WAVELET-VAGUELETTE DECOMPOSITIONS AND HOMOGENEOUS EQUATIONS

A Thesis
Submitted to the Faculty
of
Purdue University
by
Namyong Lee

In Partial Fulfillment of the
Requirements for the Degree
of
Doctor of Philosophy
December 1997

To my family

ACKNOWLEDGEMENTS

I wish to extend my most sincere gratitude and appreciation to my thesis advisor Bradley J. Lucier for his invaluable suggestions, encouragement, and patience throughout the writing of this thesis. I am truly fortunate to have had the opportunity to build the foundations of my mathematical career under his guidance.

I am grateful to Professors Rodrigo Bañuelos, Tony Cai, and Charles Bouman for their willingness to serve on my thesis committee.

I would also like to thank Youngjoon Cha, Darrel Czubala, Miyoung Lee, Seongan Lim, Heon Joo Kim, Roy Osawa, Dongwoo Sheen, Chang Hyun Song, and Ikwon Yie for their friendship. I am glad that we were all afforded the opportunity to get to know each other while at Purdue.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Lee, Namyong. Ph.D., Purdue University, December 1997. Wavelet-Vaguelette Decompositions and Homogeneous Equations. Major Professor: Bradley J. Lucier.

We describe the wavelet-vaguelette decomposition (WVD) for solving a homogeneous equation $Y = Af + Z$, where A satisfies $\widehat{A^*Af}(\xi) = |\xi|^{-2\alpha}\widehat{f}(\xi)$ for some $\alpha \geq 0$. We find a sufficient condition on functions to have a WVD. This result generalizes Daubechies's work on the discrete wavelet transform. We examine the relation between the WVD-based method and variational problems for solving a homogeneous equation. Algorithms are derived as exact minimizers of variational problems of the form; given observed function Y, minimize over all g in the Besov space $B^{\beta_0}_{1,1}(R^d)$ the functional $\|Y - Ag\|^2_{\mathcal{Y}} + 2\gamma|g|_{B^{\beta_0}_{1,1}}$, where $\mathcal{Y}$ is a separable Hilbert space. We use the theory of nonlinear wavelet approximation in $L^2(R^d)$ to derive accurate error bounds for recovering f through wavelet shrinkage applied to observed data Y corrupted with independent and identically distributed mean zero Gaussian noise Z. We give a new proof of the rate of convergence of wavelet shrinkage that allows us to estimate rather sharply the best shrinkage parameter. We conduct tomographic reconstruction computations that support the hypothesis that near-optimal shrinkage parameters can be derived if one knows (or can estimate) only two parameters about a phantom image f: the largest β for which $f \in B^{\beta}_{p,p}(R^2)$, $p = \frac{3}{\beta+3/2}$, and the seminorm $|f|_{B^{\beta}_{p,p}}$. Both theoretical and experimental results indicate that our choice of shrinkage parameters yields uniformly better results than Kolaczyk's procedure and classical filtered backprojection method.

CHAPTER 1
INTRODUCTION

For a given linear operator $A : L^2(\mathbb{R}^d) \longrightarrow \mathcal{Y}$ between two separable Hilbert spaces $L^2(\mathbb{R}^d)$ and $\mathcal{Y}$, where throughout this thesis A always satisfies

$$\widehat{A^*Af}(\xi) = |\xi|^{-2\alpha}\widehat{f}(\xi) \tag{1.1}$$

for some $\alpha \geq 0$, we wish to recover $f \in L^2(\mathbb{R}^d)$ from Af. In practice we are only able to observe noisy data of the form

$$Y = Af + Z, \tag{1.2}$$

where Z represents a perturbed data error in our observation procedure. We assume that Z is a mean zero Gaussian noise.

Such linear inverse problems arise in various scientific fields. The noise removal problem in image processing takes the form of $Y = f + Z$. In this case the underlying operator is the identity operator, which obviously satisfies (1.1) with $\alpha = 0$. The Radon transform, which plays an important role in modeling the measurement procedure of medical imaging, satisfies (1.1) with $\alpha = 1/2$. On the other hand, the 2π multiple of the one-dimensional integration operator also satisfies (1.1) with $\alpha = 1$.

