithms with the techniques of nonlinear wavelet approximation [47], [48], obtaining a class of (adaptive) wavelet methods, whose general form is

$$\underline{u}^{n+1} = \mathbb{P}_{N_{i+1}}(\underline{u}^n - \mathcal{B}_n \mathcal{R}(\underline{u}^n)) \tag{0.0.5}$$

where $\mathbb{P}_N$ is a nonlinear projector retaining the N largest, in absolute value, wavelet coefficients. The introduction of the nonlinear projection will result in an implicit form of adaptivity, in which no specific approximation space is fixed, but the finite number of degrees of freedom to be used is determined at each stage by the nonlinear projector itself.

For different choices of the matrix $\mathcal{B}_n$ we will obtain different nonlinear wavelet methods, namely:

(nl.1) for $\mathcal{B}_n = \theta\mathcal{I}$ we will obtain the **Nonlinear Richardson** scheme (Chapter 3),

(nl.2) for $\mathcal{B}_n = [\mathcal{R}'(\underline{u}^n)]^{-1}$, where $\mathcal{R}'(\underline{u})$ is the Frèchet derivative of $\mathcal{R}$ at $\underline{u}$, we will obtain the **Nonlinear Newton** scheme (Chapters 4).

In this thesis we will perform an analysis of the two schemes, dealing with the issues of stability and convergence. We will also deal, in a direct way, with the problem of the evaluation-compression of linear operators and, in an indirect way, with the same problem for nonlinear operators, resulting in a recipe for the choice, at each step, of the involved tolerances.

The outline of the thesis is as follows: in Part I (Chapter 1 and Chapter 2) we recall for the sake of completeness some results about linear and nonlinear wavelet approximation: almost all the material of these chapters is borrowed from [27]. In Part II we design and study adaptive wavelet methods for linear equations (Chapter 3) and for nonlinear equations (Chapter 4). Finally in Chapter 5 we discuss adaptive wavelet packets methods for linear equations.

Part I

Wavelets, smoothness & approximation

Chapter 1

LINEAR WAVELET APPROXIMATION

Days passed and the Day drew nearer. An odd-looking waggon laden with odd-looking packages rolled into Hobbiton one evening and toiled up the Hill to Bag End. An old man was driving it all alone. He wore a tall pointed blue hat, a long grey cloak, and a silver scarf. He had a long white beard and bushy eyebrows that stuck out beyond the brim of his hat. At Bilbo's front door the old man began to unload: there were great bundles of fireworks of all sorts and shapes, each labelled with a large red G and the elf-rune ψ. That was Gandalf's mark, of course, and the old man was Gandalf the Wizard, whose fame in the Shire was due mainly to his skill with fires, smokes and light. His real business was far more difficult and dangerous, but the Shire-folk knew nothing about it.

(J.R.R. Tolkien, The Fellowship of the Ring)

1.1 Some functional spaces

In this section we want to describe the smoothness spaces that we shall need in what follows. There are two important ways to describe smoothness spaces: the first way is through notions such as differentiability and moduli of smoothness, the second way is to expand functions into a series of building blocks (for instance Fourier or wavelet) and describe smoothness as decay conditions on the coefficients in such expansions. We shall give both descriptions. The first is given below, while the second will be given when we discuss wavelet decompositions.

The most natural way of measuring the smoothness of a multivariate function f is certainly the order of differentiability, i.e. the maximal m such that $\partial^\alpha f, \quad |\alpha| := \alpha_1 + \alpha_2 + \ldots + \alpha_d \leq m$ is continuous. For $\Omega \subseteq \mathbb{R}$, we define $C^m(\Omega)$ to be the space of continuous functions which have bounded and continuous partial derivatives ∂^α, $|\alpha| \le m$. This space is equipped with the norm

$$\|f\|_{C^m(\Omega)} := \sup_{x\in\Omega} |f(x)| + \sum_{|\alpha|=m} \sup_{x\in\Omega} |\partial^\alpha f(x)|,$$

for which it is a Banach space.

In order to measure the smoothness properties of a function in an average sense, it is also natural to introduce Sobolev spaces $W^{m,p}(\Omega)$ consisting of all functions $f \in L^p$, with partial derivatives up to order m in L^p, $p \in [1,\infty]$. This space is also a Banach space, when equipped with the norm

$$\|f\|_{W^{m,p}} := \|f\|_{L^p} + |f|_{W^{m,p}}, \quad |f|_{W^{m,p}} := \sum_{|\alpha|=m} \|\partial^\alpha f\|_{L^p},$$

where we used the notation $|\cdot|$ to denote the corresponding semi-norm. All the above spaces share the common feature that the regularity index m is an integer.

How to generalize describing the regularity of a function in a more precise way, through fractional order of smoothness?

In the case of L^2-Sobolev spaces $H^m := W^{m,2}$ and when $\Omega = \mathbb{R}$, we can define an equivalent formula based on the Fourier transform

$$\|f\|_{H^m} \simeq \int_{\mathbb{R}} (1+|\omega|)^{2m} |\hat{f}(\omega)|^2 d\omega.$$

For a non-integer $s \ge 0$, it is thus natural to define the space H^s as the set of all L^2 functions such that

$$\|f\|_{H^s} \simeq \int_{\mathbb{R}} (1+|\omega|)^{2s} |\hat{f}(\omega)|^2 d\omega$$

is finite.

In the case of C^m spaces, we note that $\sup_{x\in\Omega} |f(x+h) - f(x)| \le (\sup|f'|)|h|$ if $f \in C^1$ for any $h \in \mathbb{R}$, whereas for an arbitrary $f \in C^0$, $\sup_{x\in\Omega} |f(x+h) - f(x)|$ might go to zero arbitrarily slow as $|h| \to 0$. This motivates the definition of the Hölder space C^s, $0 < s < 1$ consisting of those $f \in C^0$ such that

$$\sup_{x\in\Omega} |f(x+h) - f(x)| \le C|h|^s.$$

If $m < s < m+1$, a natural definition of C^s is given by $f \in C^m$ and $\partial^\alpha f \in C^{s-m}$, $|\alpha| = m$. It is not difficult to prove that this property can also be expressed by

$$\sup_{x\in\Omega} |\Delta_h^n f(x)| \le C|h|^s,$$

where $n > s$ and Δ_h^n is the n-th order finite difference operator defined recursively by $\Delta_h^1 f(x) = f(x+h) - f(x)$ and $\Delta_h^n f(x) = \Delta_h^1(\Delta_h^{n-1})f(x)$.

Let us now consider the generalization of "s order of smoothness in L^p" for s non-integer and p different from 2 and ∞. In particular we consider two classes of function spaces: Sobolev and Besov spaces.

Sobolev spaces $W^{s,p}$ are defined (if $m < s < m+1$) by $\|f\|_{W^{s,p}} = \|f\|_{L^p} + |f|_{W^{s,p}}$ with

$$|f|_{W^{s,p}} := \sum_{|\alpha|=m} \int_{\Omega^2} \frac{|\partial^\alpha f(x) - \partial^\alpha f(y)|^p}{|x-y|^{(s-m)p+d}} dxdy.$$

We refer to [1] for a general introduction.

Besov spaces $B_{p,q}^s$, involve an extra parameter q and can be defined through finite differences. These spaces include most of those we have listed so far as particular cases for certain ranges of indices. As we will show in the next Chapter, these spaces are also produced by general "interpolation techniques" between function spaces of integer smoothness, and they can be exactly characterized by the rate of multiresolution approximation error, as well as from the size properties of the wavelet coefficients. For these reasons we briefly recall their definitions and properties.

We define the n-th order L^p modulus of smoothness of f by

$$\omega_n(f,t,\Omega)_p = \sup_{|h|\le t} \|\Delta_h^n f\|_{L^p(\Omega_{h,n})},$$

(h is a vector in $\mathbb{R}$ of Euclidean norm less than t), where $\Omega_{h,n} := \{x \in \Omega : x + kh \in \Omega, k = 0,\dots,n\}$. For $p,q \ge 1$, $s \ge 0$, the Besov spaces $B_{p,q}^s(\Omega)$ consists of those functions $f \in L^p(\Omega)$, such that

$$(2^{sj}\omega_n(f,2^{-j})_p)_{j\ge 0} \in \ell^q,$$

where n is an integer such that $s < n$. A natural norm for such a space is then given by

$$\|f\|_{B_{p,q}^s} := \|f\|_{L^p} + |f|_{B_{p,q}^s}, \quad |f|_{B_{p,q}^s} := \|(2^{sj}\omega_n(f,2^{-j})_p)_{j\ge 0}\|_{\ell^q}.$$

The space $B_{p,q}^s$ represents "s order of smoothness measured in L^p", with the parameter q allowing a finer tuning on the degree of smoothness - one has $B_{p,q_1}^s \subset B_{p,q_2}^s$ if $q_1 \le q_2$ - but plays a minor role in comparison to s since

$$B_{p,q_1}^{s_1} \subset B_{p,q_2}^{s_2}, \text{if } s_1 \ge s_2,$$

regardless of the value of q_1 and q_2.

More generally, it can be proved [78] that $W^{s,p} = B^s_{p,p}$, when s is not an integer. Indeed the spaces $W^{m,p}$ are not Besov spaces for $m \in \mathbb{N}$ and $p \neq 2$.

Let us now state the so-called Sobolev embedding theorem [1]:

$$W^{s_1,p_1} \subset W^{s_2,p_2} \quad \text{if } s_1 - s_2 \geq d(1/p_1 - 1/p_2),$$

except in the case where $p_2 = +\infty$ and $s_1 - d(1/p_1 - 1/p_2)$ is an integer, for which one needs to assume that $s_1 - s_2 > d(1/p_1 - 1/p_2)$.

In the case of Besov spaces, a similar embedding relation [78] is given by

$$B^{s_1}_{p_1,p_1} \subset B^{s_2}_{p_2,p_2} \quad \text{if } s_1 - s_2 \geq d(1/p_1 - 1/p_2)$$

with no restriction on the indices $s_1, s_2 \geq 0$ and $p_1, p_2 \geq 1$.
The Besov spaces can also be defined for p and q less than 1. This extension, which will be of particular importance in the study of nonlinear and adaptive approximation, is the source of additional difficulties which go beyond the scope of this introduction [73].

Let us now discuss the topic of characterizing functional spaces through wavelet coefficients.

1.2 Wavelets: an overview

We recall some general notations and features for wavelet bases [69], [45]. They are usually associated with multiresolution approximation spaces $\{V_j\}_{j\geq 0}$:

Definition *A multiresolution analysis (MRA) is defined as a sequence of closed subspaces V_j of $L^2(\mathbb{R})$, $j \in \mathbb{Z}$, with the following properties*

1. $V_j \subset V_{j+1}$,
2. $v(x) \in V_j \Leftrightarrow v(2x) \in V_{j+1}$,
3. $v(x) \in V_0 \Leftrightarrow v(x+1) \in V_0$,
4. $\cup_{j=-\infty}^{+\infty} V_j$ *is dense in* $L^2(\mathbb{R})$ *and* $\cap_{j=-\infty}^{+\infty} V_j = \{0\}$,
5. *A compactly support scaling function* $\varphi \in V_0$, *with a non-vanishing integral exists such that the collection* $\{\varphi(x-k) : \quad k \in \mathbb{Z}\}$ *is a Riesz basis of* V_0.

It is immediate to note that the collection of functions $\{\varphi_{j,k} : \quad k \in \mathbb{Z}\}$, with $\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k)$ is a Riesz basis of V_j. Defining $\lambda = (j,k)$, $|\lambda| = j$ and $\Gamma_j := \{(j,k) : \quad k \in \mathbb{Z}\}$, we have that V_j is generated by a local basis $\{\varphi_\lambda\}_{\lambda\in\Gamma_j}$, whose supports are controlled by

$$|\text{supp}(\varphi_\lambda)| \leq C2^{-j}, \tag{1.2.1}$$

if $\lambda \in \Gamma_j$ and satisfy

$$\#\{\mu \in \Gamma_j : \quad \text{supp}(\varphi_\lambda) \cap \text{supp}(\varphi_\mu) \neq 0\} \leq C, \tag{1.2.2}$$

with C independent of λ and j.

We will use W_j to denote a space complementing V_j in V_{j+1}, i.e. a space that satisfies

$$V_{j+1} = V_j \oplus W_j,$$

where the symbol $\oplus$ stands for direct sum.

The complement space W_j, which contains the "detail" information needed to go from an approximation at resolution j to an approximation $j+1$, is generated by a similar local basis $\{\psi_\lambda\}_{\lambda\in\Lambda_j}$, $\Lambda_j = \Gamma_{j+1} \setminus \Gamma_j$, with $\psi_\lambda := \psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$.

The full multiscale wavelet basis $\{\psi_\lambda\}_{\lambda\in\Lambda}$, where $\Lambda := \cup_{j\geq 0}\Lambda_j$, is a Riesz basis for $L^2(\mathbb{R})$: it allows to expand an arbitrary function f into

$$f = \sum_{\lambda\in\Lambda} d_\lambda \psi_\lambda,$$

where $\Lambda := \cup_{j\geq 0}\Lambda_j$ with the convention that we incorporate the functions $\{\varphi_\lambda\}_{\lambda\in\Gamma_0}$ into $\{\psi_\lambda\}_{\lambda\in\Lambda_0}$ and for all sequences $\{d_\lambda\}_{\lambda\in\Lambda}$ we have the norm equivalence

$$\|\sum_{\lambda\in\Lambda} d_\lambda\psi_\lambda\|^2_{L^2} \sim \sum_{\lambda\in\Lambda} |d_\lambda|^2, \tag{1.2.3}$$

where the coefficients d_λ in the expansion of f are named wavelet coefficients.

In the case of biorthogonal wavelets [32] the coefficients d_λ are obtained by an inner product $d_\lambda = (f, \tilde{\psi}_\lambda)$, where the dual wavelet $\tilde{\psi}_\lambda$ is an L^2 function. In the standard biorthogonal constructions, a dual scaling function $\tilde{\varphi}$ and a dual wavelet $\tilde{\psi}$ exist and generate a dual multiresolution analysis with subspaces $\tilde{V}_j$ and $\tilde{W}_j$, such that

$$\tilde{V}_j \perp W_j \qquad V_j \perp \tilde{W}_j.$$

Moreover the dual functions also have to satisfy

$$(\tilde{\varphi}, \varphi(\cdot - l)) = \delta_l \qquad (\tilde{\psi}, \psi(\cdot - l)) = \delta_l,$$

where $\delta_l = 1$ if $l = 0$, zero otherwise.
The dual wavelet system $\{\tilde{\psi}_\lambda\}_{\lambda\in\Lambda}$ (risp. the dual scaling system $\{\varphi_\lambda\}_{\lambda\in\Lambda}$) has similar local support properties as the primal wavelets ψ_λ (risp. the primal scaling functions φ_λ).

It is useful to introduce the following projection operators:

$$\begin{aligned} P_j &: L^2 \to V_j \\ &: f \to P_j f = \sum_{k\in\mathbb{Z}} (f, \tilde{\varphi}_{j,k})\varphi_{j,k}, \\ Q_j &: L^2 \to W_j \\ &: f \to Q_j f = \sum_{k\in\mathbb{Z}} (f, \tilde{\psi}_{j,k})\psi_{j,k} \end{aligned} \tag{1.2.4}$$

and the corresponding dual projectors:

$$\begin{aligned} P_j^* &: L^2 \to \tilde{V}_j \\ &: f \to P_j f = \sum_{k\in\mathbb{Z}} (f, \varphi_{j,k})\tilde{\varphi}_{j,k}, \\ Q_j^* &: L^2 \to \tilde{W}_j \\ &: f \to Q_j^* f = \sum_{k\in\mathbb{Z}} (f, \psi_{j,k})\tilde{\psi}_{j,k}. \end{aligned}$$

As far as we have only considered the univariate case; in the standard constructions of wavelets on the Euclidean space $\mathbb{R}$, the scaling functions have the form $\varphi_\lambda = \varphi_{j,k} = 2^{jd/2}\varphi(2^j \cdot -k)$, $k \in \mathbb{Z}^d$ and similarly for the wavelets $\psi_\lambda = \psi_{j,k} = 2^{jd/2}\psi(2^j \cdot -k)$, $k \in \mathbb{Z}^d$, so that Γ_j is naturally viewed as the uniform mesh $2^{-j}\mathbb{Z}^d$. In the case of a general domain $\Omega \in \mathbb{R}$, special adaptations of the basis functions are required near the boundary $\partial\Omega$ (see e.g. [68], [33], [3], [26], [40], [56], [70], [13]).

The practical advantage of such a wavelet setting is the possibility of switching between the standard discretization of $f \in V_j$ in the basis $\{\varphi_\lambda\}_{\lambda\in\Gamma_j}$ and its multiscale representation in the basis $\{\psi_\lambda\}_{|\lambda|<j}$, by means of fast $\mathcal{O}(N)$ decomposition reconstruction algorithms, where $N \sim 2^{dj}$ denotes the dimension of V_j in the case where Ω is bounded.

An important feature of wavelet bases is the possibility of characterizing the smoothness of a function f through its wavelet coefficients or the linear approximation $\|f - P_j f\|$. Never less in next section we will discuss extensively this topic, we would recall here a particular result: for functions of d variables, Sobolev spaces are characterized by

$$\|f\|_{H^s} \sim \|P_0 f\|_{L^2}^2 + \sum_{j\geq 0} 2^{2sj}\|f - P_j f\|_{L^2}^2 \sim \sum_{\lambda\in\Lambda} 2^{2s|\lambda|}|d_\lambda|^2, \tag{1.2.5}$$

which reflects the intuitive idea that the linear approximation error decays like $\mathcal{O}(2^{-sj})$ or $\mathcal{O}(N^{-s/d})$, provided that f has s derivatives in L^2.

1.3 Approximation & smoothness

One of the goal of the approximation theory is to relate the analytical properties of arbitrary functions (in particular smoothness) with the accuracy of their approximation by simpler functions, such as polynomials, trigonometric series or finite elements.

As an instances we recall the following result of finite element approximation in L^2-Sobolev spaces: if Ω is a polygonal domain, $\mathcal{T}_h$, $0 < h < h_{max}$, a family of regular triangulation with mesh size h, and V_h a finite element space built from $\mathcal{T}_h$ that contains polynomials up to degree $n-1$ and is contained in $W^{s,2}$, then for $s \leq t \leq n$ one has the estimate

$$\inf_{g\in V_h} \|f - g\|_{W^{s,2}} \leq Ch^{t-s}|f|_{W^{t,2}}, \tag{1.3.1}$$

where C does not depend on f and h. For the multiresolution spaces $V_j \sim V_{h=2^{-j}}$ the above inequality takes the form

$$\inf_{g\in V_j} \|f - g\|_{W^{s,2}} \leq C2^{-(t-s)j}|f|_{W^{t,2}}. \tag{1.3.2}$$

These results express that a smoothness property implies an approximation rate. It turns out that a large number of smoothness classes, including L^2-Sobolev spaces, can be characterized from the rate of decay of the approximation error in the spaces V_j, or also from the summability and decay properties of the wavelet coefficients.

Let us now focus on the characterization from the approximation error: given a sequence of approximation spaces V_j and introduced

$$\mathrm{dist}_{L^p}(f, V_j) := \inf_{g\in V_j} \|f - g\|_{L^p},$$

we would like to relate the property

$$\mathrm{dist}_{L^p}(f, V_j) \leq \mathcal{O}(2^{-sj}),$$

to some classical notion of smoothness satisfied by f. To be more precise we introduce the following definition

Definition *If X is a Banach space and $(V_j)_{j\geq 0}$ a nested sequence of subspaces of X such that $\cup_{j\geq 0} V_j$ is dense in X. For $s>0$ and $1\leq q\leq\infty$, we define the approximation space $\mathcal{A}^s_q(X)$ related to the sequence V_j by*

$$\mathcal{A}^s_q(X) := \{f\in X : (2^{sj}\, dist_X(f,V_j))_{j\geq 0}\in \ell^q\}.$$

In the case where X is a Lebesgue space, we use the notation $\mathcal{A}^s_{p,q} := \mathcal{A}^s_q(L^p)$.

Roughly speaking, the space $\mathcal{A}^s_q(X)$ contains those functions such that $\mathrm{dist}_X(f,V_j)\leq \mathcal{O}(2^{-sj})$, with a tuned information provided by the extra parameter q. One can check that it is a proper subspace of X and it is a Banach space when equipped with the norm

$$\|f\|_{\mathcal{A}^s_q(X)} := \|f\|_X + \|(2^{sj}\mathrm{dist}_X(f,V_j))_{j\geq 0}\|_{\ell^q}.$$

What we actually want is to prove that under specific assumptions on the multiresolution approximation spaces, the identity

$$\mathcal{A}^s_{p,q} = B^s_{p,q} \tag{1.3.3}$$

holds together with the norm equivalences

$$\|f\|_{B^s_{p,q}} \simeq \|f\|_{\mathcal{A}^s_{p,q}} \quad \text{and} \quad |f|_{B^s_{p,q}} \simeq |f|_{\mathcal{A}^s_{p,q}}. \tag{1.3.4}$$

Following [27], in order to prove (1.3.3) and (1.3.4) we need direct and inverse estimates. This is the topic of what follows.

1.3.1 Direct estimates

Let us first collect a preliminary result, originally due to Lebesgue,

Lemma 1.3.1: *If P is a bounded projector from a Banach space X to a closed subspace Y, then for all $f\in X$,*

$$\inf_{g\in Y}\|f-g\|_X \leq \|f-Pf\|_X \leq (1+\|P\|))\inf_{g\in Y}\|f-g\|_X,$$

with $\|P\| = \sup_{\|f\|_X=1}\|Pf\|_X$.

Now we consider the following result about the L^p stability of the projector $P_j : L^p\to V_j$.

Theorem 1.3.1: *Let $1\leq p\leq\infty$. Assume that $\varphi\in L^p$ and $\tilde\varphi\in L^{p'}$ where $1/p+1/p'=1$. Then the projectors P_j are uniformly bounded in L^p. Moreover the basis $\varphi_{j,k}$ is L^p-stable, in the sense that the equivalence*

$$\|\sum_{k\in\mathbb{Z}^d} c_k\varphi_{j,k}\|_{L^p} \sim 2^{dj(1/2-1/p)}\|(c_k)_{k\in\mathbb{Z}^d}\|_{\ell^p}, \tag{1.3.5}$$

holds with constants that do not depend on j.

According to Lemma 1.3.1 it follows that if P_j is L^p-stable uniformly in j, then

$$\inf_{g \in V_j} \|f - g\|_{L^p} \sim \|f - P_j f\|_{L^p}, \tag{1.3.6}$$

i.e. the error estimate $\|f - P_j f\|_{L^p}$ is optimal in V_j.
Now we are ready state a direct estimate for this particular approximation process [27]:

Theorem 1.3.2: *Let $1 \le p \le \infty$. Assume that $\varphi \in L^p$ and $\tilde{\varphi} \in L^{p'}$ where $1/p + 1/p' = 1$. Then we have*

$$\|f - P_j f\|_{L^p} \lesssim 2^{-nj} |f|_{W^{n,p}}, \tag{1.3.7}$$

where $n-1$ is the order of polynomial exactness in V_j.

An important variant of the direct estimate 1.3.7 is the Whitney estimate, which involves the modulus of smoothness [27]:

Theorem 1.3.3: *Let $1 \le p \le \infty$. Assume that $\varphi \in L^p$ and $\tilde{\varphi} \in L^{p'}$ where $1/p + 1/p' = 1$. Then we have*

$$\|f - P_j f\|_{L^p} \lesssim \omega_n(f, 2^{-j})_p, \tag{1.3.8}$$

where $n-1$ is the order of polynomial exactness in V_j.

A simple corollary of the Whitney estimate is a direct estimate for general Besov spaces [27].

Corollary 1.3.1: *Let $1 \le p \le \infty$. Assume that $\varphi \in L^p$ and $\tilde{\varphi} \in L^{p'}$ where $1/p + 1/p' = 1$. Then we have*

$$\|f - P_j f\|_{L^p} \lesssim 2^{-js} |f|_{B^s_{p,q}}, \tag{1.3.9}$$

for $0 < s < n$, where $n-1$ is the order of polynomial exactness in V_j.

1.3.2 Inverse estimates

Inverse estimate takes into account the smoothness properties of the approximation spaces V_j [27]:

Theorem 1.3.4: *Let $1 \le p \le \infty$. Assume that $\varphi \in L^p$ and $\tilde{\varphi} \in L^{p'}$ where $1/p + 1/p' = 1$ and that $\varphi \in W^{n,p}$. Then*

$$\|f\|_{W^{n,p}} \le C 2^{nj} \|f\|_{L^p}, \quad \text{if } f \in V_j, \tag{1.3.10}$$

with a constant C that does not depend on j.

Another type of inverse estimate involves the modulus of smoothness [27]:

Theorem 1.3.5: *Let* $1 \le p \le \infty$. *Assume that* $\varphi \in L^p$ *and* $\tilde{\varphi} \in L^{p'}$ *where* $1/p + 1/p' = 1$ *and that* $\varphi \in W^{n,p}$. *Then we have*

$$\omega_n(f,t)_p \le C[\min\{1, 2^j t\}]^n \|f\|_{L^p}, \quad \text{if } f \in V_j, \tag{1.3.11}$$

with a constant C *that does not depend on* j.

Our last inverse estimate deals with general Besov spaces of integer or fractional order [27].

Theorem 1.3.6: *Let* $1 \le p \le \infty$. *Assume that* $\varphi \in L^p$ *and* $\tilde{\varphi} \in L^{p'}$ *where* $1/p + 1/p' = 1$ *and that* $\varphi \in B^s_{p,q}$. *Then*

$$\|f\|_{B^s_{p,q}} \le C 2^{sj} \|f\|_{L^p} \quad \text{if } f \in V_j, \tag{1.3.12}$$

with a constant C *that does not depend on* j.

By combining the above direct and inverse estimates, we obtain more general direct and inverse estimates involving Besov norms on both side of the inequalities [27]:

Corollary 1.3.2: *Let* $1 \le p, q_1, q_2 \le \infty$ *and* $0 < s < t$. *Assume that* $\varphi \in L^p$ *and* $\tilde{\varphi} \in L^{p'}$. *If* $\varphi \in B^s_{p,q_1}$ *and* $t < n$ *where* $n - 1$ *is the degree of polynomial reproduction in* V_j, *one has the direct estimate*

$$\|f - P_j f\|_{B^s_{p,q_1}} \lesssim 2^{-j(t-s)} |f|_{B^t_{p,q_2}}. \tag{1.3.13}$$

When s *and/or* t *are integers, these estimates also hold with the classical Sobolev space* $W^{s,p}$ *and/or* $W^{t,p}$ *and* t *up to* n.

Proof: [27] First of all we note that:

$$\|f - P_j f\|_{B^s_{p,q_1}} \le \sum_{l \ge j} \|P_{l+1} f - P_l f\|_{B^s_{p,q_1}}. \tag{1.3.14}$$

Combining the inverse estimate of Theorem 1.3.6 and the direct estimate of Corollary 1.3.1, we obtain

$$\|P_{l+1} f - P_l f\|_{B^s_{p,q_1}} \lesssim 2^{sl} \|P_{l+1} f - P_l f\|_{L^p} \lesssim 2^{-l(t-s)} |P_{l+1} f - P_l f|_{B^t_{p,q_2}}. \tag{1.3.15}$$

This together with (1.3.14) yields the thesis. The case of classical Sobolev spaces is treated in the same way, using the direct and inverse estimate of Theorem 1.3.2 and 1.3.4. □

Remark 1.3.1: *Inequality (1.3.13) tells us that an error estimate* $\mathcal{O}(2^{-j(t-s)})$ *in* $B^s_{p,p}$ *is achieved by a linear method for functions in* $B^t_{p,p}$*, with* $0 < s < t$*. We will see in the next Chapter that the same error estimate in* $B^s_{p,p}$ *can be achieved by a nonlinear method for less regular functions.*

Corollary 1.3.3: *Let* $1 \le p, q_1, q_2 \le \infty$ *and* $0 < s < t$*. Assume that* $\varphi \in L^p$ *and* $\tilde{\varphi} \in L^{p'}$*. If* $\varphi \in B^t_{p,q_2}$ *and* $s < n$ *where* $n-1$ *is the degree of polynomial reproduction in* V_j*, one has the inverse estimate*

$$\|f\|_{B^t_{p,q_2}} \lesssim 2^{j(t-s)}\|f\|_{B^s_{p,q_1}} \quad \text{if } f \in V_j. \tag{1.3.16}$$

When s *and/or* t *are integers, these estimates also hold with the classical Sobolev space* $W^{s,p}$ *and/or* $W^{t,p}$ *and* s *up to* n*.*

Proof: [27] First we note that:

$$\|f\|_{B^t_{p,q_2}} \le \|P_0 f\|_{B^t_{p,q_2}} + \sum_{l=0}^{j-1} \|P_{l+1}f - P_l f\|_{B^t_{p,q_2}}. \tag{1.3.17}$$

We clearly have $\|P_0 f\|_{B^t_{p,q_2}} \lesssim \|f\|_{L^p} \lesssim \|f\|_{B^s_{p,q_1}}$. For the remaining terms we combine the inverse estimate of Theorem 1.3.6 and the direct estimate of Corollary 1.3.1 and obtain

$$\|P_{l+1}f - P_l f\|_{B^t_{p,q_2}} \lesssim 2^{tl}\|P_{l+1}f - P_l f\|_{L^p} \lesssim 2^{-l(t-s)}|f|_{B^s_{p,q_1}}. \tag{1.3.18}$$

This together with (1.3.17) yields the inverse estimate of the thesis. The case of classical Sobolev spaces is treated in the same way, using the direct and inverse estimate of Theorem 1.3.2 and 1.3.4.

□

1.3.3 Smoothness

Now we are ready to prove [27] the characterization (1.3.3) of the approximation space $\mathcal{A}^s_{p,q}$ in terms of the Besov smoothness:

Theorem 1.3.7: *Let* $1 \le p \le \infty$*. Assume* $\phi \in L^p$ *and* $\tilde{\phi} \in L^{p'}$*, where* $\frac{1}{p} + \frac{1}{p'} = 1$*, then we have the norm equivalences*

$$\|f\|_{B^t_{p,q}} \sim \|P_0 f\|_{L^p} + \|(2^{tj}\|Q_j f\|_{L^p})_{j\ge 0}\|_{\ell^q} \tag{1.3.19}$$

and

$$\|f\|_{\mathcal{A}^t_{p,q}} \sim \|f\|_{B^t_{p,q}}, \tag{1.3.20}$$

for all $t < min\{n, s\}$*, where* $n-1$ *is the order of polynomial reproduction of the* V_j *space and* s *is such that* $\phi \in B^s_{p,q_o}$ *for some* q_0*.*

In order to prove the Theorem we need the following two Lemmas (discrete Hardy inequalities):

Lemma 1.3.2: *If $(a_j)_{j\geq 0}$ is a positive sequence and $b_j = 2^{-mj}\sum_{\ell=0}^{j} 2^{m\ell}a_\ell$, with $0 < s < m$, one has*

$$\|(2^{sj}b_j)_{j\geq 0}\|_{\ell^q} \lesssim \|(2^{sj}a_j)_{j\geq 0}\|_{\ell^q},$$

for all $q \in [1,\infty]$

Proof: Let us first consider the case $q = +\infty$. Assuming $a_j \leq C_a 2^{-sj}$ we have $b_j \leq C_a 2^{mj}\sum_{\ell=0}^{j} 2^{(m-s)\ell} \lesssim C_a 2^{-sj}$.
For $q < \infty$, we define q' such that $\frac{1}{q} + \frac{1}{q'} = 1$ and $\epsilon = \frac{m-j}{2} > 0$. Using Hölder inequality we obtain

$$\begin{aligned}
\sum_{j\geq 0}(2^{sj}b_j)^q &= \sum_{j\geq 0} 2^{(s-m)qj}\left(\sum_{\ell=0}^{j} 2^{m\ell}a_\ell\right)^q \leq \\
&\leq \sum_{j\geq 0} 2^{(s-m)qj}\left[\sum_{\ell=0}^{j}\left(2^{(m-\epsilon)\ell}a_\ell\right)^q\right]\left[\sum_{\ell=0}^{j}\left(2^{\epsilon\ell}\right)^{q'}\right]^{q/q'} \\
&\lesssim \sum_{j\geq 0} 2^{-\epsilon qj}\left[\sum_{\ell=0}^{j}\left(2^{(m-\epsilon)\ell}a_\ell\right)^q\right] \\
&= \sum_{\ell\geq 0}\left(2^{(m-\epsilon)\ell}a_\ell\right)^q \sum_{j\geq\ell} 2^{-\epsilon qj} \\
&\lesssim \sum_{\ell\geq 0}\left(2^{m\ell}a_\ell\right)^q.
\end{aligned}$$

□

Lemma 1.3.3: *If $(a_j)_{j\geq 0}$ is a positive sequence and $b_j = \sum_{\ell\geq j} a_\ell$, with $s > 0$, one has*

$$\|(2^{sj}b_j)_{j\geq 0}\|_{\ell^q} \lesssim \|(2^{sj}a_j)_{j\geq 0}\|_{\ell^q},$$

for all $q \in [1,\infty]$

Proof: The case $q = +\infty$ is treated in the same way as in the previous Lemma. For $q < +\infty$, we define q' such that $\frac{1}{q} + \frac{1}{q'} = 1$ and $s' = s/2$. Using

Hölder inequality we have

$$\begin{aligned}
\sum_{j\geq 0}(2^{sj}b_j)^q &= \sum_{j\geq 0} 2^{sqj}\left(\sum_{\ell\geq j} a_\ell\right)^q \leq \\
&\leq \sum_{j\geq 0} 2^{sqj}\left[\sum_{\ell\geq j}\left(2^{s'\ell}a_\ell\right)^q\right]\left[\sum_{\ell\geq j}\left(2^{-s'lq'}\right)\right]^{q/q'} \\
&\lesssim \sum_{j\geq 0} 2^{s'qj}\left[\sum_{\ell\geq j}\left(2^{s'\ell}a_\ell\right)^q\right] \\
&= \sum_{\ell\geq 0}\left(2^{s'\ell}a_\ell\right)^q \sum_{j=0}^{\ell} 2^{s'qj} \\
&\lesssim \sum_{\ell\geq 0}\left(2^{s\ell}a_\ell\right)^q.
\end{aligned}$$

□

Proof of Theorem 1.3.7: [27] Here we shall directly compare the modulus of smoothness which is involved in the definition of the Besov spaces $B^t_{p,q}$ and the quantities $\|Q_j f\|_{L^p}$. In one direction, from Theorem 1.3.6, we have

$$\mathrm{dist}_{L^p}(f, V_j) \leq \|f - P_j f\|_{L^p} \lesssim \omega_n(f, 2^{-j})_p, \tag{1.3.21}$$

and thus $\|Q_j f\|_{L^p} \lesssim \omega_n(f, 2^{-j})_p$. It follows that the $\mathcal{A}^s_{p,q}$ norm and the right hand side of 1.3.19 are both controlled by the $B^s_{p,q}$ norm. In order to prove the converse result, we remark that the inverse estimate of Theorem 1.3.6 implies the simpler inverse estimate

$$\omega_n(f, t)_p \lesssim [\min\{1, t2^j\}]^s \|f\|_{L^p}, \quad \text{if } f \in V_j. \tag{1.3.22}$$

Indeed this property holds for the values $t = 2^{-l}, l \geq j$, by Theorem 1.3.6 and the other values of t are treated by the monotonicity of $\omega_n(f, t)_p$.

For $f \in L^p$, we let $f_j \in V_j$ be such that

$$\|f - f_j\|_{L^p} \leq 2\mathrm{dist}_{L^p}(f, V_j). \tag{1.3.23}$$

We then have

$$\begin{aligned}
\omega_n(f,2^{-j}) &\leq \omega_n(f_0,2^{-j}) + \sum_{l=0}^{j} \omega_n(f_{l+1}-f_l,2^{-j})_p + \omega_n(f-f_j,2^{-j})_p \\
&\lesssim 2^{-sj}\|f_0\|_{L^p} + 2^{-sj}[\sum_{l=0}^{j-1} 2^{sl}\|f_{l+1}-f_l\|_{L^p}] + \|f-f_j\|_{L^p} \\
&\lesssim 2^{-sj}\|f_0\|_{L^p} + 2^{-sj}[\sum_{l=0}^{j} 2^{sl}\|f-f_l\|_{L^p}] \\
&\lesssim 2^{-sj}\|f\|_{L^p} + 2^{-sj}[\sum_{l=0}^{j} \mathrm{dist}_{L^p}(f,V_l)],
\end{aligned}$$

where we used the inverse estimate 1.3.23.

In order to conclude the proof, we apply Lemma 1.3.2 with $a_l = \mathrm{dist}_{L^p}(f,V_l)$ and we can thus conclude that the $B^s_{p,q}$ norm is controlled by the $\mathcal{A}^s_{p,q}$ norm. We can do the same reasoning with P_jf instead of f_j and replace $\mathrm{dist}_{L^p}(f,V_j)$ by $\|f-P_jf\|_{L^p}$ for the characterization of $B^s_{p,q}$. Finally, once we note that $\|f-P_jf\|_{L^p} \leq \sum_{l\geq j}\|Q_lf\|_{L^p}$, we can use the Lemma 1.3.3 with $a_l = \|Q_lf\|_{L^p}$ to replace $\|f-P_jf\|_{L^p}$ by $\|Q_jf\|_{L^p}$ and conclude the proof. □

1.4 Approximation & interpolation spaces

1.4.1 Interpolation theory: an overview

In the present section we shall describe a more general mechanism that allows to identify the approximation spaces $\mathcal{A}^s_{p,q}(X)$ with spaces obtained by interpolation theory. Although this mechanism can be avoided when proving (1.3.3) and (1.3.4), its usefulness will appear in particular in the nonlinear context.

Interpolation spaces arise in the study of the following problem of analysis. Given two spaces X and Y, for which spaces Z it is true that each linear operator T mapping X and Y boundedly into themselves automatically maps Z boundedly into itself? Such spaces Z are called interpolation spaces for the pair X, Y and the problem is to construct and to characterize the space Z. The classical result in this direction is the Riesz-Thorin theorem, which states that the spaces L^p, $1 < p < \infty$ are interpolation spaces for the pair L^1, L^∞. There are two primary methods for constructing interpolation spaces Z: the complex method as developed by J.L. Lions and A.P. Calderón and the real method of J.L. Lions and J. Peetre. We shall focus on the latter approach and we give below some of its main features. A detailed treatment can be found in [9], [8].

Let X and Y be a pair of Banach function spaces. To such a pair, we associate the so-called K-functional defined for $f \in X+Y$ and $t \geq 0$ by

$$K(f,t) = K(f,t,X,Y) := \inf_{a\in X, b\in Y, a+b=f}[\|a\|_X + t\|b\|_Y].$$

The functional has some elementary properties:

- it is continuous, nondecreasing and concave with respect to t.
- if $X\cap Y$ is dense in Y, then $K(f,0) := 0$. Similarly, if $X \cap Y$ is dense in X, then the limit $\lim_{t\to+\infty} K(f,t)/t = 0$.

For $\theta \in]0,1[$ and $1 \leq q \leq +\infty$, we define a family of intermediate spaces $X\cap Y \subset [X,Y]_{\theta,q} \subset X+Y$ as follows: $[X,Y]_{\theta,q}$ consists of those functions such that

$$\|f\|_{[X,Y]_{\theta,q}} := \|t^{-q}K(f,t)\|_{L^q(]0,+\infty[,dt/t)}, \tag{1.4.1}$$

is finite. One easily checks that the above defined intermediate spaces inherit the Banach spaces structure of X and Y.
In the present context we shall be interested in interpolation between spaces of representing various degrees of smoothness. In particular we shall always work in the situation where $Y \subset X$ with a continuous embedding and Y is dense in X. A typical example is $X = L^p$ and $Y = W^{m,p}$. In this specific situation, we write

$$K(f,t) = \inf_{g\in Y} \|f-g\|_X + t\|g\|_Y,$$

and make a few additional remarks:

- the K functional is bounded at infinity since $K(f,t) \leq \|f\|_X$. Therefore, the finiteness of (1.4.1) is equivalent to

$$\|t^{-\theta}K(f,t)\|_{L^q(]0,A[,dt/t)} < +\infty, \tag{1.4.2}$$

  for some fixed $A > 0$ and we can use this modified expression as an equivalent norm for $[X,Y]_{\theta,q}$.

- due to the monotonicity of $K(f,t)$ in t, we also have an equivalent discrete norm given by

$$\|f\|_{[X,Y]_{\theta,q}} := \|(\rho^{jq}K(f,\rho^{-j}))_{j\geq 0}\|_{\ell^q},$$

  for any fixed $\rho > 1$.

1.4.2 Smoothness via interpolation spaces

The main result of this section connects approximation spaces $\mathcal{A}^s_{p,q}$ and interpolation spaces $[X,Y]_{\theta,q}$ by means of direct and inverse estimates [27]:

Theorem 1.4.1: *Assume V_j is a sequence of approximation spaces*

$$V_j \subset V_{j+1} \subset \ldots \subset Y \subset X,$$

such that for some $m > 0$, one has a Jackson-type estimate

$$dist_X(f, V_j) = \inf_{g \in V_j} \|f - g\|_X \lesssim 2^{-mj} \|f\|_Y, \tag{1.4.3}$$

and a Bernstein-type estimate

$$\|f\|_Y \lesssim 2^{mj} \|f\|_X \quad \text{if } f \in V_j. \tag{1.4.4}$$

Then, for $s \in]0, m[$, one has the norm equivalence

$$\|(2^{js} K(f, 2^{-mj}))_{j \geq 0}\|_{\ell^q} \sim \|f\|_X + \|(2^{js} dist_X(f, V_j))_{j \geq 0}\|_{\ell^q}, \tag{1.4.5}$$

and thus $[X,Y]_{\theta,q} = \mathcal{A}^s_q(X)$ for $s = \theta m$.

Proof: [27] We need to compare the K-functional $K(f, 2^{-mj})$ and the error of best approximation $\text{dist}_X(f, V_j)$. In one direction, this comparison is simple: for all $f \in X$, $g \in Y$ and $g_j \in V_j$, we have

$$\text{dist}_X(f, V_j) \leq \|f - g_j\|_X \leq \|f - g\|_X + \|g - g_j\|_X. \tag{1.4.6}$$

Minimizing $\|g - g_j\|_X$ over $g_j \in V_j$ and using a Jackson-type estimate, we obtain

$$\text{dist}_X(f, V_j) \lesssim \|f - g\|_X + 2^{-mj} \|g\|_Y. \tag{1.4.7}$$

Finally, we minimize over $g \in Y$ to obtain

$$\text{dist}_X(f, V_j) \lesssim K(f, 2^{-mj}). \tag{1.4.8}$$

Since $\|f\|_X \lesssim K(f, 1)$ (by the continuous embedding of Y into X and the triangle inequality), we thus have proved that $\|f\|_{\mathcal{A}^s_q(X)} \lesssim \|f\|_{[X,Y]_{\theta,q}}$.
In the other direction, we let $f_j \in V_j$ be such that

$$\|f - f_j\|_X \leq 2\text{dist}_X(f, V_j), \tag{1.4.9}$$

and we write

$$
\begin{aligned}
K(f,2^{-mj}) &\le \|f-f_j\|_X + 2^{-mj}\|f\|_Y \\
&\le \|f-f_j\|_X + 2^{-mj}[\|f_0\|_Y + \|f_1-f_0\|_Y + \ldots + \|f_j-f_{j-1}\|_Y] \\
&\lesssim \|f-f_j\|_X + 2^{-mj}[\|f_0\|_X + \sum_{l=0}^{j-1} 2^{ml}\|f_{l+1}-f_l\|_X] \\
&\lesssim 2^{-mj}\|f_0\|_X + 2^{-mj}[\sum_{l=0}^{j} 2^{ml}\mathrm{dist}_X(f,V_l)], \qquad (1.4.10)
\end{aligned}
$$

where we have used the inverse inequality (together with the fact that $f_{l+1} - f_l \in V_{l+1}$) and the inequality $\|f\|_X \le \|f\|_X + 2\mathrm{dist}_X(f,V_0) \le 3\|f\|_X$.
In order to conclude the proof, we first remark that the term $2^{-mj}\|f\|_X$ satisfies

$$\|(2^{sj}2^{-mj}\|f\|_X)_{j\ge0}\|_{\ell^q} \lesssim \|f\|_X \qquad (1.4.11)$$

and we concentrate on the second term. Using Lemma 1.3.2 (discrete Hardy inequality) with $a_j = \mathrm{dist}_X(f,V_j)$ allows to estimate the weighted ℓ^q norm of the second term and to conclude that $\|f\|_{[X,Y]_{\theta,q}} \lesssim \|f\|_{\mathcal{A}^s_q(X)}$. □

The following variant [27] of the previous theorem deals with similar norm equivalences involving specific approximation operators P_j, rather than the error of best approximation $\mathrm{dist}_X(f,V_j)$

Theorem 1.4.2: *Assume V_j is a sequence of approximation spaces*

$$V_j \subset V_{j+1} \subset \ldots \subset Y \subset X,$$

suppose that we have

$$\|P_jf - f\|_X \lesssim 2^{-mj}\|f\|_Y,$$

for a family of linear operators $P_j : X \to V_j$ which is uniformly bounded in X. Then, for $s = \theta m$, $s \in]0,m[$, the $\mathcal{A}^s_q(X)$ and the $[X,Y]_{\theta,q}$ norms are equivalent to

$$\|P_0f\|_X + \|(2^{sj}\|f-P_jf\|_X)_{j\ge0}\|_{\ell^q}, \qquad (1.4.12)$$

and to

$$\|P_0f\|_X + \|(2^{sj}\|Q_jf\|_X)_{j\ge0}\|_{\ell^q} \qquad (1.4.13)$$

where $Q_j = P_{j+1} - P_j$.

Proof: [27] We first consider (1.4.12). In one direction, since

$$\mathrm{dist}_X(f,V_j) \le \|f-P_jf\|_X, \qquad (1.4.14)$$

and

$$\|f\|_X \leq \|P_0 f\|_X + \|f - P_0 f\|_X, \tag{1.4.15}$$

we clearly have that the norm of $\mathcal{A}^s_q(X)$ is controlled by (1.4.12). In the other direction, we operate as in the inequality (1.4.10) in the proof of Theorem 1.4.1: replacing f_j by $P_j f$ proves that the $[X, Y]_{\theta,q}$ norm is controlled by (1.4.12).

We then turn to (1.4.13). In one direction we have

$$\|Q_j f\|_X \leq \|f - P_{j+1} f\|_X + \|f - P_j f\|_X, \tag{1.4.16}$$

which shows that (1.4.13) is controlled by (1.4.12). In the other direction, we write

$$\|f - P_j f\|_X \leq \sum_{l \geq j} \|Q_l f\|_X. \tag{1.4.17}$$

Using Lemma 1.3.3 (discrete Hardy inequality) with $a_j = \|Q_j f\|_X$ allows to conclude that (1.4.12) is controlled by (1.4.13). $\square$

Now we want to prove the equivalence between $\mathcal{A}^s_{p,q}$ and $B^s_{p,q}$ and the norm equivalence 1.3.4, by using the general interpolation results obtained above. The key observation is that Besov spaces are obtained by the real interpolation applied to Sobolev spaces. For example the following result [62] is true on general Lipschitz domains $\Omega \in \mathbb{R}$:

Theorem 1.4.3: *It holds*

$$B^s_{p,q} = [L^p, W^{n,p}]_{\theta,q},$$

with $s = \theta n$, $\theta \in]0, 1[$.

1.5 Wavelets & functional spaces

In this section we show how a large number of smoothness classes can be characterized from the summability and decay properties of the wavelet coefficients.

1.5.1 Besov spaces with $p \geq 1$

The following result gives an equivalent norm of Besov spaces $B^s_{p,q}$, $p, q \geq 1$, in terms of the wavelet coefficients, by using characterization (1.3.19) (obtained through the "direct" way) or equivalently by using characterization (1.4.12) (obtained through the "interpolation approach"):

Corollary 1.5.1: *If $f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda$, then we have the norm equivalence*

$$\|f\|_{B^s_{p,q}} \sim \| \left(2^{sj} 2^{d(\frac{1}{2}-\frac{1}{p})j} \|(c_\lambda)_{\lambda \in \Lambda_j}\|\right)_j \|_{\ell^q}, \tag{1.5.1}$$

under the assumptions of Theorem 1.3.7.

Proof: It suffices to remark that for $j \geq 0$, we have the equivalence

$$\|Q_j f\|_{L^p} \sim 2^{d(1/2-1/p)j} \|(c_\lambda)_{\lambda \in \Lambda_j}\|_{\ell^p}, \tag{1.5.2}$$

by the same arguments as for the proof of Theorem 1.3.1. □

This last result also shows that wavelet bases are unconditional bases for all the Besov spaces in the above range: the convergence of the series holds in the corresponding norm without being affected by a rearrangement or a change of sign of the coefficients, since it only depends on the finiteness of the right-hand side of (4.2.4).

1.5.2 Besov spaces with $0 < p < 1$

So far, we have only considered values of p in the range $[1, \infty]$, whereas Besov spaces can be defined for $0 < p < 1$. In particular the case $0 < p < 1$ turns out to be crucial for nonlinear approximation. Here we sketch, following [27], how to extend the above results to the case $0 < p < 1$.
A first result is that, although we do not have the L^p boundness of P_j we still have some L^p stability for the scaling function basis. Here we continue to assume that $(\varphi, \tilde{\varphi})$ are a pair of compactly supported biorthogonal scaling functions, with $\varphi \in L^r$ and $\tilde{\varphi} \in L^{r'}$, for some $r \geq 1$, $1/r' + 1/r = 1$.

Theorem 1.5.1: *Assuming that $\varphi \in L^p$, for $p > 0$, one has the L^p stability property*

$$\| \sum_k c_k \varphi_{j,k} \|_{L^p} \sim 2^{dj(1/2-1/p)} \|(c_k)_k\|_{\ell^p}, \tag{1.5.3}$$

with constants that do not depend on j.

An immediate consequence of Theorem 1.5.1 is that we can extend the inverse inequalities of Theorems 1.3.5 and 1.3.6 to the case $p < 1$.

Theorem 1.5.2: *Under the assumption of the Theorem 1.5.1 and if $\varphi \in B^n_{p,q}$, for some $q > 0$, one has*

$$\omega_n(f,t)_p \leq C[\min\{1, t2^j\}]^n \|f\|_{L^p} \quad \textit{if } f \in V_j. \tag{1.5.4}$$

If $\varphi \in B^s_{p,q}$ one has

$$\|f\|_{B^s_{p,q}} \lesssim 2^{sj} \|f\|_{L^p} \quad \textit{if } f \in V_j. \tag{1.5.5}$$

If we want now to extend the Whitney estimate (1.3.8) to the case $p < 1$, we are facing the problem that L^p functions are not necessarily distributions and that the operator P_j is not a-priori well defined in these spaces (unless we put restrictions on s). One way to circumvent this problem is to consider the error of best-approximation $\text{dist}_{L^p}(f, V_j)$ rather than $\|f - P_j f\|_{L^p}$.

Theorem 1.5.3: *Under the same assumptions as in Theorem 1.5.1, we have*

$$dist_{L^p}(f, V_j) \lesssim \omega_n(f, 2^{-j})_p, \tag{1.5.6}$$

where $n-1$ is the order of polynomial exactness in V_j.

By using Theorems 1.5.2 and 1.5.3 it is possible to extend [27] the identity $\mathcal{A}^s_{p,q} = B^s_{p,q}$ to all possible values of p and q:

Theorem 1.5.4: *Assuming that $\varphi \in L^p$, for $p > 0$, we have the norm equivalence*

$$\|f\|_{\mathcal{A}^t_{p,q}} \sim \|f\|_{B^t_{p,q}}, \tag{1.5.7}$$

for all $t < \min(n, s)$, where $n-1$ is the order of polynomial reproduction of the V_j spaces and s is such that $\varphi \in B^s_{p,q_0}$, for some $q_0 > 0$.

If we now want to use the specific projectors P_j or the wavelet coefficients to characterize the $B^s_{p,q}$-norm for $p < 1$, we are obliged to impose condition s such that $P_j f$ will at least be well defined on the corresponding space. We shall now see that such characterizations are feasible if s is large enough so that $B^s_{p,q}$ is embedded in some L^r, $r \geq 1$. By using the above results it is possible to prove [27] the following result which extends the characterization of Besov spaces to the case $0 < p < 1$:

Theorem 1.5.5: *Assume that $\varphi \in L^r$ and $\varphi \in L^{r'}$ for some $r \in [1, \infty]$, $1/r + 1/r' = 1$ or that $\phi \in C^0$ and $\tilde{\varphi}$ is a Radon measure, in which case we set $r = \infty$. Then, for $0 < p \leq r$, one has the norm equivalence*

$$\|f\|_{B^s_{p,q}} \sim \|P_0 f\|_{L^p} + \|(2^{sj}\|Q_j f\|_{L^p})_{j\geq 0}\|_{\ell^q}, \tag{1.5.8}$$

for all $s > 0$ such that $d(1/p - 1/r) < s < \min(t, n)$, where $n-1$ is the order of polynomial reproduction in V_j and t is such that $\phi \in B^t_{p,q_0}$, for some q_0. If $f = \sum_{\lambda\in\Lambda} c_\lambda \psi_\lambda$ is the decomposition of f into the corresponding wavelet basis, we also have the norm equivalence

$$\|f\|_{B^s_{p,q}} \sim \| \left(2^{sj} 2^{d(\frac{1}{2}-\frac{1}{p})j} \|(c_\lambda)_{\lambda\in\Lambda_j}\|\right)_{j\geq -1} \|_{\ell^q},$$

under the same assumptions.

1.5.3 Characterization of negative smoothness

For $s < 0$, Besov spaces are usually defined by duality for $p, q \geq 1$:

$$B^{-s}_{p',q'} := (B^s_{p,q})^*, \tag{1.5.9}$$

with $1/p + 1/p' = 1$ and $1/q + 1/q' = 1$. The characterization of such dual spaces relies on the characterization of the corresponding primal space by the dual multiscale decomposition:

Theorem 1.5.6: *Assuming that the dual projectors are such that*

$$\|f\|_{B^s_{p,q}} \sim \|P_0^* f\|_{L^p} + \|(2^{sj}\|Q_j^* f\|_{L^p})_{j\geq 0}\|_{\ell^q} \tag{1.5.10}$$

for some $s > 0$ and $p, q \geq 1$, we then have

$$\|f\|_{B^{-s}_{p',q'}} \sim \|P_0 f\|_{L^{p'}} + \|(2^{-sj}\|Q_j f\|_{L^{p'}})_{j\geq 0}\|_{\ell^{q'}} \tag{1.5.11}$$

with $1/p + 1/p' = 1$. We also have the norm equivalence

$$\|f\|_{B^{-s}_{p',q'}} \sim \|(2^{-sj}2^{d(1/2-1/p')j}\|(c_\lambda)_{\lambda\in\Lambda_j}\|_{\ell^{p'}})_{j\geq -1}\|_{\ell^{q'}}, \tag{1.5.12}$$

if $f = \sum_{\lambda\in\Lambda} c_\lambda \psi_\lambda$.

1.5.4 The Hilbert case

In the particular case of Sobolev spaces, we have thus proved the norm equivalence for a regularity index which is either strictly positive or strictly negative:

$$\|f\|^2_{H^s} \sim \|P_0 f\|^2_{L^2} + \sum_{j\geq 0} 2^{2js}\|Q_j f\|^2_{L^2} \sim \sum_{\lambda\in\Lambda} 2^{2|\lambda|s}|c_\lambda|^2. \tag{1.5.13}$$

The above equivalence for $s = 0$, which corresponds to L^2, means that $\{\psi_\lambda\}_{\lambda\in\Lambda}$, and by duality $\{\tilde{\psi}_\lambda\}_{\lambda\in\Lambda}$, are Riesz bases for L^2:

$$\|f\|_{L^2} \simeq \sum_{\lambda\in\Lambda} |c_\lambda|^2 \simeq \sum_{\lambda\in\Lambda} |\tilde{c}_\lambda|^2 \tag{1.5.14}$$

with $\tilde{c}_\lambda = (f, \tilde{\psi}_\lambda)$. Different methods exist to prove norm equivalence (1.5.14): by using for example interpolation theory [27], or stable multiscale transformations [38].

1.5.5 Characterization of L^p spaces

We know that L^2 identifies with $B^0_{2,2}$. By using elementary interpolation properties of weighted ℓ^p spaces [27], one obtains the norm equivalence

$$\|f\|_{B^0_{p,q}} \sim \|(2^{d(1/2-1/p)j}\|(c_\lambda)_{\lambda\in\Lambda_j}\|_{\ell^p})_{j\geq -1}\|_{\ell^q}, \tag{1.5.15}$$

provided that the wavelet basis allows to characterize B^{ε}_{p,q_1} and $B^{-\varepsilon}_{p,q_2}$ for some $\epsilon > 0$ and $q_1, q_2 > 0$.
However we cannot identify L^p with $B^0_{p,q}$, for any $q > 0$, if $p \neq 2$. This reflects the fact that L^p spaces do not belong to the scale of Besov spaces when $p \neq 2$. Concerning such L^p spaces, we can formulate two basic questions:

1. does the wavelet expansion of an arbitrary function $f \in L^p$ converge unconditionally in L^p?
2. is there a simple characterization of L^p by the size of the wavelet coefficients?

The answer is negative for L^∞, since it is not separable, and for L^1, which is known to possess no unconditional basis. For the case $1 < p < \infty$, a positive answer to the first question is provided by the real value theory developed by Calderon and Zygmund in order to study the continuity properties of operators.

Finally the characterization of L^p norms from the size properties of the wavelet coefficients is also possible for $1 < p < \infty$, by means of a square function, defined for $f = \sum_{\lambda\in\Lambda} c_\lambda \psi_\lambda$ by

$$Sf(x) = [\sum_{\lambda\in\Lambda} |c_\lambda|^2 |\psi_\lambda(x)|^2]^{1/2}. \tag{1.5.16}$$

Clearly we have $\|f\|_{L^2} \sim \|Sf\|_{L^2}$. Moreover, using the classical Khinchine inequality [76] allows to prove [27] the following norm equivalence:

$$\|f\|_{L^p} \sim \|Sf\|_{L^p} \tag{1.5.17}$$

for $1 < p < \infty$.

1.5.6 Bounded domains and boundary conditions

It is possible to extend the results of previous sections, by characterizing functions spaces related to a bounded domain $\Omega \subset \mathbb{R}$, with prescribed boundary conditions, in terms of their multiscale decomposition. Following [27] we fix some general assumptions on our domain: Ω should have a simple geometry, expressed by a conformal partition

$$\Omega = \cup_{i=1}^{n} S_i,$$

into simplicial subdomains: each S_i is the image of the unit simplex

$$S := \{0 \leq x_1 + \ldots + x_d \leq 1\},$$

by an affine transformation. By "conformal" we mean that a face of an S_i is either part of the boundary Γ or coincides with a face of another S_j. We also assume that Ω is connected in the sense that for all $j, l \in \{1, \ldots, n\}$ there exists a sequence $i_0, \ldots, i_m$, such that $i_0 = j$ and $i_m = l$ and such that S_{i_k} and $S_{i_{k+1}}$ have a common face. Clearly, polygons and polyedrons fall in this category. More general curved domains or manifolds are also concerned here, provided that they can be smoothly parametrized by such a simple reference domains.
We denote by $C^\infty(\Gamma, m)$ the space of smooth functions defined on Ω which vanish at order m on Γ. This means that $f \in C^\infty(\Gamma, m)$ if and only if $f \in C^\infty(\Omega)$ and $|f(x)| \leq C[\mathrm{dist}(x, \Gamma)]^{m+1}$. We then define the spaces $W^{s,p}(\Gamma, m)$(resp. $B^s_{p,q}(\Gamma, m)$) as the closure of $C^\infty(\Gamma, m)$ in $W^{s,p}$(resp. $B^s_{p,q}$). We have the following result [27].

Theorem 1.5.7: *Assume that $\varphi \in L^r$ and $\tilde{\varphi} \in L^{r'}$ for some $r \in [1, \infty]$, $1/r + 1/r' = 1$, or that $\varphi \in C^0$ and $\tilde{\varphi}$ is a Radon measure, in which case we set $r = \infty$. Also assume that V_j reproduces polynomials of degree $n-1$ with flatness m at the boundary and that $\varphi \in B^s_{p,q_0}$. Then for $0 < p \leq r$, one has the norm equivalence*

$$\|f\|_{B^t_{p,q}} \sim \|P_0 f\|_{L^p} + \|(2^{tj}\|Q_j f\|_{L^p})_{j \geq 0}\|_{\ell^q}, \tag{1.5.18}$$

for all $f \in B^t_{p,q}(\Gamma, m)$ with $t > 0$ such that $d(1/p - 1/r) < t < \min(s, n)$ and $t - 1/p$ is not an integer among $0, \ldots, m$. If $f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda$ is the decomposition of f into the corresponding wavelet basis, we also have the norm equivalence

$$\|f\|_{B^s_{p,q}} \sim \| \left(2^{sj} 2^{d(\frac{1}{2} - \frac{1}{p})j} \|(c_\lambda)_{\lambda \in \Lambda_j}\|\right)_{j \geq -1} \|_{\ell^q},$$

under the same assumptions.

Finally we note [27] that also the characterizations of negative smoothness and L^p spaces extends to the type of domains that we are considering here.

Chapter 2

NONLINEAR WAVELET APPROXIMATION

"Well!" said Gandalf at last. "What are you thinking about? Have you decided what to do?". "I suppose I must keep the Ring and guard it, at least for the present, whatever it may do to me" answered Frodo. "But I feel very small, and very uprooted, and well – desperate. The Enemy is so strong and terrible".

In the black abyss there appeared a single Eye that slowly grew, until it filled nearly all the Mirror. The Eye was rimmed with fire, but was itself glazed, yellow as a cat's, watchful and intent, and the black slit of its pupil opened on a pit, a window into nothing.

(J.R.R. Tolkien, The Fellowship of the Ring)

Wavelet bases allow an efficient representation to characterise isolated singularities of functions, thanks to a particularly good location both in space and frequency of each element of the basis. This amounts to say that the decomposition of a function with isolated singularities is lacunary, in the sense that very few coefficients of its wavelet decomposition are non negligible. Then, given such a function, a simple strategy for building a compressed approximation is possible by getting rid of the coefficients that are smaller than a prescribed threshold, or equivalently by choosing the N largest, in absolute value, coefficients. In other words for functions which are not uniformly regular, possibly better approximations are obtained by choosing the approximation space depending on the function itself. This means the we look for a space $V_E := \mathrm{Span}\{\psi_\lambda : \quad \lambda \in E\}$, where $E = E(f) \subset \Lambda$ is a finite subset of indices which depends on the function f itself, and for an approximation $\tilde{f}$ to f belonging to V_E. Typically E could result to be the union of two subsets of indexes: the first allowing a coarse approximation of f and the second aiming to resolve the local singularities of f. If the target function f is smooth on a region we can use a course resolution on that region, by putting terms in the approximation corresponding to low frequency-terms. On regions where the target function is not smooth we use higher resolution, by taking in the approximation more wavelet functions corresponding to higher-frequencies. The questions that arise from these observations are:

(i) How does one practically build the set $E(f)$ and the approximation $\tilde{f}$?

(ii) Is there a precise characterization of the functions that can be approximated with a given rate of approximation, by this adaptive strategy?

Let us introduce the space

$$\Sigma_N = \{\sum_{\lambda\in E} c_\lambda\psi_\lambda : \quad \#(E) \le N\}, \tag{2.0.1}$$

of all possible N-term combinations of wavelets, and the error of best N-term approximation in some norm $\|\cdot\|_X$ defined by

$$\text{dist}_X(f, \Sigma_N) = \inf_{E\subset\Lambda, \#(E)\le N} \inf_{(c_\lambda)_{\lambda\in E}} \|f - \sum_{\lambda\in E} c_\lambda\psi_\lambda\|_X. \tag{2.0.2}$$

It is well understood that Σ_N is not a linear space: if f and g are in Σ_N we can only conclude that $f + g \in \Sigma_{2N}$

Suppose now one has access to the wavelet expansion $f = \sum_{\lambda\in\Lambda} c_\lambda\psi_\lambda$ of the function f to be approximated: a natural N-term approximation in X is provided by the choice

$$f_N = \sum_{\lambda\in E_N} c_\lambda\psi_\lambda, \tag{2.0.3}$$

where $E_N = E_N(f, X)$ is the set of indices corresponding to the N largest contributions $\|c_\lambda\psi_\lambda\|_X$. Then, we shall see that for several interesting choice of the space X, we have

$$\|f - f_N\|_X \lesssim \text{dist}_X(f, \Sigma_N), \tag{2.0.4}$$

i.e. a simple thresholding of the largest contributions in the wavelet decomposition provides a near optimal N-term approximation.

2.1 Nonlinear approximation in L^2

Let us consider for the moment N-term approximation in the L^2-norm: we are interested in the behaviour of $\text{dist}_{L^2}(f, \Sigma_N)$ as N goes to infinity, where Σ_N is defined as above. In order to simplify this example, we assume that $\{\psi_\lambda\}_{\lambda\in\Lambda}$ is an orthonormal basis for L^2. Thus any $f \in L^2$ can be decomposed into

$$f = \sum_{\lambda\in\Lambda} c_\lambda\psi_\lambda, \quad c_\lambda = (f, \psi_\lambda),$$

and we can define the set $\Lambda_N = \Lambda_N(f) \subset \Lambda$ of the N largest coefficients of f, i.e. such that $\#(\Lambda_N) = N$ and

$$\lambda \in \Lambda_N, \lambda' \notin \Lambda_N \Rightarrow |c_{\lambda'}| \leq |c_\lambda|. \tag{2.1.1}$$

If the modulus of several coefficients of f take the same value, we simply take for Λ_N any of the sets of cardinality N that satisfies 2.1.1. From the orthonormality of the basis, we clearly have

$$\text{dist}_{L^2}(f, \Sigma_N) = \|f - \sum_\lambda c_\lambda \psi_\lambda\|_{L^2} = \left(\sum_{\lambda \notin \Lambda_N} |c_\lambda|^2\right)^{1/2}. \tag{2.1.2}$$

Let us now consider the spaces $B^s_{q,q}$, where $s > 0$ and q is such that $1/q = 1/2 + s/d$. We assume here that $B^s_{q,q}$ is characterized by (4.2.4):

$$\|f\|_{B^s_{p,q}} \sim \| \left(2^{sj} 2^{d(\frac{1}{2}-\frac{1}{p})j} \|(c_\lambda)_{\lambda\in\Lambda_j}\|\right)_j \|_{\ell^q}.$$

For such indices, we note that this equivalence can be simplified

$$\|f\|_{B^s_{q,q}} \sim \|(c_\lambda)_{\lambda\in\Lambda}\|_{\ell^q}. \tag{2.1.3}$$

A first immediate consequence of (2.1.3) is the embedding of $B^s_{q,q}$ in L^2, since ℓ^q is trivially embedded in ℓ^2. Note that that such an embedding is not compact: the canonical sequence $(s_n)_{n\geq 0} = ((\delta_{n,k})_{k\geq 0})_{n\geq 0}$ is uniformly bounded in ℓ^q but does not contain any subsequence that converges in ℓ^2.

In order to characterize Besov spaces in terms of the L^2-error of nonlinear wavelet approximation, we follow the same strategy as in the linear context and we try to obtain three ingredients: a direct estimate, an inverse estimate and a result of interpolation theory.

If we now define by $(c_n)_{n\geq 1}$ any rearrangement of the coefficients $(c_\lambda)_{\lambda\in\Lambda}$ with decreasing moduli, i.e. such that $|c_{n+1}| \leq |c_n|$, we also have

$$n|c_n|^q \leq \sum_{k\geq 1} |c_k|^q \sim \|f\|^q_{B^s_{q,q}}, \tag{2.1.4}$$

which yields

$$|c_n| \lesssim \|f\|_{B^s_{q,q}} n^{-1/q}. \tag{2.1.5}$$

Taking the decreasing rearrangement c_n to be such that $\{c_n : \quad n \leq N\} = \{c_\lambda : \quad \lambda \in \Lambda_N\}$, it follows that we have

$$\begin{aligned}\mathrm{dist}_{L^2}(f, \Sigma_N) &= \|f - \sum_{\lambda \in \Sigma_N} c_\lambda \psi_\lambda\|_{L^2} \\ &= \left(\sum_{n>N} |c_n|^2\right)^{1/2} \\ &\lesssim N^{1/2-1/q} \|f\|^q_{B^s_{q,q}}.\end{aligned}$$

We thus have obtained a Jackson-type estimate

$$\mathrm{dist}_{L^2}(f, \Sigma_N) \lesssim N^{-s/d} \|f\|^q_{B^s_{q,q}}, \tag{2.1.6}$$

with respect to the non linear spaces Σ_N.
On the other hand, if $f \in \Sigma_N$, we also have by Hölder inequality

$$\|f\|_{B^s_{q,q}} \lesssim \|(c_\lambda)_{\lambda\in\Lambda}\|_{\ell^q} \leq N^{1/q-1/2} \|(c_\lambda)_{\lambda\in\Lambda}\|_{\ell^2} = N^{s/d} \|f\|_{L^2}, \tag{2.1.7}$$

i.e. a Bernstein-type estimate.
The equivalence (2.1.3) also shows that for $0 < t < s$ and $1/r = 1/2 + t/d$, we have the interpolation identity

$$B^t_{r,r} = [L^2, B^s_{q,q}]_{\theta,r}, \quad \theta = t/s, \tag{2.1.8}$$

which is a simple re-expression of $[\ell^2, \ell^q]_{\theta,r} = \ell^r$.

Now if we were in the linear context we could characterize functions spaces by the error of linear approximation. It turns out that a similar result as in linear case also holds in the present nonlinear context using the following Theorem [27].

Theorem 2.1.1: *Assume that X and Y are quasi-normed spaces and that $S_j, j \geq 0$, is a sequence of non linear approximation spaces*

$$S_j \subset S_{j+1} \subset \ldots \subset Y \subset X, \tag{2.1.9}$$

such that for some $m > 0$, one has a Jackson type estimate

$$dist_X(f, S_j) = \inf_{g \in S_j} \|f - g\| \lesssim 2^{-mj} \|f\|_Y, \tag{2.1.10}$$

and a Bernstein-type estimate

$$\|f\|_Y \lesssim 2^{mj} \|f\|_X \quad \text{if } f \in S_j. \tag{2.1.11}$$

Moreover assume that there exists a fixed integer a such that

$$S_j + S_j \subset S_{j+a}, \quad j \geq 0. \tag{2.1.12}$$

Then, for $t \in]0, m[$, one has the norm equivalence

$$\|(2^{jt}K(f, 2^{-mj}))_{j\geq 0}\|_{\ell^q} \sim \|f\|_X + \|(2^{jt}dist_X(f, S_j))_{j\geq 0}\|_{\ell^q}, \tag{2.1.13}$$

and thus $[X, Y]_{\theta,q} = \mathcal{A}^t_q(X)$ for $t = \theta m$.

From the monotonicity of the sequence $\mathrm{dist}_{L^2}(f, \Sigma_N)$, we have the equivalence

$$\sum_{j\geq 0}[2^{jt}\mathrm{dist}_{L^2}(f, S_j)]^q \sim \sum_{N\geq 1} N^{-1}[N^{t/d}\mathrm{dist}_{L^2}(f, \Sigma_N)]^q. \tag{2.1.14}$$

The finiteness of the above quantities for $q < \infty$ is a slightly stronger property than $\mathrm{dist}_{L^2}(f, \Sigma_N) \lesssim N^{-t/d}$ which was initially obtained. According to the above theorem, with $X = L^2$ and $Y = B^s_{q,q}$, this last property characterizes the intermediate space

$$[L^2, B^s_{q,q}]_{\theta,\infty} = \mathcal{A}^t_{2,\infty}, \tag{2.1.15}$$

with $t = \theta s$, which cannot be thought as a Besov space. One can check that this space is also characterized by the property that $(c_\lambda)_{\lambda\in\Lambda}$ belongs to the weak space ℓ^r_w, i.e.

$$\#\{\lambda \in \Lambda : \quad |c_\lambda| \geq \varepsilon\} \leq C\varepsilon^{-t}. \tag{2.1.16}$$

Hence the property $f \in B^t_{r,r}$, $1/r = t/d + 1/2$ is almost equivalent to the rate $\mathrm{dist}_{L^2}(f, \Sigma_N) \lesssim N^{-t/d}$, while the exact characterization passes through the stronger property

$$\sum_{N\geq 1} N^{-1}[N^{t/d}\mathrm{dist}_{L^2}(f, \Sigma_N)]^r < \infty,$$

or the weak space ℓ^r_w.

2.2 Nonlinear approximation in $B^s_{p,p}$

It is not difficult to extend the above results to the case where the error is measured in more general Besov space of the type $B^s_{p,p}$. Following [27] we cancel the orthonormality assumption which is irrelevant at this stage of generality and we prove the following result.

Lemma 2.2.1: *Let assume that $B^s_{p,p}$ admits a wavelet characterization of the type (4.2.4). If $f = \sum_{\lambda\in\Lambda} c_\lambda\psi_\lambda$, then*

$$\|f - \sum_{\lambda\in\Lambda_N} c_\lambda\psi_\lambda\|_{B^s_{p,p}} \lesssim dist_{B^s_{p,p}}(f, \Sigma_N), \tag{2.2.1}$$

where $\Lambda_N = \Lambda_N(f, B^s_{p,p})$ is the set of the indices corresponding to the N largest contributions $\|c_\lambda\psi_\lambda\|_{B^s_{p,p}}$ or equivalently the N largest $2^{(s+d/2-d/p)|\lambda|}|c_\lambda|$.

Proof: [27] The norm equivalence (4.2.4) shows that $\|\psi_\lambda\|_{B^s_{p,p}} \sim 2^{(s+d/2-d/p)|\lambda|}$. It can thus be reformulated as

$$\|f\|_{B^s_{p,p}} \sim \|(\|c_\lambda\psi_\lambda\|_{B^s_{p,p}})_{\lambda\in\Lambda}\|_{\ell^p}. \tag{2.2.2}$$

Clearly the N-term approximation $\sum_{\lambda\in\Lambda_N} c_\lambda\psi_\lambda$ minimizes the distance between f and Σ_N when measured in this equivalent norm for $B^s_{p,p}$. It is thus a near minimizer for the $B^s_{p,p}$ norm in the sense of (2.2.1). □

Now we are ready to state and prove [27] the following result dealing with the characterization of Besov spaces $B^t_{q,q}$, with $1/q = t/d + 1/2$ in terms of the nonlinear approximation error measured in the norm of $B^s_{p,p}$, with $1/p = s/d + 1/2$

Theorem 2.2.1: *Assume that the spaces $B^t_{q,q}$, $t - s = d(1/q - 1/p)$ admit a wavelet characterization of the type (4.2.4) for $t \in [s, s']$, $s' > s$. Then for $t \in]s, s'[, t - s = d(1/q - 1/p)$, we have the norm equivalence*

$$\|f\|_{B^t_{q,q}} \sim \|f\|_{B^s_{p,p}} + \|(2^{j(t-s)} dist_{B^s_{p,p}}(f, S_j))_{j\geq 0}\|_{\ell^q}. \tag{2.2.3}$$

Proof: [27] If $t > s$ and $t - s = d/q - d/p$, the norm equivalence (4.2.4) can be rewritten as

$$\|f\|_{B^t_{q,q}} \sim \|(\|c_\lambda\psi_\lambda\|_{B^s_{p,p}})_{\lambda\in\Lambda}\|_{\ell^q}, \tag{2.2.4}$$

where c_λ are the wavelet coefficients of f. We can then proceed in a similar way as in the particular case of approximation in L^2. Denoting by $(\varepsilon_n)_{n\geq 1}$ the decreasing rearrangement of the sequence $(\|c_\lambda\psi_\lambda\|_{B^s_{p,p}})_{\lambda\in\Lambda}$, we remark that since

$$n\varepsilon_n^q \leq \sum_{k\geq 1} \varepsilon_k^q \lesssim \|f\|^q_{B^t_{q,q}}, \tag{2.2.5}$$

we have the estimate

$$\varepsilon_n \lesssim n^{-1/q}\|f\|_{B^t_{q,q}}. \tag{2.2.6}$$

Denoting by Λ_N a set of indices defined as above, we obtain

$$\begin{aligned} \text{dist}_{B^s_{p,p}}(f, \Sigma_N) &\le \|f - \sum_{\lambda \in \Lambda_N} c_\lambda \psi_\lambda\|_{B^s_{p,p}} \\ &\lesssim (\sum_{n \ge N} \varepsilon_n^p)^{1/p} \\ &\lesssim N^{1/p-1/q} \|f\|_{B^t_{q,q}}. \end{aligned}$$

We have thus established the Jackson type estimate

$$\text{dist}_{B^s_{p,p}}(f, \Sigma_N) \le N^{-(t-s)/d} \|f\|_{B^t_{q,q}}. \tag{2.2.7}$$

If $f \in \Sigma_N$, then using Hölder inequality and (2.2.4) yields

$$\|f\|_{B^t_{q,q}} \lesssim N^{1/q-1/p} \|(\|c_\lambda \psi_\lambda\|_{B^s_{p,p}})_{\lambda \in \Lambda}\|_{\ell^p} = N^{(t-s)/d} \|f\|_{B^s_{p,p}}, \tag{2.2.8}$$

i.e. the corresponding Bernstein-type estimate.

Finally it remains to observe that Besov spaces $B^t_{q,q}$, $t - s = d/q - d/p$ are interpolation spaces: for $s < t < s'$, $t - s = d/q - d/p$ and $s' - s = d/p' - d/p$, we have the interpolation identity

$$B^t_{q,q} = [B^s_{p,p}, B^{s'}_{p',p'}]_{\theta,q}, \quad \theta = (t-s)/(s'-s), \tag{2.2.9}$$

which is a re-expression of $[\ell^p, \ell^{p'}]_{\theta,q} = \ell^q$, with $1/p = s/d + 1/2$ and $1/p' = s'/d + 1/2$.
By Theorem (2.1.1), with $X = B^s_{p,p}$ and $Y = B^t_{q,q}$, we thus obtained the norm equivalence (2.2.3). □

2.2.1 Linear versus Nonlinear

Let us consider (2.2.7). The analog linear result (1.3.13) tells us that the same error rate $\mathcal{O}(N^{-(t-s)/d})$ in $B^s_{p,p}$ is achieved by a linear method (i.e. with $N = N_j = \dim(V_j)$) for functions in $B^t_{p,p}$, which is a smaller space than $B^t_{q,q}$. It should be noted that, as t becomes large, the functions in the space $B^t_{p,p}$ become smooth in the classical sense, while $B^t_{q,q}$ might still contain discontinuous functions.

2.3 Nonlinear approximation in $L^p, 1 < p < \infty$

It is important to note that the result of previous section are easy to prove, due to the simple link existing between Besov spaces and ℓ^p spaces through

wavelet decomposition. The study of non linear approximation in L^p norm is more difficult since we cannot identify L^p to a $B^0_{p,p}$ for $p \neq 2$. Anyway switching from $B^0_{p,p}$ to L^p does not seriously affect the results of N-terms approximation for $1 < p < \infty$.
Following [27] we first recall two lemmas, due to Temlyakov [77], that allow to estimate the L^p norm of a linear combination of wavelets according to the size of the coefficients:

Lemma 2.3.1: *Let $1 < p < \infty$ and $\{\psi_\lambda\}_{\lambda\in\Lambda}$ a wavelet basis constructed from scaling functions in L^p. If E is a finite subset of Λ of cardinality $\#(E) < \infty$, then*

$$\|\sum_{\lambda\in E} c_\lambda\psi_\lambda\|_{L^p} \leq C\#(E)^{1/p} \sup_\lambda \|c_\lambda\psi_\lambda\|_{L^p}, \tag{2.3.1}$$

where C is independent of $\#(E)$.

Lemma 2.3.2: *Let $1 < p < \infty$ and $\{\psi_\lambda\}_{\lambda\in\Lambda}$ (resp. $\{\check{\psi}_\lambda\}_{\lambda\in\Lambda}$) a (resp. dual) wavelet basis constructed from scaling functions in L^p (resp. in $L^{p'}$, with $1/p' + 1/p = 1$). If E is a finite subset of Λ of cardinality $\#(E) < \infty$, then*

$$\|\sum_{\lambda\in E} c_\lambda\psi_\lambda\|_{L^p} \geq C\#(E)^{1/p} \inf_\lambda \|c_\lambda\psi_\lambda\|_{L^p}, \tag{2.3.2}$$

where C is independent of $\#(E)$.

Using the above two Lemmas it is possible to prove [77] that a near-best N-term approximation in L^p can be achieved by a simple thresholding procedure:

$$\|f - \sum_{\lambda\in\Lambda_N} c_\lambda\psi_\lambda\|_{L^p} \lesssim \mathrm{dist}_{L^p}(f, \Sigma_N),$$

where Λ_N is the set of indices corresponding to the N largest contributions $\|c_\lambda\psi_\lambda\|_{L^p}$.

2.3.1 Jackson and Bernstein result in L^p, $1 < p < \infty$

Now we prove [27] Jackson and Bernstein estimate for N-term approximation in L^p.

Theorem 2.3.1: *Let $1 < p < \infty$ and $\{\psi_\lambda\}_{\lambda\in\Lambda}$ a wavelet basis constructed from scaling functions in L^p. Assuming that the space $B^s_{q,q}$, $1/q = 1/p + s/d$ admits a wavelet characterization of the type (4.2.4), we have the Jackson estimate*

$$dist_{L^p}(f, \Sigma_N) \lesssim N^{-s/d}\|f\|_{B^s_{q,q}}, \tag{2.3.3}$$

and for $f \in \Sigma_N$, the Bernstein estimate

$$\|f\|_{B^s_{q,q}} \lesssim N^{s/d}\|f\|_{L^p}. \tag{2.3.4}$$

Proof: [27] Let $f \in B^s_{q,q}$. We remark that the norm equivalence (4.2.4) also writes

$$\|f\|_{B^s_{q,q}} \sim \|(\|c_\lambda \psi_\lambda\|_{L^p})_{\lambda\in\Lambda}\|_{\ell^q}. \tag{2.3.5}$$

In particular, we have

$$\#\{\lambda : \quad \|c_\lambda\psi_\lambda\|_{L^p} \geq \varepsilon\} \lesssim \varepsilon^{-q}\|f\|^q_{B^s_{q,q}}. \tag{2.3.6}$$

It follows that there exists a constant $C > 0$ (depending on the constant in the equivalence (2.3.5)), such that if we define

$$A_j = \{\lambda : \quad C2^{-j/q}\|f\|_{B^s_{q,q}} \leq \|c_\lambda\psi_\lambda\|_{L^p} \leq C2^{-(j-1)/q}\|f\|_{B^s_{q,q}}\}, \tag{2.3.7}$$

we then have

$$\#(A_j) \leq 2^j. \tag{2.3.8}$$

From (2.3.1) we can evaluate the L^p norm of $T_{A_j} f = \sum_{\lambda\in A_j} c_\lambda\psi_\lambda$ by

$$\|T_{A_j}\|_{L^p} \lesssim 2^{-j/q}\|f\|_{B^s_{q,q}}\#(A_j)^{1/p} \leq 2^{j(1/p-1/q)}\|f\|_{B^s_{q,q}}. \tag{2.3.9}$$

Now define $B_j = \cup_{l=0}^{j-1} A_l$. By (2.3.8), we have $\#(B_j) \leq 2^j$. For $S_j := \Sigma_N$, we thus have

$$\begin{aligned}
\mathrm{dist}_{L^p}(f, S_j) &\leq \|t - T_{B_j} f\|_{L^p} \\
&\leq \sum_{l\geq j} \|T_{A_l} f\|_{L^p} \\
&\lesssim \sum_{l\geq j} 2^{l(1/p-1/q)}\|f\|_{B^s_{q,q}} \\
&\lesssim 2^{j(1/p-1/q)}\|f\|_{B^s_{q,q}} = 2^{-js/d}\|f\|_{B^s_{q,q}}.
\end{aligned}$$

By the monotonicity of $\mathrm{dist}_{L^p}(f, S_N)$, this implies the direct estimate (2.3.3) for all N.

In order to prove the Bernstein estimate, we distinguish two cases: $p \geq 2$ and $p \leq 2$.
If $p \geq 2$, we have

$$\|f\|_{B^0_{p,p}} \sim \|(\|c_\lambda\psi_\lambda\|_{L^p})_{\lambda\in\Lambda}\|_{\ell^p} \lesssim \|f\|_{L^p}. \tag{2.3.10}$$

One way of checking (2.3.10) is to use an interpolation argument: the property holds when $p = 2$ for which one actually has the equivalence $\|f\|_{B^0_{2,2}} \sim \|f\|_{L^2}$ and $p = \infty$ since

$$\|c_\lambda\psi_\lambda\|_{L^\infty} \lesssim 2^{|\lambda|d/2}|c_\lambda| = 2^{|\lambda|d/2}|(f, \tilde{\psi}_\lambda)| \lesssim \|f\|_{L^\infty}. \tag{2.3.11}$$

In the case where $p \le 2$, let $f = \sum_{\lambda\in E} c_\lambda \psi_\lambda \in \Sigma_N$. We then estimate $\|f\|_{B^s_{q,q}}$ as follows:

$$\begin{aligned}
\|f\|_{B^s_{q,q}} &\lesssim \sum_{\lambda\in E} \|c_\lambda\psi_\lambda\|^q_{L^p} \\
&= \sum_{\lambda\in E} |c_\lambda|^q \|\psi_\lambda\|^p_{L^p} \|\psi_\lambda\|^{q-p}_{L^p} \\
&\lesssim \int_\Omega \sum_{\lambda\in E} |c_\lambda|^q |\psi_\lambda|^p 2^{d(1/2-1/p)(q-p)|\lambda|} \\
&\lesssim \int_\Omega \sum_{\lambda\in E} |c_\lambda|^q |\psi_\lambda|^p [2^{d(1/p-1/2)|\lambda|} |\psi_\lambda|]^{p-q} \\
&\lesssim \int_\Omega [Sf(x))]^q R_E(x) dx,
\end{aligned}$$

where we applied Hölder's inequality on sequences to obtain the last line. Here $Sf(x)$ is the square function and $R_E(x)$ can be estimated as follows:

$$\begin{aligned}
R_E(x) &= \Big(\sum_{\lambda\in E} [2^{d(1/p-1/2)|\lambda|} |\psi_\lambda(x)|]^{2(p-q)/(2-q)} \Big)^{(2-q)/2} \\
&\lesssim \Big(\sum_{\lambda\in E, \psi_\lambda(x)\neq 0} 2^{2d|\lambda|(p-q)/(2p-qp)} \Big)^{(2-q)/2} \\
&\lesssim 2^{j(x)d(1-q/p)},
\end{aligned}$$

where $j(x) = \max\{j \ge -1 : \quad x \in \mathrm{supp}(\psi_\lambda),\ \text{for some } \lambda \in E_j\}$. Using Hölder's inequality, we thus obtain

$$\begin{aligned}
\|f\|^q_{B^s_{q,q}} &\lesssim \|Sf\|^{q/p}_{L^p} (\int_\Omega 2^{j(x)d} dx)^{1-q/p} \\
&\lesssim \|f\|_{L^p} [\sum_{j\ge -1} \#(\Omega_j) 2^{jd}]^{1-q/p} \\
&\lesssim \|f\|_{L^p} [\sum_{j\ge -1} N_j]^{1-q/p} \\
&= N^{1-q/p} \|f\|_{L^p},
\end{aligned}$$

where $\Omega_j := \{x \in \Omega : \quad j(x) = j\}$. □

2.3.2 The main result

By using Theorem 2.3.1 and interpolation properties for the $\mathcal{A}^s_{p,r}$ spaces and for the Besov spaces $B^s_{q,q}$, $1/q = 1/p + s/d$, it is possible to prove [27] the analog of the Theorem 2.2.1 for Besov spaces $B^s_{q,q}$, with $s < 1$:

Theorem 2.3.2: *Let $1 < p < \infty$ and $\{\psi_\lambda\}_{\lambda\in\Lambda}$ a wavelet basis constructed from scaling functions in L^p. Assuming that for $0 < s < t$ the space $B^s_{q,q}$, $1/q = 1/p + s/d$ admits a wavelet characterization of the type (4.2.4), then one also has the norm equivalence*

$$\|f\|_{B^s_{q,q}} \sim \|f\|_{L^p} + \|(2^{js} dist_{L^2}(f, S_j))_{j\geq 0}\|_{\ell^q}. \tag{2.3.12}$$

Remark 2.3.1: *We refer to [27] and references therein for the cases $p \leq 1$ and $p = \infty$ which are not covered by the results of this section.*

2.4 Nonlinear approximation of sequences

In this section we re-obtain explicitly a result of nonlinear approximation in the Sobolev space H^s, by combining nonlinear approximation of sequences and the norm equivalences in terms of the summability properties of the wavelet coefficients.

We saw that in nonlinear approximation in a wavelet framework, a function $u \in L^2(\Omega)$, whose wavelet decomposition is $u = \sum_\lambda u_\lambda \psi_\lambda$, can be approximated by a lacunary series; that is by an approximation v to u, belonging to the non linear space

$$\Sigma_N = \{v = \sum_{\lambda\in\Lambda} v_\lambda \psi_\lambda : \quad \underline{v} = \{v_\lambda\}_{\lambda\in\Lambda} \in \sigma_N\}, \tag{2.4.1}$$

containing all the functions of $L^2(\Omega)$, whose wavelet coefficients belong to the set

$$\sigma_N = \{\underline{v} \in \ell^2(\Lambda) : \quad \#\{\lambda : v_\lambda \neq 0\} \leq N\}$$

of sequences with at most N elements different from zero. The set Σ_N contains the functions of $L^2(\Omega)$, which can be expressed as a linear combination of at most N wavelets. A nonlinear projector

$$\mathbb{P}_N : L^2(\Omega) \to \Sigma_N$$

can be built as follows: given $u = \sum_\lambda u_\lambda \psi_\lambda$, let us sort the sequence $\{|u_\lambda|\}_{\lambda\in\Lambda}$ in decreasing order. We denote $\{|u_{\lambda(k)}|\}_{k\in\mathbb{N}}$ the coefficient of rank k:

$$|u_{\lambda(k)}| \geq |u_{\lambda(k+1)}|, \quad \text{with } k > 0.$$

Hence the image $\mathbb{P}_N(u)$ is defined by:

$$\mathbb{P}_N(u) = \sum_{n=1}^{N} u_{\lambda(n)} \psi_{\lambda(n)},$$

that is only the N greatest (in absolute value) coefficients of u are retained. By abuse of notation we will also indicate by

$$\mathbb{P}_N : \ell^2 \to \sigma_N$$

the operator associating to the sequence $\underline{u} = \{u_\lambda\}$, the coefficients of the function $\mathbb{P}_N(\sum_\lambda u_\lambda \psi_\lambda)$. The accuracy of the corresponding approximation is directly related to ℓ^τ_w regularity of the sequence of coefficients of u, as stated by the following theorem [47], [48]:

Theorem 2.4.1: *Let $u = \sum_{\lambda\in\Lambda} u_\lambda \psi_\lambda$. If $\underline{u} = \{u_\lambda\}_\lambda \in \ell^\tau_w$, with τ such that $0 < \tau < 2$, then*

$$\|\underline{u} - \mathbb{P}_N \underline{u}\|_{\ell^2} \lesssim \inf_{\underline{w}\in\sigma_N} \|\underline{u} - \underline{w}\|_{\ell^2} \lesssim N^{-(\frac{1}{\tau}-\frac{1}{2})} \|\underline{u}\|_{\ell^\tau},$$

where the implicit constants in the bounds depend only on τ.

Proof: We sketch the proof. We have

$$\#\{\lambda \in \Lambda : \quad |c_\lambda| \geq \varepsilon\} \leq M\varepsilon^{-\tau}.$$

Let $\Lambda_j := \{\lambda : \quad 2^{-j} < |c_\lambda| < 2^{-j+1}\}$. Then for each $k = 1, 2, \ldots$, we have

$$\sum_{j=-\infty}^{k} \#\Lambda_j \leq CM2^{k\tau} \tag{2.4.2}$$

with C depending only on τ.

Let $S_j := \sum_{\lambda\in\Lambda_j} c_\lambda \psi_\lambda$ and $T_k := \sum_{j=-\infty}^{k} S_j$. Then $T_k \in \Sigma_N$ with $N = CM2^{k\tau}$. We have

$$\|f - T_k\|_{L^2} \leq \sum_{j=k+1}^{\infty} \|S_j\|_{L^2}. \tag{2.4.3}$$

We fix $j > k$ and estimate $\|S_j\|_{L^2}$. Since $|c_\lambda| \leq 2^{-j+1}$ for all $\lambda \in \Lambda_j$, we have from Lemma 2.3.1 and (2.4.2),

$$\|S_j\|_{L^2} \leq C2^{-j}\#\Lambda_j^{1/2} \leq CM^{1/2}2^{j(\tau/2-1)}.$$

We therefore conclude from (2.4.3) that

$$\|f - T_k\|_{L^2} \leq CM^{1/2} \sum_{j=k+1}^{\infty} 2^{j(\tau/2-1)} \leq CM(M^{1/\tau}2^k)^{\tau/2-1},$$

because $\tau/2 - 1 < 0$. In other words for $N \simeq M2^{k\tau}$, we have

$$\inf_{\underline{w}\in\sigma_N} \|\underline{u} - \underline{w}\|_{\ell^2} \leq CMN^{1/2-1/\tau}.$$

To conclude it is sufficient to recall the near optimality of the nonlinear projector $\mathbb{P}_N$ in L^2:

$$\|\underline{u} - \mathbb{P}_N \underline{u}\|_{\ell^2} \lesssim \inf_{\underline{w}\in\sigma_N} \|\underline{u} - \underline{w}\|_{\ell^2}.$$

□

In particular, as $\ell^\tau \subset \ell^\tau_w$, if τ is such that $\frac{1}{\tau} = \frac{r}{d} + \frac{1}{2}$, using norm equivalence (4.2.4), we obtain

$$\|\sum_\lambda u_\lambda \psi_\lambda\|_{B^r_{\tau,\tau}(\Omega)} \simeq \|\underline{u}\|_{\ell^\tau} \tag{2.4.4}$$

and from Theorem 2.4.1 we recover the above result of nonlinear approximation in L^2: if u belongs to $B^r_{\tau,\tau}(\Omega)$, with τ such that $\frac{1}{\tau} = \frac{r}{d} + \frac{1}{2}$, then

$$\inf_{w\in\Sigma_N} \|u - w\|_{L^2(\Omega)} \lesssim \|u - \mathbb{P}_N u\|_{L^2(\Omega)} \lesssim N^{-(\frac{1}{\tau}-\frac{1}{2})}\|u\|_{B^s_{\tau,\tau}(\Omega)}.$$

In other words, once we normalise in $L^2(\Omega)$ the wavelet basis $\{\psi_\lambda\}$, the natural functional setting of nonlinear approximation in $L^2(\Omega)$ is the scale of Besov spaces $B^r_{\tau,\tau}(\Omega)$.

2.4.1 Nonlinear approximation in H^s

Let us now consider a rescaled version $\{\check{\psi}_\lambda\}_\lambda$ of the wavelet basis $\{\psi_\lambda\}_\lambda$, where $\check{\psi}_\lambda = \rho^{-js}\psi_\lambda$, for $\lambda \in \Lambda_j$. If τ is such that $\frac{1}{\tau} = \frac{r}{d} + \frac{1}{2}$, from norm equivalence (4.2.4) we obtain:

$$\|\sum_\lambda u_\lambda \check{\psi}_\lambda\|_{B^{r+s}_{\tau,\tau}(\Omega)} \simeq \|\underline{u}\|_{\ell^\tau}. \tag{2.4.5}$$

Applying now Theorem 2.4.1 and norm equivalence (1.5.13) for Sobolev spaces to the normalised sequence $\underline{u}$, we obtain the following result of non linear approximation in $H^s(\Omega)$:

Corollary 2.4.1: *Let $u \in B^{s+r}_{\tau,\tau}(\Omega)$, with τ such that $1/\tau = r/d + 1/2$, then*

$$\|u - \mathbb{P}_N u\|_{H^s(\Omega)} \lesssim \inf_{w\in\sigma_N} \|u - w\|_{H^s(\Omega)} \lesssim N^{-(\frac{1}{\tau}-\frac{1}{2})}\|u\|_{B^{s+r}_{\tau,\tau}(\Omega)},$$

where the implicit constants in the bounds depend only on τ.

That is when we consider nonlinear approximation in $H^s(\Omega)$ the natural functional setting is the scale of Besov spaces $B^{r+s}_{\tau,\tau}(\Omega)$, where τ is defined by the relation $\frac{1}{\tau} = \frac{r}{d} + \frac{1}{2}$.

2.5 Towards adaptive wavelet methods

Wavelet bases are being increasingly used in the numerical solution of partial differential and integral equations. There are many aspects in a discretization procedure for such equations that can benefit from the features of these bases. Wavelets share with other multilevel methods the capability of easily preconditioning the discrete realizations of symmetric positive definite operators. More typical of wavelets is their orthogonality to certain classes of smooth functions (e.g. polynomials), a feature that can be exploited in the compression of dense matrices and, in a more general context, in the design of adaptive discretization strategies. The finite-dimensional space, which is used in a Galerkin-type approximation, is adaptively constructed by including in it precisely those wavelet basis functions that have the potential of representing the most significant structures of the solution. From this point of view, adaptive wavelet methods can be viewed as meshless methods or space refinement methods, with a highly flexible mechanism for adding and removing degrees of freedom. Nonlinear approximation provides a natural benchmark for an adaptive scheme: if $\|\cdot\|_X$ is the norm where we measure the error between the solution u of a PDE and its numerical approximation and if it holds $\|u-u_N\|_X \lesssim N^{-s}$ for a family of N-term wavelet approximations $u_N \in \Sigma_N$, then an optimal adaptive scheme should provide N-terms approximate solutions $\tilde{u}_N \in \Sigma_N$, such that one also has $\|u-\tilde{u}_N\|_X \lesssim N^{-s}$.

2.5.1 Wavelet preconditioning

One of the interest of multiscale discretizations is the possibility of preconditioning large systems which arise from elliptic operator equations. A general setting for such equations is the following: H is a Hilbert space embedded in $L^2(\Omega)$ and $a(\cdot,\cdot)$ is a bilinear form on $H \times H$ such that

$$a(u,u) \sim \|u\|_H^2. \tag{2.5.1}$$

Let H' be the dual space of H. Given $f \in H'$, we search for $u \in H$ such that

$$a(u,v) = (f,v), \quad \text{for all } v \in H. \tag{2.5.2}$$

It is well known that from Lax-Milgram lemma, this problem has a unique solution. If we define the operator A by

$$(Au,v) = a(u,v), \quad \text{for all } v \in H,$$

the equivalence (2.5.1) implies that A is an isomorphism from H to H', so that u is also the unique solution in H of

$$Au = f. \tag{2.5.3}$$

If V_h is a subspace of H, the Galerkin approximation of u in V_h is classically defined by $u_h \in V_h$ such that

$$a(u_h, v_h) = (f, v_h), \quad \text{for all } v_h \in V_h. \tag{2.5.4}$$

If V_h is finite dimensional, the approximated problem (2.5.4) amounts in solving a linear system. Here we are interested in the situation where H is an L^2 Sobolev space: classical instances are given by the Poisson equation $-\Delta u = f$ with Dirichelet conditions, where $H = H_0^1$, or the Helmholtz equation $u - \Delta u = f$, with Neumann boundary conditions, where $H = H^1$. For such equations it is known that the matrices resulting from Galerkin discretizations in the finite element spaces are ill-conditioned, i.e. their conditions number grows like h^{-2s}, where h is the mesh size and s is the order of the corresponding Sobolev space H, where $2s$ is the order of the elliptic operator. We remark that elliptic equations involving integral operators of negative order also enter the above class of problems [41].

The use of multilevel methods for preconditioning such matrices is linked to the possibility of characterizing the L^2 Sobolev space H^s (possibly with boundary conditions) by means of wavelet coefficients:

$$\|f\|_{H^s}^2 \sim \sum_{\lambda \in \Lambda} 2^{2s|\lambda|} |c_\lambda|^2. \tag{2.5.5}$$

Let us consider the Galerkin discretization (2.5.4) on a multiresolution approximation space $V_J \subset H$, corresponding to a mesh size 2^{-J} and we denote by u_J the corresponding solution. For the computation of u_J we use the multiscale basis $\{\psi_\lambda\}_{|\lambda|<J}$ and we obtain a system

$$A_J U_J = F_J, \tag{2.5.6}$$

where U_J is the coordinate vector of u_J in the basis $\{\psi_\lambda\}_{|\lambda|<J}$, $F_J = \{(f, \psi_\lambda)\}_{|\lambda|<J}$ and $A_J = \{(A\psi_\lambda, \psi_\mu)\}_{|\lambda|,|\mu|<J}$ is the stiffness matrix.
A result relating the norm equivalence to wavelet preconditioning was first proposed in [59]:

Theorem 2.5.1: *Consider the diagonal matrix D_J with $(D_J)_{\lambda,\mu} = (2^{2s|\lambda|}\delta_{\lambda,\mu})$, where $|\lambda|, |\mu| < J$. The two following statements are equivalent:*

(i) H is characterized by a norm equivalence

$$\|f\|_H^2 \sim \sum_{\lambda \in \Lambda} 2^{2s|\lambda|} |c_\lambda|^2, \tag{2.5.7}$$

(ii) The condition number $\mathcal{K}(D_J^{-1}A_J) = \mathcal{K}((D_J^{-1/2}A_J D^{-1/2})$ is bounded independently of J.