Despite the simple description (1.1) of A via the Fourier transform, linear filtering methods based on (1.1) such as *filtered backprojection* often exhibit degradation in recovering f from noisy data. The Fourier transform diagonalizes any convolution-type operator, and this property has been an advantage of Fourier transform methods in deconvolution problems such as (1.2). However, poor representation of nonsmooth function via the Fourier transform often yields an unacceptable decision rule, which considers any low-frequency structure to be information and any high-frequency structure to be noise, in recovering f from (1.2). This is largely due to the lack of *smoothness characterization* in the Fourier transform.

Very recently, there has been great interest in the use of wavelet bases to represent functions, and many important advantages of wavelet bases have been discovered. One of the key features of wavelet bases is smoothness characterization for various function spaces in terms of wavelet coefficients. This property has been used in several image processing and statistical applications, such as data compression [13], [14], [33], noise removal [4], [17], and nonparametric estimation in statistics [3], [23], [25], [26], etc.

Even though the convolution-type operator, in general, is not diagonalizable with respect to wavelet bases, the operator satisfying (1.1) can be diagonalized with the help of a *wavelet-vaguelette system*. In [22] Donoho proved that there exists a wavelet-vaguelette system $\{\psi_\lambda, U_\lambda, \widetilde{U}_\lambda\}$ such that $\{\psi_\lambda\}$ is an orthogonal wavelet basis of $L^2(\mathbb{R}^d)$, $\{U_\lambda\}$ and $\{\widetilde{U}_\lambda\}$ are biorthogonal Riesz bases of $\mathcal{Y}$, and for any $f \in L^2(\mathbb{R}^d)$,

$$f = \sum_\lambda c_\lambda [Af, U_\lambda]\psi_\lambda \tag{1.3}$$

for known scalars $\{c_\lambda\}$, where $[\cdot,\cdot]$ is the inner product of $\mathcal{Y}$.

Applying the *wavelet-vaguelette decomposition* (1.3) in solving (1.2), Donoho suggested the *wavelet shrinkage* method (see, e.g., [22], [23], and [31]), which shrinks $c_\lambda[Y, U_\lambda]$ towards zero by a certain amount, and proved that if the true solution f is known to lie in Besov space $B^\beta_{q,p}(\mathbb{R}^d)$ (for the definition, see [15] and Theorem 3.2.1), where $\beta > (2\alpha + d)(1/p - 1/2)$, then the suggested method converges to f with the optimal rate if one uses the optimal shrinkage parameter (although Donoho did not give a method for finding that parameter). For details, see [22] or Theorem 9.1.1 in this thesis.

In this thesis we consider a family of variational problems for solving (1.2). These variational problems take the form: Given a positive parameter γ and a function space B, find a function $\widetilde{f}$ that minimizes over all possible functions g in B the functional

$$\|Y - Ag\|_{\mathcal{Y}}^2 + 2\gamma |g|_B, \tag{1.4}$$

where $\|\cdot\|_{\mathcal{Y}}$ is the norm defined by the inner product $[\cdot,\cdot]$ of $\mathcal{Y}$ and $|\cdot|_B$ is the semi-norm of B.

The function space B differs from usual smoothness subspaces in regularization techniques (see, e.g., [42]), because B could be a function space which is not *embedded* in $L^2(\mathbb{R}^d)$. The parameter γ balances importances between the difference $\|Y - Ag\|_{\mathcal{Y}}^2$ in $\mathcal{Y}$ and the smoothness $|g|_B$ of g in B.

A fast way of solving (1.4) is required for practical algorithms. We suggest using vaguelettes and wavelets in (1.4), to characterize $\|Y - Ag\|_{\mathcal{Y}}$ and $|g|_B$, respectively. Using L^2-*stability* of vaguelettes and smoothness characterization of wavelets, we derive an equivalent expression involving wavelet coefficients to (1.4), which can be minimized quickly. In particular, if we choose $B = B^{\beta_0}_{1,1}(\mathbb{R}^d)$ in (1.4), then the exact minimizer, denoted by $\widetilde{f}^*_{\gamma,\beta_0}$, of the resulting equivalent sequence minimization problem has the same form as the wavelet shrinkage method proposed by Donoho [22] with possibly different shrinkage parameters.