Proof: The property (ii) is equivalent to

$$(D_J U, U) \sim (A_J U, U), \tag{2.5.8}$$

with constants independent of the vector U and the scale level J. From the definition of A_J, this can also be expressed by

$$a(v_J, v_J) \sim \sum_{|\lambda|<J} 2^{2s|\lambda|}|c_\lambda|^2, \tag{2.5.9}$$

for all $v_J = \sum_{|\lambda|<J} c_\lambda \psi_\lambda$ in V_J. Since $a(u,u) \sim \|u\|_H^2$, (2.5.9) is equivalent to (2.5.7) for all $f \in V_J$. By density of the V_J multiresolution spaces, this is equivalent to (2.5.7) for all $f \in H$.

□

2.5.2 Compression of operators

Another advantage of the wavelet basis is the sparse structure that results from the multiscale discretization of most operators involved in partial differential and integral equations and the good properties such operators exhibit when apply on functions that also have a sparse multiscale representation. Given an operator A acting on functions defined on a domain $\Omega \subset \mathbb{R}$ and a wavelet basis $\{\psi_\lambda\}_{\lambda\in\Lambda}$, we are interested in evaluating the entries

$$m_{\lambda,\mu} = (A\psi_\lambda, \psi_\mu). \tag{2.5.10}$$

In order to treat different examples within a unified framework [27], we shall now introduce general classes of matrices associated to operators through wavelet bases.

Definition *Let $s \in \mathbb{R}$ and $\alpha, \beta > 0$. A matrix M belongs to the class $\mathcal{M}^s_{\alpha,\beta}$ if and only if its entries satisfy the estimate*

$$|m_{\lambda,\mu}| \le C_M 2^{s(|\lambda|+|\mu|)} 2^{-(d/2+\alpha)||\lambda|-|\mu||} d(\lambda,\mu)^{-(d+\beta)} \tag{2.5.11}$$

where $d(\lambda,\mu) := 1 + 2^{\min\{|\lambda|,|\mu|\}} dist(supp(\psi_\lambda), supp(\psi_\mu))$. We denote by $\mathcal{M}_{\alpha,\beta}$ this class when $s = 0$.

The factor $2^{s(|\lambda|+|\mu|)}$ describes the growth or decay (depending on the sign of s, hence depending on the order of the operator) of the entries of M along the diagonal, i.e. the multiplicative effect of the operator on the different scales. The parameter s thus indicates the order of the operator: for instance $s = 2$ if $A = \Delta$

Remark 2.5.1: *Note that the diagonal bi-infinite matrix D_s, with $(D_s)_{\lambda,\mu} = 2^{s|\lambda|}\delta_{\lambda,\mu}$ allows to renormalize M, in the sense that $\tilde{M} = D_s^{-1} M D_s^{-1}$ satisfies the estimate (2.5.11) with $s = 0$, i.e. belongs to the class $\mathcal{M}_{\alpha,\beta}$. Such a renormalization is exactly the preconditioning process on finite matrix described in the previous section.*

The factor $2^{-||\lambda|-|\mu||(d/2+\alpha)}$ describes the decay of the entries away from the diagonal blocks corresponding to $|\lambda| = |\mu|$. Finally the factor $(1 + d(\lambda, \lambda'))^{-\beta}$ describes the decay of the entries away from the diagonal within each blocks corresponding to fixed values of $|\lambda|$ and $|\mu|$.

A basic tool for the study of the classes $\mathcal{M}^s_{\alpha,\beta}$ is the Schur lemma that we recall below.

Lemma 2.5.1: *Let $M = (m_{\lambda,\mu})_{\lambda,\mu\in\Lambda}$ be a matrix indexed by Λ. Assume that there exists a sequence of positive numbers $(\omega_\lambda)_{\lambda\in\Lambda}$ and a constant C such that*

$$\sum_{\mu\in\Lambda} \omega_\mu |m_{\lambda,\mu}| + \sum_{\mu\in\Lambda} \omega_\mu |m_{\mu,\lambda}| \leq C\omega_\lambda \tag{2.5.12}$$

for all $\lambda \in \Lambda$. Then M defines a bounded operator in $\ell^2(\Lambda)$ with $\|M\| \leq C$.

A first application of the Schur lemma is the following result [27].

Theorem 2.5.2: *If $\alpha, \beta > 0$, then any $M \in \mathcal{M}_{\alpha,\beta}$ defines a bounded operator in $\ell^2(\Lambda)$. In turn, any matrix $M \in \mathcal{M}_{\alpha,\beta}$ together with a Riesz basis $(\psi_\lambda)_\lambda \in \Lambda$ of L^2 defines an L^2 bounded operator A represented by M in this basis.*

Proof: [27] We shall use the Schur Lemma with $\omega_\lambda = 2^{-d|\lambda|/2}$. From (2.5.11), we first obtain

$$\begin{aligned}
\omega_\lambda^{-1} \sum_{\mu\in\Lambda} \omega_\mu |m_{\lambda,\mu}| &\lesssim 2^{d|\lambda|/2} \sum_{\mu\in\Lambda} 2^{-d|\mu|/2} 2^{-(d/2+\alpha)||\lambda|-|\mu||} d(\lambda,\mu)^{-(d+\beta)} \\
&\lesssim 2^{d|\lambda|/2} \sum_{j\geq 0} 2^{-dj/2} 2^{-(d/2+\alpha)||\lambda|-j|} \sum_{|\mu|\in\Lambda_j} d(\lambda,\mu)^{-(d+\beta)}.
\end{aligned}$$

Since $\beta > 0$ the last factor $\sum_{|\mu|\in\Lambda_j} d(\lambda,\mu)^{-(d+\beta)}$ is bounded by a uniform constant if $j \leq |\lambda|$ and by $2^{d(j-|\lambda|)}$ if $j \geq |\lambda|$. Splitting the sum in j according to these two cases, we finally obtain

$$\begin{aligned}
\omega_\lambda^{-1} \sum_{\mu\in\Lambda} \omega_\mu |m_{\lambda,\mu}| &\lesssim \sum_{j\geq 0} 2^{d||\lambda|-j|/2} 2^{-(d/2+\alpha)||\lambda|-j|} \\
&\lesssim 2 \sum_{l\geq 0} 2^{-\alpha l} < \infty,
\end{aligned}$$

which shows that (2.5.12) holds with such weights. □

Let us now introduce the weighted spaces

$$\ell_t^2(\Lambda) := \{(c_\lambda)_{\lambda\in\Lambda} : \quad \|(c_\lambda)_{\lambda\in\Lambda}\|_{\ell_t^2}^2 := \sum_{\lambda\in\Lambda} 2^{2t|\lambda|}|c_\lambda|^2 < \infty\}.$$

We already noted that for $M \in \mathcal{M}_{\alpha,\beta}^s$, for $s \neq 0$, the preconditioned matrix

$$\tilde{M} = D_s^{-1} M D_s^{-1},$$

with $D_s = (2^{s|\lambda|}\delta_{\lambda,\mu})_{\lambda,\mu}$ belongs to the class $\mathcal{M}_{\alpha,\beta}$. We remark that D_s defines an isomorphism from ℓ_t^2 to ℓ_{t+s}^2. Combining these remarks with the above Theorem, we can describe the action of $M = D_s\tilde{M}D_s$, as follows [27]:

Corollary 2.5.1: *If $\alpha, \beta > 0$, then any $M \in \mathcal{M}_{\alpha,\beta}^s$ defines a bounded operator from ℓ_s^2 to ℓ_{-s}^2. In turn, any matrix $M \in \mathcal{M}_{\alpha,\beta}^s$ together with a wavelet basis $(\psi_\lambda)_\lambda \in \Lambda$ which characterize $H^s(\Omega)$ and $H^{-s}(\Omega)$ (possibly with boundary conditions) defines a bounded operator A from $H^s(\Omega)$ to $H^{-s}(\Omega)$, represented by M in this basis.*

The next step is to show that the estimate (2.5.11) allows to compress the matrices in the class $\mathcal{M}_{\alpha,\beta}^s$ by discarding certain entries. We first consider the case $s = 0$ [27].

Theorem 2.5.3: *Let $\mathcal{M}_{\alpha,\beta}$ and $t < \inf(\alpha/d, \beta/d)$. For all $N \geq 0$ one can discard the entries of M in such a way that the resulting matrix M_N has N nonzero entries per rows and columns and satisfies*

$$\|M - M_N\| \lesssim N^{-t}, \tag{2.5.13}$$

in the operator norm $\ell^2(\Lambda)$.

Proof: [27] We first truncate the matrix M in scale: for a given $J > 0$, we discard $m_{\lambda,\mu}$ if $||\lambda| - |\mu|| \geq J$. Denoting by A_J the resulting matrix, we can use the same technique as in the proof of the above Theorem (Schur lemma with weights $2^{d|\lambda|/2}$) to measure the error $\|M - A_J\|$ in the operator norm. By a very similar computation, we obtain

$$\|M - A_J\| \lesssim \sum_{l\geq J} 2^{-\alpha l} \lesssim 2^{-\alpha J}.$$

We next truncate A_J in space, by preserving in each remaining block of A_J the entries $m_{\lambda,\mu}$ such that $d(\lambda,\mu) \leq k(||\lambda| - |\mu||)$ where the function k is to be determined. We denote by B_J the resulting matrix. Using again the Schur lemma in the same way as in the proof of the above Theorem, we evaluate the error $\|A_J - B_J\|$ by the supremum in λ of

$$\omega_\lambda^{-1} \sum_{j=|\lambda|-J}^{|\lambda|+J} \sum_{\mu\in\Lambda_j} \omega_\mu |b^J_{\lambda,\mu} - m_{\lambda,\mu}|,$$

and we obtain an estimated contribution of $2^{-\alpha J}$ for each term in j by taking $k(l) = 2^{J\alpha/\beta}2^{l(1-\alpha/\beta)}$. The total error is thus estimated by

$$\|M - B_J\| \lesssim J2^{-\alpha J},$$

with the number of nonzero entries per rows and columns in B_J is estimated by $N(J) \lesssim \sum_{l=0}^{J} k(l)^d$. In the case where $\alpha > \beta$ (resp. $\beta < \alpha$) this sum is bounded by the first term $K(0)^d = 2^{Jd\alpha/\beta}$ (resp. last term $K(J) = 2^{dJ}$). In the case $\alpha = \beta$, we obtain $N(J) \lesssim J2^{dJ}$. In all cases, it follows from the evaluation of $N(J)$ and the error $\|M - B_J\| \lesssim J2^{-\alpha J}$ that

$$\|M - B_J\| \lesssim N(J)^{-t},$$

if t is such that $t < \inf\{\alpha/d, \beta/d\}$. Since J ranges over all positive integers, this is enough to conclude the proof. □

We can derive simple consequences of this result concerning the sparsity of the operators in the classes $\mathcal{M}^s_{\alpha,\beta}$ by the same considerations as for the study of boundedness properties: for $M \in \mathcal{M}^s_{\alpha,\beta}$, we apply the compression process of the above Theorem to the preconditioned matrix $\tilde{M} = D_s^{-1} M D_s^{-1}$. Denoting by $\tilde{M}_N$ the compressed version of the matrix $\tilde{M}$, we then define $M_N = D_s \tilde{M}_N D_s$. This new matrix has also N entries per rows and columns and approximates M in the sense expressed by the following Corollary.

Corollary 2.5.2: *Let $\mathcal{M}_{\alpha,\beta}$ and $t < \inf(\alpha/d, \beta/d)$. For all $N \geq 0$ one can discard the entries of M in such a way that the resulting matrix M_N has N nonzero entries per rows and columns and satisfies*

$$\|M - M_N\| \lesssim N^{-t}, \tag{2.5.14}$$

in the norm of operators from $\ell^2_s(\Lambda)$ to $\ell^2_{-s}(\Lambda)$.

The last result [27] concerns the application of sparse matrices of the type that we have introduced in this section on sparse vectors, possibly resulting of adaptive multiscale discretizations for the solution of PDE's. In the context of nonlinear approximation theory, the sparsity of such a vector is precisely described by the rate of decay of the error of N-term approximation: an infinite vector U has a degree of sparsity $t > 0$ in some metric X, if there exists a sequence of vectors $(U_N)_{N\geq 0}$ such that U_N has N nonzero coordinates and such that

$$\|U - U_N\|_X \lesssim N^{-t}. \tag{2.5.15}$$

In the case where $X = \ell^2$ the vectors U_N are simply obtained by retaining the N largest coordinates of U and property (2.5.15) is equivalent to $U \in \ell^p_w$ with $1/p = 1/2 + t$.

Theorem 2.5.4: *The matrices $M \in \mathcal{M}_{\alpha,\beta}$ define bounded operator in $\ell^2 \cap \ell^p_w$, for $1/p = 1/2+t$ and $t < \min\{\alpha/d, \beta/d\}$. In other words a vector U of sparsity in ℓ^2 is mapped by M onto a vector $V = MU$ with the same property.*

Proof: Following [27] we will directly construct an N-term approximation to $V = MU$ from the N-term approximation of U. For $j \geq 0$ we denote by U_j the vector that consists of the 2^j largest coordinates of U. From the assumptions we know that

$$\|U - U_j\|_X \lesssim 2^{-tj}.$$

Fixing $r \in]t, \min\{\alpha/d, \beta/d\}[$, we can define, according to the above Theorem, truncated operators M_j such that M_j has at most $2^{(1-\varepsilon)j}$ nonzero entries per rows and columns with $\varepsilon > 0$ and

$$\|M - M_j\| \lesssim 2^{-rj}.$$

We define an approximation to $V = MU$ by

$$V_j := A_jU_0 + A_{j-1}(U_1 - U_0) + \ldots + A_0(U_j - U_{j-1}) = A_jU_0 + \sum_{l=1}^{j} A_{j-l}(U_l - U_{l-1}).$$

It is possible to evaluate the number of nonzero entries of V_j by

$$N(j) \leq 2^{(1-\varepsilon)j} + \sum_{l=1}^{j} 2^{(1-\varepsilon)j-l}2^{l-1} \lesssim 2^j.$$

Finally we can evaluate the error of approximation as follows.

$$\begin{aligned}
\|V - V_j\| &= \|M(U - U_j) + \sum_{l=0}^{j-1}(M - M_l)(U_{j-l} - U_{j-l-1}) + (M - M_j)U_0\| \\
&\leq \|M(U - U_j)\| + \sum_{l=0}^{j-1} \|M - M_l\|\|U_{j-l} - U_{j-l-1}\| + \|M - M_j\|\|U_0\| \\
&\lesssim 2^{-tj} + 2^{-tj}\sum_{l=0}^{j-1} 2^{(t-r)l} + 2^{-rj} \\
&\lesssim 2^{-tj}.
\end{aligned} \tag{2.5.16}$$

Since j ranges over all possible integers, we have thus proved that $V \in \ell^p_w$, with $1/p = 1/2 + t$. □

We can again derive [27] an immediate Corollary.

Corollary 2.5.3: *Let $M \in \mathcal{M}^s_{\alpha,\beta}$ and U a vector of sparsity t in ℓ^2_s with $s < \inf\{\alpha/d, \beta/d\}$. Then $V = MU$ has sparsity t in the dual space ℓ^2_{-s}.*

Part II

Adaptive Wavelet Methods

Chapter 3

ADAPTIVE SCHEMES FOR LINEAR EQUATIONS

And Minas Morgul answered. There was a flare of livid lightnings: forks of blue flame springing up from the tower and from the encircling hills into the sullen clouds. The earth groaned; and out of the city there came a cry.(...) As the terrible cry ended, falling back through a long sickening wail to silence, Frodo slowly raised his head. Across the narrow valley the walls of the evil city stood, and its cavernous gate, shaped like an open mouth with gleaming teeth, was gaping wide. And out of the gate an army came. All that host was clad in sable, dark as the night. Against the wan walls and the luminous pavement of the road Frodo could see them, small black figures in rank upon rank, marching swiftly and silently, passing outwards in an endless stream. Before them went a great cavalry of horsemen moving like ordered shadows, and at their head was one greater than all the rest: a Rider all black, save that on his hooded head he had an helm like a crown that flickered with a perilous light.(...) Frodo waited, and as he waited, he felt, more urgent than ever before, the command that he should put on the Ring. But he knew that the Ring would only betray him, and that he had not, even if he put it on, the Power to face the Morgul-king – not yet.

(J.R.R. Tolkien, The Two Towers)

3.1 Introduction

In the study of numerical algorithms for the solution of PDE's, adaptive methods are commonly used when the solution u exhibits localized singularities. Typically such methods use informations at a given step to produce, for the next iteration, a new approximation to u with a finer resolution near the singularities. Adaptive procedures are particular form of nonlinear approximation of the unknown solution u. The approximation space in which we look for the numerical solution is not a linear space, since the degrees of freedom are not chosen a priori, but depend on the solution u. In this context, what one would like to do is to fix the number N of degrees of freedoms and to design an algorithm able to find the best possible approximation of the solution u among all possible approximations (of a give type) with N degrees of freedoms, and this, with a computational cost growing only linearly with N.

In the wavelet framework, thanks to nonlinear approximation techniques, it is possible to approximate a given function with an element of the nonlinear space Σ_N containing functions, whose wavelet expansions have at most N non-vanishing coefficients. When the function to be approximated is known, it is easy to define a nonlinear projection $\mathbb{P}_N$ which allows to get, given u, the best N-terms wavelet approximation, by simply keeping the N biggest (in absolute value) coefficient (properly rescaled, depending on the norm in which the approximation is measured).

Unfortunately, in the case we are interested in, the function to be approximated is not known. If we consider for instance the problem of building a numerical scheme to approximate the solution u to the partial differential equation

$$-\mathrm{div}(\mathbf{a}\nabla u) = f \quad \text{in } \Omega \subset \mathbb{R}^d, \qquad\qquad u = 0 \quad \text{on } \partial\Omega, \tag{3.1.1}$$

one possibility is to look for iterative approximation schemes in which by definition the iterates belong to the nonlinear space Σ_N.

On an abstract level, we write down a convergent iterative scheme for the continuous problem, and then we force the iterate to belong to Σ_N, by simply projecting it onto such space using the simple approximation strategy for known functions described above.

On a practical point of view, this is achieved by following the "new approach" (see Introduction):

1. transform the given PDE into an equivalent infinite linear system whose unknown is the infinite vector of wavelet coefficients of the unknown solution.

2. write down a convergent iterative scheme for the infinite linear system.

3. at each iteration approximate (possibly adaptively) the **infinite** matrix-vector multiplication by a **finite** matrix-vector, by performing a prevision step, aiming at individuating a priori finite number of relevant coefficients, which will be possibly picked up by the nonlinear projection step, while the remaining will be most certainly discarded.

Convergence of algorithms of this type strongly relies on a key feature of wavelets, namely what is usually referred to as *wavelet preconditioning* (see Section 2.5.1): equation (3.1.1) can be rewritten in an equivalent form as infinite system:

$$\mathcal{A}\underline{u} = \underline{g}, \qquad \text{with } \mathcal{A}, \mathcal{A}^{-1} \in \mathcal{L}(\ell^2, \ell^2), \tag{3.1.2}$$

where $\underline{u}$ is the infinite array of the coefficients u_λ of the unknown solution $u = \sum_\lambda u_\lambda \check{\psi}_\lambda$, expressed with respect to the basis $\{\check{\psi}_\lambda\}_\lambda$ obtain by suitably renormalizing the basis $\{\psi_\lambda\}_\lambda$.

Based on these ideas, we present a computable scheme, which we prove to be convergent to an approximate solution with almost the same approximation rate as the one which is achieved, (under the same Besov smoothness assumptions) by the nonlinear approximation of a given function. In particular, as mentioned in the Introduction, we deal directly with the problem of the approximate application of a certain class of linear operators in wavelet coordinates, providing an explicit strategy.

Let us for the moment assume that we have selected two sequences N_n and ϵ_n

$$N_0 < N_1 < \cdots N_{n_1} < N_{n_2} < \cdots < N_{\bar{n}} = N$$
$$\epsilon_0 > \epsilon_1 > \cdots \epsilon_{n_1} > \epsilon_{n_2} > \cdots > \epsilon_{\bar{n}}$$

We are interested in schemes of the following type, where, by abuse of notation, we will also denote by $\mathbb{P}_N$, the operator that associates to the coefficients of a function u, the vector of coefficients of its nonlinear projection $\mathbb{P}_N(u)$:

Nonlinear Richardson

Step 1. Initialization: set $\underline{u}^{(0)} = 0$.

Step 2. Until $n \leq \bar{n}$, repeat

Step 2.1 Prevision: select a finite dimensional set $V^{(n)} \subset \Lambda$ such that, denoted by S_n the linear subspace of ℓ^2 of the form $S_n = \{\underline{v} \in \ell^2 : \; v_\lambda = 0, \; \lambda \notin V^{(n)}\}$, we have

$$\inf_{\underline{v} \in S_n} \|\underline{u}^{(n)} + \theta \underline{r}^{(n)} - \underline{v}\|_{\ell^2} \leq C\epsilon_n,$$

where $\underline{r}^{(n)} = \underline{g} - \mathcal{A}\underline{u}^{(n)}$ denotes the residual and where the constant C depends only on initial data.

Step 2.2 Compute an approximation

$$\underline{\tilde{r}}^{(n)} \in \{\underline{v} = (v_\lambda)_{\lambda\in\Lambda} \in \ell^2 : \ \lambda \notin V^{(n)} \Rightarrow v_\lambda = 0\}$$

of the residual $\underline{r}^{(n)}$, in such a way that $\|\underline{r}^{(n)} - \underline{\tilde{r}}^{(n)}\|_{\ell^2} \leq C\epsilon_n$.

Step 2.3 Projection: set

$$\underline{u}^{(n+1)} = \mathbb{P}_{N_n}(\underline{u}^{(n)} + \theta\underline{\tilde{r}}^{(n)}).$$

Step 2.4 Update: $n+1 \to n$

The construction of a suitable set $V^{(n)}$, which is possible thanks to the good space frequency localization properties of wavelets, is a necessary step for an algorithm of this type to be practically implemented. Neverless we postpone such an issue and we first concentrate on the study of the influence of the nonlinear projection step in the Richardson type algorithm.

The outline is as follows: in section 3.2 we recall some useful results about wavelets and nonlinear approximation, in section 3.3 we state the problem to solve, in section 3.4 we discuss a non-computable abstract scheme to study the influence of the nonlinear projector step and finally in section 3.6.3 we analyze the **Nonlinear Richardson** scheme and the construction of the prevision set $\Lambda^{(n)}$.

3.2 Notations and Preliminary results

In the following we will employ the notation $A \lesssim B$ to indicate that the quantity A is bounded from above by a positive constant times the quantity B, while $A \simeq B$ will stand for $A \lesssim B \lesssim A$.

For simplicity let us fix the following functional setting: let $\Omega \subseteq \mathbb{R}$ be a bounded domain, and suppose we are given a Riesz basis $\{\psi_\lambda\}_{\lambda\in\Lambda}$, $\Lambda = \cup_{j=0}^\infty \Lambda_j$, for $L^2(\Omega)$, such that, for some parameter $\rho > 1$, the following norm equivalence for the Besov spaces $B^s_{p,q}(\Omega)$ holds for all s, p, q, $0 \leq s \leq S$, $0 < p < \infty$, $q > 0$:

$$\|\sum_\lambda u_\lambda\psi_\lambda\|^q_{B^s_{p,q}(\Omega)} \simeq \sum_j \rho^{q(s+\frac{d}{2}-\frac{d}{p})j}\left(\sum_{\lambda\in\Lambda_j}|u_\lambda|^p\right)^{q/p}, \tag{3.2.1}$$

The splitting of the index set Λ as $\Lambda = \cup_{j=0}^\infty \Lambda_j$, corresponds to distinguishing functions "living" at different scales ($\lambda \in \Lambda_j \leftrightarrow (\mathrm{supp}\psi_\lambda) \sim 2^{-\mathrm{j}}$). It is beyond the goal of this paper to describe how and under which conditions on Ω such bases $\{\psi_\lambda\}_\lambda$ are constructed (see, among others, [24], [39]).

Since $H^s(\Omega) = B^s_{2,2}(\Omega)$, from equivalence (3.2.1) we deduce that for all s, $0 \le s \le S$:

$$\|\sum_\lambda u_\lambda \check{\psi}_\lambda\|_{H^s(\Omega)} \simeq \|\underline{u}\|_{\ell^2}, \quad \text{with } \check{\psi}_\lambda = \rho^{-js}\psi_\lambda. \tag{3.2.2}$$

Moreover, when considering nonlinear approximation in H^s, the scale of Besov spaces $B^{r+s}_{\tau,\tau}(\Omega)$ – where $\tau = \tau(r)$ is defined by the relation $\frac{1}{\tau} = \frac{r}{d} + \frac{1}{2}$ – will naturally appear. For these spaces the norm equivalences in terms of wavelet coefficients are quite simple; indeed using again equivalence (3.2.1), we obtain:

$$\|\sum_\lambda u_\lambda \check{\psi}_\lambda\|_{B^{r+s}_{\tau,\tau}(\Omega)} \simeq \|\underline{u}\|_{\ell^\tau}, \quad \lambda \in \Lambda_j, \qquad \text{with } \frac{1}{\tau} = \frac{r}{d} + \frac{1}{2} \tag{3.2.3}$$

where again $\check{\psi}_\lambda = \rho^{-js}\psi_\lambda$.

In the following it will also be useful to consider the space of functions whose coefficients, with respect to the rescaled basis $\{\check{\psi}_\lambda\}_\lambda$, are in the *weak–*$\ell^\tau$ space ℓ^τ_w, which can be defined as the space of sequences $\underline{u} = \{u_\lambda\}_\lambda$ for which there exists a constant C such that

$$\#\{\lambda : \quad |u_\lambda| \ge \epsilon\} \le C\epsilon^{-\tau}, \tag{3.2.4}$$

the norm $\|\underline{u}\|^\tau_{\ell^\tau_w}$ being defined as the smallest C which verifies relation (3.2.4). It is possible to prove that $\ell^\tau \subseteq \ell^\tau_w$, which implies that the coefficients $\{u_\lambda\}_\lambda$ of a function $u \in B^{r+s}_{\tau,\tau}$ verify $\{u_\lambda\}_\lambda \in \ell^\tau_w$.

Let us now recall some facts about nonlinear wavelet approximation: the space $\Sigma_N \subseteq V$,

$$\Sigma_N = \{u = \sum_\lambda c_\lambda \psi_\lambda : \quad \underline{c} = \{c_\lambda\}_{\lambda\in\Lambda} \in \sigma_N\}$$

with

$$\sigma_N = \{\underline{c} \in \ell^2(\Lambda) : \quad \#\{\lambda \in \Lambda : \ c_\lambda \ne 0\} \le N\},$$

is a nonlinear space containing functions in $L^2(\Omega)$ which can be represented as the linear combination of at most N elements of the basis $\{\check{\psi}_\lambda\}_\lambda$. A nonlinear projector $\mathbb{P}_N : L^2(\Omega) \to \Sigma_N$ can be defined as follows: given $u = \sum_\lambda u_\lambda \check{\psi}_\lambda$, let us introduce a decreasing rearrangement $\{|u_{\lambda(n)}|\}_{n\in\mathbb{N}}$ of the sequence $\{|u_\lambda|\}_{\lambda\in\Lambda}$, where the application $n \in \mathbb{N} \longrightarrow \lambda(n) \in \Lambda$ is bijective and verifies $n < m \Longrightarrow |u_{\lambda(n)}| \geq |u_{\lambda(m)}|$; $\mathbb{P}_N(u)$ is then defined by:

$$\mathbb{P}_N(u) = \sum_{n=1}^{N} u_{\lambda(n)} \check{\psi}_{\lambda(n)},$$

that is only the N greatest (in absolute value) coefficients of u are retained. We recall that by abuse of notation we will also indicate by $\mathbb{P}_N : \ell^2 \to \sigma_N$ the operator associating to the sequence $\underline{u}$ the coefficients of the function $\mathbb{P}_N(\sum_\lambda u_\lambda \psi_\lambda)$. The accuracy of the corresponding approximation is directly related to ℓ^τ_w regularity of the sequence of coefficients of u, as stated by the following theorem [47], [48].

Theorem 3.2.1: *Let $u = \sum_{\lambda \in \Lambda} u_\lambda \check{\psi}_\lambda$, with $\check{\psi}_\lambda = \rho^{-js} \psi_\lambda$, $s \leq S$. If $\{u_\lambda\}_\lambda \in \ell^\tau_w$ then*

$$\|\underline{u} - \mathbb{P}_N \underline{u}\|_{\ell^2} \lesssim \inf_{\underline{w} \in \sigma_N} \|\underline{u} - \underline{w}\|_{\ell^2} \lesssim N^{-(\frac{1}{\tau} + \frac{1}{2})} \|\underline{u}\|_{\ell^w_\tau}$$

where the implicit constants in the bounds depend only on τ.

3.3 The Problem

Let us now consider a linear operator $A : H^s(\Omega) \to H^{-s}(\Omega)$, $0 < s \leq S$, ($H^{-s}(\Omega)$ denoting here the dual of $H^s(\Omega)$) and let the corresponding bilinear form $a : H^s(\Omega) \times H^s(\Omega) \to \mathbb{R}$ be defined as:

$$a(u, v) := < Au, v >, \qquad \forall u, v \in H^s(\Omega),$$

where $< ., . >$ denotes the duality pairing between H^{-s} and H^s. We assume that the bilinear form a is continuous coercive, that is; $\forall u, v \in H^s(\Omega)$:

$$a(u, v) \leq M \|u\|_{H^s(\Omega)} \|v\|_{H^s(\Omega)}, \qquad a(u, u) \geq \alpha \|u\|^2_{H^s(\Omega)}.$$

We consider the following problem: given $g \in H^{-s}(\Omega)$, find $u \in H^s(\Omega)$ such that:

$$Au = g. \tag{3.3.1}$$

Under our assumptions it is well known that for any $g \in H^{-s}(\Omega)$ equation (3.3.1) has a unique solution; this is also the unique solution of the equivalent variational problem: find $u \in H^s$ such that

$$a(u, v) = < g, v >, \quad \forall v \in H^s(\Omega). \tag{3.3.2}$$

3.4 Nonlinear Richardson I: the basic scheme

Depending on the regularity of the data and on the domain Ω, the solution of problem (3.3.2) may be smooth, or it may present some singularity (see e.g. [66], [55], [61], [37]). In the last case, the fact that using some adaptive technique – in which the approximating space is tailored to the function u itself – is necessary in order to get a good approximation rate, is well accepted. The results in Section 3.2 on nonlinear approximation allow to rigorously formalize such fact and provide, in the wavelet context, a simple and efficient strategy for adaptively approximating u, if this was given. However, in the partial differential equations framework the function that one needs to approximate is not known. We would then like a strategy for designing an approximation space for the (unknown) solution u of the given PDE with the same approximation property that one would get if the solution was known.

In order to do so, according to the abstract approach described in Section 4.1, the first step is to transform the given continuous problem into an ∞-dimensional problem: we express u in terms of the rescaled basis $\{\check{\psi}_\lambda\}_\lambda$, $\check{\psi}_\lambda = 2^{-js}\psi_\lambda$, and we rewrite the initial continuous problem (3.3.2) in terms of the Fourier coefficients $\underline{u} = \{u_\lambda\}_\lambda$ of the unknown solution

$$u = \sum_\lambda u_\lambda \check{\psi}_\lambda,$$

thus obtaining an ∞-dimensional linear system of equations:

$$\mathcal{A}\underline{u} = \underline{g} \tag{3.4.1}$$

where

$$\mathcal{A} = (a_{\mu,\lambda})_{\mu,\lambda\in\Lambda}, a_{\mu,\lambda} =< A\check{\psi}_\lambda, \check{\psi}_\mu >, \qquad \underline{g} = \{g_\mu\}_\mu =< f, \check{\psi}_\mu >,$$

are a bi-infinite matrix and an infinite array respectively. It is not difficult to check (see Sections 2.5.1 and 2.5.2) that $\mathcal{A} \in \mathcal{L}(\ell^2, \ell^2)$ and that it is boundedly invertible, that is:

$$\|\mathcal{A}\|_{\mathcal{L}(\ell^2,\ell^2)} \lesssim C_1, \qquad \|\mathcal{A}^{-1}\|_{\mathcal{L}(\ell^2,\ell^2)} \lesssim C_2.$$

The second step is to write down a convergent numerical scheme for the ∞-dimensional problem: we design a method finding an approximate solution to the ∞-dimensional problem (3.4.1) in Σ_N. To this end let us assume that the basis $\{\check{\psi}_\lambda\}_\lambda$ and the operator $\mathcal{A}$ are such that it holds for some $\tau_0 < 2$:

$$\mathcal{A} \in \mathcal{L}(\ell^{\tau_0}_w, \ell^{\tau_0}_w). \tag{3.4.2}$$

We remark that under suitable space-frequency localization properties of the wavelet basis $\{\check{\psi}_\lambda\}_\lambda$, condition (3.6) holds for a wide class of differential and pseudo-differential operators ([28]).

The abstract scheme we want here to discuss is the following, where the tolerance *tol* clearly depends on the number N of degrees of freedom:

basic nonlinear Richardson

begin

Input: N, tol
$\underline{u}^{(0)} = \underline{0}$

while $\|\underline{r}^{(n)}\|_{\ell^2} > tol$ **do**

compute $\underline{r}^{(n)}$
as $\underline{r}^{(n)} = \underline{g} - \mathcal{A}\underline{u}^{(n)}$

update
$\underline{u}^{(n+1)} = \mathbb{P}_N(\underline{u}^{(n)} + \theta \underline{r}^{(n)}) \in \sigma_N$

end

Output: $u_N = \sum_\lambda u_\lambda^{(n+1)} \check{\psi}_\lambda$

end

This is not a computable numerical scheme since it involves operations on infinite matrices and vectors. Such scheme will be coupled in section 3.6.3 with suitable compression steps applied both to the operator $\mathcal{A}$ and to the right hand side $\underline{g}$, which will allow to actually implement it efficiently. Nevertheless it is interesting to consider such a scheme in order to analyze the influence of the nonlinear operator $\mathbb{P}_N$. In particular the main result of this section is the following Theorem:

Theorem 3.4.1: *Let $\mathcal{A} \in \mathcal{L}(\ell^{\tau_0}_w, \ell^{\tau_0}_w) \cap \mathcal{L}(\ell^2, \ell^2)$ for some $\tau_0 < 2$. Then there exists a $\tilde{\tau} < 2$ and a $\theta_0 > 0$ such that, for all θ, $0 < \theta < \theta_0$, it holds (the implicit constants in the inequalities depending on θ)*

(i) *stability: if $\underline{g} \in \ell^2$, we have*

$$\|\underline{u}^n\|_{\ell^2} \lesssim \|\underline{g}\|_{\ell^2}, \quad \forall n \in \mathbb{N}; \tag{3.4.3}$$

(ii) *approximation error estimate: if $\underline{g} \in \ell^\tau_w$, $\tilde{\tau} < \tau \leq 2$ then, setting $\underline{e}^n = \underline{u}^n - \underline{u}$, it holds for some $\mu < 1$:*

$$\|\underline{e}^n\|_{\ell^2} \lesssim \mu^n \|e_o\|_{\ell^2} + \frac{1}{1-\mu} N^{-(\frac{1}{\tau}+\frac{1}{2})}; \tag{3.4.4}$$

In order to prove theorem 3.4.1 we need to recall the following lemma [28].

Lemma 3.4.1: *Let $\mathcal{A} \in \mathcal{L}(\ell^{\tau_0}_w, \ell^{\tau_0}_w) \cap \mathcal{L}(\ell^2, \ell^2)$, for some $\tau_0 < 2$. Then there exists a constant $\theta_0 > 0$ and a $\tilde{\tau}$, $\tau_0 \leq \tilde{\tau} < 2$, such that $\forall \theta$ with $0 \leq \theta \leq \theta_0$ it holds:*

$$\|I - \theta\mathcal{A}\|_{\mathcal{L}(\ell^2,\ell^2)} \leq \mu < 1, \tag{3.4.5}$$

$$\|I - \theta\mathcal{A}\|_{\mathcal{L}(\ell^\tau_w,\ell^\tau_w)} \leq \gamma < 1, \qquad \textit{for all } \tau, \quad \tilde{\tau} \leq \tau \leq 2 \tag{3.4.6}$$

Proof of Theorem 3.4.1: First of all we observe that $\mathbb{P}_N$ is ℓ^2-contractive. Then, since $\underline{u}^n \in \ell^2$ (only N coefficients are non zero by definition), using Lemma 3.4.1, by (3.4.5) one has:

$$\|\underline{u}^{n+1}\|_{\ell^2} = \|P_N(\underline{u}^n + \theta(\underline{g} - \mathcal{A}\underline{u}^n))\|_{\ell^2} \leq \|\underline{u}^n + \theta(\underline{g} - \mathcal{A}\underline{u}^n)\|_{\ell^2} \leq \|\theta\underline{g}\|_{\ell^2} + \mu\|\underline{u}^n\|_{\ell^2}.$$

By iterating this bound we obtain:

$$\|\underline{u}^{n+1}\|_{\ell^2} \leq \left(\sum_{i=0}^{n} \mu^i\right) \|\theta\underline{g}\|_{\ell^2} + \mu^{n+1}\|\underline{u}^0\|_{\ell^2},$$

which, since $\mu < 1$ and $\underline{u}^0 = 0$, yields (3.4.3).

Now let

$$\underline{\varepsilon}^n = P_N(\underline{u}^n + \theta(\underline{g} - \mathcal{A}\underline{u}^n)) - (\underline{u}^n + \theta(\underline{g} - \mathcal{A}\underline{u}^n)).$$

A simple calculation yields:

$$\underline{e}^{n+1} - \underline{e}^n + \theta\mathcal{A}\underline{e}^n = \underline{\varepsilon}^n,$$

which, taking the ℓ^2 norm and using (3.4.5) again, yields

$$\|\underline{e}^{n+1}\|_{\ell^2} \leq \mu\|\underline{e}^n\|_{\ell^2} + \|\underline{\varepsilon}^n\|_{\ell^2}. \tag{3.4.7}$$

Iterating (3.4.7), we then obtain:

$$\|\underline{e}^{n+1}\|_{\ell^2} = \sum_{k=0}^{n} \mu^{n-k}\|\underline{\varepsilon}^k\|_{\ell^2} + \mu^{n+1}\|e_0\|_{\ell^2} \leq \left(\max_{0\leq k\leq n} \|\underline{\varepsilon}^k\|_{\ell^2}\right) \sum_{i=0}^{n} \mu^i + \mu^{n+1}\|e_0\|_{\ell^2}. \tag{3.4.8}$$

The sum on the right hand side of (3.4.8) converges ($\mu < 1$) and then we can write:

$$\|\underline{e}^{n+1}\|_{\ell^2} \leq \frac{1}{1-\mu} \max_k \|\underline{\varepsilon}^k\|_{\ell^2} + \mu^{n+1}\|e_0\|_{\ell^2}. \tag{3.4.9}$$

To conclude, we only need to give an uniform bound on $\|\underline{\varepsilon}^k\|_{\ell^2}$. Using Lemma 3.4.1 it is not difficult to show that $\underline{g} \in \ell^\tau_w$ implies $\underline{u}^k + \theta(\underline{g} - \mathcal{A}\underline{u}^k) \in \ell^\tau_w$ with:

$$\|\underline{u}^k + \theta(\underline{g} - \mathcal{A}\ \underline{u}^k)\|_{\ell^\tau_w} \leq C, \tag{3.4.10}$$

uniformly in k. Indeed $\|I - \theta\mathcal{A}\|_{\mathcal{L}(\ell^\tau_w, \ell^\tau_w)} \leq \gamma < 1$ and since $\mathbb{P}_N$ is ℓ^τ_w contractive, setting

$$\underline{w}^{k+1} = \underline{u}^k + \theta(\underline{g} - \mathcal{A}\underline{u}^k)$$

we can write:

$$\begin{aligned}
\|\underline{w}^{k+1}\|_{\ell^\tau_w} &= \|(I - \theta\mathcal{A})\underline{u}^k + \theta\underline{g}\|_{\ell^\tau_w} \leq \|(I - \theta\mathcal{A})\underline{u}^k\|_{\ell^\tau_w} + \|\theta\underline{g}\|_{\ell^\tau_w} \leq \gamma\|\underline{u}^k\|_{\ell^\tau_w} + \|\theta\underline{g}\|_{\ell^\tau_w} \\
&= \gamma\|\mathbb{P}_N(\underline{w}^k)\|_{\ell^\tau_w} + \|\theta\underline{g}\|_{\ell^\tau_w} \leq \gamma\|\underline{w}^k\|_{\ell^\tau_w} + \|\theta\underline{g}\|_{\ell^\tau_w}.
\end{aligned}$$

By iterating this bound, we obtain:

$$\|\underline{w}^{k+1}\|_{\ell^\tau_w} \leq \gamma^k\|\underline{w}^1\|_{\ell^\tau_w} + \left(\sum_{i=0}^k \gamma^i\right)\|\theta\underline{g}\|_{\ell^\tau_w} \leq \frac{\theta}{1-\gamma}\|\underline{g}\|_{\ell^\tau_w}, \quad \forall k,$$

which yields (3.4.10). Thanks to (3.4.10), by applying Theorem 3.2.1, we have that:

$$\max_k \|\underline{\varepsilon}^k\|_{\ell^2} \lesssim N^{-(\frac{1}{\tau}+\frac{1}{2})} \max_k \|\underline{u}^k + \theta(\underline{g} - \mathcal{A}\underline{u}^k)\|_{\ell^\tau_w} \leq N^{-(\frac{1}{\tau}+\frac{1}{2})} C(g).$$

Combining such bound with (3.4.9), implies the thesis. □

Using norm equivalence (3.2.2), this yields the following corollary.

Corollary 3.4.1: *Let u be the solution of (3.3.2) and let g belong to $B^{r+s}_{\tau,\tau}(\Omega)$, with r such that $0 < r + s \leq \min\{S, d/\tilde{\tau} - d/2\}$ and with τ given by $1/\tau = r/d + 1/2$. If $u_N^{(n)} = \sum_\lambda u_\lambda^{(n)}\check{\psi}_\lambda$ is the non linear approximation of u at step n given by the non-linear Richardson scheme with $\theta < \theta_0$, then it holds, for some $\mu < 1$:*

$$\|u - u_N^{(n)}\|_{H^s(\Omega)} \leq \mu^n\|u - u_N^{(0)}\|_{H^s(\Omega)} + \frac{C}{1-\mu}N^{-\frac{r}{d}}, \quad \forall n \in \mathbb{N}.$$

Remark 3.4.1: *Though for simplicity we set $\underline{u}^0 = 0$ throughout this section, it is not difficult to realize that the result holds unchanged also for any initial guess in Σ_N.*

3.5 Numerical results

In this section, we test **basic nonlinear Richardson** scheme on a very simple 1D model problem, namely, let $\mathbb{T}$ be the unit circle,

Problem 3.5.1: *find* $u \in H^1(\mathbb{T})$ *such that*

$$a(u, v) = (f, v) \quad \textit{for all } v \in H^1(\mathbb{T}),$$

where $a(u, v) = \int_{\mathbb{T}} u'v' + \int_{\mathbb{T}} uv$ *and* $(f, v) = \int_{\mathbb{T}} fv$.

The tests we will show in the following aim at studying only the influence of the nonlinear projector $\mathbb{P}_N$ at each iteration of the scheme and they won't face the problem of the effective construction of the prevision sets. The tests are referred to different choices of the function f. In particular we will study the behavior of the error in L^2 and H^1 norm as a function of

1. the number n of iterations,
2. the number N of degrees of freedom to be retained.

3.5.1 Two cases

We performed the numerical tests, by using the pair $\{\psi_\lambda, \tilde{\psi}_\lambda\}$ of biorthogonal B-spline wavelets B2.2 on $\mathbb{T}$. We will refer to biorthogonal B-splines wavelets B$\tilde{N}$.N to signify that V_j and $\tilde{V}_j$ are the subspaces of B-splines of order $N-1$ and $\tilde{N}-1$ respectively (with $N, \tilde{N} \geq 1$) defined on the uniform grid obtained by splitting the unit interval into 2^j equal segments.

Let us denote by $u_J = \sum_{|\lambda|<J} u_\lambda \psi_\lambda$ the wavelet decomposition of the exact solution u at resolution J and by u_N^{nmax} the N-terms approximation to u_J, built after $nmax$ iterations of the scheme.

(A) The exact solution of the Problem 3.5.1 is

$$u(x) = e^{\cos 4\pi x} + \sin 4\pi x - e \tag{3.5.1}$$

and f is defined by $f = -u'' + u$.

The numerical results, performed with $J = 9$ e $nmax = 50$, are presented in the following tables

Error in norm L^2	
N	$\|u_J - u_N^{nmax}\|_{L^2}$
10	32.92885
20	1.33461
30	0.51570
40	0.31117
51	0.22131
61	0.11242
71	0.07977
81	0.10276
92	0.09832
102	0.06135

Table 3.1. L^2-approximation error.

Error in norm H^1	
N	$\|u_J - u_N^{nmax}\|_{H^1}$
10	39.88046
20	11.95111
30	8.39278
40	5.94793
51	4.83391
61	4.05235
71	3.32526
81	2.88048
92	2.56872
102	2.30032

Table 3.2. H^1-approximation error.

Plot of the exact solution (3.5.1)

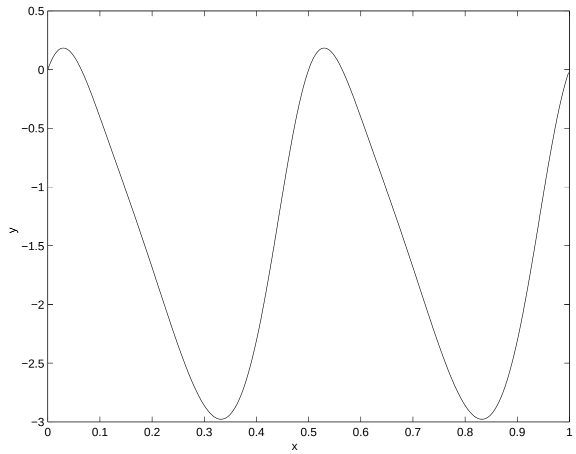

Figure 3.1. Exact solution.

In the following we show some plots dealing with the behavior of the approximation error.

(*) Behavior of $\|u_J - u_N^{nmax}\|_{L^2}$ as a function of N

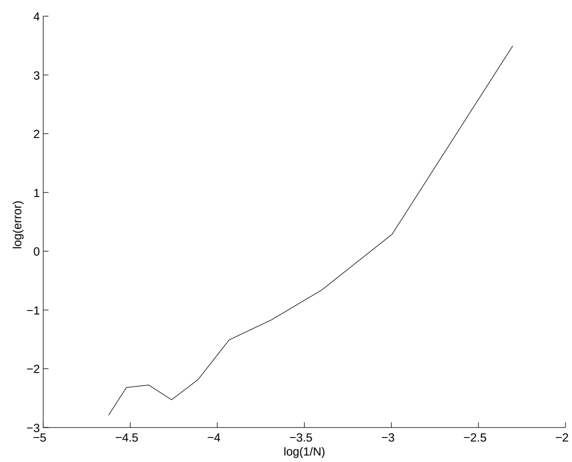

Figure 3.2. Logarithmic plot of the L^2-approximation error.

(**) Behavior of $\|u_J - u_N^{nmax}\|_{H^1}$ as a function of N

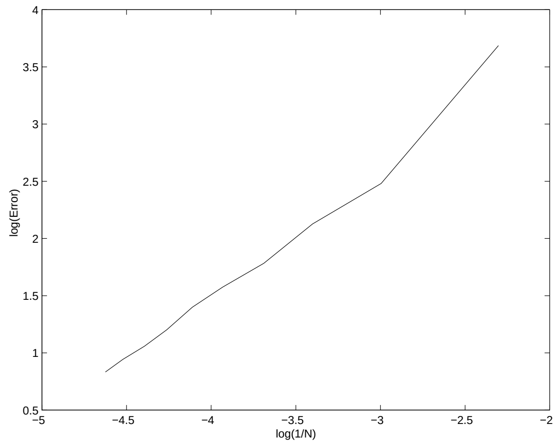

Figure 3.3. Logarithmic plot of the H^1-approximation error.

(***) Behavior of $\|u_J - u_N^n\|_{L^2}$ as a function of the number n of the iterations, for different values of N

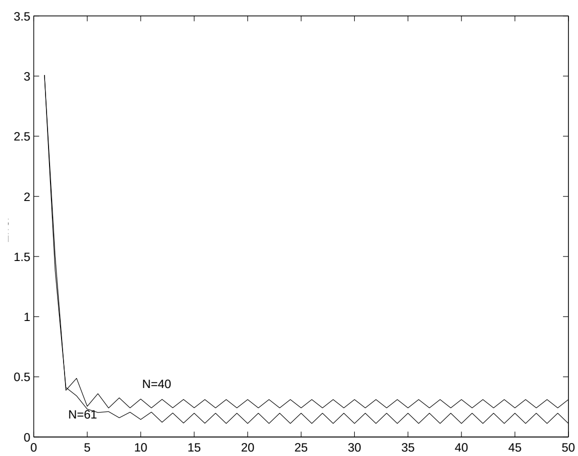

Figure 3.4. Plot of the error $\|u_J - u_N^n\|_{L^2}$ as a function of n, with $N = 40, N = 61$.

Having in mind error estimate (3.4.4)

$$\|u_J - u_N^n\|_{L^2} \lesssim \mu^n \|u_J - u_N^0\|_{L^2} + \frac{1}{1-\mu} N^{-(\frac{1}{\tau}+\frac{1}{2})},$$

from Figure 3.5.1 it is clear that going beyond the efficient number of iterations, which can be obtained by "balancing" the two terms on the right-hand side of the above inequality, does not bring any further reduction of the approximation error.