We prove a sufficient condition on regularity and *vanishing moments* (for the definition, see Chapter 3) of functions to have a wavelet-vaguelette system. This sufficient condition is weaker than in [22] (see Theorem 6.1.4 and Theorem 6.3.1), and allows more wavelets such as Daubechies's compactly supported orthogonal wavelets [8] and symmetric biorthogonal wavelets [5] to be used for solving (1.2). These compactly supported wavelets are generated from *scaling functions* by *refinement equations* (see (4) of Definition 3.1.2); thus vaguelette coefficients can be obtained in a recursive manner as wavelet coefficients can be.

We give a new proof of the rate of convergence of $\widetilde{f}^*_{\gamma,\beta_0}$ that allows us to estimate rather sharply the best wavelet shrinkage parameter for solving (1.2) in the presence of Gaussian noise. Our analysis reveals that the parameter β_0 depends only on the ill-posedness of (1.2), and for given β_0 the parameter $\gamma 2^{k(\beta_0 - d/2 + 2\alpha)}$ represents the wavelet shrinkage amount.

We show through tomographic reconstruction experiments that with a *rotational averaging technique* (see Chapter 10) the wavelet shrinkage method $\widetilde{f}^*_{\gamma,\beta_0}$, where parameters γ and β_0 are determined from our analysis, leads to a better reconstruction in the presence of Gaussian noise than the traditional filtered backprojection method. As compared with the shrinkage parameter suggested by Kolaczyk [31], which is motivated by the VisuShrink method by Donoho and Johnstone, our shrinkage parameter leads to a better reconstructioni, which removes less noise, but keeps more image features.

CHAPTER 2
PRELIMINARIES

In this chapter we review some notations, definitions, theorems, and inequalities that are needed in this thesis.

2.1 Definitions and Notations

The *translation* operator T_h: For $h \in \mathbb{R}^d$,

$$T_h f(x) = f(x-h).$$

The *dilation* operator D_a: For $a > 0$,

$$D_a f(x) = f(x/a).$$

The *convolution* operator T associated with h:

$$Tf(x) = \int_{\mathbb{R}^d} f(x-y)h(y)\,dy. \tag{2.1}$$

The *shrinkage* operator $S_\mu : \mathbb{R} \longrightarrow \mathbb{R}$: For $\mu \geq 0$,

$$S_\mu(x) = \begin{cases} x-\mu, & \text{if } x > \mu, \\ 0, & \text{if } |x| \leq \mu, \\ x+\mu, & \text{if } x < -\mu. \end{cases} \tag{2.2}$$

Let $\mathcal{H}$ be a separable Hilbert space. For real-valued functions S_1 and S_2 defined on $\mathcal{H}$, we denote $S_1(f) \asymp S_2(f)$ if there are positive constants C_1 and C_2 such that for all $f \in \mathcal{H}$

$$C_1 S_1(f) \leq S_2(f) \leq C_2 S_1(f).$$

A collection of functions $\{\varphi_n\}$ in $\mathcal{H}$ is said to be L^2-*stable* if

$$\|\varphi\|_{\mathcal{H}}^2 \asymp \sum_n |\langle \varphi, \varphi_n \rangle|^2.$$

An L^2-stable basis of $\mathcal{H}$ is also called a *Riesz basis.*

Let X_1 and X_2 be normed vector spaces. The space X_1 is said to be *embedded* in the space X_2, denoted by $X_1 \hookrightarrow X_2$, if for each $f \in X_1$, $f \in X_2$ and there is a constant C such that for all $f \in X_1$

$$\|f\|_{X_2} \leq C\|f\|_{X_1}.$$

The *Kronecker delta function* δ is defined by

$$\delta_{x,x'} = \begin{cases} 1, & \text{if } x = x', \\ 0, & \text{if } x \neq x'. \end{cases}$$