(B) The exact solution of the Problem 3.5.1 is

$$u(x) = \begin{cases} 3x & \text{if } 0 \leq x \leq 1/3 \\ 2x + 3/2 & \text{if } 1/3 \leq x \leq 1. \end{cases} \tag{3.5.2}$$

and f is defined by $f = -u'' + u$. The numerical results, performed with $J = 9$ e $nmax = 50$, are shown in the following tables.

Error in norm L^2	
N	$\|u_J - u_N^{nmax}\|_{L^2}$
10	0.44666
20	0.03138
30	0.00173
40	0.00001
51	0.00002
61	0.00008
71	0.00017
81	0.00029
92	0.00042
102	0.00054

Table 3.3. L^2-approximation error.

Error in norm H^1	
N	$\|u_J - u_N^{nmax}\|_{H^1}$
10	0.96855
20	0.26324
30	0.01619
40	0.00001
51	0.00004
61	0.00019
71	0.00044
81	0.00080
92	0.00117
102	0.00152

Table 3.4. H^1-approximation error.

Plot of the exact solution (3.5.2)

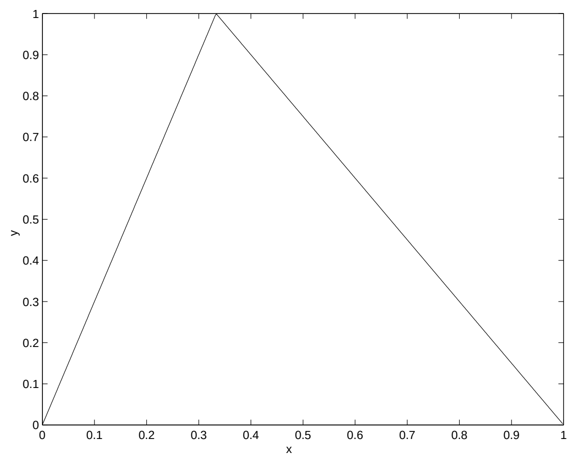

Figure 3.5. Exact solution.

In the following we show some plots dealing with the behavior of the approximation error.

(*) Behavior of $\|u_J - u_N^{nmax}\|_{L^2}$ as a function of N

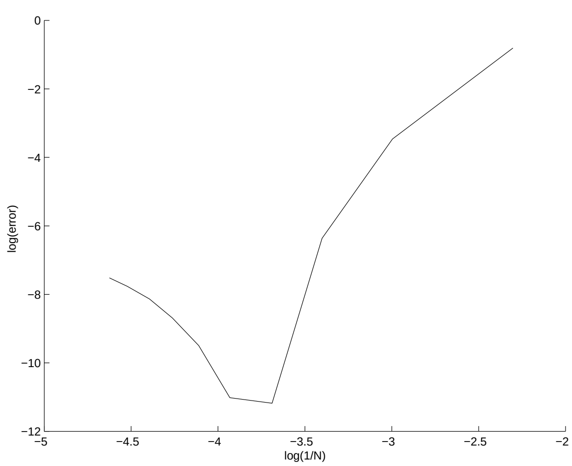

Figure 3.6. Logarithmic plot of the L^2-approximation error.

(**) Behavior of $\|u - u_N^{nmax}\|_{H^1}$ as a function of N

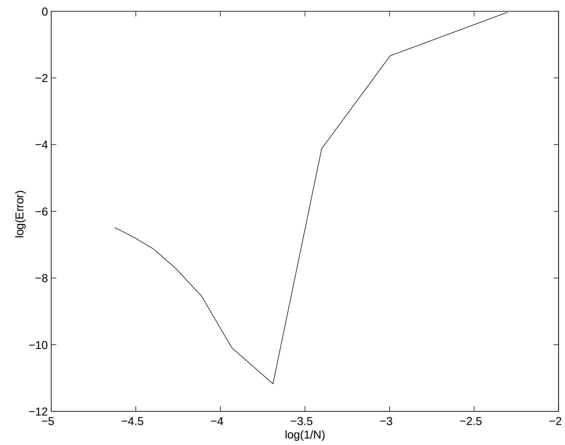

Figure 3.7. Logarithmic plot of the H^1-approximation error.

> **Remark 3.5.1:** *Due to the particular form of the exact solution, the approximation errors $\|u_J - u_N^{nmax}\|_{L^2}$ and $\|u_J - u_N^{nmax}\|_{H^1}$ decrease in a steep way, rapidly reaching the machine precision.*

(***) Behavior of $\|u_J - u_N^n\|_{L^2}$ as a function of n, for different values of N

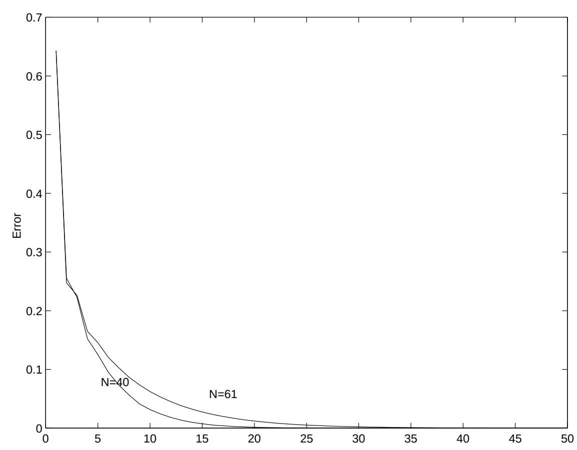

Figure 3.8. Plot of the error $\|u_J - u_N^n\|_{L^2}$, as a function of n, with $N = 40, N = 61$.

3.6 Nonlinear Richardson II: the reliable scheme

In this section we discuss computable Nonlinear Richardson-type schemes in order to find an approximate solution to the problem (3.4.1) in σ_N.

We make the following assumptions: the basis $\{\check{\psi}_\lambda\}_\lambda$ and the operator $\mathcal{A}$ are such that, setting

$$i(\lambda,\lambda') = \begin{cases} 1 & \text{supp}\psi_\lambda \cap \text{supp}\psi_{\lambda'} \neq \emptyset, \\ 0 & \text{otherwise}, \end{cases} \tag{3.6.1}$$

it holds for some $R \geq 2$, for $\lambda \in \Lambda_j$ and $\lambda' \in \Lambda_{j'}$

$$| < \mathcal{A}\psi_{\lambda'}, \psi_\lambda > | \leq \mathcal{K}_{\lambda,\lambda'} = K 2^{-(\frac{1}{2}+R)|j-j'|} i(\lambda,\lambda'). \tag{3.6.2}$$

We remark that under suitable space-frequency localization properties of the wavelet basis $\{\check{\psi}_\lambda\}_\lambda$, condition (3.6.2) holds for a wide class of differential and pseudo-differential operators [28].

Property (3.6.2) implies the boundedness of the operator $\mathcal{A}$ as an operator from ℓ^τ_w to ℓ^τ_w on a whole range of indexes τ. In particular we let $\tau_0 > 0$ be such that the matrix $\mathcal{K} = (\mathcal{K}_{\lambda,\lambda'})_{\lambda,\lambda'}$, satisfies

$$\mathcal{K} \in \mathcal{L}(\ell^\tau_w, \ell^\tau_w), \qquad \forall \tau \geq \tau_0.$$

We observe that certainly $\tau_0 = \tau_0(R)$ can actually be strictly smaller than 1, as R increases. The relation between the value of R and the actual value of τ_0 is the topic of the following Lemma which is a simple consequence of Theorem 2.5.4.

Lemma 3.6.1: *Let $\sigma > d/2$. If $s < \sigma - d/2$, a matrix $\mathcal{D}$, whose entries satisfy the estimate*

$$|\mathcal{D}_{\lambda,\lambda'}| \leq K 2^{-\sigma|j-j'|} i(\lambda,\lambda'), \tag{3.6.3}$$

defines a bounded operator in $\ell^2 \cap \ell^p_w$, for all p such that $1/p = 1/2 + s/d$.

Let us now consider the following result [28].

Lemma 3.6.2: *Let $\mathcal{A}$ satisfy property (3.6.2), and let $\tau_0 < 2$ be defined as above. Then there exist $\tilde{\tau}$, $\tau_0 \leq \tilde{\tau} < 2$ and a constant θ_0, such that $\forall\theta$ with $1 \leq \theta \leq \theta_0$ it holds:*

$$\|I - \theta\mathcal{A}\|_{\mathcal{L}(\ell^2,\ell^2)} \leq \mu < 1, \tag{3.6.4}$$

and $\forall\tau$, $\tilde{\tau} < \tau \leq 2$,

$$\|I - \theta\mathcal{A}\|_{\mathcal{L}(\ell^\tau_w,\ell^\tau_w)} \leq \gamma < 1. \tag{3.6.5}$$

In order to find an approximate solution to problem (3.3.2) in σ_N, where N is fixed *a priori*, we propose a computable nonlinear Richardson type-scheme, which couples Richardson iterative scheme and nonlinear approximation. We observe that Lemma 3.6.2 guarantees the convergence of the plain Richardson scheme for the solution of the infinite linear system (3.4.1). Nonlinearity is plugged in the scheme, by forcing, at each iteration, the approximate solution to belong to the nonlinear space σ_{N_n}, for some $N_n \leq N$. For $n_1 < n_2$, we ask that $N_{n_1} < N_{n_2}$ (and so $\sigma_{N_{n_1}} \subset \sigma_{N_{n_2}}$) and that $N_{\bar{n}} = N$ for some $\bar{n} \geq 1$. Roughly speaking at each iteration we inflate the nonlinear space σ_{N_n}, which the approximate solution $\underline{u}^{(n+1)}$ belongs to, till we reach the target nonlinear space σ_N. Since the plain Richardson scheme involves the multiplication by the bi-infinite matrix $\mathcal{A}$, in order to make such an algorithm practically feasible, we will also need to approximate such multiplication. At each iteration this will be done with a precision ϵ_n, with $\epsilon_{n+1} \leq \epsilon_n$. Suitable choices for the parameters N_n and ϵ_n will be discussed in the following. Clearly, the choice of the N_n's and of the ϵ_n's will be not made independently, if optimal performance is aimed at.

Let us for the moment assume that we have selected two sequences N_n and ϵ_n

$$N_0 < N_1 < \cdots N_{n_1} < N_{n_2} < \cdots < N_{\bar{n}} = N \tag{3.6.6}$$

$$\epsilon_0 > \epsilon_1 > \cdots \epsilon_{n_1} > \epsilon_{n_2} > \cdots > \epsilon_{\bar{n}} \tag{3.6.7}$$

The algorithm is performed in several steps.

Computable nonlinear Richardson

Step 1. Initialization: set $\underline{u}^{(0)} = 0$.

Step 2. Until $n \leq \bar{n}$, repeat

Step 2.1 Prevision: select a finite dimensional set $V^{(n)} \subset \Lambda$ such that, denoted by S_n the linear subspace of ℓ^2 of the form $S_n = \{\underline{v} \in \ell^2 : \quad v_\lambda = 0, \quad \lambda \notin V^{(n)}\}$, we have

$$\inf_{\underline{v} \in S_n} \|\underline{u}^{(n)} + \theta \underline{r}^{(n)} - \underline{v}\|_{\ell^2} \leq C\epsilon_n,$$

where $\underline{r}^{(n)} = \underline{g} - \mathcal{A}\underline{u}^{(n)}$ denotes the residual and where the constant C depends only on initial data.

Step 2.2 Compute an approximation

$$\underline{\tilde{r}}^{(n)} \in \{\underline{v} = (v_\lambda)_{\lambda \in \Lambda} \in \ell^2 : \ \lambda \notin V^{(n)} \Rightarrow v_\lambda = 0\}$$

of the residual $\underline{r}^{(n)}$, in such a way that $\|\underline{r}^{(n)} - \underline{\tilde{r}}^{(n)}\|_{\ell^2} \leq C\epsilon_n$.

Step 2.3 Projection: set
$$\underline{u}^{(n+1)} = \mathbb{P}_{N_n}(\underline{u}^{(n)} + \theta \underline{\tilde{r}}^{(n)}).$$

Step 2.4 Update: $n+1 \to n$

Even if $\mathcal{A}$ is a bi-infinite matrix, the scheme is reliable as it carries out computations, at each iteration, over a (finite) prevision set $V^{(n)}$. Moreover with a suitable choice (see Remark 3.6.3) of the number N_n of the coefficients to be retained at each iteration, it is possible to reduce the computational cost of the algorithm, without loosing accuracy in the approximate solution. Indeed when the number of iterations is small we are far from the exact solution and so we don't loose too much using a small number of degrees of freedom, instead when the number of iterations increases, the error of the iterative scheme becomes small and in order not to waste this gain, we need to use a larger number of degrees of freedom.

3.6.1 Prevision

A crucial ingredient in the above algorithm is the *a priori* construction of the finite dimensional prevision set $V^{(n)}$. As at each iteration of the scheme we restrict the action of the nonlinear projector $\mathbb{P}_{N_n}$ to the finite set $V^{(n)}$, we need to choose this set in such a way that the coefficients whose indexes we discharge a priori (i.e. not belonging to $V^{(n)}$) are sufficiently small, in order not to make the algorithm loose accuracy in the estimation of the approximate solution.

To accomplish this goal, we start by defining a sort of measure of the interactions of two indexes $\lambda = (j,k)$ and $\lambda' = (j',k')$: for each λ we define a neighborhood in Λ by:
$$I(\lambda, \epsilon) = \{\lambda' : \quad v(\lambda, \lambda') > \epsilon\},$$
where ϵ is a given tolerance and
$$v(\lambda, \lambda') = 2^{-\frac{R}{2}|j-j'|} i(\lambda, \lambda'),$$
with R depending on the regularity of wavelet basis and
$$i(\lambda, \lambda') = \begin{cases} 1 & \mathrm{supp}\psi_\lambda \cap \mathrm{supp}\psi_{\lambda'} \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases} \tag{3.6.8}$$

Let now $w \in \Sigma_N$
$$w = \sum_{\lambda \in \bar{\Lambda}} w_\lambda \psi_\lambda, \qquad \#(\bar{\Lambda}) = N.$$

We want to compute r_λ:

$$r_\lambda = g_\lambda - (\mathcal{A}\underline{w})_\lambda = g_\lambda - \sum_{\lambda' \in \bar{\Lambda}} a_{\lambda,\lambda'} w_{\lambda'}, \tag{3.6.9}$$

where $a_{\lambda,\lambda'} =< A\check{\psi}_\lambda, \check{\psi}_{\lambda'} >$, within a prescribed accuracy. In order to do so, we split the sum at the right hand side as the sum of two contributions: one, d_λ, coming from frequencies λ' belonging to the neighborhood $I(\lambda, \epsilon)$ of λ; the other, e_λ, coming from frequencies not belonging to it:

$$(\mathcal{A}\underline{w})_\lambda = d_\lambda + e_\lambda, \tag{3.6.10}$$

where

$$d_\lambda = \sum_{\lambda' \in \bar{\Lambda} \cap I(\lambda,\epsilon)} < \mathcal{A}\psi_{\lambda'}, \psi_\lambda > w_{\lambda'}, \tag{3.6.11}$$

$$e_\lambda = \sum_{\lambda' \in \bar{\Lambda} \setminus I(\lambda,\epsilon)} < \mathcal{A}\psi_{\lambda'}, \psi_\lambda > w_{\lambda'}. \tag{3.6.12}$$

The contribution e_λ to $(\mathcal{A}\underline{w})_\lambda$, of frequencies λ' not belonging to the neighborhood $I(\lambda, \epsilon)$ of λ, can be controlled by tuning the parameter ϵ, that is with a suitable choice of the size of the neighborhood $I(\lambda, \epsilon)$. To this aim we let τ_1 be such that the matrix $\check{\mathcal{K}} = (\check{\mathcal{K}}_{\lambda,\lambda'})_{\lambda,\lambda'}$, with, for $\lambda \in \Lambda_j$ and $\lambda' \in \Lambda_{j'}$ $\check{\mathcal{K}}_{\lambda,\lambda'} = 2^{\frac{R}{2}|j-j'|}\mathcal{K}_{\lambda,\lambda'}$ satisfies

$$\check{\mathcal{K}} \in \mathcal{L}(\ell^\tau_w, \ell^\tau_w), \forall \tau \geq \tau_1, \tag{3.6.13}$$

where, by using Lemma 3.6.1 with $\mathcal{D} = \check{\mathcal{K}}$ and $\sigma = 1/2 + R/2$, the value of $\tau_1 = \tau_1(R)$ can again be strictly smaller than 1, as R is sufficiently large.

We have the following Lemma.

Lemma 3.6.3: *Let $\mathcal{A}$ satisfy (3.6.2). There exist constant C_0 and C_1 depending on the operator $\mathcal{A}$ and the wavelet basis $\{\psi_\lambda\}_\lambda$, such that we have:*

$$\|\underline{e}\|_{\ell^2} \leq C_0 \epsilon \|\underline{w}\|_{\ell^2}, \tag{3.6.14}$$

and for all p, $\tau_1 \leq p < 2$

$$\|\underline{e}\|_{\ell^p_w} \leq C_1 \epsilon \|\underline{w}\|_{\ell^p_w}. \tag{3.6.15}$$

Proof: Using property (3.6.2) in the definition of e_λ, together with the definition of the set $I(\lambda, \epsilon)$, we obtain the following estimate for all λ:

$$|e_\lambda| < \epsilon K \sum_{\lambda'} 2^{-\frac{R+1}{2}|j-j'|} i(\lambda, \lambda') |w_\lambda| = \epsilon (\check{\mathcal{K}}\underline{w})_\lambda. \tag{3.6.16}$$

We then easily write

$$\|e\|_{\ell^2} \leq \epsilon \|\check{\mathcal{K}}\|_{\mathcal{L}(\ell^2,\ell^2)} \|\underline{w}\|_{\ell^2},$$

and for all p, $\tau_1 \leq p < 2$

$$\|e\|_{\ell^p_w} \leq \epsilon \|\check{\mathcal{K}}\|_{\mathcal{L}(\ell^p_w,\ell^p_w)} \|\underline{w}\|_{\ell^p_w}.$$

By using Lemma 3.6.1 it is not difficult to prove that there exist constants C_0 and C_1 such that $\|\check{\mathcal{K}}\|_{\mathcal{L}(\ell^2,\ell^2)} \leq C_0$, and $\|\check{\mathcal{K}}\|_{\mathcal{L}(\ell^p_w,\ell^p_w)} \leq C_1$, and this allows to conclude. □

Now we come to the topic of approximating the right hand side $\underline{g}$. As a matter of fact we recall the following result [47]:

Lemma 3.6.4: *Let $\underline{g} \in \ell^\infty$ and let $T^\epsilon_\tau \underline{g}$ be defined by:*

$$(T^\epsilon_\tau \underline{g})_\lambda = \begin{cases} g_\lambda & |g_\lambda| \geq \epsilon^{\frac{2}{2-\tau}} \\ 0 & |g_\lambda| < \epsilon^{\frac{2}{2-\tau}}. \end{cases}$$

The following estimate holds for every $0 < \tau \leq 2$:

$$\|\underline{g} - T^\epsilon_\tau \underline{g}\|_{\ell^2} \leq \epsilon \|\underline{g} - T^\epsilon_\tau \underline{g}\|^{\tau/2}_{\ell^\tau} \leq \epsilon \|\underline{g}\|^{\tau/2}_{\ell^\tau}.$$

Define then sets $\mathcal{B}$ and $\mathcal{C}$:

$$\mathcal{B} = \{\lambda : \quad I(\lambda,\epsilon) \cap \bar{\Lambda} \neq \emptyset\},$$

$$\mathcal{C} = \{\lambda : \quad |g_\lambda| \geq \epsilon^{\frac{2}{2-\tau}}\}.$$

We then define $\underline{\tilde{r}} = \underline{\tilde{r}}(w)$ in $\{\underline{v} = (v_\lambda) \in \ell^\infty, \;\; v_\lambda = 0 \; \forall \lambda \notin \mathcal{B} \cup \mathcal{C}\}$ as

$$\underline{\tilde{r}} = T^\epsilon_\tau \underline{g} - \underline{d}, \qquad \text{that is} \qquad \tilde{r}_\lambda = \begin{cases} g_\lambda - d_\lambda & \lambda \in \mathcal{B} \cap \mathcal{C} \\ g_\lambda & \lambda \in \mathcal{C} \setminus \mathcal{B} \\ -d_\lambda & \lambda \in \mathcal{B} \setminus \mathcal{C} \\ 0 & \textit{otherwise}. \end{cases} \tag{3.6.17}$$

Now it is easy to obtain the following two Corollaries:

Corollary 3.6.1: *Let $\underline{r} = \underline{g} - \mathcal{A}\underline{w}$, then we have*

$$\|\underline{r} - \underline{\tilde{r}}\|_{\ell^2} \leq C_0 \epsilon (\|\underline{w}\|_{\ell^2} + \|\underline{g}\|_{\ell^2}).$$

Corollary 3.6.2: *If $\underline{\tilde{r}}$ is defined as above, then it holds that:*

$$\|\underline{\tilde{r}}\|_{\ell^\tau_w} \leq C_1 (\|\underline{w}\|_{\ell^\tau_w} + \|\underline{g}\|_{\ell^\tau_w}).$$

3.6.2 Error Estimate

Let us then suppose that we are give non increasing sequence $(\epsilon_n)_n$ of positive real numbers and a non decreasing sequence $(N_n)_n$ of integers, respectively converging to 0 and to $+\infty$ as $n \longrightarrow +\infty$.

Given $\underline{u}^{(n)}$, we consider the prevision set $V^{(n)} := \mathcal{B}^{(n)} \cup \mathcal{C}^{(n)}$, where $\Lambda^{(n)} \subset \Lambda$ is the set containing the coefficients of $\underline{u}^{(n)}$:

$$\mathcal{B}^{(n)} = \{\lambda : \quad I(\lambda, \epsilon_n) \cap \Lambda^{(n)} \neq \emptyset\},$$

$$\mathcal{C}^{(n)} = \{\lambda : \quad |g_\lambda| \geq \epsilon_n^{\frac{2}{2-\tau}}\}.$$

We then define $\tilde{r}^{(n)} := \underline{\tilde{r}}(\underline{u}^{(n)}) \in \{\underline{v} \in \ell^\infty, \quad v_\lambda = 0 \ \forall \lambda \notin V^{(n)}\}$ according to (3.6.17):

$$\tilde{r}_\lambda^{(n)} = T_\tau^{\epsilon_n} \underline{g} - \sum_{\lambda' \in I(\lambda, \epsilon_n) \cap \Lambda^{(n)}} < \mathcal{A}\psi_{\lambda'}, \psi_\lambda > u_{\lambda'}^{(n)}$$

and we consider the nonlinear Richardson scheme:

$$\underline{u}^{(n+1)} = \mathbb{P}_{N_n}(\underline{u}^{(n)} + \theta \underline{\tilde{r}}^{(n)}). \tag{3.6.18}$$

The residual $\underline{\tilde{r}}^{(n)} = (\tilde{r}_\lambda^{(n)})_{\lambda \in V^{(n)}}$ is the truncated residual computed on the prevision set $V^{(n)}$.

The main result of this paper is the following Theorem, which considers the influence of the prevision step on the accuracy of the algorithm and provides an error estimate [19]:

Theorem 3.6.1: *Let $\mathcal{A}$ satisfy assumption (3.6.2). Then there exist a $\tilde{\tau} < 2$, a $\theta_0 > 0$ and an $\bar{\epsilon} > 0$ such that if $\epsilon_0 < \bar{\epsilon}$, for all θ, $0 < \theta < \theta_0$, it holds (the constants in the inequalities depending on initial data):*

(i) *stability: if $\underline{g}, \underline{u}^{(0)} \in \ell^2$ we have*

$$\|\underline{u}^{(n)}\|_{\ell^2} \leq C, \quad \forall n \in \mathbb{N}; \tag{3.6.19}$$

(ii) *approximation error estimate: if $\underline{g} \in \ell_w^\tau$, with $\max\{\tilde{\tau}, \tau_1\} < \tau \leq 2$, then it holds for some $\mu < 1$:*

$$\|\underline{u}^{(n+1)} - \underline{u}\|_{\ell^2} \lesssim \mu^{n+1} \|\underline{u}^{(0)} - \underline{u}\|_{\ell^2} + C(\|\underline{u}^{(0)}\|_{\ell^2}, \|\underline{g}\|_{\ell^2}) \left(\theta \sum_{i=0}^{n} \mu^{n-i} \epsilon_i + \sum_{i=0}^{n} \mu^{n-i} N_i^{-(\frac{1}{\tau} - \frac{1}{2})} \right). \tag{3.6.20}$$

Proof: Let $\tilde{\tau}$ and θ_0 be given by Lemma 3.6.2. Let us start by proving the stability estimate (3.6.19). First of all we observe that $\mathbb{P}_{N_n}$ is trivially both ℓ^2 and ℓ^τ_w -contractive. According to (3.6.18), by Lemma 3.6.2 we have:

$$\begin{aligned}\|\underline{u}^{(n+1)}\|_{\ell^2} &\leq \|(I-\theta\mathcal{A})\underline{u}^{(n)}\|_{\ell^2} + \theta\|\underline{g}\|_{\ell^2} + \theta\|\underline{e}\|_{\ell^2} \\ &\leq (\mu + \theta C_0\epsilon_n)\|\underline{u}^{(n)}\|_{\ell^2} + \theta\|\underline{g}\|_{\ell^2} \end{aligned} \tag{3.6.21}$$

where we have used inequality (3.6.14). Then iterating relation (3.6.21), we obtain:

$$\|\underline{u}^{(n+1)}\|_{\ell^2} \leq \left(\prod_{i=0}^{n}(\mu + C\theta\epsilon_n)\right)\|\underline{u}^{(0)}\|_{\ell^2} + \theta\left(1 + \sum_{i=0}^{n}\prod_{\ell=i}^{n}(\mu + C\theta\epsilon_\ell)\right)\|\underline{g}\|_{\ell^2}. \tag{3.6.22}$$

Similarly we have:

$$\|\underline{u}^{(n+1)}\|_{\ell^\tau_w} \leq \left(\prod_{i=0}^{n}(\gamma + C\theta\epsilon_n)\right)\|\underline{u}^{(0)}\|_{\ell^\tau_w} + \theta\left(1 + \sum_{i=0}^{n}\prod_{\ell=i}^{n}(\gamma + C\theta\epsilon_\ell)\right)\|\underline{g}\|_{\ell^\tau_w}. \tag{3.6.23}$$

Since condition (3.6.7) implies $\epsilon_n < \epsilon_0$, we then have

$$\|\underline{u}^{(n+1)}\|_{\ell^2} \leq (\mu + C\theta\epsilon_0)^{n+1}\|\underline{u}^{(0)}\|_{\ell^2} + \theta\sum_{i=0}^{n}(\mu + C\theta\epsilon_0)^i\|\underline{g}\|_{\ell^2}. \tag{3.6.24}$$

If ϵ_0 is chosen in such a way that $(\mu + C\theta\epsilon_0) < 1$, then the sum on the right hand side converges, and this yields (3.6.19). Analogously, if ϵ_0 is chosen in such a way that $(\gamma + C\theta\epsilon_0) < 1$, then one can prove that

$$\|\underline{u}^{(n+1)}\|_{\ell^\tau_w} \leq C_\tau. \tag{3.6.25}$$

Let us now consider the error. Letting

$$\underline{\varepsilon}^{(n)} = \mathbb{P}_{N_n}(\underline{u}^{(n)} + \theta\underline{\tilde{r}}^{(n)}) - (\underline{u}^{(n)} + \theta\underline{\tilde{r}}^{(n)}),$$

a simple calculation yields:

$$\underline{u}^{(n+1)} - \underline{u} = (I - \theta\mathcal{A})(\underline{u}^{(n)} - \underline{u}) + \theta(\underline{\tilde{r}}^{(n)} - \underline{r}^{(n)}) + \underline{\varepsilon}^{(n)},$$

where $\underline{r}^{(n)} = g - \mathcal{A}\underline{u}^{(n)}$ is the residual calculated on Λ, from which, taking the ℓ^2 norm and using (3.6.4),

$$\|\underline{u}^{(n+1)} - \underline{u}\|_{\ell^2} \leq \mu\|\underline{u}^{(n)} - \underline{u}\|_{\ell^2} + \theta\|\underline{\tilde{r}}^{(n)} - \underline{r}^{(n)}\|_{\ell^2} + \|\underline{\varepsilon}^{(n)}\|_{\ell^2}. \tag{3.6.26}$$

Iterating (3.6.26), we then obtain:

$$\|\underline{u}^{(n+1)} - \underline{u}\|_{\ell^2} \leq \mu^{n+1}\|\underline{u}^{(0)} - \underline{u}\|_{\ell^2} + \theta \sum_{i=0}^{n} \mu^{n-i}\|\underline{\tilde{r}}^{(i)} - \underline{r}^{(i)}\|_{\ell^2} + \sum_{i=0}^{n} \mu^{n-i}\|\underline{\varepsilon}^{(i)}\|_{\ell^2}. \tag{3.6.27}$$

The first term on the right hand side converges to zero, since, by Lemma 3.6.2 we have that $\mu < 1$. Let us then bound the remaining two terms.

By applying Corollary 3.6.1 with $\underline{w} = \underline{u}^{(i)}$, together with the stability result (3.6.19) we obtain:

$$\|\underline{\tilde{r}}^{(i)} - \underline{r}^{(i)}\|_{\ell^2} \leq C(\|\underline{u}^{(0)}\|_{\ell^2}, \|\underline{g}\|_{\ell^2})\ \epsilon_i$$

As far as $\|\underline{\varepsilon}^{(n)}\|_{\ell^2}$ is concerned, by applying Theorem 3.2.1, we have that:

$$\|\underline{\varepsilon}^{(n)}\|_{\ell^2} \lesssim N_n^{-(\frac{1}{\tau}+\frac{1}{2})}\|\underline{u}^{(n)} + \theta\underline{\tilde{r}}^{(n)}\|_{\ell^\tau_w}.$$

Let us then bound the ℓ^τ_w norm on the right hand side. Thanks to Corollary 3.6.2, we have:

$$\begin{aligned}\|\underline{u}^{(n)} + \theta\underline{\tilde{r}}^{(n)}\|_{\ell^\tau_w} &\leq \|\underline{u}^{(n)}\|_{\ell^\tau_w} + \theta\|\underline{\tilde{r}}^{(n)}\|_{\ell^\tau_w} \\ &\lesssim C(\underline{u}^{(0)}, g).\end{aligned} \tag{3.6.28}$$

which yields

$$\|\underline{\varepsilon}^{(n)}\|_{\ell^2} \lesssim N_n^{-(\frac{1}{\tau}-\frac{1}{2})}C(\underline{g}, \underline{u}^{(0)}).$$

Combining such bound with (3.6.27) and using Lemma 3.6.4 imply the thesis:

$$\|\underline{u}^{(n+1)}-\underline{u}\|_{\ell^2} \lesssim \mu^{n+1}\|\underline{u}^{(0)}-\underline{u}\|_{\ell^2}+C(\|\underline{u}^{(0)}\|_{\ell^2}, \|\underline{g}\|_{\ell^2}) \left(\theta\sum_{i=0}^{n}\mu^{n-i}\epsilon_i + \sum_{i=0}^{n}\mu^{n-i}N_i^{-(\frac{1}{\tau}-\frac{1}{2})}\right)$$

□

Remark 3.6.1: *We would like to point out that in order for the sum*

$$\sum_{i=0}^{n}\prod_{\ell=i}^{n}(\mu + C\theta\epsilon_\ell)$$

at the right hand side of (3.6.22) to converge it is sufficient that

$$(\mu + C\theta\epsilon_\ell) \leq c < 1, \qquad \forall \ell > \bar{\ell},$$

for some $\bar{\ell}$. The method is then stable also if ϵ_0 is not smaller than $\bar{\epsilon}$. Clearly, in such case the stability constant will depend on the sequence (ϵ_n). Such *dependence is however not a problem as far as the asymptotic is concerned, since such constant gets smaller if the sequence* (ϵ_n) *gets smaller element-wise. In particular one can choose a reference sequence* $(\bar{\epsilon}_n)$ *and for all sequence* (ϵ_n) *such that* $\epsilon_n < \bar{\epsilon}_n$ *the stability constant is uniformly bounded by a constant depending on* $(\bar{\epsilon}_n)$.

Using norm equivalences (4.2.5), the above result can be translated in terms of the corresponding continuous problem (3.3.1), as stated by the following corollary:

Corollary 3.6.3: *Let* u *be the solution of (3.3.2) and* g *belong to* $B^{r+s}_{\tau,\tau}(\Omega)$, *with* τ *given by* $1/\tau = r/d + 1/2$ *and with* r *such that* $0 < r + s \leq \min(S, d/(\max\{\tilde{\tau}, \tau_1\}) - d/2)$. *If* $u^{(n+1)} = \sum_\lambda u^{(n)}_\lambda \check{\psi}_\lambda$ *is the non linear approximation of* u *at step* $n+1$ *of the non-linear Richardson scheme, with* $\theta < \theta_0$, *then it holds, for some* $\mu < 1$:

$$\|u - u^{(n)}\|_{H^s(\Omega)} \lesssim \sum_{i=0}^{n} \mu^{n-i} N_i^{-\frac{r}{d}} + \frac{\theta}{2} \sum_{i=0}^{n} \mu^{n-i} \epsilon_i + \mu^{n+1} \|u^{(0)} - u\|_{H^s(\Omega)},$$

for all $n \in \mathbb{N}$, *where the implicit constants depend only on initial data.*

Remark 3.6.2: *Though for simplicity we fixed throughout this paper a variational framework corresponding to an elliptic Neumann BVP on the bounded domain* Ω, *it is not difficult to check that the whole proof of the result obtained relies on the representation (3.4.1) of the Problem and on the norm equivalences (3.2.1) and (3.2.3). Therefore such results carry over to much more general situations (Dirichlet BVP, Integral equations) where the space* $B^{r+s}_{\tau,\tau}$ *is substituted by the space for which a representation of the form (3.2.3) holds.*

3.6.3 Choosing the tolerances

We now need to choose the sequences (N_i) of the number of d.o.f. to be retained at each iteration and (ϵ_i) of the tolerances to be used in the prevision step. In order to do so we imposing a sort of "balancing" between the terms of the sums in equation (3.6.1), in such a way that all contributions to the error have roughly the same order. More precisely we ask that

$$\mu^{n+1} \sim \sum_{i=0}^{n} \mu^{n-i} \epsilon_i \sim \sum_{i=0}^{n} \mu^{n-i} N_i^{-(\frac{1}{\tau} - \frac{1}{2})}.$$

Since

$$\sum_{i=0}^{n} \mu^{n-i} \epsilon_i = \mu^{n+1} \sum_{i=0}^{n} \mu^{-i-1} \epsilon_i,$$

the tolerances ϵ_i should then be chosen in such a way that

$$\sum_{i=0}^{+\infty} \mu^{-i-1}\epsilon_i < +\infty.$$

It is not difficult to see that this holds for instance for the choice

$$\epsilon_i = \frac{\mu^{i+1}}{i \log i}.$$

Analogously, we will choose N_i

$$N_i = \left(\frac{i \log i}{\mu^{i+1}}\right)^{\frac{2\tau}{2-\tau}}.$$

If we are interested in approximating u with N degrees of freedom, we will then have to stop for

$$N = N_n = \left(\frac{n \log n}{\mu^{n+1}}\right)^{\frac{2\tau}{2-\tau}},$$

that is at iteration $n(N)$, $n(N)$ being the smallest integer such that

$$\log n + \log\log n + (n+1)|\log\mu| \geq (\frac{2\tau}{2-\tau})\log N.$$

It is not difficult to realize that

$$n \geq (1/\tau - 1/2)\frac{\log N}{2 + |\log\mu|} - 1.$$

3.7 Nonconforming domain decomposition

3.7.1 Functional Setting

Let $\Omega \subset I\!R^n$ be a bounded polygonal domain with Lipschitz boundary $\partial\Omega$. Let Ω be decomposed into a finite number of non–overlapping polygonal subdomains Ω^k, $k = 1, \ldots, K$,

$$\overline{\Omega} = \bigcup_{k=1}^{K} \overline{\Omega}^k, \qquad \Omega^k \cap \Omega^m = \emptyset, \quad k \neq m, \tag{3.7.1}$$

where

$$\Sigma = \left(\bigcup_{k=1}^{K} \partial\Omega^k\right) \setminus \partial\Omega \tag{3.7.2}$$

is called the *skeleton* of the decomposition. We will also employ the notation

$$\Gamma^k := \partial\Omega^k \setminus \partial\Omega \tag{3.7.3}$$

so that

$$\Sigma = \bigcup_{k=1}^{K} \Gamma^k. \tag{3.7.4}$$

Let $H^{1/2}(\partial\Omega^k)$ and $H_{00}^{1/2}(\Gamma^k)$ be defined as the trace space of, respectively, $H^1(\Omega^k)$ and $H^1_{\partial\Omega}(\Omega^k) := \{v^k \in H^1(\Omega^k),\ v^k = 0 \text{ on } \partial\Omega \cap \partial\Omega^k\}$, with norms

$$\begin{aligned} \|\sigma\|_{H^{1/2}(\partial\Omega^k)} &:= \inf_{v^k \in H^1(\Omega^k):\ v^k|_{\partial\Omega^k} = \sigma} \|v^k\|_{H^1(\Omega^k)}, \\ \|\sigma\|_{H_{00}^{1/2}(\Gamma^k)} &:= \inf_{v^k \in H^1_{\partial\Omega}(\Omega^k):\ v^k|_{\Gamma^k} = \sigma} \|v^k\|_{H^1(\Omega^k)}. \end{aligned} \tag{3.7.5}$$

Remark 3.7.1: *Equivalent norms for $H^{1/2}(\partial\Omega^k)$ and for $H_{00}^{1/2}(\Gamma^k)$ can be defined through the norms of $H^{1/2}(]0,1[^{n-1})$ and $H_{00}^{1/2}(]0,1[^{n-1})$ by using an atlas and partitions of unity where the constants in the equivalence depend on the diameter of Ω^k.*

For each k, let $H_{\partial\Omega}^{1/2}(\Gamma^k)$ be either $H^{1/2}(\Gamma^k)$ or $H_{00}^{1/2}(\Gamma^k)$, depending on whether Γ^k is a closed or an open set, with norm

$$|\sigma|_{1/2,\Gamma^k} := \begin{cases} \|\sigma\|_{H_{00}^{1/2}(\Gamma^k)}, & \text{if } \Gamma^k \cap \partial\Omega \neq \emptyset, \\ \|\sigma\|_{H^{1/2}(\partial\Omega^k)}, & \text{otherwise, (i.e., if } \Gamma^k = \partial\Omega^k). \end{cases} \tag{3.7.6}$$

We will denote by $(\cdot,\cdot)_{1/2,\Gamma^k}$ the corresponding inner product. Moreover, let $H_{\partial\Omega}^{-1/2}(\Gamma^k)$ be the corresponding dual, whose norm and inner product will be denoted by $|\cdot|_{-1/2,\Gamma^k}$ and $(\cdot,\cdot)_{-1/2,\Gamma^k}$. Duality between $H_{\partial\Omega}^{1/2}(\Gamma^k)$ and $H_{\partial\Omega}^{-1/2}(\Gamma^k)$ will be written as $\langle\cdot,\cdot\rangle_k$.

We can now introduce the functional setting for the domain decomposition method we are going to consider. Let V be the product space

$$V := \{(v^1,\ldots,v^K),\ v^k \in H^1(\Omega^k),\ v^k = 0 \text{ on } \partial\Omega \cap \partial\Omega^k,\ k = 1,\cdots,K\} \tag{3.7.7}$$

which is isomorphic to

$$\{v \in L_2(\Omega):\ v^k = v|_{\Omega^k} \in H^1(\Omega^k),\ v^k = 0 \text{ on } \partial\Omega \cap \partial\Omega^k,\ k = 1,\cdots,K\},$$

endowed with the norm

$$\|v\|_V^2 := \sum_{k=1}^{K} \|v^k\|^2_{H^1(\Omega^k)}, \qquad v \in V, \tag{3.7.8}$$

induced by the inner product

$$(u,v)_V := \sum_{k=1}^{K} (u^k, v^k)_{H^1(\Omega^k)}. \tag{3.7.9}$$

Moreover, let Λ and Φ be defined by

$$\Lambda := \prod_{k=1}^{K} H^{-1/2}_{\partial\Omega}(\Gamma^k) \tag{3.7.10}$$

and $\Phi = H^1_0(\Omega)|_\Sigma$, that is

$$\Phi := \{\sigma \in L_2(\Sigma) : \text{ there exists } v \in H^1_0(\Omega) \text{ such that } \sigma = v|_\Sigma\}. \tag{3.7.11}$$

Λ and Φ are endowed with the norms

$$\|\mu\|_\Lambda := \sum_{k=1}^{K} |\mu^k|^2_{-1/2,\Gamma^k}, \qquad \|\sigma\|_\Phi := \inf_{v\in H^1_0(\Omega):\ v|_\Sigma=\sigma} \|v\|_{H^1(\Omega)}, \tag{3.7.12}$$

respectively. We remark that $H^1_0(\Omega)$ can be identified with a subset of V,

$$H^1_0(\Omega) \cong \{v = (v^k)_{k=1,\cdots,K} \in V : \quad \text{there exists } \sigma \in \Phi,\ v^k = \sigma \text{ on } \Gamma^k\} \subset V.$$

In the following, when writing $v = (v^k) \in H^1_0(\Omega)$ for an element $v \in V$, we will refer to such an isomorphism. Moreover, for the sake of notational simplicity we will write (v^k) for $(v^k)_{k=1,\cdots,K}$, always assuming that, unless otherwise stated, the index k ranges from 1 to K.

An observation that will be important in the sequel is the following:

Proposition 3.7.2: *It holds*

$$\Lambda' = \prod_{k=1}^{K} H^{1/2}_{\partial\Omega}(\Gamma^k), \qquad \|\sigma\|_{\Lambda'} \sim \left(\sum_{k=1}^{K} |\sigma^k|^2_{1/2,\Gamma^k}\right)^{1/2}. \tag{3.7.13}$$

Moreover, one has that

(i) the space Φ can be identified with a proper subset of Λ', by identifying $\sigma \in \Phi$ with $(\sigma^k) \in \Lambda'$, $\sigma^k = \sigma|_{\Gamma^k}$;

(ii) for $\sigma \in \Phi$ the equivalences

$$\|\sigma\|_\Phi \sim \left(\sum_{k=1}^{K} |\sigma^k|_{1/2,\Gamma^k}\right)^{1/2} \sim \|\sigma\|_{\Lambda'}, \qquad (\sigma^k = \sigma|_{\Gamma^k}), \tag{3.7.14}$$

hold.

Proof: The proof of (3.7.13) follows by standard arguments since the dual of a Cartesian product of spaces is the product of the duals.

As far as (ii) is concerned, let now $\sigma \in \Phi$ and let $v \in H_0^1(\Omega)$ be any function such that $v = \sigma$ on Σ. Then we have

$$\sum_{k=1}^{K} |\sigma|^2_{1/2,\Gamma^k} \leq \sum_{k=1}^{K} \|v\|^2_{H^1(\Omega^k)} = \|v\|^2_{H_0^1(\Omega)}.$$

Since V is arbitrary, this yields

$$\sum_{k=1}^{K} |\sigma|^2_{1/2,\Gamma^k} \lesssim \inf_{v \in H_0^1(\Omega):\ v|_\Sigma = \sigma} \|v\|^2_{H^1(\Omega)} = \|\sigma\|^2_\Phi.$$

Let now u in $H_0^1(\Omega)$ be defined such that $-\Delta u = 0$ in Ω^k and $u = \sigma$ on Γ^k for all k. Then one has

$$\|\sigma\|^2_\Phi \leq \|u\|^2_{H_0^1(\Omega)} = \sum_{k=1}^{K} \|u\|^2_{H^1(\Omega^k)} \lesssim \sum_{k=1}^{K} |\sigma|^2_{1/2,\Gamma^k}.$$

■

3.7.2 The Three-Fields formulation

We consider the second order elliptic boundary value problem

$$\begin{aligned} -\operatorname{div} a(x) \operatorname{grad} u(x) &= f && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega, \end{aligned} \tag{3.7.15}$$

with a sufficiently smooth and uniformly positive definite matrix $a(x)$ and $f \in L_2(\Omega)$. To solve this by a domain decomposition approach, for each $k \in \{1, \ldots, K\}$ let $a^k : H^1(\Omega^k) \times H^1(\Omega^k) \to I\!R$ be the bilinear form induced by the differential operator on the subdomain Ω^k,

$$a^k(u, v) = \int_{\Omega^k} a(x)\ \operatorname{grad} u \cdot \operatorname{grad} v \, dx.$$

A composed bilinear form $a : V \times V \to I\!R$ can be defined by

$$a(\cdot,\cdot) := \sum_{k=1}^{K} a^k(\cdot,\cdot). \tag{3.7.16}$$

For all $v \in H_0^1(\Omega)$, $a(\cdot,\cdot)$ satisfies

$$a(v,v) = \sum_{k=1}^{K} \int_{\Omega^k} a(x) |\operatorname{grad} v|^2 \, dx \sim \|v\|^2_{H^1(\Omega)}. \tag{3.7.17}$$

In alternative to the standard weak formulation of the boundary value problem (3.7.15),

$$\begin{aligned} &\text{find } u^* \in H_0^1(\Omega) \text{ such that} \\ &\quad a(u^*, v) = (f, v)_{L_2(\Omega)} \qquad \text{for all } v \in H_0^1(\Omega), \end{aligned} \tag{3.7.18}$$

we can then consider the *Three Fields Formulation*: find $(u, \lambda, \varphi) \in V \times \Lambda \times \Phi$ such that

$$\begin{cases} a(u,v) - \displaystyle\sum_{k=1}^{K} \langle \lambda^k, v^k \rangle_k &= (f,v)_{L_2(\Omega)} \qquad \text{for all } v \in V, \\ \displaystyle\sum_{k=1}^{K} \langle \varphi - u^k, \mu^k \rangle_k &= 0 \qquad \text{for all } \mu \in \Lambda, \\ \displaystyle\sum_{k=1}^{K} \langle \lambda^k, \sigma \rangle_k &= 0 \qquad \text{for all } \sigma \in \Phi. \end{cases} \tag{3.7.19}$$

This formulation was introduced in [23], where it was shown that problem (3.7.19) has for every $f \in L_2(\Omega)$ a unique solution (u, λ, φ), satisfying

$$\begin{cases} u^k = u^* & \text{in } \Omega^k, \quad k = 1, \ldots, K, \\ \lambda^k = a \dfrac{\partial u^*}{\partial n^k} & \text{on } \partial\Omega^k, \quad k = 1, \ldots, K, \\ \varphi = u^* & \text{on } \Sigma, \end{cases} \tag{3.7.20}$$

where $\partial u^* / \partial n^k$ is the outward normal derivative of the restriction of u^* to Ω^k.

We can write the system (3.7.19) more conveniently in operator form as follows. Define $A : V \to V'$ by

$$\langle Av, v \rangle := a(v, v).$$

For each $k = 1, \ldots, K$, let $B^k : H^1_{\partial\Omega}(\Omega^k) \to H^{1/2}_{\partial\Omega}(\Gamma^k)$ denote the *trace operator*. Let then $B : V \to \Lambda'$ be defined by $B := \operatorname{diag}(B^1, \ldots, B^K)$. Moreover, let $C^k : H^{-1/2}_{\partial\Omega}(\Gamma^k) \to \Phi'$ be defined as

$$\langle C^k \mu^k, \sigma \rangle := \langle \mu^k, \sigma \rangle_k,$$

and let $C : \Lambda \to \Phi'$ be assembled as $C := \text{diag}(C^1, \dots, C^K)$. With these notations, the system (3.7.19) can be written as follows: find $(u, \lambda, \varphi) \in V \times \Lambda \times \Phi$ as solution of

$$\begin{pmatrix} A & -B^T & 0 \\ -B & 0 & C^T \\ 0 & C & 0 \end{pmatrix} \begin{pmatrix} u \\ \lambda \\ \varphi \end{pmatrix} = \begin{pmatrix} f \\ 0 \\ 0 \end{pmatrix}. \tag{3.7.21}$$

Remark 3.7.3: *[23] One of the interests for the Three Fields Formulation lies in the observation that, for given $\varphi \in \Phi$, the computation of u and λ reduces to solving K independent Dirichlet problems on the subdomains Ω^k. Each of these is of the form*

$$\begin{pmatrix} A^k & -(B^k)^T \\ -B^k & 0 \end{pmatrix} \begin{pmatrix} u^k \\ \lambda^k \end{pmatrix} = \begin{pmatrix} f^k \\ -(C^k)^T \varphi \end{pmatrix}. \tag{3.7.22}$$

φ can then be computed as the solution of

$$\mathbf{C}\mathbf{A}^{-1}\mathbf{C}^T \varphi = \mathbf{C}\mathbf{A}^{-1} \begin{pmatrix} f \\ 0 \end{pmatrix} \tag{3.7.23}$$

where

$$\mathbf{A} := \begin{pmatrix} A & -B^T \\ -B & 0 \end{pmatrix}, \qquad \mathbf{C} := (0 \ \ C), \tag{3.7.24}$$

Here the operator $\mathbf{S} := \mathbf{C}\mathbf{A}^{-1}\mathbf{C}^T$ is just the Poincaré-Steklov operator on Σ.

Hence, by applying a Schur complement technique, the solution of the original problem (3.7.15) or equivalently of the problem (3.7.20), is reduced to the solution of the equation on the "trace" unknown:

$$\mathbf{S}\varphi = g, \tag{3.7.25}$$

with $g = \mathbf{C}\mathbf{A}^{-1} \begin{pmatrix} f \\ 0 \end{pmatrix}$.

It has been shown in [23] the following result

Lemma 3.7.4: $\mathbf{S}$ *is an isomorphism from Φ to Φ' that satisfies in addition* ${}_{\Phi'}\langle \mathbf{S}\cdot, \cdot \rangle_{\Phi} \sim \| \cdot \|_{\Phi}^2$.

3.7.3 An adaptive wavelet method

Our aim is to solve the continuous linear problem

$$\mathbf{S}\varphi = g \tag{3.7.26}$$

by means of an adaptive wavelet method. According to the "new approach" we first have to transform the initial continuous problem (3.7.26) into an equivalent ∞-dimensional problem.

To do this let us assume that we have a couple of biorthogonal wavelet bases for $L_2(\Sigma)$

$$\begin{aligned} \boldsymbol{\psi} &:= \{\psi_{j,m},\ (j,m) \in \nabla := \cup_{j\geq j_0}\nabla_j\}, \\ \tilde{\boldsymbol{\psi}} &:= \{\tilde{\psi}_{j,m},\ (j,m) \in \nabla := \cup_{j\geq j_0}\nabla_j\}, \end{aligned} \tag{3.7.27}$$

with the following properties:

(P1) any function $\sigma \in L_2(\Sigma)$ can be expanded in terms of either $\boldsymbol{\psi}$ or $\tilde{\boldsymbol{\psi}}$,

$$\sigma = \sum_{j\geq j_0}\sum_{m\in\nabla_j} \langle \sigma, \tilde{\psi}_{j,m}\rangle \psi_{j,m} = \sum_{j\geq j_0}\sum_{m\in\nabla_j} \langle \sigma, \psi_{j,m}\rangle_k \tilde{\psi}_{j,m}, \tag{3.7.28}$$

($\langle\cdot,\cdot\rangle$ denoting here the $L_2(\Sigma)$ scalar product);

(P2) one has $\psi_{j,m} \in \Phi$, and the following *norm equivalence* holds

$$\|\sigma\|_\Phi^2 \sim \sum_{j\geq j_0} 2^j \sum_{m\in\nabla_j} |\langle \sigma, \tilde{\psi}_{j,m}\rangle|^2, \qquad \sigma \in \Phi; \tag{3.7.29}$$

(P3) $\boldsymbol{\psi}$ and $\tilde{\boldsymbol{\psi}}$ have local support, i.e.,

$$\mathrm{diam}(\mathrm{supp}\, \psi_{j,m}) \sim \mathrm{diam}(\mathrm{supp}\, \tilde{\psi}_{j,m}) \sim 2^{-j}. \tag{3.7.30}$$

There are by now a number of constructions of such biorthogonal wavelets [34], [25], [42], [43] that can be applied to the present setting. In particular we refer to [15] for a a particularly simple construction which, in the two-dimensional case, is sufficient for the present purpose.

Now we decompose the functions φ and $\mathbf{S}\varphi$, by choosing two suitable rescaled versions $\{\hat{\boldsymbol{\psi}}, \hat{\tilde{\boldsymbol{\psi}}}\}$ and $\{\check{\boldsymbol{\psi}}, \check{\tilde{\boldsymbol{\psi}}}\}$, of the given pair of biorthogonal wavelet bases $\{\boldsymbol{\psi}, \tilde{\boldsymbol{\psi}}\}$, such that

$$\varphi \in \Phi \qquad \varphi = \sum_{j\geq j_0}\sum_{m\in\nabla_j} \varphi_{j,m}\check{\psi}_{j,m}, \qquad \check{\psi}_{j,m} = 2^{-j}\psi_{j,m};$$

$$\mathbf{S}\varphi \in \Phi' \qquad \mathbf{S}\varphi = \sum_{j \geq j_0} \sum_{m \in \nabla_j} s_{j,m} \check{\psi}_{j,m}, \qquad \hat{\psi}_{j,m} = 2^j \psi_{j,m}.$$

Now we build an ∞-dimensional operator $\mathcal{S}$ which acts on wavelet coefficients as follows:

$$\mathcal{S} : \underline{\varphi} = \{\varphi_{j,m}\}_{j,m} \to \underline{s} = \{s_{j,m}\}_{j,m}. \tag{3.7.31}$$

Thanks to (P.2), the above operator $\mathcal{S}$ is an isomorphism from ℓ^2 onto ℓ^2:

$$\mathcal{S} : \ell^2 \to \ell^2.$$

Thus to solve the continuous problem (3.7.26) is equivalent to solve

$$\mathcal{S}\underline{\varphi} = \underline{g},$$

where $\underline{\varphi} \in \ell^2$ is the infinite vector of the wavelet coefficients of the unknown solution, while the infinite vector $\underline{g} = \{g_{j,m}\}_{j,m}$, contains the wavelet coefficients of the function $g \in \Phi'$.

Now let us consider the following **$\mathcal{S}$-nonlinear Richardson** scheme for the solution of the ∞-dimensional problem $\mathcal{S}\underline{\varphi} = \underline{g}$:

given $\underline{\varphi}^0$

for $i = 0, \ldots$

1. Compute $(\mathcal{S}\underline{\varphi}^i)_{\varepsilon_i}$ approximation to $\mathcal{S}\varphi^i$, with precision ε_i.
2. Compute $\underline{\varphi}^{i+1} = \mathbb{P}_{N_{i+1}}(\underline{\varphi}^i + \theta(\underline{g} - (\mathcal{S}\underline{\varphi}^i)_{\varepsilon_i})$

end

where the nonlinear projector $\mathbb{P}_{N_{i+1}}$ retains the N_{i+1} largest, in absolute value, wavelet coefficients.

It is important to remark that computing $(\mathcal{S}\underline{\varphi}^i)_{\varepsilon_i}$, approximation to $\mathcal{S}\varphi^i$, is equivalent to approximately solving K decoupled Dirichlet problems, where K is the number of the subdomains. This will in general be done by applying an adaptive solver (not necessarily of wavelet type).

We analyze such a nonlinear Richardson-type scheme, as an element of a more general class of algorithms, which we present in the next section within an abstract (not necessarily wavelet) framework.

3.8 An abstract framework for Nonlinear Richardson-type algorithm

Assume X and Y are quasi-normed spaces, with Y continuously embedded into X, and that $\{S_N\}_{j\geq 0}$ and $\{T_M\}_{j\geq 0}$ are two unrelated sequences of nonlinear approximation spaces

$$S_0 \subset \ldots \subset S_N \subset S_{N+1} \subset \ldots \subset Y \subset X,$$

$$T_0 \subset \ldots \subset T_M \subset T_{M+1} \subset \ldots \subset Y \subset X,$$

such that for some $s > 0$ (s is called rate of convergence) one has Jackson-type estimates, for all $f \in Y$

$$\mathrm{dist}_X(f, S_N) = \inf_{g\in S_N} \|f - g\|_X \lesssim N^{-s}\|f\|_Y, \tag{3.8.1}$$

$$\mathrm{dist}_X(f, T_M) = \inf_{g\in T_M} \|f - g\|_X \lesssim M^{-s}\|f\|_Y \tag{3.8.2}$$

and Bernstein-type estimates

$$\|f\|_Y \lesssim N^s\|f\|_X \quad \text{if } f \in S_N, \tag{3.8.3}$$

$$\|f\|_Y \lesssim M^s\|f\|_X \quad \text{if } f \in T_M. \tag{3.8.4}$$

Moreover assume that

$$S_N + S_{N'} \subset S_{N+N'}, \quad N > 0, \tag{3.8.5}$$
$$T_M + T_{M'} \subset T_{M+M'}, \quad M > 0, \tag{3.8.6}$$

and there exist two nonlinear projectors

$$\mathbb{P}_N : X \to S_N, \tag{3.8.7}$$

and

$$\mathbb{Q}_M : X \to T_M, \tag{3.8.8}$$

which are quasi-optimal

$$\|f - \mathbb{P}_N f\|_X \lesssim \mathrm{dist}_X(f, S_N) \lesssim N^{-s}\|f\|_Y, \quad f \in X, \tag{3.8.9}$$
$$\|f - \mathbb{Q}_M f\|_X \lesssim \mathrm{dist}_X(f, T_M) \lesssim M^{-s}\|f\|_Y, \quad f \in X, \tag{3.8.10}$$

and X, Y-contractive:

$$\|\mathbb{P}_N(f)\|_X \leq \|f\|_X, \qquad \|\mathbb{P}_N(f)\|_Y \leq \|f\|_Y, \tag{3.8.11}$$
$$\|\mathbb{Q}_M(f)\|_X \leq \|f\|_X, \qquad \|\mathbb{Q}_M(f)\|_Y \leq \|f\|_Y. \tag{3.8.12}$$

Finally we suppose that $\mathbb{P}_{N_i}$ are such that for all $f \in X$ there exist $S^f_{N_i} \subset S_{N_i}$ linear subspace of S_{N_i} and

$$\Pi_{N_i} : f \in X \to \Pi_{N_i}(f) \in \tilde{S}^f_{N_i}, \tag{3.8.13}$$

bounded linear projector, such that

$$\Pi_{N_i}(f) = \mathbb{P}_{N_i}(f).$$

Let now consider the continuous linear operator $\mathcal{L} : X \to X$ such that when it is restricted to more regular space Y, which is continuously embedded into X, it preserves such regularity:

$$\mathcal{L}_{|Y} : Y \to Y.$$

Let us assume that

(S.1) there exists θ_0 such that for all θ with $0 < \theta < \theta_0$ it holds

$$\|I - \theta\mathcal{L}\|_{X\to X} \leq \mu < 1, \tag{3.8.14}$$

$$\|I - \theta\mathcal{L}\|_{Y\to Y} \leq \gamma < 1. \tag{3.8.15}$$

(S.2) there exists an approximation strategy for $\mathcal{L}\varphi$ such that for all $\varepsilon > 0$, there exists a $(\mathcal{L}\varphi)_\varepsilon$ with

$$(\mathcal{L}\varphi)_\varepsilon \in T_{M_\varepsilon}, \qquad M_\varepsilon \text{ depending on } \varepsilon, \tag{3.8.16}$$

satisfying the following inequality

$$\|\mathcal{L}\varphi - (\mathcal{L}\varphi)_\varepsilon\|_X \leq \varepsilon\|\varphi\|_X, \tag{3.8.17}$$

for all $\varphi \in X$.

The approximation strategy for $\mathcal{L}\varphi$ can be regarded as a black-box approximation strategy satisfying inequality (3.8.17). For instance, in the context of three-field formulation, it has been studied [14] a finite element based approximation strategy for $\mathbf{S}\varphi \in \Phi'$, where $\mathbf{S}$ is the continuous Steklov-Poincaré operator, such that for every $\varepsilon > 0$ it exists a $(\mathbf{S}\varphi)_\varepsilon$ verifying

$$\|\mathbf{S}\varphi - (\mathbf{S}\varphi)_\varepsilon\|_{\Phi'} \leq \varepsilon\|\tau\|_\Phi.$$

By using norm equivalences, the above inequality can be equivalently restated in wavelet coordinates

$$\|\mathcal{S}\underline{\varphi} - (\mathcal{S}\underline{\varphi})_\varepsilon\|_{\ell^2} \leq \varepsilon\|\underline{\varphi}\|_{\ell^2}.$$

Generally no a-priori information is available about the relation between the tolerance ε and the number M_ε of degrees of freedom used to build the approximation. anyway nonlinear approximation provides a natural benchmark for this kind of relation and in view of this we give the following definition of optimal approximation strategy.

Definition *Let s be the rate of convergence associated to the sequence $\{T_M\}_M$ of nonlinear spaces of approximation. The approximation strategy in (S.2) is said to be* **optimal** *if*

$$f \in X \longrightarrow \varepsilon(M_\varepsilon)^s \leq K,$$

for all $\varepsilon > 0$.

We are interested in the following

Problem 3.8.1: *Given $f \in Y$, solve the linear equation*

$$\mathcal{L}\varphi = f. \tag{3.8.18}$$

In particular, given a number $\tilde{N} > 0$ of degrees of freedom, we are interested in finding an approximation $\varphi_{\tilde{N}} \in S_{\tilde{N}}$ to the exact solution φ^* of (3.8.18).

We consider the following **$\mathcal{L}$-nonlinear Richardson** scheme:

given $\varphi^0 \in S_{N_0}$

for $i = 0, \ldots$

1. Compute $(\mathcal{L}\varphi^i)_{\varepsilon_i} \in T_{M_{\varepsilon_i}}$
2. Compute $\varphi^{i+1} = \mathbb{P}_{N_{i+1}}(\varphi^i + \theta(f - (\mathcal{L}\varphi^i)_{\varepsilon_i})) \in S_{N_{i+1}}$

end

Remark 3.8.1: *It is important to remark that the above scheme is able to couple two eventually different approximation strategies (e.g. finite element-wavelet or finite element-finite element) into an iterative procedure, one strategy coming from the choice of the nonlinear spaces $\{S_N\}$ and the other from the choice of the nonlinear spaces $\{T_M\}$.*

Let us now prove the following result

Lemma 3.8.1: *If the sequence $\{\varepsilon_i\}$ is chosen such that*

$$\mu + \theta\varepsilon_i < 1, \tag{3.8.19}$$

then the $\mathcal{L}$-nonlinear Richardson is X-stable.

Proof: Using the X-contractivity of the nonlinear projector $\mathbb{P}_N$ and assumption (S.2) yields

$$\begin{aligned} \|\varphi^{i+1}\|_X &= \|\mathbb{P}_{N_{i+1}}(\varphi^i + \theta(f - (\mathcal{L}\varphi^i)_{\varepsilon_i}))\|_X \\ &\le \|\varphi^i + \theta(f - (\mathcal{L}\varphi^i)_{\varepsilon_i})\|_X \\ &\le \|(I - \theta\mathcal{L})\varphi^i\|_X + \theta\|(\mathcal{L}\varphi^i)_{\varepsilon_i} - \mathcal{L}\varphi^i\|_X + \theta\|f\|_X \\ &\le \mu\|\varphi^i\|_X + \theta\varepsilon_i\|\varphi^i\|_X + \theta\|f\|_X \\ &\le (\mu + \theta\varepsilon_i)\|\varphi^i\|_X + \theta\|f\|_X. \end{aligned}$$

By iterating the above inequality we obtain

$$\|\varphi^{i+1}\|_X \le \|\varphi^0\|_X \Pi_{k=0}^i(\mu + \theta\varepsilon_k)^k + \theta\|f\|_X \sum_{k=0}^{i}(\mu + \theta\varepsilon_k)^k, \tag{3.8.20}$$

which yields the inequality

$$\|\varphi^{i+1}\|_X \le C_X, \tag{3.8.21}$$

if we choose $\{\varepsilon_i\}$ such that

$$\mu + \theta\varepsilon_i \le \eta < 1.$$

■

Let us now discuss the Y-stability of the algorithm.

Lemma 3.8.2: *Under the assumption of Lemma 3.8.1, if the sequences $\{N_k\}$ and $\{\varepsilon_k\}$ are chosen in such a way that*

$$\sum_{k=0}^{\infty} \mu^{i-k} N_{k+1}^s \varepsilon_k < \infty, \tag{3.8.22}$$

then the $\mathcal{L}$-nonlinear Richardson is Y-stable.

Proof: Using the definition of the linear projector $\Pi_{N_{i+1}}$ yields

$$\begin{aligned} \|\varphi^{i+1}\|_Y &= \|\mathbb{P}_{N_{i+1}}(\varphi^i + \theta(f - (\mathcal{L}\varphi^i)_{\varepsilon_i}))\|_Y \\ &= \|\Pi_{N_{i+1}}(\varphi^i + \theta(f - (\mathcal{L}\varphi^i)_{\varepsilon_i}))\|_Y \\ &\le \|\Pi_{N_{i+1}}(\varphi^i + \theta(f - \mathcal{L}\varphi^i))\|_Y + \theta\|\Pi_{N_{i+1}}(\mathcal{L}\varphi^i - (\mathcal{L}\varphi^i)_{\varepsilon_i})\|_Y. \end{aligned}$$

As $\Pi_{i+1}(\mathcal{L}\varphi^i - (\mathcal{L}\varphi^i)_{\varepsilon_i}) \in S_{N_{i+1}}$, then by using Bernstein inequality (3.8.3), we have

$$\|\varphi^{i+1}\|_Y \le \|(I - \theta S)\varphi^i\|_Y + \theta\|f\|_Y + N_{i+1}^s\|\mathcal{L}\varphi^i - (\mathcal{L}\varphi^i)_{\varepsilon_i}\|_X.$$

Then assumptions (S.1) and (S.2) yield

$$\begin{aligned}\|\varphi^{i+1}\|_Y &\le \mu\|\varphi^i\|_Y + \theta\|f\|_Y + N_{i+1}^s\varepsilon_i\|\varphi^i\|_X \\ &\le \mu\|\varphi^i\|_Y + \theta\|f\|_Y + C_X N_{i+1}^s\varepsilon_i.\end{aligned}$$

By iterating the above inequality we obtain

$$\|\varphi^{i+1}\|_Y \le \mu^{i+1}\|\varphi^0\|_Y + C_X\sum_{k=0}^{i}\mu^{i-k}N_{k+1}^s\varepsilon_k + \frac{\theta}{1-\mu}\|f\|_Y.$$

Hence the property

$$\|\varphi^{i+1}\|_Y \le C_Y$$

follows if we choose the sequences $\{N_k\}_k$ and $\{\varepsilon_k\}$ such that

$$\sum_{k=0}^{\infty}\mu^{i-k}N_{k+1}^s\varepsilon_k < \infty.$$

■

Theorem 3.8.1: *Under the assumptions of Lemmas 3.8.1 and 3.8.2, if the approximation strategy for $\mathcal{L}\varphi$ is optimal, then the following error estimate holds*

$$\|\varphi^{i+1}-\varphi\|_X \lesssim \sum_{k=0}^{i}\mu^{i-k}N_{k+1}^{-s} + \mu^{i+1}\|\varphi^0-\varphi\|_X + \sum_{k=0}^{i}\mu^{i-k}\varepsilon_k, \quad (3.8.23)$$

with $\mu < 1$ and all the constants depending only on the initial data.

Proof: By using Jackson estimate (3.8.1), we have

$$\begin{aligned}\|\varphi^{i+1}-\varphi\|_X &\le \|(\mathbb{P}_{N_{i+1}}-I)(\varphi^i+\theta(f-(\mathcal{L}\varphi^i)_{\varepsilon_i}))\|_X \\ &\quad +\|(I-\theta\mathcal{L})(\varphi^i-\varphi)\|_X + \theta\|\mathcal{L}\varphi^i-(\mathcal{L}\varphi^i)_{\varepsilon_i}\|_X \\ &\le N_{i+1}^{-s}\|\varphi^i+\theta(f-(\mathcal{L}\varphi^i)_{\varepsilon_i})\|_Y + \mu\|\varphi^i-\varphi\|_X + \theta\varepsilon_i\|\varphi^i\|_X.\end{aligned}$$

Let us now estimate the term

$$\mathcal{U}^i := \|\varphi^i+\theta(f-(\mathcal{L}\varphi^i)_{\varepsilon_i})\|_Y.$$

We have

$$\begin{aligned}\mathcal{U}^i &\le \|(I-\theta\mathcal{L})\varphi^i\|_Y + \theta\|(\mathcal{L}\varphi^i)_{\varepsilon_i}-\mathcal{L}\varphi^i\|_Y + \theta\|f\|_Y \\ &\le \gamma\|\varphi^i\|_Y + \theta\|(\mathcal{L}\varphi^i)_{\varepsilon_i}-\mathcal{L}\varphi^i\|_Y + \theta\|f\|_Y\end{aligned}$$

If we denote by M_{ε_i} the number of degrees of freedom used to build $(\mathcal{L}\varphi^i)_{\varepsilon_i}$, then

$$\mathbb{Q}_{M_{\varepsilon_i}}(\mathcal{L}\varphi^i) \in T_{M_{\varepsilon_i}},$$

with $\#T_{i(\varepsilon_i)} = M(\varepsilon_i)$, provides a quasi-optimal M_{ε_i}-term approximation of $\mathcal{L}\varphi^i$ in the norm of X.

It follows that

$$\mathcal{U}^i \leq \gamma\|\varphi^i\|_Y + \theta\|(\mathcal{L}\varphi^i)_{\varepsilon_i} - \mathbb{Q}_{M_{\varepsilon_i}}(\mathcal{L}\varphi)\|_Y + \theta\|\mathbb{Q}_{M_{\varepsilon_i}}(\mathcal{L}\varphi^i) - \mathcal{L}\varphi^i\|_Y + \theta\|f\|_Y.$$

Remarking that $\mathbb{Q}_{M_{\varepsilon_i}}(\mathcal{L}\varphi) - (\mathcal{L}\varphi^i)_{\varepsilon_i} \in T_{2M_{\varepsilon_i}}$ allows to use Bernstein-type inequality (3.8.4) which gives

$$\|\mathbb{Q}_{M_{\varepsilon_i}}(\mathcal{L}\varphi^i) - (\mathcal{L}\varphi^i)_{\varepsilon_i}\|_Y \lesssim (2M_{\varepsilon_i})^s\|\mathbb{Q}_{M_{\varepsilon_i}}(\mathcal{L}\varphi^i) - (\mathcal{L}\varphi^i)_{\varepsilon_i}\|_X.$$

Hence, by using Y-continuity of the operator $\mathcal{L}$, we have

$$\begin{aligned}\mathcal{U}^i &\lesssim \gamma\|\varphi^i\|_Y + \theta(2M_{\varepsilon_i})^s\|(\mathcal{L}\varphi^i)_{\varepsilon_i} - \mathbb{Q}_{M_{\varepsilon_i}}(\mathcal{L}\varphi^i)\|_X \\ &\quad +\theta\|\mathbb{Q}_{M_{\varepsilon_i}}(\mathcal{L}\varphi^i) - \mathcal{L}\varphi^i\|_Y + \theta\|f\|_Y \\ &\lesssim \gamma\|\varphi^i\|_Y + \theta(2M_{\varepsilon_i})^s\|(\mathcal{L}\varphi^i)_{\varepsilon_i} - \mathbb{Q}_{M_{\varepsilon_i}}(\mathcal{L}\varphi^i)\|_X \\ &\quad +2\theta\|\mathcal{L}\varphi^i\|_Y + \theta\|f\|_Y \\ &\lesssim \gamma\|\varphi^i\|_Y + \theta(2M_{\varepsilon_i})^s\big[\|(\mathcal{L}\varphi^i)_{\varepsilon_i} - \mathcal{L}\varphi^i\|_X + \|\mathcal{L}\varphi^i - \mathbb{Q}_{M_{\varepsilon_i}}(\mathcal{L}\varphi)\|_X\big] \\ &\quad +2C\theta\|\varphi^i\|_Y + \theta\|f\|_Y.\end{aligned}$$

Now using assumption $(S.2)$ and quasi-optimality of the nonlinear projector $\mathbb{P}_N$ together with Jackson-type estimate (3.8.1) gives

$$\begin{aligned}\mathcal{U}^i &\lesssim \gamma\|\varphi^i\|_Y + \theta(2M_{\varepsilon_i})^s\varepsilon_i\|\varphi^i\|_X + \theta(2M_{\varepsilon_i})^s M_{\varepsilon_i}^{-s}\|\mathcal{L}\varphi^i\|_Y \\ &\quad 2C\theta\|\varphi^i\|_Y + \theta\|f\|_Y,\end{aligned}$$

where C is such that $\|\mathcal{L}\varphi\|_Y \leq C\|\varphi\|_Y$, for all $\varphi \in Y$.

Thanks to Y-continuity of $\mathcal{L}$ and to X-stability (3.8.20), we obtain the following inequality

$$\mathcal{U}^i \lesssim (\gamma + 2\theta C + \theta C'2^s)\|\varphi^i\|_Y + \theta C_X(2M_{\varepsilon_i})^s\varepsilon_i + \theta\|f\|_X,$$

where C' is such that $\|\mathcal{L}\varphi\|_X \leq C'\|\varphi\|_X$, for all $\varphi \in Y$.

By using Lemma 3.8.2, we finally obtain

$$\|\varphi^i + \theta(f - (\mathcal{L}\varphi^i)_{\varepsilon_i})\|_Y \lesssim (\gamma + 2\theta C + \theta C' 2^s) C_Y + \theta 2^s C_X M^s_{\varepsilon_i} \varepsilon_i + \theta \|f\|_X. \quad (3.8.24)$$

Now we go back to the error estimate.

It follows

$$\|\varphi^{i+1} - \varphi\|_X \lesssim N^{-s}_{i+1}\Big\{(\gamma + 2\theta C + \theta C' 2^s) C_Y + \theta 2^s C_X M^s_{\varepsilon_i} \varepsilon_i + \theta \|f\|_X\Big\} + \mu \|\varphi^i - \varphi\|_X + \theta \varepsilon_i \|\varphi^i\|_X.$$

Finally, by iterating the above inequality and by using X-stability, we obtain the following error estimate

$$\|\varphi^{i+1} - \varphi\|_X \lesssim \sum_{k=0}^{i} \mu^{i-k} N^{-s}_{k+1}\Big\{(\gamma + 2\theta C + \theta C' 2^s) C_Y + \theta 2^s C_X M^s_{\varepsilon_k} \varepsilon_k + \theta \|f\|_X\Big\} + \mu^{i+1} \|\varphi^0 - \varphi\|_X + \theta C_X \sum_{k=0}^{i} \mu^{i-k} \varepsilon_k.$$

Hence, as we assume that the strategy for $\mathcal{L}\varphi$ is optimal, i.e.

$$\varepsilon M^s_\varepsilon \leq K,$$

for all $\varepsilon > 0$, the thesis follows.

■

Definition *We say that the $\mathcal{L}$-nonlinear Richardson scheme is* **optimal**, *if it exhibits, after $i + 1$ iterations, an error reduction by a factor μ^{i+1}.*

We now need to choose the sequence $\{N_i\}$ of degrees of freedom to be retained at each iteration of the scheme and the sequence $\{\varepsilon_i\}$ of the tolerances related to the approximation of $\mathcal{L}\varphi^i$, in order to guarantee the optimality of the scheme. To do this, we impose a sort of "balancing" between the terms of the sums in equation (3.8.25), in such a way that all contributions to the error has roughly the same order

$$\mu^{i+1} \sim \sum_{k=0}^{i} \mu^{i-k} \varepsilon_k \sim \sum_{k=0}^{i} \mu^{i-k} N^{-s}_{k+1} \beta_k,$$

where

$$\beta_k := (\gamma + 2\theta C + \theta C' 2^s) C_Y + \theta 2^s C_X M(\varepsilon_k)^s \varepsilon_k + \theta \|f\|_X.$$

Let us suppose that the approximation strategy for $\mathcal{L}\varphi$ is optimal, then β_k is uniformly bounded, otherwise it could be unbounded.

Since

$$\sum_{k=0}^{i} \mu^{i-k}\varepsilon_k = \mu^{i+1}\sum_{k=0}^{i} \mu^{-k-1}\varepsilon_k,$$

the tolerances ε_k should then be chosen in such a way that

$$\sum_{k=0}^{+\infty} \mu^{-k-1}\varepsilon_k < +\infty.$$

It is not difficult to see that this holds for instance for the choice

$$\varepsilon_k = \frac{\mu^{k+1}}{k\log k}.$$

Analogously we will choose N_{k+1} such that

$$N_{k+1} = \left(\frac{\beta_k k \log k}{\mu^{k+1}}\right)^{1/s}.$$

Remark 3.8.2: *If the approximation strategy for $\mathcal{L}\varphi$ is not optimal, i.e. $\{\beta_k\}_k$ is possibly unbounded, then roughly speaking the choice of N_{k+1} will have to compensate, at each iteration k, the loss of optimality. The result is a more quick growth of $\{N_k\}$, than in the optimal case and in a loss of the optimality in reducing the final error.*

3.8.1 An application: the Three-Fields formulation

Now we apply the above setting to the particular case of three-field nonconforming domain decomposition method.

Assume that the strategy for computing $(S\underline{\varphi})_\varepsilon$ falls in the framework described in the previous section. This basically reduces to saying that the type of discretization spaces used for approximating the Lagrange multiplier verifies Bernstein and Jackson inequalities of the type (3.8.2) and (3.8.4) and that the adaptive strategy guarantees a prescribed error on the Lagrange multiplier. This holds for instance if free-knot splines [74] are used to approximate the Lagrange multiplier, together with a-posteriori error indicator proposed in [14].

Then using Theorem 3.8.1 yields the following Corollary

Corollary 3.8.1: *Let $\mathcal{S} : \ell^2 \to \ell^2$ be the ∞-dimensional linear operator associated to the three-field formulation. Assume there exists a τ, with $0 < \tau < 2$,* *such that $\mathcal{S}$ is an isomorphism from ℓ^τ onto ℓ^τ. Under the assumptions of Lemmas 3.8.1 and 3.8.2, if the approximation strategy for $\mathcal{S}\underline{\varphi}$ is optimal, then the $\mathcal{S}$-nonlinear Richardson scheme is ℓ^2-stable and the following error estimate holds*

$$\|\underline{\varphi}^{i+1} - \underline{\varphi}\|_X \lesssim \sum_{k=0}^{i} \mu^{i-k} N_{k+1}^{-s} + \mu^{i+1}\|\underline{\varphi}^0 - \underline{\tau}\|_X + \sum_{k=0}^{i} \mu^{i-k}\varepsilon_k, \quad (3.8.25)$$

with $\mu < 1$ and all the constants depending only on the initial data.

3.9 Open problems and perspectives

Nonlinear Richardson type algorithms could be applied in a quite wide class of situations including non-conforming domain decomposition methods, but a deeper study in this direction, including probing numerical experiments, is necessary. Such nonlinear algorithms are attractive in the sense that the user is able to control the number of degrees of freedom (and therefore the memory size and complexity) at each iteration, but they also suffer from the following drawback compared for example to the adaptive wavelet scheme proposed in [31]: all the parameters involved in the fine tuning of the algorithm depend on the number $\tau < 2$, which describes the degree of sparsity of the solution in the wavelet basis, in the sense that the coefficient sequence belongs to ℓ^τ_w, or equivalently the order of convergence of the nonlinear projection algorithm is $N^{-s/d}$, with $s/d = 1/\tau - 1/2$. This means that in order to converge with this optimal rate, the algorithm requires an a-priori knowledge on the smoothness of the solution which is somehow the opposite of adaptivity. Another drawback is that the range of the convergence rate which can be considered is limited by some condition $\tilde{\tau} < \tau$, where $\tilde{\tau}$ could be much larger than the actual degree of sparsity of the solution. This comes from the fact that all the proofs rely on a contraction property of an operator in the ℓ^τ norm, which is known to be contractive in ℓ^2 and bounded in the ℓ^τ norm for some arbitrarily small τ, and therefore contractive in ℓ^τ by interpolation, for τ sufficiently close to 2. Since there is no clear estimate on $\tilde{\tau}$, it could well be that the rates of convergence which can be achieved by the algorithm are quite deceiving compared to the optimal rate of nonlinear approximation, except when the solution is not so sparse (in which case a uniform method would work as well).

Chapter 4

ADAPTIVE SCHEMES FOR NONLINEAR EQUATIONS

"The realm of Sauron is ended!" said Gandalf. "The Ring-bearer has fulfilled his Quest." And as the Captains gazed south to the Land of Mordor, it seemed to them that, black against the pall of cloud, there rose a huge shape of shadow, impenetrable, lightning-crowned, filling all the sky. Enormous it reared above the world, and stretched out towards them a vast threatening hand, terrible but impotent: for even as it leaned over them, a great wind took it, and it was all blown away, and passed; and then a hush fell.

(J.R.R. Tolkien, The Return of the King)

4.1 Introduction

The aim of this chapter is to show how it is possible to apply nonlinear wavelet approximation to design wavelet based adaptive schemes to solve a general class of nonlinear problems.

The iterative methods, that we propose to find a solution to a given nonlinear equation

$$F(u) = 0,$$

are Inexact Newton-type methods: given u_0, u_{n+1} is computed as follows:

$$u_{n+1} = u_n + \tilde{\rho}_n,$$

with $\tilde{\rho}_n$ *approximation* to ρ_n, where ρ_n satisfies $F'(u_n)\rho_n = -F(u_n)$ and $F'(u_n)$ is the Fréchet derivative of F at u_n.

In designing such adaptive methods we follow the ∞-dimensional "new approach" (see Introduction) already used to design adaptive methods for linear equation (Chapter 3). In the nonlinear setting such an approach reads as follows:

1. transform the initial nonlinear continuous problem $F(u) = 0$ into an equivalent ∞-dimensional nonlinear problem $\mathcal{F}(\underline{u}) = \underline{0}$, whose unknown $\underline{u}$ is the infinite vector of wavelet coefficients of the solution.

2. write down a convergent Inexact Newton-type iterative scheme for the ∞-dimensional problem.

3. at each iteration approximately (possibly adaptively) apply the involved infinite dimensional operators to finite dimensional spaces.

Here, as already mentioned in the Introduction, we don't face the problem of the effective construction of the approximate application of nonlinear operators in wavelet coordinates [29]. We rather consider it as a "black-box" strategy and we provide a recipe for a dynamically choice, as the iteration procedure progresses, of the involved tolerances, in order to guarantee the efficiency and the convergence of the resulting algorithm.

4.2 Notations and Preliminary results

Let $\Omega \subseteq \mathbb{R}$ be a Lipschitz domain. We will denote by $(\cdot,\cdot)$ the $L^2(\Omega)$ scalar product. For sequences $\underline{v} \in \ell^\alpha$, we will denote by $\mathcal{I}^a_\delta(\underline{v})$ the ball of centre $\underline{v}$ and radius δ in ℓ^α topology: $\mathcal{I}^a_\delta(\underline{v}) = \{\underline{u} \in \ell^\alpha : \quad \|\underline{u}-\underline{v}\|_{\ell^\alpha} \leq \delta\}$. For functions $v \in H^a(\Omega)$, we will denote by $I^a_\delta(v)$ the ball of centre v and radius δ in $H^a(\Omega)$ topology: $I^a_\delta(v) = \{u \in H^a(\Omega) : \quad \|u - v\|_{H^a(\Omega)} \leq \delta\}$.

Now let us assume we are given a couple $\{\psi_\lambda, \quad \lambda \in \Lambda = \cup_{j\geq 0}\Lambda_j\}$, $\{\tilde{\psi}_\lambda, \quad \lambda \in \Lambda = \cup_{j\geq 0}\Lambda_j\}$ (Λ_j finite dimensional) of biorthogonal bases for $L^2(\Omega)$, satisfying the following properties:

(W0) Any function $f \in L^2(\Omega)$ can be decomposed in terms of either one of the two bases as follows:

$$f = \sum_{\lambda\in\Lambda} (f, \tilde{\psi}_\lambda)\psi_\lambda = \sum_{\lambda\in\Lambda} (f, \psi_\lambda)\tilde{\psi}_\lambda. \tag{4.2.1}$$

(W1) For any $f \in B^s_{p,q}(\Omega)$, $0 < s \leq S$, $0 < p < \infty$, $q > 0$ the following norm equivalence holds:

$$\| \sum_\lambda (f, \tilde{\psi}_\lambda)\psi_\lambda\|^q_{B^s_{p,q}(\Omega)} \simeq \sum_j 2^{q(s+d(\frac{1}{2}-\frac{1}{p}))j} \left(\sum_{\lambda\in\Lambda_j} |(f, \tilde{\psi}_\lambda)|^p \right)^{q/p} . \tag{4.2.2}$$

Remark 4.2.1: *The splitting of the index set* Λ *as* $\Lambda = \cup_{j\geq 0}\Lambda_j$ *indicates that the basis function* ψ_λ *(and* $\tilde{\psi}_\lambda$*) are "living" at different scales:* $\lambda \in \Lambda_j \Leftrightarrow \mathrm{supp}(\psi_\lambda) \sim 2^{-\mathrm{j}} \sim \mathrm{supp}(\tilde{\psi}_\lambda)$.

Remark 4.2.2: *We remark that (W0) implies that the two bases are "biorthogonal" in the following sense:*

$$(\tilde{\psi}_\lambda, \psi_{\lambda'}) = \delta_{\lambda,\lambda'}, \quad \lambda, \lambda' \in \Lambda = \cup_{j\geq 0}\Lambda_j. \tag{4.2.3}$$

Describing how bases satisfying assumptions (W0)-(W1) can be constructed is beyond the goals of this work. We want to stress out that *biorthogonal wavelets* fall in the class here described [38], [32], [25] and therefore they will be used throughout this chapter.

In the setting of biorthogonal wavelets we denote by $u_\lambda := (f, \tilde{\psi}_\lambda)$ the wavelet coefficients in the expansion $f = \sum_{\lambda\in\Lambda}(f, \tilde{\psi}_\lambda)\psi_\lambda$. Hence the norm equivalence (4.2.2) for Besov spaces $B^s_{p,q}$ rewrites as follows: for all $0 < s \leq S$, $0 < p < \infty$, $q > 0$:

$$\|\sum_\lambda u_\lambda\psi_\lambda\|^q_{B^s_{p,q}(\Omega)} \simeq \sum_j 2^{q(s+d(\frac{1}{2}-\frac{1}{p}))j}\left(\sum_{\lambda\in\Lambda_j}|u_\lambda|^p\right)^{q/p}. \tag{4.2.4}$$

In particular, since $H^s(\Omega) = B^s_{2,2}(\Omega)$, from equivalence (4.2.4) we deduce that for all $0 < s \leq S$:

$$\|\sum_j\sum_{\lambda\in\Lambda_j} u_\lambda(2^{-js}\psi_\lambda)\|_{H^s(\Omega)} \simeq \|\underline{u}\|_{\ell^2}. \tag{4.2.5}$$

where $\underline{u} = \{u_\lambda\}_\lambda$.

Now let us briefly recall some results about nonlinear approximation in a wavelet framework. In such a setting a given function $u \in L^2(\Omega)$, whose wavelet decomposition is $u = \sum_\lambda u_\lambda\psi_\lambda$, is approximated by a lacunary series: u is approximated by an element v belonging to the nonlinear space

$$\Sigma_N = \{v = \sum_{\lambda\in\Lambda} v_\lambda\psi_\lambda : \quad \underline{v} = \{v_\lambda\}_{\lambda\in\Lambda} \in \sigma_N\}, \tag{4.2.6}$$

containing all the functions of $L^2(\Omega)$, whose wavelet coefficients belong to the set

$$\sigma_N = \{\underline{v} \in \ell^2(\Lambda) : \quad \#\{\lambda : v_\lambda \neq 0\} \leq N\}$$

of sequences with at most N elements different from zero. The set Σ_N contains the functions of $L^2(\Omega)$, which can be expressed as a linear combination of at most N wavelets. A nonlinear projector

$$\mathbb{P}_N : L^2(\Omega) \to \Sigma_N$$

can be built as follows: given $u = \sum_\lambda u_\lambda \psi_\lambda$, let us sort the sequence $\{|u_\lambda|\}_{\lambda\in\Lambda}$ in decreasing order. We denote $\{|u_{\lambda(k)}|\}_{k\in\mathbb{N}}$ the coefficient of rank k:

$$|u_{\lambda(k)}| \geq |u_{\lambda(k+1)}|, \quad \text{with} \quad k > 0.$$

Hence the image $\mathbb{P}_N(u)$ is defined by:

$$\mathbb{P}_N(u) = \sum_{n=1}^{N} u_{\lambda(n)} \psi_{\lambda(n)},$$

that is only the N greatest (in absolute value) coefficients of u are retained. By abuse of notation we will also indicate by $\mathbb{P}_N : \ell^2 \to \sigma_N$ the operator associating to the sequence $\underline{u} = \{u_\lambda\}$, the coefficients of the function $\mathbb{P}_N(\sum_\lambda u_\lambda \psi_\lambda)$. The accuracy of the corresponding approximation is directly related to ℓ^τ regularity of the sequence of coefficients of u, as stated by the following theorem [47], [48]:

Theorem 4.2.1: *Let $u = \sum_{\lambda\in\Lambda} u_\lambda \psi_\lambda$. If $\underline{u} = \{u_\lambda\}_\lambda \in \ell^\tau$, with τ such that $0 < \tau < 2$, then*

$$\|\underline{u} - \mathbb{P}_N \underline{u}\|_{\ell^2} \lesssim \inf_{\underline{w}\in\sigma_N} \|\underline{u} - \underline{w}\|_{\ell^2} \lesssim N^{-(\frac{1}{\tau}-\frac{1}{2})} \|\underline{u}\|_{\ell^\tau},$$

where the constants in the bounds depend only on τ.

In particular, if τ is such that $\frac{1}{\tau} = \frac{r}{d} + \frac{1}{2}$, using norm equivalence (4.2.4), we obtain

$$\|\sum_\lambda u_\lambda \psi_\lambda\|_{B^r_{\tau,\tau}(\Omega)} \simeq \|\underline{u}\|_{\ell^\tau} \tag{4.2.7}$$

and from Theorem 4.2.1 we have that if u belongs to $B^r_{\tau,\tau}(\Omega)$, with τ such that $\frac{1}{\tau} = \frac{r}{d} + \frac{1}{2}$, then

$$\inf_{w\in\Sigma_N} \|u - w\|_{L^2(\Omega)} \lesssim \|u - \mathbb{P}_N u\|_{L^2(\Omega)} \lesssim N^{-(\frac{1}{\tau}-\frac{1}{2})} \|u\|_{B^s_{\tau,\tau}(\Omega)}.$$

In other words, once we normalise in $L^2(\Omega)$ the wavelet basis $\{\psi_\lambda\}$, the natural functional setting of nonlinear approximation in $L^2(\Omega)$ is the scale of Besov spaces $B^r_{\tau,\tau}(\Omega)$.

Let us now consider a rescaled version $\{\check{\psi}_\lambda\}_\lambda$ of the wavelet basis $\{\psi_\lambda\}_\lambda$, where $\check{\psi}_\lambda = \rho^{-js}\psi_\lambda$, for $\lambda \in \Lambda_j$. If τ is such that $\frac{1}{\tau} = \frac{r}{d} + \frac{1}{2}$, from norm equivalence (4.2.4) we obtain:

$$\|\sum_\lambda u_\lambda \check{\psi}_\lambda\|_{B^{r+s}_{\tau,\tau}(\Omega)} \simeq \|\underline{u}\|_{\ell^\tau}. \tag{4.2.8}$$

Applying now Theorem 4.2.1 and norm equivalence (4.2.5) for Sobolev spaces to the normalised sequence $\underline{u}$, we obtain the following result of nonlinear approximation in $H^s(\Omega)$:

Corollary 4.2.1: *Let $u \in B^{s+r}_{\tau,\tau}(\Omega)$, with τ such that $1/\tau = r/d + 1/2$, then*

$$\|u - \mathbb{P}_N u\|_{H^s(\Omega)} \lesssim \inf_{w\in\Sigma_N} \|u - w\|_{H^s(\Omega)} \lesssim N^{-(\frac{1}{\tau}-\frac{1}{2})}\|u\|_{B^{s+r}_{\tau,\tau}(\Omega)},$$

where the implicit constants in the bounds depend only on τ.

That is when we consider nonlinear approximation in $H^s(\Omega)$ the natural functional setting is the scale of Besov spaces $B^{r+s}_{\tau,\tau}(\Omega)$, where τ is defined by the relation $\frac{1}{\tau} = \frac{r}{d} + \frac{1}{2}$.

4.3 Inexact Newton methods

Let Ω be a Lipschitz domain in $\mathbb{R}$ and U be an open subset of the Sobolev space $H^s(\Omega)$. Given a nonlinear functional between Sobolev spaces:

$$F : U \subseteq H^s(\Omega) \to H^t(\Omega),$$

we want to solve the nonlinear equation

$$F(u) = 0.$$

A classical algorithm for solving nonlinear equations of type $F(u) = 0$ is Newton's method: given u_0,

$$u_{i+1} = u_i + x_i, \quad \text{with } F'(u_i)x_i = -F(u_i),$$

where $F'(u_i)$ is the Fréchet derivative of F at u_i.

The method is attractive because it converges rapidly, whenever the initial guess u_0 is sufficiently good. In [46], [85] a generalization of such method has been considered:

Inexact Newton methods:

begin

input: u_0

for $i = 0, 1, \dots$
find s_i which satisfies

$$F'(u_i)s_i = -F(u_i) + r_i$$

set $u_{i+1} = u_i + s_i$

end

output: $\tilde{u} = u_{i+1}$

end

in which at each iteration i the involved equation $F'(u_i)w = -F(u_i)$ is solved only approximately, because of the presence of the perturbative term r_i, which is possibly related to different sources of error.

In order to obtain the convergence of a general inexact Newton method, according to [85], we assume that F satisfies the following conditions:

(A.1) There exists a solution $u^* \in U$ of $F(u) = 0$, with $I^s_\delta(u^*) \subseteq U$, for some $\delta > 0$.

(A.2) On the ball $I^s_\delta(u^*)$ the functional F is Fréchet differentiable and its Fréchet derivative F' is continuous.

(A.3) At u^* the Fréchet derivative of F is not singular.

(A.4) There exist $\lambda \in [0,1]$ and $K > 0$ such that for all $u, v \in I^s_\delta(u^*)$:

$$\|(F'(u^*))^{-1}(F'(u) - F'(v))\|_{H^s(\Omega)\to H^s(\Omega)} \le K\|u - v\|^\lambda_{H^s(\Omega)}.$$

It has been proved in [85] that under assumptions (A.1)-(A.4) on the regularity of F, if the perturbations r_i are chosen in a suitable way, then the sequence $\{u_i\}$ converges in $H^s(\Omega)$ to a solution u^* of $F(u) = 0$, for any starting point $u_0 \in U$ sufficiently close to u^*:

Theorem 4.3.1: *Let $F : U \subseteq H^s(\Omega) \to H^t(\Omega)$ satisfy assumptions (A.1)-(A.4) and r_i such that:*

$$\frac{\|(F'(u_i))^{-1} r_i\|_{H^s(\Omega)}}{\|(F'(u_i))^{-1} F(u_i)\|_{H^s(\Omega)}} \le \nu < 1, \quad \textit{for all } i. \tag{4.3.1}$$

There exists a $\delta_s > 0$ *such that, if* $\|u_0 - u^*\|_{H^s(\Omega)} < \delta_s$*, then the sequence* $\{u_i\}$ *of inexact Newton method converges to* u^* *in* $H^s(\Omega)$ *and satisfies*

$$\|u_{i+1} - u^*\|_{H^s(\Omega)} \leq \rho_i \|u_i - u^*\|_{H^s(\Omega)}, \tag{4.3.2}$$

with

$$\rho_i \leq \rho = \left\{\nu + \frac{(1+\nu)\mu_\lambda(u^*)\|u_0 - u^*\|^\lambda_{H^s(\Omega)}}{(1+\lambda)(1-\mu_\lambda(u^*)\|u_0 - u^*\|^\lambda_{H^s(\Omega)})}\right\} < 1.$$

Let us define

$$\mu_\lambda(u^*) := \sup\left\{\frac{\|F'(u^*)^{-1}(F'(v) - F'(w))\|_{H^s \to H^s}}{\|v - w\|^\lambda_{H^s}} : \quad v \neq w, v, w \in I^s_\sigma(u^*)\right\}.$$

In order to prove Theorem 4.3.1 we need the following two results:

Lemma 4.3.1: *Let* $F : U \subseteq H^s(\Omega) \to H^t(\Omega)$ *satisfy assumptions (A.1)-(A.4). Then there exists a* σ*, with* $\sigma \leq \delta$*, such that* $F'(u)$ *is not singular for all* $u \in I^s_\delta(u^*)$ *and the following inequality holds*

$$\|F'(u)^{-1}(F'(v) - F'(w))\|_{H^s \to H^s} \leq \frac{\mu_\lambda(u^*)}{1 - \mu_\lambda(u^*)\|u - u^*\|_{H^s}}\|v - w\|^\lambda_{H^s}, \tag{4.3.3}$$

for all $v, w \in I^s_\sigma(u^*)$*.*

Proof: Consider the case $\lambda > 0$ and $\lambda = 0$ separately.

If $\lambda > 0$ and $\|u - u^*\|_{H^s} \leq \mu_\lambda(u^*)^{-1/\lambda}$, with $u \in I^s_\delta(u^*)$, then

$$\|I - F'(u^*)^{-1}F'(u)\|_{H^s \to H^s} = \|F'(u^*)^{-1}(F'(u^*) - F'(u))\|_{H^s \to H^s} \leq \mu_\lambda(u^*)\|u - u^*\|^\lambda_{H^s} < 1$$

and thus the Neumann series

$$F'(u)^{-1} = \sum_{n=0}^{\infty}(I - F'(u^*)^{-1}F'(u))^n F'(u^*)^{-1}$$

converges. Hence, as

$$F'(u)^{-1}(F'(v) - F'(w)) = F'(u^*)^{-1}(F'(v) - F'(w))\sum_{n=0}^{\infty}[I - F'(u^*)^{-1}F'(u)]^n,$$

inequality (4.3.3) follows from the definition of $\mu_\lambda(u^*)$.

Now consider the case $\lambda = 0$. By continuity of F' at u^* it is clear that $\mu_0(u^*)$ can be made as small as desired by making δ sufficiently small. Hence it follows that for σ sufficiently small and $u \in I^s_\sigma(u^*)$, we have $\|I - F'(u^*)^{-1}F'(u)\|_{H^s} < 1$, so the result once more follows as above. ∎

Lemma 4.3.2: *Let $F : U \subseteq H^s(\Omega) \to H^t(\Omega)$ satisfy assumptions (A.1)-(A.4). Then for any $u \in I^s_\sigma(u^*)$, where σ is chosen accordingly with Lemma 4.3.1, the following inequality holds*

$$\|F'(u)^{-1}(F(u^*) - F(u) - F'(u)(u^* - u))\|_{H^s} \leq \frac{\mu_\lambda(u^*)\|u - u^*\|^{1+\lambda}_{H^s}}{(1+\lambda)(1 - \mu_\lambda(u^*)\|u^* - u\|^\lambda_{H^s})} \tag{4.3.4}$$

Proof: Define $H : I^s_\sigma(u^*) \to H^t(\Omega)$ by $H(z) = F'(u)^{-1}F(z)$, where σ is chosen accordingly with Lemma 4.3.1. Then, by Lemma 4.3.1, we have

$$\begin{aligned}\|H'(v) - H'(w)\|_{H^s \to H^s} &= \|F'(u)^{-1}(F'(v) - F'(w))\|_{H^s \to H^s} \\ &\leq \frac{\mu_\lambda(u^*)}{1 - \mu_\lambda(u^*)\|u - u^*\|_{H^s}}\|v - w\|^\lambda_{H^s}, \quad (4.3.5)\end{aligned}$$

for all $v, w \in I^s_\sigma(u^*)$.