2.2 Theorems and Inequalities

We denote by $S(\mathbb{R}^d)$ the space of rapidly decreasing C^∞ functions on $\mathbb{R}^d$ and by $S'(\mathbb{R}^d)$ its topological dual, the space of tempered distributions. The Fourier transform $\widehat{f}$ of a function $f \in S(\mathbb{R}^d)$ is defined by

$$\widehat{f}(\xi) = \int_{\mathbb{R}^d} e^{-2\pi i\xi\cdot x} f(x)\, dx,$$

while the inverse Fourier transform gives f back from $\widehat{f}$ by

$$f(x) = \int_{\mathbb{R}^d} e^{2\pi i x\cdot\xi} \widehat{f}(\xi)\, d\xi.$$

One then, extends the Fourier transform and its inverse from $S(\mathbb{R}^d)$ to $S'(\mathbb{R}^d)$ by duality.

The Fourier transform has the following properties: For $f \in S(\mathbb{R}^d)$ and $g \in S'(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} f(x)\overline{g(x)}\, dx = \int_{\mathbb{R}^d} \widehat{f}(\xi)\overline{\widehat{g}(\xi)}\, d\xi, \tag{2.3}$$

$$\widehat{T_h g}(\xi) = e^{-2\pi i \xi \cdot h} \widehat{g}(\xi), \tag{2.4}$$

$$\widehat{D_a g}(\xi) = a^d \widehat{g}(a\xi), \tag{2.5}$$

$$\widehat{\frac{\partial f}{\partial x_k}}(\xi) = 2\pi i \xi_k \widehat{f}(\xi), \tag{2.6}$$

and

$$\frac{\partial \widehat{f}}{\partial \xi_k}(\xi) = \int_{\mathbb{R}^d} e^{-2\pi i \xi \cdot x}(-2\pi i x_k) f(x)\, dx. \tag{2.7}$$

In particular, by (2.4) and (2.5),

$$\widehat{g_{k,j}}(\xi) = 2^{-kd/2} e^{-2\pi i \xi \cdot j 2^{-k}} \widehat{g}(2^{-k}\xi), \tag{2.8}$$

if $g_{k,j} = 2^{kd/2} g(2^k \cdot - j)$.

If $p(x)$ be a function defined on $[0, L]^d$, then

$$L^{-d} \sum_{j \in \mathbb{Z}^d} \left| \int_{[0,L]^d} e^{2\pi i x \cdot j/L} p(x)\, dx \right|^2 = \int_{[0,L]^d} |p(x)|^2\, dx. \tag{2.9}$$

For the convolution operator T associated with h as (2.1),

$$\widehat{Tf}(\xi) = \widehat{h}(\xi) \widehat{f}(\xi). \tag{2.10}$$

Conversely, if T can be characterized as the right hand side of (2.10), then T is the convolution operator associated with h.

For a countable index set Λ and a sequence (a_n), $n \in \Lambda$, we define

$$\|(a_n)\|_{\ell^p} = \left(\sum_{n \in \Lambda} |a_n|^p \right)^{1/p}$$

and

$$\ell^p(\Lambda) = \{(a_n) : n \in \Lambda, \|(a_n)\|_{\ell^p} < \infty\}$$

for $0 < p \leq \infty$, where $\|(a_n)\|_{\ell^\infty}$ is understood as $\sup_n\{|a_n|\}$. We omit Λ in $\ell^p(\Lambda)$ whenever this does not make any confusion.

We define

$$\|f\|_{L^p} = \begin{cases} \left(\int_{\mathbb{R}^d} |f(x)|^p \, dx\right)^{1/p}, & \text{if } 0 < p < \infty, \\ \sup_x |f(x)|, & \text{if } p = \infty. \end{cases}$$

Throughout this thesis, $\sup_x |f(x)|$ is understood as the *essential supremum* of f (for the definition, see [39]). The space $L^p(\mathbb{R}^d)$ is the set of real-valued functions f with $\|f\|_{L^p} < \infty$.