Now, by using [[72], Lemma 3.2.12] together with an appropriate definition of the notion of integral [64], it follows

$$\begin{aligned}&\|F'(u)^{-1}(F(u^*) - F(u) - F'(u)(u^* - u))\|_{H^s} \\ &= \|H(u^*) - H(u) - H'(u)(u^* - u)\|_{H^s} \\ &= \Big\| \int_0^1 (H'(u + t(u^* - u)) - H'(u))(u^* - u)dt \Big\|_{H^s}.\end{aligned}$$

Hence using inequality 4.3.5 yields

$$\begin{aligned}&\Big\| \int_0^1 (H'(u + t(u^* - u)) - H'(u))(u^* - u)dt \Big\|_{H^s} \\ &\leq \int_0^1 \|H'(u + t(u^* - u)) - H'(u)\|_{H^s}\|u^* - u\|_{H^s}dt \\ &\leq \frac{\mu_\lambda(u^*)}{1 - \mu_\lambda(u^*)\|u - u^*\|_{H^s}} \int_0^1 \|(u + t(u^* - u)) - u\|^\lambda_{H^s}\|u^* - u\|_{H^s}dt \\ &\leq \frac{\mu_\lambda(u^*)}{1 - \mu_\lambda(u^*)\|u - u^*\|_{H^s}}\|u^* - u\|^{\lambda+1}_{H^s}\int_0^1 t^\lambda dt \\ &\leq \frac{\mu_\lambda(u^*)\|u - u^*\|^{\lambda+1}_{H^s}}{(1+\lambda)(1 - \mu_\lambda(u^*)\|u^* - u\|^\lambda_{H^s})}. \qquad (4.3.6)\end{aligned}$$

■

Proof of Theorem 4.3.1: Set $\delta_s := \sigma$, where σ is chosen accordingly to Lemma 4.3.1. The proof is by induction.

Assume $\|u_i - u^*\|_{H^s} \leq \|u_0 - u^*\|_{H^s} \leq \delta_s$, for some $i \geq 0$. Then $u_i \in I^s_{\delta_s}(u^*)$, hence, by Lemma 4.3.1, $F'(u_i)^{-1}$ exists. Thus the i-th stage of the inexact Newton method is well defined. Now $s_i = F'(u_i)^{-1}(-F(u_i) + r_i)$ and hence

$$\begin{aligned} u_{i+1} - u^* &= u_i + F'(u_i)^{-1}(-F(u_i) + r_i) - u^* \\ &= F'(u_i)^{-1}(F(u^*) - F(u_i) - F'(u_i)(u^* - u_i) + r_i). \end{aligned}$$

Since the following two inequalities hold

$$\begin{aligned} \|F'(u_i)r_i\|_{H^s} &\leq \nu\|F'(u_i)^{-1}F(u_i)\|_{H^s}, \\ \|F'(u_i)^{-1}F(u_i)\|_{H^s} &\leq \|F'(u_i)^{-1}(F(u^*) - F(u_i) - F'(u_i))(u^* - u_i)\|_{H^s} + \|u_i - u^*\|_{H^s}, \end{aligned}$$

we conclude that

$$\|u_{i+1}-u^*\|_{H^s} \leq \nu\|u_i-u^*\|_{H^s}+(1+\nu)\|F'(u_i)^{-1}(F(u^*)-F(u_i)-F'(u_i))(u^*-u_i)\|_{H^s}. \tag{4.3.7}$$

Using inequality (4.3.4) yields

$$\|u_{i+1} - u^*\|_{H^s} \leq \Big\{\nu + \frac{(1+\nu)\mu_\lambda(u^*)\|u_i - u^*\|^\lambda_{H^s}}{(1+\lambda)(1 - \mu_\lambda(u^*)\|u_i - u^*\|^\lambda_{H^s})}\Big\}\|u_i - u^*\|_{H^s}, \tag{4.3.8}$$

and by choosing

$$\delta_s := \min\Big\{\sigma, \Big(\frac{(1+\lambda)(1-\nu)}{2+\lambda(1-\nu)}\mu_\lambda(u^*)^{-1}\Big)^{1/\lambda}\Big\},$$

we have $\Big\{\nu + \frac{(1+\nu)\mu_\lambda(u^*)\|u_i-u^*\|^\lambda_{H^s}}{(1+\lambda)(1-\mu_\lambda(u^*)\|u_i-u^*\|^\lambda_{H^s})}\Big\} < 1.$

Thus it follows

$$\|u_{i+1} - u^*\|_{H^s} < \|u_i - u^*\|_{H^s},$$

which yields by induction

$$\|u_{i+1} - u^*\|_{H^s} \leq \Big\{\nu + \frac{(1+\nu)\mu_\lambda(u^*)\|u_0 - u^*\|^\lambda_{H^s}}{(1+\lambda)(1 - \mu_\lambda(u^*)\|u_0 - u^*\|^\lambda_{H^s})}\Big\}\|u_i - u^*\|_{H^s},$$

for all i. ∎

Remark 4.3.1: *From Theorem 4.3.1 one has that $u_i \in I^s_{\delta_s}(u^*)$ for all i.*

4.4 Nonlinear Newton

4.4.1 The problem

Let U be an open subset of the Sobolev space $H^s(\Omega)$, $s < S$. Consider a map between Sobolev spaces:

$$F : U \subseteq H^s(\Omega) \to F(U) \subseteq H^t(\Omega), \quad t < S. \tag{4.4.1}$$

We want to find a solution u^* to the nonlinear functional equation

$$F(u) = 0, \tag{4.4.2}$$

using an adaptive wavelet method based on an inexact Newton scheme; then we assume that F satisfies assumptions (A.1)-(A.4).
Moreover we assume that F, restricted to more regular spaces, preserves such regularity:

(A.5) For some $r > 0$ it holds

$$F_{U\cap B^{s+r}_{\tau,\tau}(\Omega)} : U \cap B^{s+r}_{\tau,\tau}(\Omega) \to B^{t+r}_{\tau,\tau}(\Omega), \quad s, t < S,$$

where $0 < \tau < 2$ is such that $1/\tau = r/d + 1/2$.

Finally we assume that:

(A.6) The solution u^* belongs to $U \cap B^{s+r}_{\tau,\tau}(\Omega)$.

Remark 4.4.1: *From Corollary 4.2.1, as by assumption u^* belongs to $B^{s+r}_{\tau,\tau}(\Omega)$, it follows that*

$$\|u^* - \mathbb{P}_N u^*\|_{H^s(\Omega)} \lesssim N^{-(\frac{1}{\tau}-\frac{1}{2})}\|u^*\|_{B^{s+r}_{\tau,\tau}(\Omega)}, \tag{4.4.3}$$

where $\mathbb{P}_N$ is the non linear projector which retains the N greatest, in absolute value, wavelet coefficients of a given function.

Fixed a number M of degrees of freedom, our aim is to provide an *approximation*

$$u^*_M \in \Sigma_M;$$

that is u^*_M is built using at most M wavelet functions.

We would like an approximation u^*_M behaving possibly as well as the best M terms wavelet approximation $\mathbb{P}_M u^*$.

To achieve this goal, according to the abstract approach described in Section 4.1, we first translate (4.4.2) in terms of wavelet coefficients, thus obtaining an ∞-dimensional problem: we decompose the involved functions u and $F(u)$, by choosing two suitable rescaled versions $\{\check{\psi}_\lambda\}$ and $\{\hat{\psi}_\lambda\}$ of the wavelet basis $\{\psi_\lambda\}$:

$$u \in H^s(\Omega), \quad u = \sum_\lambda u_\lambda \check{\psi}_\lambda, \quad \text{with} \quad \check{\psi}_\lambda = 2^{-js}\psi_\lambda, \tag{4.4.4}$$

$$F(u) \in H^t(\Omega), \quad F(u) = \sum_\lambda f_\lambda \hat{\psi}_\lambda, \quad \text{with} \quad \hat{\psi}_\lambda = 2^{-jt}\psi_\lambda, \tag{4.4.5}$$

and we build a discrete version $\mathcal{F}$ of the mapping F, acting on wavelet coefficients as follows:

$$\mathcal{F} : \underline{u} = \{u_\lambda\} \to \mathcal{F}(\underline{u}) := \underline{f} = \{f_\lambda\}. \tag{4.4.6}$$

Thanks to norm equivalence (4.2.5) for Sobolev spaces, the previous map $\mathcal{F}$ results to be a mapping between ℓ^2 spaces:

$$\mathcal{F} : D \subseteq \ell^2 \to \ell^2. \tag{4.4.7}$$

Moreover, using norm equivalence (4.2.8) for Besov spaces, assumption (A.5) implies that $\mathcal{F}$, restricted to a more regular space, preserves such regularity:

$$\mathcal{F}_{D\cap\ell^\tau} : D \cap \ell^\tau \to \ell^\tau,$$

with $0 < \tau < 2$ such that $1/\tau = r/d + 1/2$.

Then solving the ∞-dimensional problem $\mathcal{F}(\underline{u}) = \underline{0}$ is equivalent to solving the initial continuous problem $F(u) = 0$.

Thanks to assumption (A.6) there exists $\underline{u}^* \in D \cap \ell^\tau$ such that $\mathcal{F}(\underline{u}^*) = \underline{0}$. Nonlinear approximation provides with a natural benchmark for adaptive schemes; indeed if we knew $\underline{u}^*$, then it would follow, by using Theorem 4.2.1, that $\|\underline{u}^* - \mathbb{P}_M \underline{u}^*\|_{\ell^2} \lesssim M^{-(\frac{1}{\tau}-\frac{1}{2})}\|\underline{u}^*\|_{\ell^\tau}$. But $\underline{u}^*$ is the unknown of our problem, so we do not have access to $\mathbb{P}_M \underline{u}^*$ exactly. Hence, given the number of degrees of freedom M, what we actually want is to design an *adaptive scheme* which builds an approximation $\underline{u}^*_M$ to $\underline{u}^*$, with

$$\underline{u}^*_M \in \sigma_M,$$

such that $\underline{u}^*_M$ behaves almost as well as $\mathbb{P}_M \underline{u}^*$.

4.4.2 The algorithm

The adaptive scheme we propose here, namely Nonlinear Newton, is an Inexact Newton-type method written for the ∞-dimensional problem $\mathcal{F}(\underline{u}) = \underline{0}$, in which, at each iteration i, the approximation $\underline{u}_{i+1}$ is forced to belong to a non-linear space $\sigma_{N_{i+1}}$, that is it is forced to be built using at most N_{i+1} degrees of freedom, where N_{i+1} is chosen accordingly to the accuracy of the approximation $\underline{u}_i$ at the previous step. Moreover at each iteration i two further sources of inexactness are introduced to deal with the problem of the approximate (possibly adaptive) application of infinite dimensional operators: $\bar{A}_i$ will denote an approximation to $\mathcal{F}_i' := \mathcal{F}'(\underline{u}_i)$ and $\bar{\mathcal{F}}_i$ an approximation to $\mathcal{F}_i := \mathcal{F}(\underline{u}_i)$. The rate of these compressions will be adapted, at each iteration i, to the accuracy (and hence to the number N_{i+1} of d.o.f.) of the approximation $\underline{u}_{i+1}$ that we want to build. The method we propose is the following:

Nonlinear Newton

begin

input: $M, \underline{u}_0$

set $i = 0$

repeat the following steps
compute $\bar{A}_i$ approximation to $\mathcal{F}_i'$
compute $\bar{\mathcal{F}}_i$ approximation to $\mathcal{F}_i$
choose N_{i+1}
set

$$\underline{u}_{i+1} = \mathbb{P}_{N_{i+1}}(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i)$$

update $i + 1 \to i$

until $N_i = M$

output: $\underline{\tilde{u}} = \underline{u}_i$

end

where $\mathbb{P}_{N_{i+1}}$ is the non linear wavelet projector, which forces $\underline{u}_{i+1}$ to belong to the non linear space $\sigma_{N_{i+1}}$.

4.4.3 The analysis of the method

First of all we note that **Nonlinear Newton** can be rewritten as an Inexact Newton scheme:

begin

input: $M, \underline{u}_0$

set $i = 0$

repeat the following steps

choose the perturbative term $\underline{r}_i$

find $\underline{s}_i$ which satisfies

$$\mathcal{F}_i' \underline{s}_i = -\mathcal{F}_i + \underline{r}_i$$

set $\underline{u}_{i+1} = \underline{u}_i + \underline{s}_i$

update $i + 1 \rightarrow i$

until $N_{i+1} = M$

end

with

$$\underline{r}_i = \mathcal{F}_i - \mathcal{F}_i'(\bar{A}_i^{-1}\bar{\mathcal{F}}_i + (I - \mathbb{P}_{N_{i+1}})(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i)). \tag{4.4.8}$$

Indeed the following equalities hold:

$$\begin{aligned}
\underline{u}_i + \underline{s}_i &= \underline{u}_i + (\mathcal{F}_i')^{-1}(-\mathcal{F}_i + \underline{r}_i) \\
&= \underline{u}_i + (\mathcal{F}_i')^{-1}(-\mathcal{F}_i + \mathcal{F}_i - \mathcal{F}_i'(\bar{A}_i^{-1}\bar{\mathcal{F}}_i + (I - \mathbb{P}_{N_{i+1}})(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i))) \\
&= \underline{u}_i - (\bar{A}_i^{-1}\bar{\mathcal{F}}_i + \underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i - \mathbb{P}_{N_{i+1}}(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i)) \\
&= \mathbb{P}_{N_{i+1}}(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i).
\end{aligned}$$

The perturbative term $\underline{r}_i$ takes into account, at step i, the different sources of inexactness: the non linear projector $\mathbb{P}_{N_{i+1}}$, the approximation $\bar{A}_i$ and the approximation $\bar{\mathcal{F}}_i$. Step by step, we can tune such sources of inexactness in order to build adaptively the approximate solution $\underline{u}_i$. Roughly speaking when $\underline{u}_i$ is far from a solution of $F(\underline{u}) = 0$, we can perform the step of our method employing high perturbative term ($\underline{r}_i$ is large), using instead a lower perturbation when $\underline{u}_i$ is nearer.

More generally it is useless to build a bad approximation with a fine resolution or a good approximation with a coarse resolution. In both cases we use a resolution which is not of the same order as the approximation. The right choice, if we want to obtain an efficient scheme, is to adapt the resolution to the quality of the approximation, that is to its distance from the exact solution, in such a way to reduce the computational cost, but not to loose at the same time the convergence.

The advantage of the reformulation of Nonlinear Newton as an inexact Newton scheme is that we can use Theorem 4.3.1 to prove that Nonlinear Newton converges in ℓ^τ. In order to apply Theorem 4.3.1 we need to assume in ℓ^τ conditions on the regularity of $\mathcal{F}$ similar to (A.1)-(A.4):

(B.1) There exists $\underline{u}^* \in D \cap \ell^\tau$ verifying $\mathcal{F}(\underline{u}^*) = \underline{0}$, with $\mathcal{I}^\tau_\delta(\underline{u}^*) \subseteq D \cap \ell^\tau$, for some $\delta > 0$.

(B.2) On the ball $\mathcal{I}^\tau_\delta(\underline{u}^*)$, the functional $\mathcal{F}$ is Fréchet differentiable and its Fréchet derivative $\mathcal{F}'$ is continuous:

$$\mathcal{F}' \in C^0(\mathcal{I}^\tau_\delta(\underline{u}^*) \subseteq \ell^\tau, \mathcal{L}(\ell^\tau, \ell^\tau)).$$

(B.3) At $\underline{u}^*$ the Fréchet derivative $\mathcal{F}'$ is not singular.

(B.4) There exist $\lambda_\tau \in [0,1]$ and $K_\tau > 0$ such that for all $\underline{u}, \underline{v} \in \mathcal{I}^\tau_\delta(\underline{u}^*)$ it holds:

$$\|\mathcal{F}'(\underline{u}^*)^{-1}(\mathcal{F}'(\underline{u}) - \mathcal{F}'(\underline{v}))\|_{\ell^\tau \to \ell^\tau} \leq K_\tau \|\underline{u} - \underline{v}\|^{\lambda_\tau}_{\ell^\tau}.$$

We recall that convergence in ℓ^τ, for $\tau < 2$, implies convergence in ℓ^2. However in general we can hope that convergence in ℓ^2 is faster than convergence in ℓ^τ.

Assumptions (A.2)-(A.4) on F translate into the following assumptions on $\mathcal{F}$:

(B.5) On the ball $\mathcal{I}^2_\delta(\underline{u}^*)$, the functional $\mathcal{F}$ is Fréchet differentiable and its Fréchet derivative $\mathcal{F}'$ is continuous:

$$\mathcal{F}' \in C^0(\mathcal{I}^2_\delta(\underline{u}^*) \subseteq \ell^2, \mathcal{L}(\ell^2, \ell^2)).$$

(B.6) At $\underline{u}^*$ the Fréchet derivative $\mathcal{F}'$ is not singular.

(B.7) There exist $\lambda_2 \in [0,1]$ and $K_2 > 0$ such that for all $\underline{u}, \underline{v} \in \mathcal{I}^2_\delta(\underline{u}^*)$ it holds:

$$\|\mathcal{F}'(\underline{u}^*)^{-1}(\mathcal{F}'(\underline{u}) - \mathcal{F}'(\underline{v}))\|_{\ell^2 \to \ell^2} \leq K_2 \|\underline{u} - \underline{v}\|^{\lambda_2}_{\ell^2}.$$

Let us now collect the following two Lemmas: the first one, dealing with non singularity of the Fréchet derivative, is the analog of Lemma 4.3.1:

Lemma 4.4.1: *Let $\mathcal{F}$ be a mapping from $D \subseteq \ell^2$ into ℓ^2, satisfying conditions from (B.1) to (B.7), for some $\tau < 2$. Let $\mu = \beta^{-\frac{1}{\min\{\lambda_2, \lambda_\tau\}}}$ where*

$$\beta = \max\Big\{ \sup\{\frac{\|\mathcal{F}'(\underline{u}^*)^{-1}(\mathcal{F}'(\underline{u}) - \mathcal{F}'(\underline{v}))\|_{\ell^\tau \to \ell^\tau}}{\|\underline{u} - \underline{v}\|^{\lambda_\tau}_{\ell^\tau}}, \underline{u} \neq \underline{v}, \underline{u}, \underline{v} \in \mathcal{I}^\tau_{\delta_*}(\underline{u}^*)\},$$
$$\sup\{\frac{\|\mathcal{F}'(\underline{u}^*)^{-1}(\mathcal{F}'(\underline{u}) - \mathcal{F}'(\underline{v}))\|_{\ell^2 \to \ell^2}}{\|\underline{u} - \underline{v}\|^{\lambda_2}_{\ell^2}} \underline{u} \neq \underline{v}, \underline{u}, \underline{v} \in \mathcal{I}^2_{\delta_*}(\underline{u}^*)\}\Big\}.$$

Then for all $\underline{u}_i \in \mathcal{I}^\tau_\mu(\underline{u}^*)$ (which implies $\underline{u}_i \in \mathcal{I}^2_\mu(\underline{u}^*)$), it follows that $\mathcal{F}'_i := \mathcal{F}'(\underline{u}_i)$ is not singular.

The second contains a classical result about perturbation of linear operators:

Lemma 4.4.2: *Let A^* and $\mathcal{C}$ be two linear and continuous operators from ℓ^2 into ℓ^2. Moreover suppose that A^* and $I + A^{*-1}\mathcal{C}$ are not singular, then the following inequality holds:*

$$\|(A^* + \mathcal{C})^{-1} - \mathcal{A}^{-1}\|_{\ell^2\to\ell^2} \le \|(A^* + \mathcal{C})^{-1}\|_{\ell^2\to\ell^2}\|\mathcal{A}^{-1}\|_{\ell^2\to\ell^2}\|\mathcal{C}\|_{\ell^2\to\ell^2} \quad (4.4.9)$$

Proof: We have the following inequalities:

$$\begin{aligned}
& \|(A^* + \mathcal{C})^{-1} - \mathcal{A}^{-1}\|_{\ell^2\to\ell^2} \\
= & \|(I + \mathcal{A}^{-1}\mathcal{C})^{-1}\mathcal{A}^{-1} - \mathcal{A}^{-1}\|_{\ell^2\to\ell^2} \\
= & \|[(I + \mathcal{A}^{-1}\mathcal{C})^{-1} - I]\mathcal{A}^{-1}\|_{\ell^2\to\ell^2} \\
= & \|[(I + \mathcal{A}^{-1}\mathcal{C})^{-1} - (I + \mathcal{A}^{-1}\mathcal{C})^{-1}(I + \mathcal{A}^{-1}\mathcal{C})]\mathcal{A}^{-1}\|_{\ell^2\to\ell^2} \\
= & \|(I + \mathcal{A}^{-1}\mathcal{C})^{-1}(I - I - \mathcal{A}^{-1}\mathcal{C})\mathcal{A}^{-1}\|_{\ell^2\to\ell^2} \\
\le & \|(I + \mathcal{A}^{-1}\mathcal{C})^{-1}\mathcal{A}^{-1}\|_{\ell^2\to\ell^2}\|\mathcal{C}\|_{\ell^2\to\ell^2}\|\mathcal{A}^{-1}\|_{\ell^2\to\ell^2} \\
\le & \|(A^* + \mathcal{C})^{-1}\|_{\ell^2\to\ell^2}\|\mathcal{C}\|_{\ell^2\to\ell^2}\|\mathcal{A}^{-1}\|_{\ell^2\to\ell^2}.
\end{aligned}$$

■

Let us suppose that the approximation $\bar{A}_i^{-1}$ to $(\mathcal{F}_i')^{-1}$ has the following form:

$$\bar{A}_i^{-1} := (\mathcal{F}_i' + E(\mathcal{F}_i'))^{-1},$$

where $E(\mathcal{F}_i')$ is a linear functional representing the correction added to $\mathcal{F}_i'$ at each iteration i.
Now we are able to prove the following results:

Lemma 4.4.3: *Let $\mathcal{F}$ be a mapping from $D \subseteq \ell^2$ into ℓ^2, satisfying conditions from (B.1) to (B.7), for some $\tau < 2$. If $\underline{u}_i$ belongs to the ball $\mathcal{I}_\mu^\tau(\underline{u}^*)$ for all i, where μ is chosen accordingly to Lemma 4.4.1 then, for some positive constants $C_0, C_1, \dots, C_4, C_1', \dots, C_4'$, the following inequalities hold:*

$$\|\mathcal{F}_i\|_{\ell^2} \le C_0, \quad (4.4.10)$$

$$C_1 \le \|\mathcal{F}_i'\|_{\ell^2\to\ell^2} \le C_2, \quad (4.4.11)$$

$$C_1' \le \|\mathcal{F}_i'\|_{\ell^\tau\to\ell^\tau} \le C_2', \quad (4.4.12)$$

$$C_3 \le \|(\mathcal{F}_i')^{-1}\|_{\ell^2\to\ell^2} \le C_4, \quad (4.4.13)$$

$$C_3' \le \|(\mathcal{F}_i')^{-1}\|_{\ell^\tau\to\ell^\tau} \le C_4'. \quad (4.4.14)$$

Proof: Continuity of $\mathcal{F}$ yields immediately inequality (4.4.10). In particular it is not restrictive to assume $\|\mathcal{F}_i\|_{\ell^2} < 1$, for all $\underline{u}_i \in \mathcal{I}_{\delta_*}^\tau(\underline{u}^*)$. In fact it suffices to consider the ℓ^2 normalized function $\mathcal{F}(\underline{u}) := \frac{1}{C_0}\mathcal{F}(\underline{u})$, which verifies

$\|\mathcal{F}(\underline{u}_i)\|_{\ell^2} \leq 1$, for all $\underline{u}_i \in \mathcal{I}^{\tau}_{\delta_*}(\underline{u}^*)$.

From Lemma 4.4.1, i.e. $\mathcal{F}'_i$ is not singular for all $\underline{u}_i \in \mathcal{I}^{\tau}_{\delta^*}(\underline{u}^*)$ and from the continuity of $\mathcal{F}'$ we deduce inequalities (4.4.11) and (4.4.12).

Finally by using the continuity of the inverse of the Fréchet derivative and inequality $\|(\mathcal{F}'_i)^{-1}\| \geq 1/\|\mathcal{F}'_i\|$, we obtain inequalities (4.4.13) and (4.4.14). ∎

Lemma 4.4.4: *Under the same hypotheses of Lemma 4.4.3, let* $E(\mathcal{F}'_i) \in \mathcal{L}(\mathcal{I}^2_{\mu}(\underline{u}^*) \subseteq \ell^2, \ell^2) \cap \mathcal{L}(\mathcal{I}^{\tau}_{\mu}(\underline{u}^*) \subseteq \ell^{\tau}, \ell^{\tau})$ *satisfy:*

$$\|E(\mathcal{F}'_i)\|_{\ell^2\to\ell^2} \leq \frac{1}{2C_4}, \tag{4.4.15}$$

$$\|E(\mathcal{F}'_i)\|_{\ell^\tau\to\ell^\tau} < \frac{1}{2C'_4}, \tag{4.4.16}$$

then the following inequalities hold:

$$\|\bar{A}_i^{-1}\|_{\ell^2\to\ell^2} \leq 2C_4, \tag{4.4.17}$$

$$\|\bar{A}_i^{-1}\|_{\ell^\tau\to\ell^\tau} \leq 2C'_4. \tag{4.4.18}$$

Moreover, denoting by $\underline{h}_i := \underline{u}_i - \underline{u}^*$ *the error committed at step* i*, there exists* $\varepsilon > 0$ *depending on* $\mathcal{I}^{\tau}_{\mu}(\underline{u}^*)$ *such that the following inequality holds:*

$$\|\underline{h}_i\|_{\ell^2} \;\geq\; \|\mathcal{F}_i\|_{\ell^2}/(C_2+\varepsilon). \tag{4.4.19}$$

Proof: Using Lemma 4.4.2 yields:

$$\begin{aligned}
\|\bar{A}_i^{-1}\|_{\ell^2\to\ell^2} &= \|(\mathcal{F}'_i + E(\mathcal{F}'_i))^{-1}\|_{\ell^2\to\ell^2} \\
&\leq \|(\mathcal{F}'_i + E(\mathcal{F}'_i))^{-1} - (\mathcal{F}'_i)^{-1}\|_{\ell^2\to\ell^2} + \|(\mathcal{F}'_i)^{-1}\|_{\ell^2\to\ell^2} \\
&\leq \|(\mathcal{F}'_i + E(\mathcal{F}'_i))^{-1}\|_{\ell^2\to\ell^2}\|(\mathcal{F}'_i)^{-1}\|_{\ell^2\to\ell^2}\|E(\mathcal{F}'_i)\|_{\ell^2\to\ell^2} \\
&\quad + \|(\mathcal{F}'_i)^{-1}\|_{\ell^2\to\ell^2},
\end{aligned}$$

from which we deduce:

$$\|\bar{A}_i^{-1}\|_{\ell^2\to\ell^2} \leq \frac{\|(\mathcal{F}'_i)^{-1}\|_{\ell^2\to\ell^2}}{1 - \|E(\mathcal{F}'_i)\|_{\ell^2\to\ell^2}\|(\mathcal{F}'_i)^{-1}\|_{\ell^2\to\ell^2}}.$$

Using now inequalities (4.4.13) and (4.4.15), we obtain:

$$\|\bar{A}_i^{-1}\|_{\ell^2\to\ell^2} \leq 2C_4.$$

Similarly we obtain $\|\bar{A}_i^{-1}\|_{\ell^\tau\to\ell^\tau} \leq 2C'_4$.

Now let us set $\mathcal{G}(\underline{u}_i) = \mathcal{F}(\underline{u}_i) - \mathcal{F}'(\underline{u}^*)\underline{u}_i$. As $\mathcal{G}'(\underline{u}^*) = 0$ (the zero linear transformation), then there exists ε depending on $\mathcal{I}^2_{\delta_*}(\underline{u}^*)$ such that

$$\|\mathcal{G}(\underline{u}_i) - \mathcal{G}(\underline{u}^*)\|_{\ell^2} \leq \varepsilon \|\underline{u}_i - \underline{u}^*\|_{\ell^2}, \tag{4.4.20}$$

for all $\underline{u}_i \in \mathcal{I}^2_{\delta_*}(\underline{u}^*)$. By using inequality (4.4.20) we get

$$\begin{aligned}
\|\mathcal{F}(\underline{u}_i) - \mathcal{F}(\underline{u}^*)\|_{\ell^2} &= \|\mathcal{F}'(\underline{u}^*)\underline{u}_i - \mathcal{F}'(\underline{u}^*)\underline{u}^* + \mathcal{G}(\underline{u}_i) - \mathcal{G}(\underline{u}^*)\|_{\ell^2} \\
&\leq \|\mathcal{F}'(\underline{u}^*)\underline{u}_i - \mathcal{F}'(\underline{u}^*)\underline{u}^*\|_{\ell^2} + \|\mathcal{G}(\underline{u}_i) - \mathcal{G}(\underline{u}^*)\|_{\ell^2} \\
&\leq \|\mathcal{F}'(\underline{u}^*)\|_{\ell^2\to\ell^2} \|\underline{u}_i - \underline{u}^*\|_{\ell^2} + \varepsilon \|\underline{u}_i - \underline{u}^*\|_{\ell^2}
\end{aligned}$$

which yields, as $\mathcal{F}(\underline{u}^*) = \underline{0}$,

$$\|\underline{h}_i\|_{\ell^2} \geq \frac{\|\mathcal{F}(\underline{u}_i)\|_{\ell^2}}{\|\mathcal{F}'(\underline{u}^*)\|_{\ell^2\to\ell^2} + \varepsilon} \geq \frac{\|\mathcal{F}(\underline{u}_i)\|_{\ell^2}}{C_2 + \varepsilon}.$$

∎

Now we are ready to prove the following Theorem concerning the convergence of the scheme, where we denote by $\underline{h}_i := \underline{u}_i - \underline{u}^*$ the error committed at step i:

Theorem 4.4.1: *Let $\mathcal{F}$ be a mapping from $D \subseteq \ell^2$ into ℓ^2, satisfying conditions from (B.1) to (B.7), for some $\tau < 2$. Assume $E(\mathcal{F}'_i) \in \mathcal{L}(\mathcal{I}^2_\delta(\underline{u}^*) \subseteq \ell^2, \ell^2) \cap \mathcal{L}(\mathcal{I}^\tau_\delta(\underline{u}^*) \subseteq \ell^\tau, \ell^\tau)$. There exist a $\delta_* > 0$ and constants D_1, D_2, D_3 and D_4 such that if $\|\underline{u}_0 - \underline{u}^*\|_{\ell^\tau} < \delta_*$ and if, at each iteration i, we choose ordinately $E(\mathcal{F}'_i), \bar{\mathcal{F}}_i$ and N_{i+1} in order to fulfil the following conditions for some $0 < \sigma < \tau$:*

$$\|E(\mathcal{F}'_i)\|_{\ell^2\to\ell^2} \leq D_1, \tag{4.4.21}$$

$$\|E(\mathcal{F}'_i)\|_{\ell^\tau\to\ell^\tau} \leq D_2, \tag{4.4.22}$$

$$\|\bar{\mathcal{F}}_i - \mathcal{F}_i\|_{\ell^2} \leq D_3 \|\mathcal{F}_i\|_{\ell^2}, \qquad \|\bar{\mathcal{F}}_i - \mathcal{F}_i\|_{\ell^\tau} \leq D_3 \|\mathcal{F}_i\|_{\ell^\tau}, \tag{4.4.23}$$

$$N_{i+1} \geq D_4 \max\left\{ N_i \left(\frac{C(\bar{A}_i, \bar{\mathcal{F}}_i, \sigma)\|\underline{u}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^\tau}} \right)^{\frac{\tau\sigma}{\tau-\sigma}}, \left(\frac{C(\bar{A}_i, \bar{\mathcal{F}}_i, \tau)\|\underline{u}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^2}} \right)^{\frac{2\tau}{2-\tau}} \right\} \tag{4.4.24}$$

where $C(\bar{A}_i, \bar{\mathcal{F}}_i, \tau)$ and $C(\bar{A}_i, \bar{\mathcal{F}}_i, \sigma)$ are explicitly computable constants depending on $\bar{A}_i, \bar{\mathcal{F}}_i$ and τ (respectively on σ), then the sequence $\{\underline{u}_i\}$ converges to $\underline{u}^*$ in ℓ^τ and in ℓ^2 (with $\underline{u}_i \in \mathcal{I}^\tau_{\delta_*}(\underline{u}^*) \subseteq \mathcal{I}^2_{\delta_*}(\underline{u}^*)$).

In order to prove Theorem 4.4.1, we need to recall the following Lemma:

Lemma 4.4.5: *Let σ such that $0 < \sigma < \tau < 2$. If $\underline{u}_i \in \sigma_{N_i}$, then we have*

$$\|(I - \mathbb{P}_{N_{i+1}})(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i)\|_{\ell^\tau} \leq C(\bar{A}_i, \bar{\mathcal{F}}_i, \sigma) \left(\frac{N_{i+1}}{N_i}\right)^{-\frac{\tau-\sigma}{\tau\sigma}} \|\underline{u}_i\|_{\ell^\tau} \qquad (4.4.25)$$

where $C(\bar{A}_i, \bar{\mathcal{F}}_i, \sigma)$ is an explicitly computable constant depending on $\bar{A}_i, \bar{\mathcal{F}}_i$ and an σ.

Proof: First we note that $\underline{u}_i \in \sigma_{N_i}$, that is $\underline{u}_i$ has at most N_i elements different from zero, implies $\underline{u}_i \in \ell^\sigma$, for all $\sigma > 0$. By using (quasi) norm equivalences on sequences with a finite number of non zero elements, it follows that $\|\underline{u}_i\|_{\ell^\sigma} \leq N_i^{1/\sigma - 1/\tau}\|\underline{u}_i\|_{\ell^\tau}$.

Moreover, as $\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i \in \sigma_{M_i}$, for some M_i, we have that $\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i \in \ell^\sigma$. Finally, applying Theorem 4.2.1 to $\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i$, we have:

$$\begin{aligned}\|(I - \mathbb{P}_{N_{i+1}})(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i)\|_{\ell^\tau} &\leq C_\sigma N_{i+1}^{-(\frac{1}{\sigma} - \frac{1}{\tau})}\|\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i\|_{\ell^\sigma} \\ &\leq C_\sigma C_\sigma(\bar{A}_i, \bar{\mathcal{F}}_i) N_{i+1}^{-(\frac{1}{\sigma} - \frac{1}{\tau})}\|\underline{u}_i\|_{\ell^\sigma} \\ &\leq C(\bar{A}_i, \bar{\mathcal{F}}_i, \sigma) N_{i+1}^{-(\frac{1}{\sigma} - \frac{1}{\tau})} N_i^{\frac{1}{\sigma} - \frac{1}{\tau}}\|\underline{u}_i\|_{\ell^\tau},\end{aligned} \qquad (4.4.26)$$

where we used the inequality $\|\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i\|_{\ell^\sigma} \leq C_\sigma(\bar{A}_i, \bar{\mathcal{F}}_i)\|\underline{u}_i\|_{\ell^\sigma}$. ∎

Proof of Theorem 4.4.1:

Thanks to assumptions from (B.1) to (B.4), we can use Theorem 4.3.1: recalling that $\underline{r}_i = \mathcal{F}_i - \mathcal{F}_i'(\bar{A}_i^{-1}\bar{\mathcal{F}}_i + (I - \mathbb{P}_{N_{i+1}})(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i))$, if the following condition

$$\frac{\|(\mathcal{F}_i')^{-1}\underline{r}_i\|_{\ell^\tau}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^\tau}} \leq \alpha < 1 \qquad (4.4.27)$$

holds for all i, there exists a δ_τ such that if $\|\underline{u}_0 - \underline{u}^*\|_{\ell^\tau} < \delta_\tau$, then $\{\underline{u}_i\}$ converges to $\underline{u}^*$ in ℓ^τ and $\underline{u}_i \in \mathcal{I}_{\delta_\tau}^\tau(\underline{u}^*)$ for all i.

Analogously thanks to assumptions from (B.5) to (B.7), we can use Theorem 4.3.1: if the following condition

$$\frac{\|(\mathcal{F}_i')^{-1}\underline{r}_i\|_{\ell^2}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^2}} \leq \alpha < 1 \tag{4.4.28}$$

holds for all i, there exists a δ_2 such that if $\|\underline{u}_0 - \underline{u}^*\|_{\ell^2} < \delta_2$, then $\{\underline{u}_i\}$ converges to $\underline{u}^*$ in ℓ^2 and $\underline{u}_i \in \mathcal{I}^2_{\delta_2}(\underline{u}^*)$ for all i.

In particular, if both (4.4.27) and (4.4.28) hold, choosing $\delta_* = \min\{\delta_\tau, \delta_2\}$, if the initial guess $\underline{u}_0$ fulfils $\|\underline{u}_0 - \underline{u}^*\|_{\ell^\tau} < \delta_*$, which implies $\|\underline{u}_0 - \underline{u}^*\|_{\ell^2} < \delta_*$, then $\{\underline{u}_i\}$ converges to $\underline{u}^*$ in ℓ^2 and in ℓ^τ. As a consequence we also have that $\underline{u}_i \in \mathcal{I}^\tau_{\delta_*}(\underline{u}^*) \subseteq \mathcal{I}^2_{\delta_*}(\underline{u}^*)$ for all i.

Hence, in order to prove convergence in ℓ^τ and in ℓ^2, we only need to show that inequalities (4.4.27) and (4.4.28) are fulfilled, if we choose

$$\bar{A}_i^{-1} := (\mathcal{F}_i' + E(\mathcal{F}_i'))^{-1}, \bar{\mathcal{F}}_i, \text{ and } N_{i+1}$$

satisfying conditions (4.4.21)-(4.4.24), for suitable choices of the constants D_1, D_2, D_3 and D_4.

Let us now remark that the following inequalities hold:

$$\frac{\|(\mathcal{F}_i')^{-1}\underline{r}_i\|_{\ell^\tau}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^\tau}} \leq \underbrace{\frac{\|(I - \bar{A}_i^{-1}\mathcal{F}_i')((\mathcal{F}_i')^{-1}\mathcal{F}_i)\|_{\ell^\tau}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^\tau}}}_{\mathcal{T}} + \underbrace{\frac{\|\bar{A}_i^{-1}(\bar{\mathcal{F}}_i - \mathcal{F}_i)\|_{\ell^\tau}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^\tau}}}_{\mathcal{A}}$$
$$+ \underbrace{\frac{\|(I - \mathbb{P}_{N_{i+1}})(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i)\|_{\ell^\tau}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^\tau}}}_{\mathcal{U}},$$

and

$$\frac{\|(\mathcal{F}_i')^{-1}\underline{r}_i\|_{\ell^2}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^2}} \leq \underbrace{\frac{\|(I - \bar{A}_i^{-1}\mathcal{F}_i')((\mathcal{F}_i')^{-1}\mathcal{F}_i)\|_{\ell^2}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^2}}}_{T} + \underbrace{\frac{\|\bar{A}_i^{-1}(\bar{\mathcal{F}}_i - \mathcal{F}_i)\|_{\ell^2}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^2}}}_{W}$$
$$+ \underbrace{\frac{\|(I - \mathbb{P}_{N_{i+1}})(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i)\|_{\ell^2}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^2}}}_{O}.$$

We only need to prove that, for suitable choices of the constants D_1, D_2, D_3 and D_4, under assumptions (4.4.21)-(4.4.24), each term on the right-hand side of the above two inequalities is smaller or equal than $\frac{\alpha}{3}$.

First we recall that the following inequalities hold:

$$\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^\tau} \geq \frac{1}{\|\mathcal{F}_i'\|_{\ell^\tau\to\ell^\tau}}\|\mathcal{F}_i\|_{\ell^\tau}, \tag{4.4.29}$$

$$\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^2} \geq \frac{1}{\|\mathcal{F}_i'\|_{\ell^2\to\ell^2}}\|\mathcal{F}_i\|_{\ell^2}. \tag{4.4.30}$$

Let us now consider

$$\begin{aligned} T &:= \frac{\|(I-\bar{A}_i^{-1}\mathcal{F}_i')((\mathcal{F}_i')^{-1}\mathcal{F}_i)\|_{\ell^2}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^2}} \\ &\leq \|I-\bar{A}_i^{-1}\mathcal{F}_i'\|_{\ell^2\to\ell^2} \\ &\leq \|\bar{A}_i^{-1}\|_{\ell^2\to\ell^2}\|E(\mathcal{F}_i')\|_{\ell^2\to\ell^2} \end{aligned}$$

where we used the fact that $I-\bar{A}_i^{-1}\mathcal{F}_i' = \bar{A}_i^{-1}(\bar{A}_i-\mathcal{F}_i') = \bar{A}_i^{-1}E(\mathcal{F}_i')$.

From assumption (4.4.21), i.e $\|E(\mathcal{F}_i')\|_{\ell^2\to\ell^2} \leq D_1$, we have

$$T \leq D_1\|\bar{A}_i^{-1}\|_{\ell^2\to\ell^2}.$$

If D_1 verifies

$$D_1 \leq \frac{1}{2C_4},$$

then condition (4.4.21) implies that inequality (4.4.15) is fulfilled and we can use inequality (4.4.17), which yields

$$T \leq 2C_4D_1.$$

Moreover if D_1 is chosen to also satisfy $2C_4D_1 \leq \frac{\alpha}{3}$, i.e.

$$D_1 \leq \min\{\frac{1}{2C_4}, \frac{\alpha}{6C_4}\},$$

then

$$T \leq \frac{\alpha}{3}.$$

Analogously we have that

$$\begin{aligned} \mathcal{T} &:= \frac{\|(I-\bar{A}_i^{-1}\mathcal{F}_i')((\mathcal{F}_i')^{-1}\mathcal{F}_i)\|_{\ell^\tau}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^\tau}} \\ &\leq \|I-\bar{A}_i^{-1}\mathcal{F}_i'\|_{\ell^\tau\to\ell^\tau} \\ &\leq \|\bar{A}_i^{-1}\|_{\ell^\tau\to\ell^\tau}\|E(\mathcal{F}_i')\|_{\ell^\tau\to\ell^\tau}. \end{aligned}$$

If D_2 verifies $D_2 \leq \frac{1}{2C'_4}$ then assumption (4.4.22) implies that inequality (4.4.16) is satisfied and we can use inequality (4.4.18) obtaining $\mathcal{T} \leq 2C'_4 D_2$. Moreover if D_2 is chosen to also satisfy $2C'_4 D_2 \leq \frac{\alpha}{3}$, i.e.

$$D_2 \leq \min\{\frac{1}{2C'_4}, \frac{\alpha}{6C'_4}\},$$

then

$$\mathcal{T} \leq \frac{\alpha}{3}.$$

Now we consider

$$\begin{aligned} W &:= \frac{\|\bar{A}_i^{-1}(\bar{\mathcal{F}}_i - \mathcal{F}_i)\|_{\ell^2}}{\|(\mathcal{F}'_i)^{-1}\mathcal{F}_i\|_{\ell^2}} \\ &\leq \frac{\|\bar{A}_i^{-1}\|_{\ell^2\to\ell^2}\|\bar{\mathcal{F}}_i - \mathcal{F}_i\|_{\ell^2}}{\|(\mathcal{F}'_i)^{-1}\mathcal{F}_i\|_{\ell^2}}. \end{aligned}$$

Using inequalities (4.4.17) and (4.4.30) yields

$$\begin{aligned} W &\leq \frac{2C_4\|\bar{\mathcal{F}}_i - \mathcal{F}_i\|_{\ell^2}}{\|(\mathcal{F}'_i)^{-1}\mathcal{F}_i\|_{\ell^2}} \\ &\leq \frac{2C_4\|\mathcal{F}'_i\|_{\ell^2\to\ell^2}\|\bar{\mathcal{F}}_i - \mathcal{F}_i\|_{\ell^2}}{\|\mathcal{F}_i\|_{\ell^2}}. \end{aligned}$$

Now thanks to inequality (4.4.11), i.e. $\|\mathcal{F}'_i\|_{\ell^2\to\ell^2} \leq C_2$, we have

$$W \leq \frac{2C_4 C_2\|\bar{\mathcal{F}}_i - \mathcal{F}_i\|_{\ell^2}}{\|\mathcal{F}_i\|_{\ell^2}}.$$

From assumption (4.4.23), i.e. $\|\bar{\mathcal{F}}_i - \mathcal{F}_i\|_{\ell^2} \leq D_3\|\mathcal{F}_i\|_{\ell^2}$, we obtain

$$W \leq 2C_4 C_2 D_3.$$

If D_3 is chosen such that $2C_4 C_2 D_3 \leq \frac{\alpha}{3}$, i.e.

$$D_3 \leq \frac{\alpha}{6C_4 C_2},$$

then

$$W \leq \frac{\alpha}{3}.$$

Analogously we have that

$$\begin{aligned} \mathcal{A} &:= \frac{\|\bar{A}_i^{-1}(\bar{\mathcal{F}}_i - \mathcal{F}_i)\|_{\ell^\tau}}{\|(\mathcal{F}'_i)^{-1}\mathcal{F}_i\|_{\ell^\tau}} \\ &\leq \frac{\|\bar{A}_i^{-1}\|_{\ell^\tau\to\ell^\tau}\|\bar{\mathcal{F}}_i - \mathcal{F}_i\|_{\ell^\tau}}{\|(\mathcal{F}'_i)^{-1}\mathcal{F}_i\|_{\ell^\tau}}. \end{aligned}$$

Using inequalities (4.4.18) and (4.4.29) yields

$$\begin{aligned}\mathcal{A} &\leq \frac{2C_4'\|\bar{\mathcal{F}}_i - \mathcal{F}_i\|_{\ell^\tau}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^2}} \\ &\leq \frac{2C_4'\|\mathcal{F}_i'\|_{\ell^\tau\to\ell^\tau}\|\bar{\mathcal{F}}_i - \mathcal{F}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^\tau}}.\end{aligned}$$

Now thanks to inequality (4.4.12) we have

$$\mathcal{A} \leq \frac{2C_4'C_2'\|\bar{\mathcal{F}}_i - \mathcal{F}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^\tau}}.$$

From assumption (4.4.23), i.e. $\|\bar{\mathcal{F}}_i - \mathcal{F}_i\|_{\ell^\tau} \leq D_3\|\mathcal{F}_i\|_{\ell^\tau}$, we obtain

$$\mathcal{A} \leq 2C_4'C_2'D_3.$$

If D_3 is also chosen such that $2C_4'C_2'D_3 \leq \frac{\alpha}{3}$, i.e.

$$D_3 \leq \min\{\frac{\alpha}{6C_4'C_2'}, \frac{\alpha}{6C_4C_2}\},$$

then

$$\mathcal{A} \leq \frac{\alpha}{3}.$$

Now in order to prove convergence in ℓ^2 and in ℓ^τ, we only need to estimate the last two terms: O and $\mathcal{U}$.

By using Theorem 4.2.1, as $\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i \in \ell^\tau$, we have

$$\begin{aligned}O &:= \frac{\|(I-\mathbb{P}_{N_{i+1}})(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i)\|_{\ell^2}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^2}} \\ &\leq \frac{C_\tau N_{i+1}^{-\frac{2-\tau}{2\tau}}\|\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i\|_{\ell^\tau}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^2}}.\end{aligned}$$

Using inequalities (4.4.30) and (4.4.11) yields

$$\begin{aligned}O &\leq \frac{C_\tau N_{i+1}^{-\frac{2-\tau}{2\tau}}\|\mathcal{F}_i'\|_{\ell^2\to\ell^2}\|\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^2}} \\ &\leq \frac{C_2C_\tau N_{i+1}^{-\frac{2-\tau}{2\tau}}\|\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^2}} \\ &\leq \frac{C_2C_\tau C_\tau(\bar{A}_i, \bar{\mathcal{F}}_i)N_{i+1}^{-\frac{2-\tau}{2\tau}}\|\underline{u}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^2}}, \\ &\leq \frac{C_2C(\bar{A}_i, \bar{\mathcal{F}}_i, \tau)N_{i+1}^{-\frac{2-\tau}{2\tau}}\|\underline{u}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^2}},\end{aligned}$$

where we used $\|\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i\|_{\ell^\tau} \le C_\tau(\bar{A}_i, \bar{\mathcal{F}}_i)\|\underline{u}_i\|_{\ell^\tau}$.

From assumption (4.4.24), we have in particular $N_{i+1} \ge D_4 \left(\frac{C(\bar{A}_i,\bar{\mathcal{F}}_i,\tau)\|\underline{u}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^2}}\right)^{\frac{2\tau}{2-\tau}}$.
Hence

$$O \le C_2 D_4^{-\frac{2-\tau}{2\tau}}.$$

If D_4 is chosen such that $C_2 D_4^{-\frac{2-\tau}{2\tau}} \le \frac{\alpha}{3}$, i.e.

$$D_4 \ge \left(\frac{3C_2}{\alpha}\right)^{\frac{2\tau}{2-\tau}},$$

then

$$O \le \frac{\alpha}{3}.$$

Finally, by using Lemma 4.4.5, we have

$$\begin{aligned}\mathcal{U} &:= \frac{\|(I - \mathbb{P}_{N_{i+1}})(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i)\|_{\ell^\tau}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^\tau}} \\ &\le \frac{C_\sigma C_\sigma(\bar{A}_i,\bar{\mathcal{F}}_i) N_i^{\frac{\tau-\sigma}{\tau\sigma}} N_{i+1}^{-\frac{\tau-\sigma}{\tau\sigma}} \|\underline{u}_i\|_{\ell^\tau}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^\tau}}.\end{aligned}$$

Using inequalities (4.4.29) and (4.4.12) yields

$$\begin{aligned}\mathcal{U} &\le \frac{C(\bar{A}_i,\bar{\mathcal{F}}_i,\sigma) N_i^{\frac{\tau-\sigma}{\tau\sigma}} N_{i+1}^{-\frac{\tau-\sigma}{\tau\sigma}} \|\mathcal{F}_i'\|_{\ell^\tau\to\ell^\tau} \|\underline{u}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^\tau}} \\ &\le \frac{C_2' C(\bar{A}_i,\bar{\mathcal{F}}_i,\sigma) N_i^{\frac{\tau-\sigma}{\tau\sigma}} N_{i+1}^{-\frac{\tau-\sigma}{\tau\sigma}} \|\underline{u}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^\tau}}.\end{aligned}$$

From assumption (4.4.24) we have in particular $N_{i+1} \ge D_4 N_i \left(\frac{C(\bar{A}_i,\bar{\mathcal{F}}_i,\sigma)\|\underline{u}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^\tau}}\right)^{\frac{\tau\sigma}{\tau-\sigma}}$.
Hence

$$\mathcal{U} \le C_2' D_4^{-\frac{\tau-\sigma}{\tau\sigma}}.$$

If D_4 is chosen such that $C_2' D_4^{-\frac{\tau-\sigma}{\tau\sigma}}$, i.e.

$$D_4 \ge \left(\frac{3C_2'}{\alpha}\right)^{\frac{\tau\sigma}{\tau-\sigma}},$$

then

$$\mathcal{U} \le \frac{\alpha}{3}.$$

Hence, if we choose

$$D_4 \ge \max\left\{\left(\frac{3C_2}{\alpha}\right)^{\frac{2\tau}{2-\tau}}, \left(\frac{3C_2'}{\alpha}\right)^{\frac{\tau\sigma}{\tau-\sigma}}\right\},$$

then we have $O \leq \frac{\alpha}{3}$ and $\mathcal{U} \leq \frac{\alpha}{3}$.

Hence we proved that if

$$\begin{aligned} D_1 &\leq \frac{\alpha}{6C_4}, \\ D_2 &\leq \frac{\alpha}{6C_4'}, \\ D_3 &\leq \min\Big\{\frac{\alpha}{6C_4C_2}, \frac{\alpha}{6C_4'C_2'}\Big\}, \\ D_4 &\geq \max\Big\{\Big(\frac{3C_2}{\alpha}\Big)^{\frac{2\tau}{2-\tau}}, \Big(\frac{3C_2'}{\alpha}\Big)^{\frac{\tau\sigma}{\tau-\sigma}}\Big\}, \end{aligned}$$

then

$$\frac{\|(\mathcal{F}_i')^{-1}\underline{r}_i\|_{\ell^\tau}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^\tau}} \leq \mathcal{T} + \mathcal{A} + \mathcal{U} \leq \alpha,$$

and

$$\frac{\|(\mathcal{F}_i')^{-1}\underline{r}_i\|_{\ell^2}}{\|(\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^2}} \leq T + W + O \leq \alpha,$$

with $\alpha < 1$. Thus, by virtue of Theorem 4.3.1, we proved convergence in ℓ^τ and in ℓ^2.

Remark 4.4.2: *Let us consider condition (4.4.24):*

$$N_{i+1} \geq D_4 \max\Big\{N_i\left(\frac{C(\bar{A}_i, \bar{\mathcal{F}}_i, \sigma)\|\underline{u}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^\tau}}\right)^{\frac{\tau\sigma}{\tau-\sigma}}, \left(\frac{C(\bar{A}_i, \bar{\mathcal{F}}_i, \tau)\|\underline{u}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^2}}\right)^{\frac{2\tau}{2-\tau}}\Big\}.$$

Convergence in ℓ^τ guarantees that $\|\underline{u}_i\|_{\ell^\tau}$ is uniformly bounded; in this way the choice of N_{i+1} is essentially driven by N_i and $\frac{1}{\|\mathcal{F}_i\|}$.

□

We now prove the claimed result of quadratic convergence of the scheme, under some slightly stronger assumptions on the choice of $E(\mathcal{F}_i')$, $\bar{A}_i$ and N_{i+1}:

Theorem 4.4.2: *Let $\mathcal{F}$ be a mapping from $D \subseteq \ell^2$ into ℓ^2, satisfying conditions from (B.1) to (B.7), for some $\tau < 2$. Assume $E(\mathcal{F}_i') \in \mathcal{L}(\mathcal{I}_\delta^2(\underline{u}^*) \subseteq \ell^2, \ell^2) \cap \mathcal{L}(\mathcal{I}_\delta^\tau(\underline{u}^*) \subseteq \ell^\tau, \ell^\tau)$. There exist a $\delta_* > 0$ and constants D_1, D_2, D_3 and D_4, such that if $\|\underline{u}_0 - \underline{u}^*\|_{\ell^\tau} < \delta_*$ and if, at each iteration i, we choose ordinately $E(\mathcal{F}_i'), \bar{\mathcal{F}}_i$ and N_{i+1} in order to fulfil the following conditions for some $0 < \sigma < \tau$:*

$$\|E(\mathcal{F}_i')\|_{\ell^2 \to \ell^2} \leq D_1 \|\mathcal{F}_i\|_{\ell^2}, \tag{4.4.31}$$

$$\|E(\mathcal{F}_i')\|_{\ell^\tau\to\ell^\tau} \le D_2, \tag{4.4.32}$$

$$\|\bar{\mathcal{F}}_i - \mathcal{F}_i\|_{\ell^2} \le D_3\|\mathcal{F}_i\|_{\ell^2}^2, \qquad \|\bar{\mathcal{F}}_i - \mathcal{F}_i\|_{\ell^\tau} \le D_3\|\mathcal{F}_i\|_{\ell^\tau}, \tag{4.4.33}$$

$$N_{i+1} \ge D_4 \max\left\{N_i\left(\frac{C(\bar{A}_i,\bar{\mathcal{F}}_i,\sigma)\|\underline{u}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^\tau}}\right)^{\frac{\tau\sigma}{\tau-\sigma}}, \left(\frac{C(\bar{A}_i,\bar{\mathcal{F}}_i,\tau)\|\underline{u}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^2}^2}\right)^{\frac{2\tau}{2-\tau}}\right\}, \tag{4.4.34}$$

where $C(\bar{A}_i,\bar{\mathcal{F}}_i,\tau)$ and $C(\bar{A}_i,\bar{\mathcal{F}}_i,\sigma)$ are explicitly computable constants depending only on $\bar{A}_i,\bar{\mathcal{F}}_i$ and on τ (respectively on σ), then the sequence $\{\underline{u}_i\}$ converges quadratically to $\underline{u}^*$ in ℓ^2:

$$\|\underline{h}_{i+1}\|_{\ell^2} \le \|\underline{h}_i\|_{\ell^2}^2 \quad \textit{for all } i \in \mathbb{N}. \tag{4.4.35}$$

Proof: As it is not restrictive to assume $\|\mathcal{F}_i\|_{\ell^2} \le C_0 < 1$, for all $\underline{u}_i \in \mathcal{I}^\tau_{\delta_*}(\underline{u}^*)$, trivially we have $\|\mathcal{F}_i\|_{\ell^2}^2 \le \|\mathcal{F}_i\|_{\ell^2}$. Now it is simple to see that choosing D_1, D_2, D_3 and D_4 according to the proof of Theorem 4.4.1, yields convergence in ℓ^τ and ℓ^2. By the way we remark that with such choices of the constants we can use inequalities (4.4.17) and (4.4.18).

If in addition we impose some further conditions on the choice of D_1, D_2, D_3 and D_4 we will finally obtain quadratic convergence in ℓ^2.

Let us now introduce the following notation:

$$\check{\underline{h}}_i = \check{\underline{u}}_i - \underline{u}^* \quad \text{with } \check{\underline{u}}_{i+1} = \underline{u}_i - (\mathcal{F}_i')^{-1}\mathcal{F}_i,$$

that is $\check{\underline{u}}_{i+1}$ is the result of the application of a step of the classical Newton method to $\underline{u}_i$.

From now on, we choose $\delta_* = \min\{\delta_2, \delta_\tau, \mu\}$, where μ is chosen according to Lemma 4.4.1. It follows that $\mathcal{F}_i' := \mathcal{F}'(\underline{u}_i)$ is not singular for all $\underline{u}_i \in \mathcal{I}^\tau_{\delta^*}(\underline{u}^*)$. Moreover, as by hypothesis the operator $E(\mathcal{F}_i')$ belongs to $\mathcal{L}(\mathcal{I}^2_{\delta_*}(\underline{u}^*) \subseteq \ell^2, \ell^2)$ and it is chosen such that $\|E(\mathcal{F}_i')\|_{\ell^2\to\ell^2} < 1/\|(\mathcal{F}_i')^{-1}\|_{\ell^2\to\ell^2}$, we also have that $I + (\mathcal{F}_i')^{-1}E(\mathcal{F}_i')$ is not singular for all $\underline{u}_i \in \mathcal{I}^2_{\delta_*}(\underline{u}^*)$.

Hence we can apply Lemma 4.4.2 with $A^* = \mathcal{F}_i'$ and $\mathcal{C} = E(\mathcal{F}_i')$, obtaining the following inequality:

$$\begin{aligned}\|(\mathcal{F}_i' + E(\mathcal{F}_i'))^{-1} &- (\mathcal{F}_i')^{-1}\|_{\ell^2\to\ell^2} \\ &\le \|(\mathcal{F}_i' + E(\mathcal{F}_i'))^{-1}\|_{\ell^2\to\ell^2}\|(\mathcal{F}_i')^{-1}\|_{\ell^2\to\ell^2}\|E(\mathcal{F}_i')\|_{\ell^2\to\ell^2}.\end{aligned} \tag{4.4.36}$$

As $\bar{A}_i^{-1} = (\mathcal{F}_i' + E(\mathcal{F}_i'))^{-1}$ and recalling (4.4.8), we have:

$$\begin{aligned}
\underline{h}_{i+1} &= \underline{h}_i - (\mathcal{F}_i')^{-1}\mathcal{F}_i + (\mathcal{F}_i')^{-1}\underline{r}_i \\
&= \underline{h}_i - (\mathcal{F}_i')^{-1}\mathcal{F}_i + (\mathcal{F}_i')^{-1}\{\mathcal{F}_i - \mathcal{F}_i'[\bar{A}_i^{-1}\bar{\mathcal{F}}_i + (I - \mathbb{P}_{N_{i+1}})(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i)]\} \\
&= \underline{h}_i - (\mathcal{F}_i')^{-1}\bar{\mathcal{F}}_i + [(\mathcal{F}_i')^{-1}\mathcal{F}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i] - (I - \mathbb{P}_{N_{i+1}})(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i) \\
&= \underline{h}_i - (\mathcal{F}_i')^{-1}\mathcal{F}_i + [\bar{A}_i^{-1}(\mathcal{F}_i - \bar{\mathcal{F}}_i) + ((\mathcal{F}_i')^{-1} - \bar{A}_i^{-1})\mathcal{F}_i] \\
&\quad -(I - \mathbb{P}_{N_{i+1}})(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i),
\end{aligned}$$

which, taking ℓ^2 norm and using inequality (4.4.36), yields:

$$\begin{aligned}
\|\underline{h}_{i+1}\|_{\ell^2} &\le \|\underline{h}_i - (\mathcal{F}_i')^{-1}\mathcal{F}_i\|_{\ell^2} + \|\bar{A}_i^{-1}(\mathcal{F}_i - \bar{\mathcal{F}}_i)\|_{\ell^2} + \|((\mathcal{F}_i')^{-1} - \bar{A}_i^{-1})\mathcal{F}_i\|_{\ell^2} \\
&\quad + \|(I - \mathbb{P}_{N_{i+1}})(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i)\|_{\ell^2} \\
&\le \|\check{\underline{h}}_{i+1}\|_{\ell^2} + \|\bar{A}_i^{-1}\|_{\ell^2\to\ell^2}\|\mathcal{F}_i - \bar{\mathcal{F}}_i\|_{\ell^2} + \|(\mathcal{F}_i')^{-1} - \bar{A}_i^{-1}\|_{\ell^2\to\ell^2}\|\mathcal{F}_i\|_{\ell^2} \\
&\quad + \|(I - \mathbb{P}_{N_{i+1}})(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i)\|_{\ell^2} \\
&\le \underbrace{\|\check{\underline{h}}_{i+1}\|_{\ell^2}}_{Q} + \underbrace{\|\bar{A}_i^{-1}\|_{\ell^2\to\ell^2}\|\mathcal{F}_i - \bar{\mathcal{F}}_i\|_{\ell^2}}_{U} \\
&\quad + \underbrace{\|\bar{A}_i^{-1}\|_{\ell^2\to\ell^2}\|(\mathcal{F}_i')^{-1}\|_{\ell^2\to\ell^2}\|E(\mathcal{F}_i')\|_{\ell^2\to\ell^2}\|\mathcal{F}_i\|_{\ell^2}}_{A} \\
&\quad + \underbrace{\|(I - \mathbb{P}_{N_{i+1}})(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i)\|_{\ell^2}}_{D}.
\end{aligned}$$

In order to obtain quadratic convergence, we only need to prove that under assumptions (4.4.31)-(4.4.34), each term on the right-hand side of the above inequality is smaller or equal than $\frac{1}{4}\|\underline{h}_i\|_{\ell^2}^2$.

First we remark that

$$Q := \|\check{\underline{h}}_{i+1}\|_{\ell^2} \le \frac{1}{4}\|\underline{h}_i\|_{\ell^2}^2$$

holds thanks to classical results on quadratic convergence of Newton method, for a starting point sufficiently near to $\underline{u}^*$.

Consider now

$$U := \|\bar{A}_i^{-1}\|_{\ell^2\to\ell^2}\|\mathcal{F}_i - \bar{\mathcal{F}}_i\|_{\ell^2}.$$

Using inequality (4.4.17) yields

$$U \le 2C_4\|\mathcal{F}_i - \bar{\mathcal{F}}_i\|_{\ell^2}.$$

From assumption (4.4.33), i.e. $\|\bar{\mathcal{F}}_i - \mathcal{F}_i\|_{\ell^2} \le D_3\|\mathcal{F}_i\|_{\ell^2}^2$, we have that

$$U \le 2C_4D_3\|\mathcal{F}_i\|_{\ell^2}^2.$$

By using inequality (4.4.19), i.e. $\|\underline{h}_i\|_{\ell^2} \geq \|\mathcal{F}_i\|_{\ell^2}/(C_2+\varepsilon)$, we obtain

$$U \leq 2C_4 D_3 (C_2+\varepsilon)^2 \|\underline{h}_i\|_{\ell^2}^2.$$

If D_3 is chosen such that $2C_4D_3(C_2+\varepsilon)^2 \leq \frac{1}{4}$, i.e.

$$D_3 \leq \frac{1}{8C_4(C_2+\varepsilon)^2},$$

then

$$U \leq \frac{1}{4}\|\underline{h}_i\|_{\ell^2}^2.$$

Now let us estimate

$$A := \|\bar{A}_i^{-1}\|_{\ell^2\to\ell^2}\|(\mathcal{F}_i')^{-1}\|_{\ell^2\to\ell^2}\|E(\mathcal{F}_i')\|_{\ell^2\to\ell^2}\|\mathcal{F}_i\|_{\ell^2}.$$

Using inequalities (4.4.17) and (4.4.11) yields

$$A \leq 2C_4^2\|E(\mathcal{F}_i')\|_{\ell^2\to\ell^2}\|\mathcal{F}_i\|_{\ell^2}.$$

From assumption (4.4.31), i.e. $\|E(\mathcal{F}_i')\|_{\ell^2\to\ell^2} \leq D_1\|\mathcal{F}_i\|_{\ell^2}$, we obtain

$$A \leq 2C_4^2 D_1 \|\mathcal{F}_i\|_{\ell^2}^2,$$

which yields, by using inequality (4.4.19),

$$A \leq 2C_4^2 D_1 (C_2+\varepsilon)^2 \|\underline{h}_i\|_{\ell^2}^2.$$

If D_1 is chosen such that $2C_4^2D_1(C_2+\varepsilon)^2 \leq \frac{1}{4}$, i.e.

$$D_1 \leq \frac{1}{8C_4^2(C_2+\varepsilon)^2},$$

then

$$A \leq \frac{1}{4}\|\underline{h}_i\|_{\ell^2}^2.$$

Finally we consider

$$\begin{aligned} D &:= \|(I-\mathbb{P}_{N_{i+1}})(\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i)\|_{\ell^2} \\ &\leq C_\tau N_{i+1}^{-\frac{2-\tau}{2\tau}}\|\underline{u}_i - \bar{A}_i^{-1}\bar{\mathcal{F}}_i\|_{\ell^\tau} \\ &\leq C(\bar{A}_i,\bar{\mathcal{F}}_i,\tau) N_{i+1}^{-\frac{2-\tau}{2\tau}}\|\underline{u}_i\|_{\ell^\tau}. \end{aligned} \tag{4.4.37}$$

From assumption (4.4.34), we have $N_{i+1} \geq D_4\left(\frac{C(\bar{A}_i,\bar{\mathcal{F}}_i,\tau)\|\underline{u}_i\|_{\ell^\tau}}{\|\mathcal{F}_i\|_{\ell^2}^2}\right)^{\frac{2\tau}{2-\tau}}$. Hence

$$D \leq D_4^{-\frac{2-\tau}{2\tau}}\|\mathcal{F}_i\|_{\ell^2}^2.$$

By using inequality (4.4.19), we have

$$D \le D_4^{-\frac{2-\tau}{2\tau}} (C_2+\varepsilon)^2 \|\underline{h}_{i+1}\|_{\ell^2}^2.$$

If D_4 is chosen such that $D_4^{-\frac{2-\tau}{2\tau}} (C_2+\varepsilon)^2 \le \frac{1}{4}$, i.e.

$$D_4 \ge \left(4(C_2+\varepsilon)^2\right)^{\frac{2\tau}{2-\tau}},$$

then

$$D \le \frac{1}{4}\|\underline{h}_i\|_{\ell^2}^2.$$

Hence we proved

$$\|\underline{h}_{i+1}\|_{\ell^2} \le Q+U+A+D \le \|\underline{h}_i\|_{\ell^2}^2. \tag{4.4.38}$$

∎

Remark 4.4.3: *Given the number of degrees of freedom M, what we actually obtain is an approximation $\underline{u}_M^*$ to $\underline{u}^*$, which satisfies, thanks to (4.4.37) and (4.4.38):*

$$\|\underline{u}^* - \underline{u}_M^*\|_{\ell^2} \le \frac{3}{4}\|\underline{h}_i\|_{\ell^2}^2 + C(\bar{A}_i, \bar{\mathcal{F}}_i, \tau) M^{-(\frac{1}{\tau}-\frac{1}{2})}\|\underline{u}_M^*\|_{\ell^\tau}.$$

Heuristically this means that we are further, but not so far, thanks to the above result of convergence, from the natural benchmark provided by non linear approximation:

$$\|\underline{u}^* - \mathbb{P}_M \underline{u}^*\|_{\ell^2} \le C_\tau M^{-(\frac{1}{\tau}-\frac{1}{2})}\|\underline{u}^*\|_{\ell^\tau}.$$

Remark 4.4.4: *Conditions (4.4.31)-(4.4.34) are not completely reliable from a computational point of view, because they imply the knowledge of quantities, like $\|\mathcal{F}_i\|$, $\|\mathcal{F}_i - \bar{\mathcal{F}}_i\|$ or $\|E(\mathcal{F}_i')\|$, that we do not want to compute at any steps of our algorithm. Hence, given $\underline{u}_i$, we need a strategy for building $\bar{\mathcal{F}}_i$ and $\bar{A}_i$, which provides estimates for $\|\mathcal{F}_i\|$, $\|\mathcal{F}_i - \bar{\mathcal{F}}_i\|$ and $\|E(\mathcal{F}_i')\|$ involving only available quantities such as $\bar{\mathcal{F}}_i$ and $\bar{A}_i$. Results in this direction can be found in [30].*

4.5 Open problems and perspectives

We proposed an extension of the Nonlinear Richardson algorithm to nonlinear problems. The Richardson scheme has been replaced by an inexact Newton scheme, where the "inexactness" comes both from the approximate application of the involved operators (this issue is somehow taken for granted: it is assumed that there is a procedure which applies these operators up to any prescribed accuracy) and from the thresholding error coming from the application of the nonlinear projector. Under some regularity assumptions on the nonlinear operators, results similar to the linear case are obtained. What it is indeed necessary for a deeper understanding of the reliability of the method are concrete examples of nonlinear problems to be treated by such an approach and numerical tests.

Chapter 5

ADAPTIVITY & WAVELET PACKETS

"Well, this is the end, Sam Gamgee," said a voice by his side. And there was Frodo, pale and worn, and yet himself again; and in his eyes there was peace now, neither strain of will, nor madness, nor any fear. His burden was taken away. "Yes," said Frodo. "But do you remember Gandalf's words: Even Gollum may have something yet to do? So let us forgive him! For the Quest is achieved, and now all is over. I am glad you are here with me. Here at the end of all things, Sam."

(J.R.R. Tolkien, The Return of the King)

5.1 Introduction

In this chapter we introduce two space-frequency adaptive strategies for the numerical approximation of the solutions of quantum hydrodynamic (QHD) model for semiconductors, based respectively on wavelets and wavelet packets. The two strategies have been compared in [18] on a test case, and wavelet packets perform better in approximating with fewer degrees of freedom the higher frequency dispersive oscillations of the solution.

Motivated by such a result, we want to provide a way of optimizing wavelet packets adaptive schemes based on Galerkin discretizations, by extending the techniques of wavelet compression of certain operators to the case in which such operators are represented in terms of wavelet packets. The essential observation is that certain operators have an almost sparse representation in wavelet coordinates thanks to the good properties of localization both in space and in frequency of wavelet bases. Thus discarding all entries below a certain threshold will be then give rise to a sparse matrix that can be further processed by efficient linear algebra tools.

As wavelet packets provide better properties of localization both in space and in frequency, the representation of such operators in terms of wavelet packets should be in principle more sparse than in the wavelet case. Thus a wavelet-like compression techniques should give rise to a much more sparse matrix, possibly reducing the computational cost in solving the linear systems resulting from the Galerkin discretization of the problem.

5.2 Motivation: the QHD Equations

The quantum hydrodynamic model (QHD) for semiconductors has been recently introduced (see e.g. [2], [51], [52]) in order to describe with macroscopic fluid-type unknowns phenomena, such as negative differential resistance in a resonant tunneling diode, which are due to quantum effects and cannot be modeled with classical or semi–classical descriptions. Mathematically, the QHD system is a dispersive regularization of the so–called hydrodynamic equations (HD) for semiconductors (a hyperbolic system of conservation laws coupled self–consistently with the Poisson equation). As usual in macroscopic semiconductor models, the electron position density or some of its derivatives may present strong variation or even blow up in some points. Moreover, the dispersive character of the QHD system implies that the solution may develop high frequency oscillations, which are localized in regions not a priori known. Therefore, numerical simulations of the QHD system with uniform discretizations require an extremely high number of grid points, also when the "pathology" of the solution, which enforces the mesh size, is localized in a small percentage of the simulation domain. This leads to unnecessarily time consuming computations.

Due to the possible dispersive oscillations, an efficient approximation demands the use of a discretization where not only the spatial grid, but also the frequency distribution is adaptively adjusted to the behavior of the solution. One way of achieving such a goal is to use bases with good localization both in space and frequency. Wavelet type bases, which display such a property, have already been successfully used in the design of efficient adaptive schemes in various application fields (see e.g. [16], [12], [11], [28], [50], [58], [67], [63]). Due to their characteristics, the definition of criteria for driving the adaptive procedure (refining and coarsening) *both in space and in frequency* is quite natural. In particular, wavelet based adaptive algorithms have been introduced in ([18]) for the semiconductor hydrodynamic model, where, after performing a diffusive regularization, the adaptive strategy is aimed at well approximating solutions with steep gradients.

Here we recall [18] the feasibility of an adaptive algorithm based on wavelets and wavelet packets for the QHD model. Wavelet packets have better frequency localization properties and consequently they are superior to wavelets in drastically diminishing the required number of degrees of freedom for well approximating solutions which exhibit high frequency oscillations.

We consider the isothermal, stationary, one dimensional quantum hydrodynamic (QHD) equations in the domain $(0,1)$

$$\left\{\begin{array}{rcl} J^\varepsilon_x &=& 0, \\ \left(\dfrac{(J^\varepsilon)^2}{u^\varepsilon} + u^\varepsilon\right)_x + u^\varepsilon V_x - \dfrac{\varepsilon^2}{2} u^\varepsilon \left(\dfrac{\sqrt{u^\varepsilon_{xx}}}{\sqrt{u^\varepsilon}}\right)_x &=& -\dfrac{J^\varepsilon}{\tau}, \\ -\lambda^2 V &=& u^\varepsilon - C(x). \end{array}\right. \qquad (P_\varepsilon)$$

Here u^ε (which we will also at times denote by $u(\varepsilon)$) denotes the electron density, J^ε the current density, V the electrostatic potential. The parameter ε is the scaled Planck constant, the function $C(x)$ represents the (prescribed) doping profile of the semiconductor device, the parameter λ is the scaled Debye length and τ is the relaxation time.

As pointed out before, equations (P_ε) are a dispersive regularization of the classical isothermal hydrodynamic (HD) equations

$$\left\{\begin{array}{rcl} J_x &=& 0, \\ \left(\dfrac{J^2}{u} + u\right)_x + uV_x &=& -\dfrac{J}{\tau}, \\ -\lambda^2 \Delta V &=& u - C(x), \end{array}\right.$$

and, in the **formal** limit, the QHD equations tend to the HD equations. However, due to the dispersive term and the non-linearity, if the HD system exhibits a shock discontinuity, the solution of the QHD system is expected to develop dispersive oscillations, which are not damped as ε goes to zero. In that case only a weak convergence can hold as ε goes to zero and the limiting system is not expected to be the HD system. In [75] a numerical study shows evidence of this fact. A complete theory on the small dispersion limit for the QHD system is still an open problem. We refer to [53], [54], [57] for partial answers in special cases.

Figures 1–3 present the solution of (P_ε) for different values of ε ($\varepsilon = 0.01$, $\varepsilon = 0.005$ and $\varepsilon = 0.0026$, resp.). The pictures clearly show that the oscillation amplitude and location does not change as ε decreases. Changes in ε affect only the oscillation frequency, which is about the double when ε is halved.

An efficient numerical scheme to solve problem (P_ε) when dispersive oscillations occur is a challenging issue. The oscillations must be well resolved in

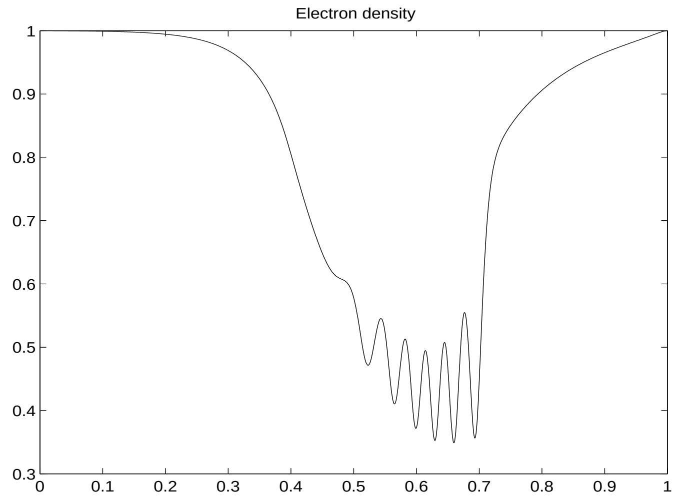

Figure 1. Solution of (P_ε) for $\varepsilon = .01$ ($\lambda = 0.1$, $\tau = 1/8$ and $V_0 = 6.5$, with doping profile C as in figure 6).

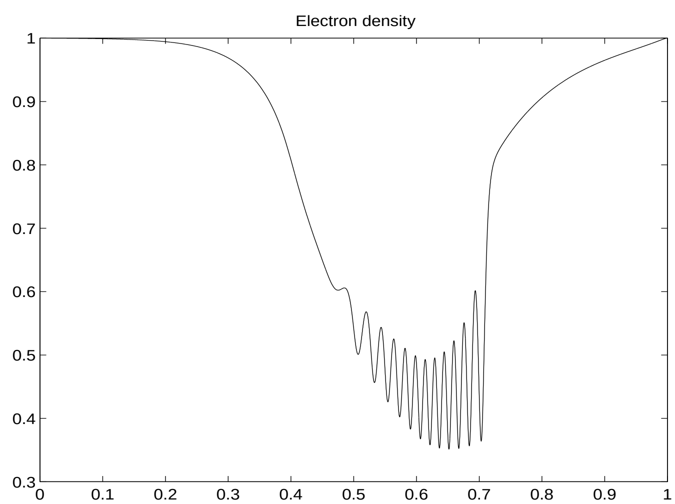

Figure 2. Solution of (P_ε) for $\varepsilon = .005$ ($\lambda = 0.1$, $\tau = 1/8$ and $V_0 = 6.5$, with doping profile C as in figure 6).

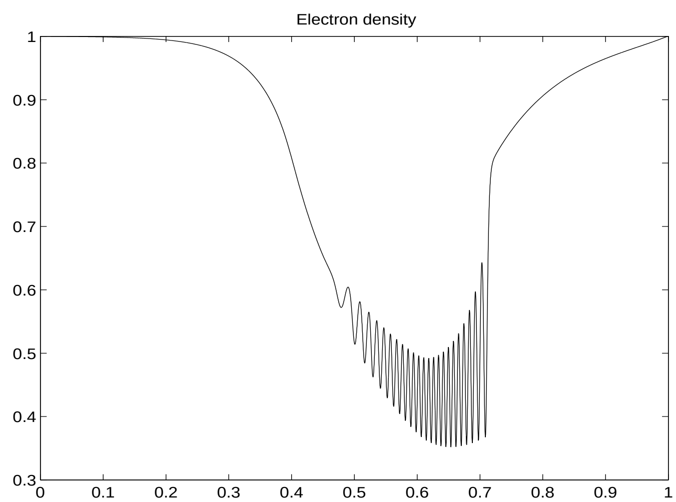

Figure 3. Solution of (P_ε) for $\varepsilon = .0026$ ($\lambda = 0.1$, $\tau = 1/8$ and $V_0 = 6.5$, with doping profile C as in figure 6).

order to keep the correct limiting behavior (for $\varepsilon \longrightarrow 0$) and it is clear that a discretization on a uniform grid requires too many degrees of freedom.

We point out that due to the presence of the dispersive regularization, no special treatment is required to approximate the convection terms in (P_ε). On the contrary, the use of upwind–type schemes would introduce spurious numerical damping of the oscillations and, consequently, further strong restriction on the mesh size in order to describe correctly the oscillations. We refer to [75] for a discussion on this issue.

5.3 Adaptive Solution of QHD Equations

Problem (P_ε) is highly non–linear, due to the non–linear terms in the second equation of (P_ε) and to the strong coupling to the Poisson equation. A continuation procedure in the parameter ε is an efficient strategy to deal with such problems and it can be coupled, for instance, to a (possibly damped) Newton algorithm for solving the non–linear system for a given ε of the continuation procedure. More precisely, for solving problem (P_ε), with a prescribed $\bar{\varepsilon}$, we define a finite decreasing sequence $\{\varepsilon_n, n = 0, ..., N\}$ with $\varepsilon_0 \sim 1$ and $\varepsilon_N = \bar{\varepsilon}$, and we solve the sequence of problems (P_{ε_n}). For the solution of problem $(P_{\varepsilon_{n+1}})$, the knowledge of the solution of (P_{ε_n}) is exploited in several ways for enhancing the efficiency of the algorithm, for instance it can be used as an initial guess for the Newton scheme. This procedure has been used in [75] for solving (P_ε), there discretized with a finite difference scheme on a uniform grid.

Here, we are interested in designing and testing some strategies for taking advantage of the knowledge of the computed solution of (P_{ε_n}) for reducing the number of degrees of freedom to be used for numerically solving problem $(P_{\varepsilon_{n+1}})$. We aim at an algorithm of the following form.

- Choose a class $\mathcal{B}$ whose elements B are the L^2-orthonormal bases of finite dimensional subspaces V_B of L^2:

  $$\mathcal{B} = \{B: \ B \text{ finite orthonormal basis of } V_B = < B >^{span} \subset L^2\}. \quad (5.3.1)$$

  We assume that the basis functions considered are sufficiently smooth.

- Compute an approximation u_0 to the solution $u(\varepsilon_0)$ of equations (P_ε) for $\varepsilon = \varepsilon_0$ using a coarse uniform grid. This is possible since for $\varepsilon \sim 1$ the solution of (P_ε) is smooth.

- Given the approximation u_n to the solution $u(\varepsilon_n)$ of (P_{ε_n}), an approximation u_{n+1} to the solution $u(\varepsilon_{n+1})$ of $(P_{\varepsilon_{n+1}})$ is obtained as follows:

– By analyzing u_n select a basis $B_{n+1} \in \mathcal{B}$, *as small as possible*, well suited for approximating $u(\varepsilon_{n+1})$.
– Compute u_{n+1} in $V_{n+1} =< B_{n+1} >^{span}$ approximate solution of $(P_{\varepsilon_{n+1}})$, by a suitable numerical method (for instance, Galerkin approximation of $(P_{\varepsilon_{n+1}})$).

Due to the onset of high frequency oscillations in a possible large (though localized) portion of the domain, in order to be able to approximate the solutions of (P_ε) with few degrees of freedom, it is not enough to work with methods that are adaptive only with respect to the space. We will then rather work with basis functions which display good localization properties also in the frequency domain. In particular we will consider bases B whose elements will be *phase atoms*. *Phase atoms* ([84]) are smooth functions which are well localized in both position and momentum in the sense of quantum mechanics. More precisely, a phase atom ψ needs to satisfy the following properties.

- Finite Energy. Possibly after a re-normalization, it holds

$$\|\psi\|_{L^2} = 1.$$

- Smoothness and decay. Both ψ and $\hat{\psi}$ are smooth ($\hat{\psi}$ being the Fourier transform of ψ).

- Finite *position* and *momentum*.

$$x_0 := \int x|\psi(x)|^2 \, dx < \infty,$$
$$\xi_0 := \int \xi|\hat{\psi}(\xi)|^2 \, d\xi < \infty,$$

are respectively called *position* and *momentum* (or *frequency*) of ψ.

- Localization in position and momentum. We have

$$\Delta x := \left(\int (x - x_0)^2 |\psi(x)|^2 \, dx \right)^{1/2} < \infty,$$
$$\Delta \xi := \left(\int (\xi - \xi_0)^2 |\hat{\psi}(\xi)|^2 \, d\xi \right)^{1/2} < \infty.$$

Δx and $\Delta \xi$ are also called *position* and *momentum uncertainty* respectively.

In the following two sections we will consider two classes of phase atoms, namely *wavelets* and *wavelet packets*. In particular we will analyze the performances of such two classes in the framework of adaptive approximation of the solution of (P_ε).

5.4 Wavelets

The first possibility that we consider is the one of wavelet adaptivity. Since the notation that we will use in this chapter differs slightly from the notation used up to now, let us fix it. In the literature it is possible to find a large number of orthonormal wavelet bases for L^2:

$$L^2 = < \psi_{jk}, j \in \mathbb{Z}^+, k = 1, \cdots, 2^j >^{span} .$$

Such bases consist of functions with the following localization properties:

- The position of ψ_{jk} is $x_{jk} = k/2^j$.
- The momentum of ψ_{jk} is $\xi_{jk} = \alpha 2^j$ ($\alpha \neq 0$ independent of j and k).
- The localization in position of ψ_{jk} is $\Delta x_{jk} \sim 2^{-j}$.
- The localization in frequency (or momentum) of ψ_{jk} is $\Delta \xi_{jk} \sim 2^j$

We remark that, by the Heisenberg uncertainty principle it is not possible to localize a function arbitrarily well both in position and momentum ($\Delta x \cdot \Delta \xi \geq 1$). Therefore the functions ψ_{jk} are localized in the phase space nearly as well as possible.

The construction of such bases is originally performed on $\mathbb{R}$, but it can be carried out also on the interval ([33]) with boundary conditions of different type (homogeneous Dirichlet, periodic, ...). However in this paper we will not explicitly deal with the issue of boundary condition, since the phenomenology we are interested in is in general concentrated far from the boundaries.

Remark 5.4.1: *Though for simplicity we consider here only orthonormal wavelet bases, the strategy that we are going to present could be applied, without major modifications, in the more general framework of biorthogonal wavelets. [32]*

We recall that the above localization properties imply that a norm equivalence of the form

$$\|f\|_{H^r} \sim \left(\sum_{j,k} 2^{2jr} | < f, \psi_{jk} > |^2 \right)^{1/2}$$

holds for all $f \in H^r$, $r \in (-R, R)$, with the parameter $R > 0$ depending on the particular wavelet basis under consideration. For $f \in H^{-r}$, $r > 0$ the notation $< \cdot, \cdot >$ is to be intended as the duality relation between H^{-r} and H^r.

In the following it will be useful to represent each basis function ψ_{jk} with the rectangle $]k2^{-j}, (k+1)2^{-j}[\times]2^j, 2^{j+1}[$ in the (x, ξ) plane. Using such a

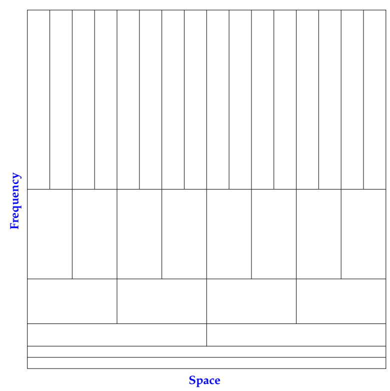

Figure 4. "tiling" of the phase-space corresponding to a wavelet basis

representation yields a "tiling" of the phase space which well represents the localization features of such bases.

When considering adaptive wavelet methods the class of bases $\mathcal{B}$ takes the form

$$\mathcal{B} = \{B = \{\psi_{jk},\ (j,k) \in \Lambda\},\ \Lambda \text{ finite subset of } \mathbb{Z}^+ \times \mathbb{Z}\}.$$

A simple, yet effective, adaptive strategy based on wavelet bases is the following [67], [12], [18]. Let $u(\varepsilon_n)$ be given:

$$u(\varepsilon_n) = \sum_{j,k} u^n_{j,k} \psi_{jk}.$$

We can construct a basis $B_{n+1} \in \mathcal{B}$ for approximating $u(\varepsilon_{n+1})$ by simply looking at the size of the coefficients $u^n_{j,k}$. If a coefficient is big, the corresponding function is included in the basis B_{n+1}, as well as some "neighboring" (in the phase-space) functions. If, on the other hand, a coefficient is very small, the corresponding function will not belong to the basis B_{n+1}.

More precisely, we define B_{n+1} as follows. We choose two tolerances δ_r and δ_a, as well as a number N_{add} of "relevant neighbors", and we set

$$B^r_{n+1} = \left\{ \psi_{jk} : \quad 2^{\frac{3}{2}j} |u^n_{j,k}| > \delta_r \right\}. \tag{5.4.1}$$

Moreover, we set

$$\begin{aligned} I^{add}_{n+1} &= \{(j,k) :\ 2^{\frac{3}{2}j} |u^n_{j,k}| > \delta_a\}, & (5.4.2)\\ N_{jk} &= \{(j+\zeta, 2^\zeta k + \eta), \zeta = 0,1, \quad \eta = -N_{add}, \cdots, N_{add}\}, & (5.4.3)\\ B^a_{n+1} &= \bigcup_{(j,k)\in I^{add}_{n+1}} \{\psi_{jk},\ (j,k) \in N_{jk}\}. & (5.4.4) \end{aligned}$$

The basis B_{n+1} is then defined as

$$B_{n+1} = B^r_{n+1} \cup B^a_{n+1}. \tag{5.4.5}$$

We stress out that the above refining and de-refining strategy is tuned in order to give a good approximation in $H^{\frac{3}{2}}$ rather than in L^2. This is reflected by the presence of factor $2^{\frac{3}{2}j}$ in equations (5.4.1) and (5.4.2). The choice [18] of such a norm is heuristically motivated by the fact that we are dealing with a third order operator.

5.5 Wavelet Packets

A "Wavelet Packet dictionary" $\mathcal{D}$ is a overly redundant (non linearly independent) set of functions, which, by abuse of notation we will call "basis functions". Each basis function $w_{p,\omega,s}$ in the set is identified by three parameters:

the (scaled) position p, the (scaled) wave number ω and the scale s. Each of these functions is constructed in such a way that its space-localization is $\Delta x \sim 2^{-s}$ with center $p/2^s$ and its frequency-localization is $\Delta\xi \sim 2^s$ with center $\omega 2^s$. Again, in view of the Heisenberg uncertainty principle, the functions $w_{p,\omega,s}$ are localized in the phase space nearly as well as possible.

For a given $f \in L^2$ one can define the *wavelet packet transform* of f:

$$d_{p,\omega,s}(f) = \int f w_{p,\omega,s}.$$

Out of a given "Wavelet Packet dictionary" it is possible to extract many different orthonormal bases for L^2 of the form $B_\Lambda = \{w_{p,\omega,s},\ (p,\omega,s) \in \Lambda\}$, by selecting suitable subsets Λ of the index set $\{(p,\omega,s)\}$. For such subsets the inversion formula holds for any $f \in L^2$:

$$f = \sum_{(p,\omega,s)\in\Lambda} d_{p,\omega,s}(f) w_{p,\omega,s}.$$

In particular, for the choice $\Lambda = \{(p,1,s),\ p \in \mathbb{Z}, s \in \mathbb{Z}^+\}$ one obtains the usual wavelet orthonormal basis described in the previous section.

The "basis functions" $w_{p,\omega,s}$ can be constructed as follows. We start by choosing two *quadrature mirror filters*, two finite sequences $\{h_n\}$ and $\{g_n\}$, satisfying the following relations:

$$\sum_n h_{2n} = \sum_n h_{2n+1} = \frac{1}{\sqrt{2}}, \qquad g_n = (-1)^{n-1} h_{1-n} \qquad \forall n \in \mathbb{Z}, \tag{5.5.1}$$

$$\sum_n h_n h_{n+2m} = \sum_n g_n g_{n+2m} = \begin{cases} 1, & \text{if } m = 0, \\ 0, & \text{otherwise}, \end{cases} \tag{5.5.2}$$

$$\sum_n h_n g_n + 2m = 0, \qquad \forall m \in \mathbb{Z}. \tag{5.5.3}$$

We can then define a family of functions, depending on an integer parameter $\ell \geq 0$ by

$$W_{2\ell}(x) \ = \ \sqrt{2}\sum_n h_n W_\ell(2x-n), \tag{5.5.4}$$

$$W_{2\ell+1}(x) \ = \ \sqrt{2}\sum_n g_n W_\ell(2x-n). \tag{5.5.5}$$

We remark that W_0 satisfies a dilation equation. Conditions (5.5.1 – 5.5.3) guarantee the existence of a compactly supported solution of such a dilation equation (and therefore they imply the well posedness of definition (5.5.4)). W_0 and W_1 are, respectively, the scaling and wavelet functions of the corresponding wavelet basis. The pair $\{h_n\}$ and $\{g_n\}$ can be chosen in such a way that the functions W_ℓ have any prescribed smoothness.

Lemma 5.5.1: *Let $\{h_n\}$ and $\{g_n\}$ be two families of QMFs satisfying conditions (5.5.1 – 5.5.3). Let $\{W_\ell\}_\ell$ the family of wavelet packets associated with the filters $\{h_n\}$ and $\{g_n\}$. Then there is a $K < \infty$, such that supp$(W_\ell) \subseteq [-K, K]$, for all $\ell \geq 1$.*

For $p \in \mathbb{Z}, \omega \in \mathbb{Z}^+, s \in \mathbb{Z}^+$, wavelet packets are then defined by

$$w_{p,\omega,s} = 2^{s/2} W_\omega(2^s x - p), \tag{5.5.6}$$

where ω is the (integer) scaled wave number.

Orthonormal sets can be extracted out of the wavelet packet dictionary by selecting index subsets Λ for which the dyadic intervals $\{[\omega 2^s, (\omega+1)2^s[: (p, \omega, s) \in \Lambda\}$ form a disjoint cover of the positive semi-axis.

More precisely, we will say that an index set Λ is *admissible* if the following condition is satisfied: for Ξ defined by

$$\Xi = \{(\omega, s) : \ \exists p, (p, \omega, s) \in \Lambda\},$$

it holds for each $(\omega, s), (\omega', s') \in \Xi$

$$(\omega, s) \neq (\omega', s') \Rightarrow [\omega 2^s, (\omega+1)2^s[\cap [\omega' 2^{s'}, (\omega'+1)2^{s'}[= \emptyset.$$

It is possible to prove that if the index set Λ is admissible, then $B_\Lambda = \{w_{p,\omega,s}, (p, \omega, s) \in \Lambda\}$ forms an orthogonal system.

We say that an admissible index set Λ,

$$\Lambda = \{(p, \omega, s), \ (\omega, s) \in \Xi, \ p \in I_{(\omega,s)}\},$$

is *complete* at scale S on the domain $\mathbb{T}$ if Ξ satisfies

$$\bigcup_{(\omega,s)\in\Xi} [\omega 2^s, (\omega+1)2^s[= [0, 2^S[, \tag{5.5.7}$$

and if, for all $(\omega, s) \in \Xi$, the set $I_{(\omega,s)} = \{p : (p, \omega, s) \in \Lambda\}$ satisfies

$$\bigcup_{p \in I_{(\omega,s)}} [p 2^{-s}, (p+1)2^{-s}] \supseteq \mathbb{T}. \tag{5.5.8}$$

Roughly speaking, a complete index set identifies the orthonormal basis of a discrete subspace of $L^2(\mathbb{T})$ corresponding to a uniform discretization with mesh size 2^{-S}. Condition (5.5.7) assures that all the frequencies below 2^S are covered, and condition (5.5.8) guarantees that, for all frequency ranges, all the spatial positions are present.

We can then define a class $\mathcal{B}$ of bases as follows:

$$\mathcal{B} = \{B_\Lambda \subset \mathcal{D}, \Lambda \in \mathcal{L}\}, \qquad B_\Lambda := \{w_{p,\omega,s}, (p,\omega,s) \in \Lambda\}, \tag{5.5.9}$$

with

$$\mathcal{L} = \{\Lambda\ : \Lambda \text{ is } \ \text{admissible} \ \text{ and } \#(\Lambda) < +\infty\}.$$

Again, it is useful to visualize the phase-space localization property of each basis function by means of a "tiling" (obtained by representing $w_{p,\omega,s}$ by the rectangle $]p2^{-s},(p+1)2^{-s}[\times]2^s\omega, 2^s(\omega+1)[$.

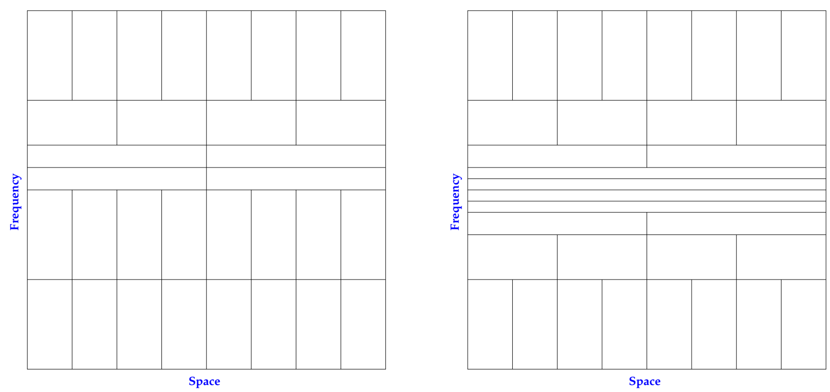

Figure 5. "tilings" corresponding to different wavelet packets orthonormal bases

The redundancy of the wavelet packet dictionary allows for a greater flexibility as far as space frequency localization is concerned. On the other hand, the computation of a good approximation of a function by means of few degrees of freedom needs for a more sophisticated approach.

Let us at first consider the problem of approximating a known function f with as few as possible degrees of freedom. The extraction of a basis well suited for approximating the given function f with few degrees of freedom needs now to be performed in two steps:

(i) Select a complete (at scale S if the function is sampled with sampling rate 2^{-S}) index set $\hat{\Lambda} \in \mathcal{L}$.

(ii) Select a subset $\hat{\Lambda}_\delta \subset \hat{\Lambda}$ such that

$$\|f - \sum_{(p,\omega,s)\in\hat{\Lambda}_\delta} w_{p,\omega,s} w_{p,\omega,s}\|_{H^{3/2}} \quad \text{is small.}$$

Task (ii) can be performed quite easily thanks to the observation that for all functions $f \in H^{3/2}$ and for any orthonormal basis $B_{\hat{\Lambda}} \in \mathcal{B}$ it holds

$$\|f\|_{H^{3/2}} \sim \left(\sum_{(p,\omega,s)\in\Lambda} |(2^s\omega)^{3/2} w_{p,\omega,s}|^2 \right)^{1/2}.$$

As far as task (i) is concerned, the optimal choice is provided by an index set $\hat{\Lambda}$ such that the corresponding basis $B_{\hat{\Lambda}}$ minimizes a suitable additive entropy:

$$H(f, B_{\hat{\Lambda}}) = \min_{B_\Lambda} H(f, B_\Lambda).$$

The basis $B_{\hat{\Lambda}}$ is usually referred to as *best basis* for the function f. If the goal is, as in our case, to approximate f with as few as possible degrees of freedom, then the entropy can be for instance chosen of the following form

$$H(f, B_\Lambda) = \#(\{(p,\omega,s) \in \Lambda : \ |(2^s\omega)^{3/2}\, w_{p,\omega,s}| \geq \delta\}). \tag{5.5.10}$$

Once $B_{\hat{\Lambda}}$ has been selected, the subset $\hat{\Lambda}_\delta$ is clearly defined as

$$\hat{\Lambda}_\delta = \{(p,\omega,s) \in \hat{\Lambda} : \ |(2^s\omega)^{3/2}\, w_{p,\omega,s}| \geq \delta\}. \tag{5.5.11}$$

The implementation of the wavelet packet transform and of the best basis search algorithm for a given function f is described in detail in [35]. If f is sampled with step h, the entire procedure has complexity $\frac{1}{h}\log\frac{1}{h}$.

5.6 Wavelets vs Wavelet Packets

In [18] the effectiveness of wavelet and wavelet packets adaptive scheme has been compared. We report here the results of the tests.

Let us start with the the following test on the effectiveness of wavelet methods. Let $(\varepsilon_n)_{n=0,\cdots,N}$ be given.

- Compute a reference solution $u(\varepsilon_n)$, $(n = 0, \cdots, N)$, by solving (P_ε) for $\varepsilon = \varepsilon_n$ with a finite difference scheme on a very fine grid ("overkill"). In our case we used a discretization step $2.5\ 10^{-4}$.

- For each $n = 0, \cdots, N-1$, perform the following procedure:

Step 1. given $\check{u}_n$ (computed at the previous step)

$$\check{u}_n = \sum_{j,k \in \Lambda_n} u^n_{j,k} \psi_{jk},$$

($\check{u}_n$ approximation of $u(\varepsilon_n)$), define $B_{n+1} = \{\psi_{jk}, \quad (j,k) \in \Lambda_{n+1}\}$ and the corresponding space $V_{n+1} =< B_{n+1} >^{span}$ by the preceding adaptive strategy (5.4.1)–(5.4.5).

Step 2. define $\check{u}_{n+1}$ as the L^2 orthogonal projection of $u(\varepsilon_{n+1})$ onto V_{n+1}

$$\check{u}_{n+1} = \sum_{(j,k) \in \Lambda_{n+1}} < u(\varepsilon_{n+1}), \psi_{jk} > \psi_{jk}.$$

- For all $n = 1, \cdots, N$ evaluate the relative error

$$e_n = \frac{\|u(\varepsilon_n) - \check{u}_n\|_{H^{3/2}}}{\|u(\varepsilon_n)\|_{H^{3/2}}}.$$

The above test has been performed for the following data in (P_ε): the doping profile $C(x)$ is chosen as in figure 6, with max(C)=1 and min(C)=0.1, the Debye length is $\lambda = 0.1$, the relaxation time is $\tau = 1/8$ and the applied voltage is $V_0 = 6.5$. The solutions for these values of the parameters are the ones depicted in pictures 1–3.