Let $1 \leq p \leq \infty$. If $1/p + 1/p' = 1$, then

$$\left|\sum_n a_n b_n\right| \leq \|(a_n)\|_{\ell^p} \|(b_n)\|_{\ell^{p'}} \tag{2.11}$$

and

$$\left|\int_{\mathbb{R}^d} f(x)\overline{g(x)}dx\right| \leq \|f\|_{L^p} \|g\|_{L^{p'}}. \tag{2.12}$$

Let $0 < p \leq p' \leq \infty$. Then for any (a_n),

$$\left(\sum_n |a_n|^{p'}\right)^{1/p'} \leq \left(\sum_n |a_n|^p\right)^{1/p}. \tag{2.13}$$

CHAPTER 3
WAVELETS AND BESOV SPACES

In this chapter we briefly review basic wavelet theory. Throughout this thesis we consider only compactly supported wavelets such as Daubechies' orthogonal wavelets [8], symmetric biorthogonal wavelets by Cohen, Daubechies, and Feauveau [5] and Herley and Vetterli [28], and modified wavelets [6] and [30], which are designed to deal with functions defined on a bounded domain.

The main advantage of wavelets is smoothness characterization. It means that we can determine the membership of a function in many different function spaces by examining its wavelet coefficients. For details, see, e.g., [15], [20], [27], [30], [32], and [35]. We utilize this property to define Besov spaces.

3.1 Biorthogonal Wavelets

We begin with biorthogonal wavelets on $\mathbb{R}$.

Definition 3.1.1. *Let ψ and $\widetilde{\psi}$ be bounded and compactly supported functions on $\mathbb{R}$. We define $\psi_{k,j}(x) = 2^{k/2}\psi(2^k x - j)$, and similarly for $\widetilde{\psi}_{k,j}$. The collections of functions $\{\psi_{k,j}\}_{k\in\mathbb{Z},j\in\mathbb{Z}}$ and $\{\widetilde{\psi}_{k,j}\}_{k\in\mathbb{Z},j\in\mathbb{Z}}$ are called biorthogonal wavelets if they are biorthogonal Riesz bases of $L^2(\mathbb{R})$, i.e., for any $f \in L^2(\mathbb{R})$,*

(1)

$$f = \sum_{k\in\mathbb{Z}}\sum_{j\in\mathbb{Z}} \langle f, \widetilde{\psi}_{k,j}\rangle \psi_{k,j} = \sum_{k\in\mathbb{Z}}\sum_{j\in\mathbb{Z}} \langle f, \psi_{k,j}\rangle \widetilde{\psi}_{k,j},$$

(2)

$$\|f\|_{L^2}^2 \asymp \sum_{k\in\mathbb{Z}}\sum_{j\in\mathbb{Z}} \left|\langle f, \widetilde{\psi}_{k,j}\rangle\right|^2 \asymp \sum_{k\in\mathbb{Z}}\sum_{j\in\mathbb{Z}} |\langle f, \psi_{k,j}\rangle|^2,$$

(3) *they are biorthogonal, i.e., for all* (k, j) *and* (k', j'),

$$\langle \psi_{k,j}, \widetilde{\psi}_{k',j'} \rangle = \delta_{(k,j),(k',j')}.$$

For orthogonal wavelets, $\psi = \widetilde{\psi}$ and we have equality in (2) in Definition 3.1.1. The equations in (1) of Definition 3.1.1 are called *wavelet decompositions* of f.

While wavelet theory about the existence of ψ and $\widetilde{\psi}$ can be found in [5] and [10], we need only the so-called *fast wavelet transform* in this thesis. To explain it, we need a *scaling function* ϕ and a *dual scaling function* $\widetilde{\phi}$.

Definition 3.1.2. *Let* ϕ *and* $\widetilde{\phi}$ *be bounded and compactly supported functions on* $\mathbb{R}$. *Let* $\phi_{k,j}(x) = 2^{k/2}\phi(2^k x - j)$, *and similarly for* $\widetilde{\phi}_{k,j}$. *We call* ϕ *a scaling function and* $\widetilde{\phi}$ *a dual scaling function of biorthogonal wavelets* $\{\psi_{k,j}\}$ *and* $\{\widetilde{\psi}_{k,j}\}$ *if*

(1) *for all* j *and* j',

$$\langle \phi_{k,j}, \widetilde{\phi}_{k,j'} \rangle = \delta_{j,j'},$$

(2) *for any* $f \in L^2(\mathbb{R}^d)$,

$$f = \sum_{k \geq k_0} \sum_{j \in \mathbb{Z}} \langle f, \widetilde{\psi}_{k,j} \rangle \psi_{k,j} + \sum_{l \in \mathbb{Z}} \langle f, \widetilde{\phi}_{k_0,l} \rangle \phi_{k_0,l},$$

(3)

$$\|f\|_{L^2}^2 \asymp \sum_{k \geq k_0} \sum_{j \in \mathbb{Z}} \left| \langle f, \widetilde{\psi}_{k,j}^{(i)} \rangle \right|^2 + \sum_{l \in \mathbb{Z}} \left| \langle f, \widetilde{\phi}_{k_0,l} \rangle \right|^2,$$

(4) *there exist finite sequences* (h_n) *and* $(\widetilde{h}_n)$ *such that* $\sum_n h_n \widetilde{h}_{n+2k} = \delta_{k,0}$ *and*

$$\phi = \sum_n h_n \phi_{1,n}, \qquad \widetilde{\phi} = \sum_n \widetilde{h}_n \widetilde{\phi}_{1,n},$$
$$\psi = \sum_n g_n \phi_{1,n}, \qquad \widetilde{\psi} = \sum_n \widetilde{g}_n \widetilde{\phi}_{1,n},$$

where $g_n = (-1)^{n+1}\widetilde{h}_{-n+1}$ *and* $\widetilde{g}_n = (-1)^{n+1} h_{-n+1}$.

For orthogonal wavelets, $\phi = \widetilde{\phi}$ and $\psi = \widetilde{\psi}$, and we have equality in (3) of Definition 3.1.2. We call (2) of Definition 3.1.2 an *inhomogeneous wavelet decomposition* of f. We also call four equations in (4) of Definition 3.1.2 *refinement equations* associated with ϕ, $\widetilde{\phi}$, ψ, and $\widetilde{\psi}$.

Notice that by (4) of Definition 3.1.2, one can have

$$\begin{aligned}\widetilde{\psi}_{k,j}(x) &= 2^{k/2}\widetilde{\psi}(2^k x - j) \\ &= 2^{k/2}\sum_n \widetilde{g}_n 2^{1/2}\widetilde{\phi}(2(2^k x - j) - n).\end{aligned}$$

Thus we have

$$\begin{aligned}\widetilde{\psi}_{k,j}(x) &= \sum_n \widetilde{g}_n 2^{(k+1)/2}\widetilde{\phi}(2^{k+1}x - (n+2j)) \\ &= \sum_n \widetilde{g}_n \widetilde{\phi}_{k+1,n+2j} \\ &= \sum_n \widetilde{g}_{n-2j}\widetilde{\phi}_{k+1,n},\end{aligned} \tag{3.1}$$

and similarly

$$\widetilde{\phi}_{k,j}(x) = \sum_n \widetilde{h}_{n-2j}\widetilde{\phi}_{k+1,n}. \tag{3.2}$$

With (3.1) and (3.2), we have

$$\langle f, \widetilde{\psi}_{k,j}\rangle = \sum_n \widetilde{g}_{n-2j}\langle f, \widetilde{\phi}_{k+1,n}\rangle \tag{3.3}$$

and

$$\langle f, \widetilde{\phi}_{k,j}\rangle = \sum_n \widetilde{h}_{n-2j}\langle f, \widetilde{\phi}_{k+1,n}\rangle. \tag{3.4}$$

Using (3.3) and (3.4) successively, we can get $\{\langle f, \widetilde{\psi}_{k,j}\rangle\}_{\{k_0 \le k < m, j\}}$ and $\{\langle f, \widetilde{\phi}_{k_0,n}\rangle\}_n$ from $\{f, \widetilde{\phi}_{m,l}\}_l$. This is the *fast wavelet transform* associated with $\{\widetilde{\psi}_{k,j}\}$.

The fast wavelet transform is reversible. To show this, we need the following lemma.