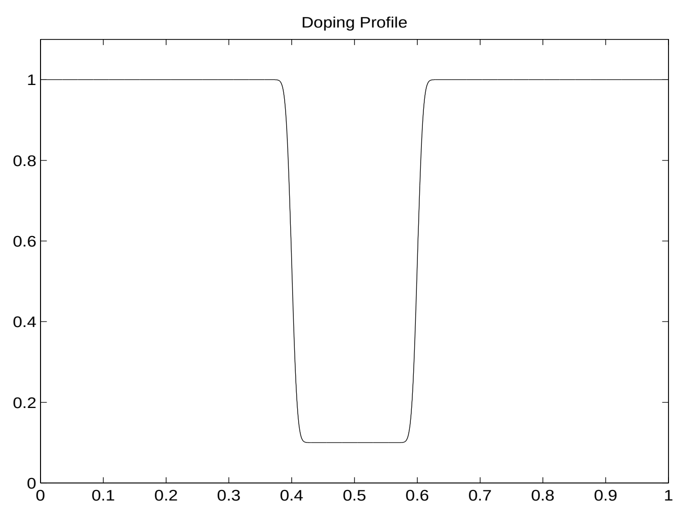

Figure 6. Doping profile $C(x)$

In the following table we give [18] for each n, the value ε_n, the relative error e_n, the cardinality N_{adapt} of the adaptively selected basis B_n, as well as the "compression ratio" (CR) N_{adapt}/N_{unif} (N_{unif} ($\sim 2^j$ for some j) being the number of degrees of freedom which would be necessary to represent $u(\varepsilon_n)$ with a uniform discretization with the same accuracy).

ε_n	e_n	N_{adapt}	CR
.0075	.1535	120	4.27
.007	.0218	156	3.28
.006	.0377	160	3.20
.005	.0202	175	2.93
.004	.0208	195	5.25
.0025	.0230	256	4.00
.002	.0137	353	5.70
.0016	.0111	446	4.51
.0011	.0123	596	6.71
.00102	.0020	858	4.66
.001	.0014	969	4.13

Table 5.6.1. Wavelet Adaptive Strategy

If one wants to use wavelet packet dictionaries and the concept of best basis in the framework of the adaptive type algorithm described in section 5.3, for each ε_{n+1} one must be able to perform tasks (i) and (ii), for $u(\varepsilon_{n+1})$ by analyzing the solution $u(\varepsilon_n)$ at the previous continuation step. We will not deal here with the problem relative to task (ii), the extraction out of the best basis of a small subset well suited to well approximate $u(\varepsilon_{n+1})$. We refer to [63], where a possible strategy has been proposed, based on a suitable definition of "neighbors" (in the phase space) of a given basis function $w_{p,\omega,s}$. We will rather concentrate here on the problem of selecting the best basis - or a close enough basis - for $u(\varepsilon_{n+1})$, by analyzing the solution $u(\varepsilon_n)$ at the previous continuation step.

In order to verify to what extent this is feasible, the following test has been performed

- Compute $u(\varepsilon_n)$ ($n = 0, \ldots, N$) by solving the QHD equations for $\varepsilon = \varepsilon_n$ by finite differences on a very fine grid ("overkill"). In our case we used a discretization step 2.5 10^{-4} (which corresponds to a scale $S \sim 12$)

- Compute the wavelet packet transform $w^n_{p,\omega,s}$ of $u(\varepsilon_n)$

$$w^n_{p,\omega,s} = \int u(\varepsilon_n) w_{p,\omega,s},$$

and select the best basis $\hat{B}^n = \{w_{p,\omega,s}, (p,\omega,s) \in \hat{\Lambda}^n\}$:

$$H(u(\varepsilon_n), \hat{B}^n) = \min_{B_\Lambda} H(u(\varepsilon_n), B_\Lambda).$$

- Compute the number of coefficients needed for approximating $u(\varepsilon_{n+1})$ using elements of the best basis for $u(\varepsilon_n)$:

$$H(u(\varepsilon_{n+1}), \hat{B}^n)) = \#(\{(p,\omega,s) \in \hat{\Lambda}^n : \ |(2^s\omega)^{3/2}\, w^{n+1}_{p,\omega,s}| \geq \delta\}),$$

and compare $H(u(\varepsilon_{n+1}), \hat{B}^n))$ with the optimal number of coefficients $H(u(\varepsilon_{n+1}), \hat{B}^{n+1}))$.

- Define an approximation to $u(\varepsilon_{n+1})$, by selecting an index set $\hat{\Lambda}^n_\delta$

$$\hat{\Lambda}^n_\delta = \{(p,\omega,s) \in \hat{\Lambda}^n : \ |(\omega 2^s)^{\frac{3}{2}} w^{n+1}_{p,\omega,s}| \geq \delta\}$$

and computing

$$\bar{u}_{n+1} = \sum_{(p,\omega,s)\in\hat{\Lambda}^n_\delta} w^{n+1}_{p,\omega,s} w_{p,\omega,s}.$$

- Evaluate the relative error

$$e^{n+1}_{wp} = \frac{\|u(\varepsilon_{n+1}) - \bar{u}_{n+1}\|_{H^{3/2}}}{\|u(\varepsilon_{n+1})\|_{H^{3/2}}}.$$

The following table [18] summarizes the results of such test, performed in the same case as in the previous section. We report the values of ε_{n+1}, of the number N_{opt} of significant degrees of freedom when approximating $u(\varepsilon_{n+1})$ by means of the best basis $\hat{B}^{n+1}_\Lambda$, the number N_{est} when approximating $u(\varepsilon_{n+1})$ by means of the best basis $\hat{B}^n_\Lambda$ obtained from the analysis of $u(\varepsilon_n)$ and the error e^{n+1}_{wp}. The table also displays the two compression ratios $CR_{opt} = N_{unif}/N_{opt}$ and $CR_{est} = N_{unif}/N_{est}$ (N_{unif} is defined, as in the previous section, as the number of degrees of freedom which would be necessary to represent the solution with the same accuracy by means of a uniform discretization of wavelet type).

If we analyse the results of the previous tests we realise that wavelet adaptivity for the numerical solution of QHD system is not entirely satisfactory. This is mainly due to the fact that wavelet bases approximate high frequencies with basis functions which are highly localized in space. In other words, when approximating an highly oscillating function, the use of wavelets correspond to using (in the region where oscillations occur) an uniform discretization. By a comparison it is clear that, especially at very high frequencies, wavelet packets perform better, allowing to almost double the compression ratio.

ε_n	N_{opt}	N_{est}	CR_{opt}	CR_{est}	e^n_{wp}
.0075	95	116	5.3	4.4	.0080
.007	102	105	5.0	4.8	.0074
.006	112	122	4.6	4.2	.0067
.005	124	137	4.1	3.7	.0080
.004	150	163	6.8	6.3	.0045
.0025	192	168	6.1	5.3	.0028
.002	253	292	8.0	6.9	.0023
.0016	331	366	6.0	5.5	.0024
.0011	480	585	8.3	6.8	.0012
.00102	525	550	7.6	7.3	.0010
.001	496	503	8.1	8.0	.0011

Table 5.6.2. Performance of an adaptive WP algorithm

5.7 Wavelet Packet adaptive methods & Compression techniques

Wavelet packets, due to their better frequency localization property, are then superior to wavelets in drastically diminishing the required number of degrees of freedom for well approximating the solution. Clearly they are more costly when used to solve problem (P_ε), due to the higher cost of the WP transform with respect to the FWT, and to the additional cost of the best basis search.

The intrinsic difficulties connected with the application of wavelets in the framework of non–linear problems [44], [30] are even harder when dealing with Wavelet Packets. On the other hand the computation of the wavelet packet coefficients of a given function obtained by applying a nonlinear functional to a wavelet packet lacunary sum, often requires the computation of the corresponding scaling coefficients at the finest scale, hence making the realization of a fully adaptive wavelet packet scheme unfeasible from a computational point of view.

However, the strong non–linearity of the problem requires a continuation algorithm in the parameter ε and for each ε_n of the ε-sequence a non–linear system must be solved (for instance with a Newton algorithm). It is then clear that, when (P_ε) is solved for a small ε, a linearized QHD system must be solved many times and the much lower number of degrees of freedom selected with the wavelet packet procedure is expected to largely compensate the higher cost of the basis selection and of the non–linearity treatment and to provide an over all more efficient numerical scheme.

1. Compute an approximation u_0 to the solution $u(\varepsilon_0)$ of equations (P_ε) for $\varepsilon = \varepsilon_0$ using a coarse uniform grid. This is possible since for $\varepsilon \sim 1$ the solution of (P_ε) is smooth.

2. Given the approximation u_n to the solution $u(\varepsilon_n)$ of (P_{ε_n}), an approximation u_{n+1} to the solution $u(\varepsilon_{n+1})$ of $(P_{\varepsilon_{n+1}})$ is obtained as follows:

   - By analyzing u_n select a wavelet packets basis B_{n+1}, well suited for approximating $u(\varepsilon_{n+1})$.
   - Compute u_{n+1} in $V_{n+1} =< B_{n+1} >^{span}$ approximate solution of $(P_{\varepsilon_{n+1}})$, by a suitable numerical method (for instance, Galerkin approximation of $(P_{\varepsilon_{n+1}})$).

When we deal with the computation of u_{n+1} in $V_{n+1} =< B_{n+1} >^{\text{span}}$, approximate solution of $(P_{\varepsilon_{n+1}})$, by, for instance, Galerkin approximation, we need to solve infinite linear systems of the type

$$Ru_{n+1} = g,$$

where u_{n+1} has only N nonzeros elements (we say $u_{n+1} \in \Sigma_N$), due to the particular choice of the basis B_{n+1} and $R = (r_{\lambda,\lambda'})$, with$r_{\lambda,\lambda'} = \int w_\lambda w_{\lambda'}^{(f)}$, $\lambda = (p, \omega, s)$, for some $f > 0$ and $w_\lambda, w_{\lambda'}$ belonging to B_{n+1}.

Since, as we will see in the next section, the entries of the matrices R involved in the wavelet packet discretization of the operators have good decay properties, we can apply such operators to sparse vectors in a quite effective way, i.e. we can exploit the sparsity of the stiffness matrix R and of the vector u_{n+1} in such a way to perform, with a low computational cost, an approximation of the matrix-vector product Ru_{n+1}.

Non the less one can still hope to take advantage of the sparsity of the solution of the QHD equation within the framework of a solution in the full space!

5.7.1 First compression

In the wavelet context it is well known [20], [41] that a large class of operators have an almost sparse representation in wavelet coordinates thanks to the good properties of localization both in space and in frequency of wavelet bases. Thus discarding all entries below a certain threshohld will then give rise to a sparse matrix that can be further processed by efficient linear algebra tools. Motivated by such a result, as wavelet packets provide better properties of localization both in space and in frequency, the representation of such operators in terms of wavelet packets should be in principle sparser than in the wavelet case. Thus applying wavelet-like compression techniques [41], [28], [27] in such a wavelet packets framework, should give rise to much sparser matrices, hence reducing the computational cost in solving the linear systems resulting from the Galerkin wavelet packets discretization of the problem.

To do that we need to estimate objects of the type

$$|\int w_{p,\omega,s}(x) w_{p',\omega',s'}^{(f)}(x)dx|,$$

where $w_{p',\omega',s'}^{(f)}$ is the $f-$th derivative of $w_{p',\omega',s'}$. It is not restrictive to suppose $s > s'$, otherwise we will integrate by parts f-times. Let us denote by $i(w_{p,\omega,s}, w_{p',\omega',s'})$ the function whose value is one if the supports of $w_{p,\omega,s}$ and $w_{p',\omega',s'}$ are disjoint, zero otherwise, then the following estimate holds:

Theorem 5.7.1: *Let $\{W_\omega\}_{\omega=0}^\infty$ be a family of wavelet packets, then it holds*

$$\left|\int w_{p,\omega,s}(x) w_{p',\omega',s'}^{(f)}(x)dx\right| \lesssim 2^{(s+s')/2} 2^{-t(s-s')+s'f-s+[\log_2 \omega'](f+t)} i(w_{p,\omega,s}, w_{p',\omega',s'}).$$

In order to prove Theorem 5.7.1 we need the following result [60]:

Lemma 5.7.1: *For every function W_ω it is possible to find a function $g(x)$ such that $g^{(t)}(x) = W_\omega(x)$, with $\|g\|_\infty \leq C$, where t equals the number of null momenta of W_ω and g has the same support as W_ω.*

Proof: Using the above Lemma and integrating by parts yield:

$$\begin{aligned}
|r_{\lambda,\lambda'}| &:= \left| \int w_{p,\omega,s}(x) w^{(f)}_{p',\omega',s'}(x) dx \right| \\
&\leq 2^{(s+s')/2} \left| \int [2^{-st} g(2^s x - p)]^{(t)} [W_{\omega'}(2^{s'}x - p')]^{(f)} dx \right| \\
&\leq 2^{(s+s')/2} \left| \int 2^{-st} g(2^s x - p) [W_{\omega'}(2^{s'}x - p')]^{(f+t)} dx \right| \\
&\leq 2^{(s+s')/2} 2^{-st} 2^{s'(f+t)} \left| \int g(2^s x - p) W^{(f+t)}_{\omega'}(2^{s'}x - p') dx \right| \\
&\leq 2^{(s+s')/2} 2^{-st} 2^{s'(f+t)} \|W_{\omega'}\|_{W^{f+t,\infty}} i(w_{p,\omega,s}, w_{p',\omega',s'}) \int |g(2^s x - p)| dx.
\end{aligned}$$

By using Bernstein-type inequality

$$\|W_{\omega'}\|_{W^{f+t,\infty}} \leq 2^{[\log_2 \omega'](f+t)} \|W_{\omega'}\|_\infty,$$

where $[x]$ denotes the smaller integer larger or equal than x, we have

$$|r_{\lambda,\lambda'}| \leq 2^{(s+s')/2} 2^{-st} 2^{s'(f+t)} 2^{[\log_2 \omega'](f+t)} i(w_{p,\omega,s}, w_{p',\omega',s'}) \|g(x)\|_\infty 2^{-s} |\mathrm{supp}(W_\omega)|.$$

Finally, by using Lemma 5.5.1 which yields $|\mathrm{supp}(W_\omega)| \leq K$ for all ω, one obtains:

$$|r_{\lambda,\lambda'}| \leq C 2^{(s+s')/2} 2^{-t(s-s')+s'f-s+[\log_2 \omega'](f+t)} K i(w_{p,\omega,s}, w_{p',\omega',s'})$$

■

Now we are ready to apply our first compression to the matrix R.

Definition *Let us fix $J > 0$. We apply first truncation defining $\tilde{R}_J = (\tilde{r}_{\lambda,\lambda'})_{\lambda,\lambda'}$ as follows*

$$\tilde{r}_{\lambda,\lambda'} = \begin{cases} r_{\lambda,\lambda'}, & \text{if } \lambda' \in I_\lambda(J) \cap \Sigma_N \\ 0, & \text{otherwise} \end{cases}$$

where $I_\lambda(J) = I^{(1)}_\lambda(J-1) \cup I^{(2)}_\lambda(J-1)$, with

$$I^{(1)}_\lambda(J-1) = \{\lambda' : \quad i(w_\lambda, w_{\lambda'}) \neq 0, s > s', 2^{l_1((s',\omega'))} \leq 2^{-(J-1)}\}$$

$$I_\lambda^{(2)}(J-1) = \{\lambda' : \quad i(w_\lambda, w_{\lambda'}) \neq 0, s' > s, 2^{l_2((s',\omega'))} \leq 2^{-(J-1)}\};$$

while functions $l_i : \mathbb{N}^2 \to \mathbb{R}$ *are defined as follows*

$$l_1((s',\omega')) := (s+s')/2 - |s-s'|(t-1) + s'f - s + [\log_2 \omega'](f+t)$$

and

$$l_2((s',\omega')) := (s+s')/2 - |s-s'|t + sf - s' + [\log_2 \omega](f+t).$$

Theorem 5.7.2: *Let* $\Lambda_N \subset \Lambda$ *be a fixed subset of* Λ *of cardinality* $\#\Lambda_N \leq N$, *then the following estimate holds*

$$\|R - \tilde{R}_J\|_{\ell^2(\Lambda_N)\to\ell^2} \leq 3MKN2^{-J}. \tag{5.7.1}$$

Proof: In order to prove the Theorem we use Schur Lemma and we reduce to estimate

$$\sum_{\lambda' \in \Lambda_N} |r_{\lambda,\lambda'} - \tilde{r}_{\lambda,\lambda'}|.$$

Using the definition of $\tilde{r}_{\lambda,\lambda'}$ yields

$$\sum_{\lambda' \in \Lambda_N} |r_{\lambda,\lambda'} - \tilde{r}_{\lambda,\lambda'}| \leq \sum_{\lambda' \in \Lambda_N \setminus I_\lambda} |r_{\lambda,\lambda'}|.$$

Hence, by using Theorem 5.7.1, it follows

$$\begin{aligned}
&\sum_{\lambda' \in \Lambda_N \setminus I_\lambda} |r_{\lambda,\lambda'}| \\
\leq & \sum_{\substack{\lambda' : s > s' \\ \lambda' \in \Lambda_N \setminus I_\lambda^{(1)}}} |r_{\lambda,\lambda'}| + \sum_{\substack{\lambda' : s < s' \\ \lambda' \in \Lambda_N \setminus I_\lambda^{(2)}}} |r_{\lambda,\lambda'}| \\
\leq & \sum_{(p',\omega',s') \in \Lambda_N \setminus I_\lambda^{(1)}} C 2^{(s+s')/2} 2^{-t(s-s') + s'f - s + [\log_2 \omega'](f+t)} K i(w_{p,\omega,s}, w_{p',\omega',s'}) + \\
& \sum_{(p',\omega',s') \in \Lambda_N \setminus I_\lambda^{(2)}} C 2^{(s+s')/2} 2^{-t(s'-s) + sf - s' + [\log_2 \omega](f+t)} K i(w_{p,\omega,s}, w_{p',\omega',s'}) \\
= & \sum_{(p',\omega',s') \in \Lambda_N \setminus I_\lambda^{(1)}} C 2^{(s+s')/2} 2^{-t|s-s'| + s'f - s + [\log_2 \omega'](f+t)} K i(w_{p,\omega,s}, w_{p',\omega',s'}) + \\
& \sum_{(p',\omega',s') \in \Lambda_N \setminus I_\lambda^{(2)}} C 2^{(s+s')/2} 2^{-t|s'-s| + sf - s' + [\log_2 \omega](f+t)} K i(w_{p,\omega,s}, w_{p',\omega',s'}).
\end{aligned}$$

Using the fact that the set $\{p' : i(w_{p',\omega',s'}, w_{p,\omega,s}) \neq 0\}$ has cardinality not greater than $3M \max(1, 2^{s-s'})$, where M depends on the length of the support of W_ω yields

$$\begin{aligned}
&\sum_{\lambda'\in\Lambda_N} |r_{\lambda,\lambda'} - \tilde{r}_{\lambda,\lambda'}| \\
\leq &\sum_{(\omega',s')\in\Lambda_N\setminus I_\lambda^{(1)}} 3MK\max(1,2^{s-s'})C2^{(s+s')/2}2^{-t|s-s'|+s'f-s+[\log_2\omega'](f+t)} + \\
&\sum_{(\omega',s')\in\Lambda_N\setminus I_\lambda^{(2)}} 3MK\max(1,2^{s-s'})C2^{(s+s')/2}2^{-t|s'-s|+sf-s'+[\log_2\omega](f+t)} \\
= &\sum_{(\omega',s')\in\Lambda_N\setminus I_\lambda^{(1)}} 3MKC2^{(s+s')/2}2^{-(t-1)|s-s'|+s'f-s+[\log_2\omega'](f+t)} + \\
&\sum_{(\omega',s')\in\Lambda_N\setminus I_\lambda^{(2)}} 3MK2^{(s+s')/2}2^{-t|s'-s|+sf-s'+[\log_2\omega](f+t)} \\
\leq &\ 3MK(\sum_{(\omega',s')\in\Lambda_N\setminus I_\lambda^{(1)}} 2^{l_1((\omega',s'))} + \sum_{(\omega',s')\in\Lambda_N\setminus I_\lambda^{(2)}} 2^{l_2((\omega',s'))})
\end{aligned}$$

Hence, by using the definitions of $I_\lambda^{(1)}$ and $I_\lambda^{(2)}$, it follows

$$\sum_{\lambda'\in\Lambda_N} |r_{\lambda,\lambda'} - \tilde{r}_{\lambda,\lambda'}| \leq 3MKN2^{-J}.$$

In complete analogy one proves an analogous estimate for the rows sums

$$\sum_{\lambda\in\Lambda_N} |r_{\lambda,\lambda'} - \tilde{r}_{\lambda,\lambda'}| \leq \sum_{\lambda\in\Lambda_N\setminus I_{\lambda'}} |r_{\lambda,\lambda'}| \leq 3MKN2^{-J}.$$

■

Remark 5.7.1: *We obtained that*

$$\|R - \tilde{R}_J\|_{\ell^2(\Lambda_N)\to\ell^2} \leq 3MKN2^{-J}$$

and so for each $u_{n+1} \in \Lambda_N$ we are able to estimate the error we commit by using $\tilde{R}_J$ instead of R:

$$\|(R - \tilde{R}_J)u_{n+1}\|_{\ell^2} \leq \|R - \tilde{R}_J\|_{\ell^2(\Lambda_N)\to\ell^2}\|u_{n+1}\|_{\ell^2} \leq 3MKN2^{-J}\|u_{n+1}\|_{\ell^2}.$$

This is a pessimistic estimate and quite rough if we compare it with the one that can be achieved in wavelet context, but this huge difference can be compensated by the fact that wavelet packets perform better than wavelet in approximating with fewer degrees of freedom the high frequency dispersive oscillations of the solution of our particular problem.

5.7.2 Second compression

By exploiting good properties of localization even in frequency of wavelet packets, we can further compress the matrix $\tilde{R}_J$, obtaining a second matrix $\tilde{\tilde{R}}_J$ such that

$$\|R - \tilde{\tilde{R}}_J\|_{\ell^2(\Lambda_N)\to\ell^2}$$

remains little.

From equation (5.5.4) it follows that

$$\hat{W}_\omega(\xi) = m_{\varepsilon_1}(\xi/2)m_{\varepsilon_2}(\xi/4)\dots m_{\varepsilon_j}(\xi/2^j)\hat{W}_0(\xi/2^j), \tag{5.7.2}$$

where $\varepsilon_i \in \{0,1\}$, $m_0(\xi) = \sum_n h_n e^{in\xi}$, $m_1(\xi) = e^{-i\xi}\overline{m_0(\xi+\pi)}$ and

$$|m_0(\xi)|^2 + |m_0(\xi+\pi)|^2 = 1. \tag{5.7.3}$$

From (5.7.3) and from the fact that $m_0(\xi)$ and $m_1(\xi)$ approximate better and better respectively the ideal low pass filter and the ideal high pass filter as the lengths of the filter increase, it is not difficult to prove the following:

Lemma 5.7.2: *For every $0 < \alpha < 1$ it exists a couple of quadrature mirror filters of lengths depending on α such that for every ξ it holds:*

$$|m_0(\xi)m_1(\xi)| \le \alpha. \tag{5.7.4}$$

Proof: We show only that if $|m_0(\xi)m_1(\xi)| \le \alpha$, then $\alpha < 1$. In fact if $\alpha = 1$, it would follow $|m_0(\bar{\xi})| = |m_1(\bar{\xi})| = 1$, for some $\bar{\xi}$, which, used together with (5.7.3), would give $|m_1(\bar{\xi})| = 0$. Absurd. ∎

Now we are interested in estimating objects of the type:

$$|\int \hat{W}_\omega(\xi)\hat{W}_{\omega'}(\xi)d\xi|.$$

Lemma 5.7.3: *The following estimate holds*

$$|\int \hat{W}_\omega(\xi)\hat{W}_{\omega'}(\xi)d\xi| \le C\alpha^{\gamma_{\omega,\omega'}},$$

where $\alpha < 1$, $\gamma_{\omega,\omega'} = \sum_{n=1}^{j'} |\omega_n - \omega'_n|$ and C depends only on $\hat{W}_0$.

Proof: From Plancherel's theorem it follows that

$$\int W_\omega(x)W_{\omega'}(x)dx = 1/(2\pi)\int \hat{W}_\omega(\xi)\hat{W}_{\omega'}(\xi)d\xi.$$

Let us suppose $j > j'$. Recalling (5.7.2) yields

$$\begin{aligned}
&|\int \hat{W}_\omega(\xi)\hat{W}_{\omega'}(\xi)d\xi| \\
&= |\int m_{\varepsilon_1}(\xi/2)m_{\varepsilon_2}(\xi/4)\dots m_{\varepsilon_j}(\xi/2^j)\hat{W}_0(\xi/2^j) \\
&\quad m_{\varepsilon'_1}(\xi/2)m_{\varepsilon'_2}(\xi/4)\dots m_{\varepsilon'_{j'}}(\xi/2^{j'})\hat{W}_0(\xi/2^{j'})d\xi| \\
&= |\int m_{\varepsilon_1}(\xi/2)m_{\varepsilon'_1}(\xi/2)\dots m_{\varepsilon_{j'}}(\xi/2^{j'})m_{\varepsilon'_{j'}}(\xi/2^{j'}) \\
&\quad m_{\varepsilon_{j'+1}}(\xi/2^{j'+1})\dots m_{\varepsilon_j}(\xi/2^j)\hat{W}_0(\xi/2^j)\hat{W}_0(\xi/2^{j'})d\xi|.
\end{aligned}$$

Now, by using Lemma 5.7.2, one obtains

$$\begin{aligned}
&|\int \hat{W}_\omega(\xi)\hat{W}_{\omega'}(\xi)d\xi| \\
&\le \int |m_{\varepsilon_1}(\xi/2)m_{\varepsilon'_1}(\xi/2)|\dots|m_{\varepsilon_{j'}}(\xi/2^{j'})m_{\varepsilon'_{j'}}(\xi/2^{j'})| \\
&\quad |m_{\varepsilon_{j'+1}}(\xi/2^{j'+1})|\dots|m_{\varepsilon_j}(\xi/2^j)||\hat{W}_0(\xi/2^j)||\hat{W}_0(\xi/2^{j'})|d\xi \\
&\le C\alpha^{\gamma_{\omega,\omega'}},
\end{aligned}$$

where $\alpha < 1$, $\gamma_{\omega,\omega'} = \sum_{n=1}^{j'} |\varepsilon_n - \varepsilon'_n|$ and C depends only on $\hat{W}_0$. ■

Let $\lambda = (p, \omega, s)$. Suppose for the moment $s' > s$. As we want to improve the compression of matrix $\tilde{R}_J$, we are interested in estimating in particular objects of the type

$$\left| \int w_{p,\omega,s}(x) w^{(f)}_{p',\omega',s'}(x)dx \right|.$$

Lemma 5.7.4: *The following estimate holds*

$$\left| \int w_{p,\omega,s}(x) w^{(f)}_{p',\omega',s'}(x)dx \right| \le 2^{-1/2(2(f-1/2)s'-s)}(2\pi)^{-1}C_1C_f\alpha^{\gamma_{\lambda,\lambda'}}, \qquad (5.7.5)$$

where C_1, C_f *depend only on* $\hat{w}_0$ *and its Sobolev regularity.* $\alpha < 1$ *and* $\gamma_{\omega,\omega'} = \sum_{n=1}^{j'} |\omega_n - \omega'_n|$.

Proof: In complete analogy with the proof of Lemma 5.7.3, one obtains

$$\begin{aligned}
&\left|\int w_{p,\omega,s}(x) w^{(f)}_{p',\omega',s'}(x)dx\right| \\
&= \frac{2^{(s+s')/2}}{2\pi}\left|\int \left(\frac{|\xi|}{2^{s'}}\right)^f e^{i\frac{\xi}{2^{s'}}p'}\hat{W}_{\omega'}(\xi/2^{s'}) e^{i\frac{\xi}{2^s}p}\hat{W}_{\omega}(\xi/2^s)d\xi\right| \\
&\leq \frac{2^{(s+s')/2}}{2^{s'f}2\pi}\int |\xi|^f|\hat{W}_{\omega'}(\xi/2^{s'})\hat{W}_{\omega}(\xi/2^s)|d\xi \\
&\leq 2^{-1/2(2(f-1/2)s'-s)}(2\pi)^{-1}\int |\xi|^f||m_{\varepsilon'_1}(\xi/2^{s'+1})m_{\varepsilon'_2}(\xi/2^{2+s'})\dots m_{\varepsilon'_{j'}}(\xi/2^{j'+s'})\hat{W}_0(\xi/2^{j'+s'}) \\
&\quad m_{\varepsilon_1}(\xi/2^{s+1})m_{\varepsilon_2}(\xi/2^{2+s})\dots m_{\varepsilon_j}(\xi/2^{s+j})\hat{W}_0(\xi/2^{s+j})|d\xi.
\end{aligned}$$

Using Lemma 5.7.2 yields

$$\left|\int w_{p,\omega,s}(x) w^{(f)}_{p',\omega',s'}(x)dx\right| \leq 2^{-1/2(2(f-1/2)s'-s)}(2\pi)^{-1}C_1C_f\alpha^{\gamma_{\lambda,\lambda'}},$$

where $\alpha < 1$, $\gamma_{\omega,\omega'} = \sum_{n=1}^{j'}|\varepsilon_n - \varepsilon'_n|$, while C_1, C_f depend only on $\hat{W}_0$ and its Sobolev regularity. ∎

Now we apply to $\tilde{R}_J$ a second compression defined as follows

Definition *Let us define* $\tilde{\tilde{R}}_J = (\tilde{\tilde{r}}_{\lambda,\lambda'})_{\lambda,\lambda'}$ *where*

$$\tilde{\tilde{r}}_{\lambda,\lambda'} = \begin{cases} \tilde{r}_{\lambda,\lambda'}, & \text{if } \lambda' \in \Lambda_\lambda(J)\cap\Lambda_N \\ 0, & \text{otherwise} \end{cases}$$

where $\Lambda_\lambda(J) = \Lambda^{(1)}_\lambda(J-1)\cup\Lambda^{(2)}_\lambda(J-1)$, *with*

$$\Lambda^{(1)}_\lambda(J-1) = \{\lambda' : s < s', \quad 2^{-1/2(2(f-1/2)s'-s)}\alpha^{\gamma_{\lambda,\lambda'}} > 2^{-J}\},$$

$$\Lambda^{(2)}_\lambda(J-1) = \{\lambda' : s > s', \quad 2^{|s-s'|-1/2(2(f-1/2)s-s')}\alpha^{\gamma_{\lambda,\lambda'}} > 2^{-J}\}.$$

Theorem 5.7.3: *Let $\Lambda_N \subset \Lambda$ be a fixed subset of Λ of cardinality $\#\Lambda_N \leq N$, then the following estimate holds*

$$||\tilde{R}_J - \tilde{\tilde{R}}_J||_{\ell^2(\Lambda_N)\to\ell^2} \lesssim N2^{-J}. \tag{5.7.6}$$

Proof: In order to prove the Theorem we use Schur's Lemma. Hence we reduce to estimate $\sum_{\lambda'\in\Lambda_N}|\tilde{r}_{\lambda,\lambda'} - \tilde{\tilde{r}}_{\lambda,\lambda'}|$. By means of the definition of the set

Λ_λ we obtain

$$\begin{aligned}
&\sum_{\lambda' \in \Lambda_N} |\tilde{r}_{\lambda,\lambda'} - \tilde{\tilde{r}}_{\lambda,\lambda'}| \\
\leq &\sum_{\lambda' \in (\Lambda_N \cap I_\lambda) \backslash \Lambda_\lambda} |\tilde{r}_{\lambda,\lambda'}| \\
\leq &\sum_{\substack{\lambda' : s < s' \\ \lambda' \in (\Lambda_N \cap I_\lambda) \backslash \Lambda_\lambda^{(1)}}} |\tilde{r}_{\lambda,\lambda'}| + \sum_{\substack{\lambda' : s > s' \\ \lambda' \in (\Lambda_N \cap I_\lambda) \backslash \Lambda_\lambda^{(2)}}} |\tilde{r}_{\lambda,\lambda'}|.
\end{aligned}$$

Now we proceed as in the proof of Theorem 5.7.2. Using Lemma 5.7.4 and the fact that the set $\{p' : i(w_{p',\omega',s'}, w_{p,\omega,s}) \neq 0\}$ has cardinality not greater than $3M \max(1, 2^{s-s'})$, where M depends on the length of the support of W_ω, yields

$$\begin{aligned}
&\sum_{\lambda' \in \Lambda_N} |\tilde{r}_{\lambda,\lambda'} - \tilde{\tilde{r}}_{\lambda,\lambda'}| \\
\leq &\sum_{(\omega',s') \in (\Lambda_N \cap I_\lambda) \backslash \Lambda_\lambda^{(1)}} (2\pi)^{-1} 3MK \max(1, 2^{s-s'}) 2^{-1/2(2(f-1/2)s'-s)} (2\pi)^{-1} C_1 C_f \alpha^{\gamma_{\lambda,\lambda'}} + \\
&\sum_{(\omega',s') \in (\Lambda_N \cap I_\lambda) \backslash \Lambda_\lambda^{(2)}} (2\pi)^{-1} 3MK \max(1, 2^{s-s'}) 2^{-1/2(2(f-1/2)s-s')} (2\pi)^{-1} C_1 C_f \alpha^{\gamma_{\lambda,\lambda'}} \\
= \; &3MK(2\pi)^{-1} C_1 C_f \Bigg(\sum_{(\omega',s') \in (\Lambda_N \cap I_\lambda) \backslash \Lambda_\lambda^{(1)}} (2\pi)^{-1} 2^{-1/2(2(f-1/2)s'-s)} \alpha^{\gamma_{\lambda,\lambda'}} + \\
&\sum_{(\omega',s') \in (\Lambda_N \cap I_\lambda) \backslash \Lambda_\lambda^{(2)}} 2^{|s-s'|} 2^{-1/2(2(f-1/2)s-s')} (2\pi)^{-1} \alpha^{\gamma_{\lambda,\lambda'}} \Bigg) \\
\leq \; &(3MK(2\pi)^{-1} C_1 C_f) N 2^{-J}.
\end{aligned}$$

In complete analogy one proves an analogous estimate for the row sums. ■

5.8 Open problems and perspectives

The use of the adaptive treatment of a nonlinear PDE (describing the electron density in a quantum hydrodynamic model for semi-conductors) by wavelet packets discretization is motivated by the locally oscillating patterns of the solution which can be better compressed by a few wavelet packets than by wavelets or adaptive finite elements. This property is illustrated by some numerical examples. Even if in this Chapter we focus mainly on the issue of the sparsity of the matrix resulting from the wavelet discretization of differential operators, there are however several new problems which are raised by the use of wavelet packets for the adaptive discretization: appropriate quadrature rules, treatment of the nonlinear terms and conditioning of the resulting matrices.

As a useful step toward a reliable implementation of such an adaptive wavelet packet scheme, it would be interesting to study, in a simple model case, the computational effectiveness of the compression procedures presented in Section 5.7, by comparing the numerical results with the theoretical estimates (5.7.1) and (5.7.6).

Index

BIBLIOGRAPHY

[1] R. A. Adams. *Sobolev spaces.* Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.

[2] M.G. Ancona and G.J. Iafrate. Quantum correction to the equation of state of an electron gas in a semiconductor. *Phys. Rev. B*, 39, 1989.

[3] L. Anderson, N. Hall, B. Jawerth, and G. Peters. Wavelets on closed subsets on the real line. In L.L. Schumaker and G. Webb eds, editors, *Topics in the Theory and Applications of Wavelets*, pages 1–61, Boston, MA, 1993. Academic Press.

[4] I. Babuska. A posteriori error estimates for the finite element methods. *J. Numer. Methods Engrg.*

[5] R. E. Bank and D. J. Rose. Analysis of a multilevel iterative method for nonlinear finite element equations. *Math. Comp.*, 39(160):453–465, 1982.

[6] R. E. Bank and A. Weiser. Some a posteriori error estimators for elliptic partial differential equations. *Math. Comp.*, 44(170):283–301, 1985.

[7] R.E. Bank and R.K. Smith. A posteriori error estimates based on hierarchical bases. *SIAM J. Numer. Anal.*, 30(4):921–935, 1993.

[8] C. Bennett and R. Sharpley. *Interpolation of operators.* Academic Press Inc., Boston, MA, 1988.

[9] J. Bergh and J. Löfström. *Interpolyatsionnye prostranstva.* “Mir”, Moscow, 1980. Vvedenie. [An introduction], Translated from the English by V. S. Krjučkov and P. I. Lizorkin.

[10] S. Bertoluzza. A posteriori error estimates for the wavelet Galerkin method. *Appl. Math. Lett.*, 8(5):1–6, 1995.

[11] S. Bertoluzza. Adaptive wavelet collocation method for the solution of Burgers equation. In *Proceedings of the Second International Workshop on Nonlinear Kinetic Theories and Mathematical Aspects of Hyperbolic Systems (Sanremo, 1994)*, volume 25, pages 339–352, 1996.

[12] S. Bertoluzza. An adaptive collocation method based on interpolating wavelets. In *Multiscale wavelet methods for partial differential equations*, pages 109–135. Academic Press, San Diego, CA, 1997.

[13] S. Bertoluzza and S. Falletta. Building wavelets on]0,1[at large scales. *Publ. IAN 1212*, 2001.

[14] S. Bertoluzza, Urban K., and M. Verani. Wavelets and adaptivity for non-conforming domain decomposition. in preparation.

[15] S. Bertoluzza and A. Kunoth. Wavelet stabilization and preconditioning for domain decomposition. *IMA Jour. Numer. Anal.*, 20, 2000.

[16] S. Bertoluzza, Y. Maday, and J.-C. Ravel. A dynamically adaptive wavelet method for solving partial differential equations. *Comput. Methods Appl. Mech. Engrg.*, 116(1-4):293–299, 1994. ICOSAHOM'92 (Montpellier, 1992).

[17] S. Bertoluzza, S. Mazet, and M. Verani. A non linear Richardson algorithm for the solution of elliptic PDE's. Technical Report 1227, Istituto di Analisi Numerica del Consiglio Nazionale delle Ricerche, Pavia, 2001.

[18] S. Bertoluzza and P. Pietra. Space- frequency adaptive approximation for quantum hydrodynamic models. *Trans. Theory and Stat. Phys*, 29, 2000.

[19] S. Bertoluzza and M Verani. Convergence of a non-linear wavelet algorithm for the solution of PDE's. *Appl. Math. Lett.*, 2001. to appear.

[20] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms. I. *Comm. Pure Appl. Math.*, 44(2):141–183, 1991.

[21] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element method with convergence rates. Technical report, IGPM-RWTH, 2002.

[22] A. Brandt. Multi-level adaptive solutions to boundary-value problems. *Math. Comp.*, 31(138):333–390, 1977.

[23] F. Brezzi and D. Marini. Macro hybrid elements and domain decomposition methods. In *Optimisation et controle*, pages 89–96. Cepadues Editions, 1993. Actes du colloque organise en l'honneur du soixantieme anniversaire de Jean Céa, Sophia-Antipolis, France, Avril 8-10, 1992, Toulouse.

[24] C. Canuto and A. Tabacco. Multilevel decomposition of functional spaces. *J. Fourier Anal. and Appl.*, 3:715–742, 1997.

[25] C. Canuto, A. Tabacco, and K. Urban. The wavelet element method. part i: Construction and analysis. *Appl. Comp. Harm. Anal.*, 6:1–52, 1999.

[26] G. Chiavassa and J. Liandrat. On the effective construction of compactly supported wavelets satisfying homogeneous boundary conditions on the interval. *ACHA*, 4:pp. 62–73, 1997.

[27] A. Cohen. Wavelet methods in numerical analysis. In *Handbook of numerical analysis, Vol. VII*, pages 417–711. North-Holland, Amsterdam, 2000.

[28] A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet methods for elliptic operator equations: convergence rates. *Math. Comp.*, 70(233):27–75 (electronic), 2001.

[29] A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet schemes for nonlinear variational problems. Technical report, RWTH Aachen, 2002.

[30] A. Cohen, W. Dahmen, and R. DeVore. Sparse evaluation of compositions of functions using multiscale expansions. Technical report, RWTH Aachen, 2002.

[31] A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet methods ii - beyond the elliptic case. Technical report, RWTH Aachen, 200o.

[32] A. Cohen, I. Daubechies, and J.C. Feauveau. Biorthogonal bases of compactly supported wavelets. *Comm. Pure Appl. Math.*, 45(5):485–560, 1992.

[33] A. Cohen, I. Daubechies, and P. Vial. Wavelets on the interval and fast wavelet transforms. *ACHA*, 1, 1993.

[34] A. Cohen and R. Masson. Adaptive wavelet methods for second order elliptic problems, boundary conditions and domain decomposition. Technical report, Laboratoire d'Analyse Numérique, Univ. Paris VI, 1998.

[35] R.R. Coifman and M.V. Wickerhauser. Entropy based algorithms for best basis selection. *IEEE Trans. Inform. Theory*, 1992.

[36] S. Dahlke, W. Dahmen, R. Hochmuth, and R. Schneider. Stable multiscale bases and local error estimation for elliptic problems. *Appl. Numer. Math.*, 23(1):21–47, 1997. Multilevel methods (Oberwolfach, 1995).

[37] S. Dahlke and R. DeVore. Besov regularity for elliptic boundary value problems. *Comm. Partial Differential Equations*, 22(1-2):1–16, 1997.

[38] W. Dahmen. Stability of multiscale transformations. *J. Fourier Anal. Appl.*, 2(4):341–361, 1996.

[39] W. Dahmen. Stability of multiscale transformations. *Journal of Fourier Analysis and Applications*, 2(4):341–361, 1996.

[40] W. Dahmen, A. Kunoth, and K. Urban. Biorthogonal spline wavelets on the interval - stability and moment conditions. *Appl. Comput. Harmon. Anal.*, 6(2):132–196, 1999.

[41] W. Dahmen, S. Proessdorf, and R. Schneider. Wavelet approximation methods for pseudodifferential operators: I stability and convergence. *Math. Z.*, 1994.

[42] W. Dahmen and R. Schneider. Composite wavelet bases for operator equations. Preprint, TU Chemnitz, 1997. to appear in Math. Comp.

[43] W. Dahmen and R. Schneider. Wavelets on manifolds i: Construction and domain decomposition. Report 149, IGPM, 1998.

[44] W. Dahmen, R. Schneider, and Y. Xu. Nonlinear functionals of wavelet expansions—adaptive reconstruction and fast evaluation. *Numer. Math.*, 86(1):49–101, 2000.

[45] I. Daubechies. *Ten lectures on wavelets.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[46] R. S. Dembo, S. C. Eisenstat, and T. Steihaug. Inexact Newton methods. *SIAM J. Numer. Anal.*, 19(2):400–408, 1982.

[47] R. A. DeVore. Nonlinear approximation. In *Acta numerica, 1998*, pages 51–150. Cambridge Univ. Press, Cambridge, 1998.

[48] R. A. DeVore, B. Jawerth, and V. Popov. Compression of wavelet decompositions. *Amer. J. Math.*, 114(4):737–785, 1992.

[49] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, 33(3):1106–1124, 1996.

[50] J. Fröhlich and K. Schneider. An adaptive wavelet Galerkin algorithm for one- and two-dimensional flame computations. *European J. Mech. B Fluids*, 13(4):439–471, 1994.

[51] C.L. Gardner. The quantum hydrodynamic model for semiconductor devices. *SIAM J. Appl. Math.*, 54(2):409–427, 1994.

[52] C.L. Gardner and C. Ringhofer. Smooth quantum hydrodynamic model simulation of the resonant tunneling diode. *VLSI Design*, 1998.

[53] I. Gasser and P. A. Markowich. Quantum hydrodynamics, Wigner transforms and the classical limit. *Asymptot. Anal.*, 14(2):97–116, 1997.

[54] E. Grenier. Limite semi-classique de l'équation de Schrödinger non linéaire en temps petit. *C. R. Acad. Sci. Paris Sér. I Math.*, 320(6):691–694, 1995.

[55] P. Grisvard. *Elliptic problems in nonsmooth domains.* Pitman (Advanced Publishing Program), Boston, MA, 1985.

[56] S. Grivet Talocia and A. Tabacco. Wavelets on the interval with optimal localization. *Math. Mod. Meth. Appl. Sci.*, 10:441 – 462, 2000.

[57] M. T. Gyi and A. Jüngel. A quantum regularization of the one-dimensional hydrodynamic model for semiconductors. *Adv. Differential Equations*, 5(4-6):773–800, 2000.

[58] A. Harten. Multiresolution algorithms for the numerical solution of hyperbolic conservation laws. *Comm. Pure Appl. Math.*, 48(12):1305–1342, 1995.

[59] S. Jaffard. Wavelet methods for fast resolution of elliptic problems. *SIAM J. Numer. Anal.*, 29(4):965–986, 1992.

[60] S. Jaffard. Analyse par ondelettes d'un problème elliptique singulier. *J. Math. Pures Appl. (9)*, 72(2):121–143, 1993.

[61] D. Jerison and C. E. Kenig. The inhomogeneous Dirichlet problem in Lipschitz domains. *J. Funct. Anal.*, 130(1):161–219, 1995.

[62] H. Johnen and K. Scherer. On the equivalence of the K-functional and moduli of continuity and some applications. In *Constructive theory of functions of several variables (Proc. Conf., Math. Res. Inst., Oberwolfach, 1976)*, pages 119–140. Lecture Notes in Math., Vol. 571. Springer, Berlin, 1977.

[63] P. Joly, Y. Maday, and V. Perrier. A dynamical adaptive concept based on wavelet packet best bases: application to convection diffusion partial differential equations. In *Multiscale wavelet methods for partial differential equations*, pages 199–235. Academic Press, San Diego, CA, 1997.

[64] L. V. Kantorovich and G. P. Akilov. *Functional analysis in normed spaces.* The Macmillan Co., New York, 1964.

[65] J. Liandrat, V. Perrier, and Ph. Tchamitchian. Numerical resolution of the regularized Burgers equation using the wavelet transform. In *Wavelets and applications (Marseille, 1989)*, pages 420–433. Masson, Paris, 1992.

[66] J.-L. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications.* Springer-Verlag, New York, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.

[67] Y. Maday, V. Perrier, and J.-C. Ravel. Adaptativité dynamique sur bases d'ondelettes pour l'approximation d'équations aux dérivées partielles. *C. R. Acad. Sci. Paris Sér. I Math.*, 312(5):405–410, 1991.

[68] Y. Meyer. Ondelettes et Opérateurs. Hermann, 1990.

[69] Y. Meyer. *Ondelettes, Vol I, Ondelettes et Opèrateur.* Hermann, Paris, 1990.

[70] P. Monasse and V. Perrier. Construction d'ondelettes sur l'intervalle pour la prise en compte de conditions aux limites. *C. R. Acad. Sci. Paris,t.321*, 1:pp. 1163–1169, 1995.

[71] P. Morin, R.H. Nochetto, and K.G. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, 38(2):466–488 (electronic), 2000.

[72] J. M. Ortega and W. C. Rheinboldt. *Iterative solution of nonlinear equations in several variables.* Academic Press, New York, 1970.

[73] J. Peetre. *New thoughts on Besov spaces.* Mathematics Department, Duke University, Durham, N.C., 1976. Duke University Mathematics Series, No. 1.

[74] P. Petrushev. Direct and converse theorems for spline and rational approximation and Besov spaces. In *Function spaces and applications (Lund, 1986)*, pages 363–377. Springer, Berlin, 1988.

[75] P. Pietra and C Pohl. Weak limits of the quantum hydrodynamic model. *IEEE VLSI Design*, 1999.

[76] E. M. Stein. *Singular integrals and differentiability properties of functions.* Princeton University Press, Princeton, N.J., 1970.

[77] V. N. Temlyakov. The best m-term approximation and greedy algorithms. *Adv. Comput. Math.*, 8(3):249–265, 1998.

[78] H. Triebel. *Theory of function spaces.* Birkhäuser Verlag, Basel, 1983.

[79] M. Verani. An adaptive wavelet method for nonlinear equations based on an adaptive solver for linear equations. Manuscript.

[80] M. Verani. Metodi wavelet non-lineari per la soluzione di equazioni differenziali. Master's thesis, University of Pavia, 1998.

[81] M. Verani. An adaptive newton scheme based on nonlinear wavelet approximation. Technical report, I.A.N., Pavia, 2002.

[82] M. Verani. A wavelet adaptive newton method for the solution of nonlinear equations. *Appl. Math. Lett.*, 2002. to appear.

[83] R. Verfurth. *A Review of a Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques.* Teubner Verlag and J. Wiley, 1996.

[84] M.V. Wickerhauser, M. Farge, and E. Goirand. Theoretical dimension and the complexity of simulated turbulence. In *Multiscale wavelet methods for partial differential equations*, pages 473–492. Academic Press, San Diego, CA, 1997.

[85] T. J. Ypma. Local convergence of inexact Newton methods. *SIAM J. Numer. Anal.*, 21(3):583–590, 1984.

[86] H. Yserentant. On the multi-level splitting of finite element spaces. *Numer. Math.*, 1986.

[87] O. C. Zienkiewicz, D. W. Kelly, and I. Gago, J.and Babuska. Hierarchical finite element approaches, error estimates and adaptive refinements. In *The mathematics of finite elements and applications.* Academic Press, New York, 1982.

Acknowledgements

That Gandalf should be late, does not bode well. But it is said: "Do not meddle in the affairs of Wizards, for they are subtle." The choice is yours: to go or wait.

"But I don't think you need to go alone" exclaimed Gandalf. "Not if you know of anyone you can trust, and who would be willing to go by your side – and that you would be willing to take into unknown perils. But if you look for a companion, be careful in choosing! And be careful of what you say, even to your closest friends! The enemy has many spies and many ways of hearing."

At last the companions turned away, and never again looking back they rode slowly homewards; and they spoke no word to one other until they came back to the Shire, but each had great comfort in his friends on the long grey road.

In a chair, at the far side of the room facing the outer door, sat a woman. Her long black hair rippled down her shoulders. "Come dear folk!" she said, taking Frodo by the hand. "Laugh and be merry! I am Goldberry, daughter of the River."

"For you" she said to Sam, "I have only a small gift." She put into his hand a little box of plain grey wood, unadorned save for a single silver rune upon the lid. "In this box there is earth from my orchard. Though you should find all barren and laid waste, there will be few gardens in Middle-earth that will bloom like your garden, if you sprinkle this earth there. Then you may remember Galadriel."

But Sam turned to Bywater, and so came back up the Hill, as the day was ending once more. And he went on, and there was yellow light, and fire within; and he was expected. He drew a deep breath. "Well, I'm back," he said.