Lemma 3.1.3. *$\langle \widetilde{\phi}_{m,l}, \psi_{k,j} \rangle = 0$ if $k \geq m$.*

Proof. We represent $\widetilde{\phi}_{m,l}$ as

$$\widetilde{\phi}_{m,l} = \sum_{k \geq m} \sum_j \langle \widetilde{\phi}_{m,l}, \psi_{k,j} \rangle \widetilde{\psi}_{k,j} + \sum_n \langle \widetilde{\phi}_{m,l}, \phi_{m,n} \rangle \widetilde{\phi}_{m,n}$$

by using (2) of Definition 3.1.2 with $k_0 = m$. (Here we switched roles of (ϕ, ψ) and $(\widetilde{\phi}, \widetilde{\psi})$.) By (1) of Definition 3.1.2, we have

$$\widetilde{\phi}_{m,l} = \sum_{k \geq m} \sum_j \langle \widetilde{\phi}_{m,l}, \psi_{k,j} \rangle \widetilde{\psi}_{k,j} + \widetilde{\phi}_{m,l}.$$

Thus

$$\sum_{k \geq m} \sum_j \langle \widetilde{\phi}_{m,l}, \psi_{k,j} \rangle \widetilde{\psi}_{k,j} = 0.$$

Since $\{\widetilde{\psi}_{k,j}\}$ is a Riesz basis of $L^2(\mathbb{R})$, $\langle \widetilde{\phi}_{m,l}, \psi_{k,j} \rangle = 0$ if $k \geq m$. ∎

Notice that for any $f \in L^2(\mathbb{R})$, using (2) of Definition 3.1.2 with $k_0 = m - 1$, we can compute $\langle f, \widetilde{\phi}_{m,l} \rangle$ by

$$\begin{aligned} &\left\langle \sum_{k \geq m-1} \sum_j \langle f, \widetilde{\psi}_{k,j} \rangle \psi_{k,j} + \sum_n \langle f, \widetilde{\phi}_{m-1,n} \rangle \phi_{m-1,n} \ , \ \widetilde{\phi}_{m,l} \right\rangle \\ &\quad = \sum_j \langle f, \widetilde{\psi}_{m-1,j} \rangle \langle \psi_{m-1,j}, \widetilde{\phi}_{m,l} \rangle + \sum_n \langle f, \widetilde{\phi}_{m-1,n} \rangle \langle \phi_{m-1,n}, \widetilde{\phi}_{m,l} \rangle, \end{aligned}$$

where we have used Lemma 3.1.3. With a similar argument used for $\widetilde{\psi}_{k,j}$ in (3.1), we have

$$\psi_{m-1,j} = \sum_{l'} g_{l'-2j} \phi_{m,l'} \qquad \text{and} \qquad \phi_{m-1,j} = \sum_{l'} h_{l'-2j} \phi_{m,l'}.$$

Thus $\langle f, \widetilde{\phi}_{m,l} \rangle$ is

$$\sum_j \langle f, \widetilde{\psi}_{m-1,j} \rangle \langle \sum_{l'} g_{l'-2j} \phi_{m,l'}, \widetilde{\phi}_{m,l} \rangle + \sum_n \langle f, \widetilde{\phi}_{m-1,n} \rangle \langle \sum_{l'} h_{l'-2n} \phi_{m,l'}, \widetilde{\phi}_{m,l} \rangle.$$

Therefore, using (1) of Definition 3.1.2, we have

$$\langle f, \widetilde{\phi}_{m,l}\rangle = \sum_j g_{l-2j}\langle f, \widetilde{\psi}_{m-1,j}\rangle + \sum_n h_{l-2n}\langle f, \widetilde{\phi}_{m-1,n}\rangle. \tag{3.5}$$

Thus we can get $\{\langle f, \widetilde{\phi}_{m,l}\rangle\}_l$ from $\{\langle f, \widetilde{\psi}_{k,j}\rangle\}_{\{k_0 \geq k < m, j\}}$ and $\{\langle f, \widetilde{\phi}_{k_0,n}\rangle\}_n$, by using (3.5) successively. This is the *inverse fast wavelet transform* associated with $\{\widetilde{\psi}_{k,j}\}$.

For biorthogonal wavelet basis on $L^2(\mathbb{R}^d)$ with $d > 1$, we use the tensor product of ψ, ϕ, $\widetilde{\psi}$, and $\widetilde{\phi}$. We shall explain it for $d = 2$. The general case will follow in the same manner. We define

$$\psi^{(1)}(x_1, x_2) = \phi(x_1)\psi(x_2),$$
$$\psi^{(2)}(x_1, x_2) = \psi(x_1)\phi(x_2),$$
$$\psi^{(3)}(x_1, x_2) = \psi(x_1)\psi(x_2),$$

and similarly for $\widetilde{\psi}^{(1)}$, $\widetilde{\psi}^{(2)}$, and $\widetilde{\psi}^{(3)}$. We define $\psi^{(i)}_{k,j} = 2^k\psi^{(i)}(2^k \cdot - j)$, and similarly for $\widetilde{\psi}^{(i)}_{k,j}$. Then sets of functions $\{\psi^{(i)}_{k,j}\}$ and $\{\widetilde{\psi}^{(i)}_{k,j}\}$ form biorthogonal Riesz bases of $L^2(\mathbb{R}^2)$. Thus any function $f \in L^2(\mathbb{R}^2)$ can be written as

$$f = \sum_{k\in\mathbb{Z}} \sum_{j\in\mathbb{Z}^2} \sum_{i=1,2,3} \langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle \psi^{(i)}_{k,j} = \sum_{k\in\mathbb{Z}} \sum_{j\in\mathbb{Z}^2} \sum_{i=1,2,3} \langle f, \psi^{(i)}_{k,j}\rangle \widetilde{\psi}^{(i)}_{k,j}.$$

We define Φ by

$$\Phi(x_1, x_2) = \phi(x_1)\phi(x_2),$$

similarly for $\widetilde{\Phi}$. As we did for $\psi^{(i)}_{k,j}$, we define $\Phi_{k,j}$ and $\widetilde{\Phi}_{k,j}$. The two-dimensional fast wavelet transform and its inverse transform associated with $\{\widetilde{\psi}^{(i)}_{k,j}\}$ can be obtained by the same methods used for one-dimensional ones.

Lemma 3.1.4. *Let* $j=(j_1,j_2)$ *and* $n=(n_1,n_2)$.
Fast wavelet transform:

$$\begin{aligned}\langle f,\widetilde{\Phi}_{k,j}\rangle &= \sum_{n_1,n_2}\widetilde{h}_{n_1-2j_1}\widetilde{h}_{n_2-2j_2}\langle f,\widetilde{\Phi}_{k+1,n}\rangle,\\ \langle f,\widetilde{\psi}^{(1)}_{k,j}\rangle &= \sum_{n_1,n_2}\widetilde{h}_{n_1-2j_1}\widetilde{g}_{n_2-2j_2}\langle f,\widetilde{\Phi}_{k+1,n}\rangle,\\ \langle f,\widetilde{\psi}^{(2)}_{k,j}\rangle &= \sum_{n_1,n_2}\widetilde{g}_{n_1-2j_1}\widetilde{h}_{n_2-2j_2}\langle f,\widetilde{\Phi}_{k+1,n}\rangle,\\ \langle f,\widetilde{\psi}^{(3)}_{k,j}\rangle &= \sum_{n_1,n_2}\widetilde{g}_{n_1-2j_1}\widetilde{g}_{n_2-2j_2}\langle f,\widetilde{\Phi}_{k+1,n}\rangle.\end{aligned}$$

Inverse fast wavelet transform:

$$\begin{aligned}\langle f,\widetilde{\Phi}_{k+1,j}\rangle =&\sum_{n_1}\widetilde{h}_{j_1-2n_1}\left(\sum_{n_2}\widetilde{h}_{j_2-2n_2}\langle f,\widetilde{\Phi}_{k,n}\rangle+\sum_{n_2}\widetilde{g}_{j_2-2n_2}\langle f,\widetilde{\psi}^{(1)}_{k,n}\rangle\right)\\ &+\sum_{n_1}\widetilde{g}_{j_1-2n_1}\left(\sum_{n_2}\widetilde{h}_{j_2-2n_2}\langle f,\widetilde{\psi}^{(2)}_{k,n}\rangle+\sum_{n_2}\widetilde{g}_{j_2-2n_2}\langle f,\widetilde{\psi}^{(3)}_{k,n}\rangle\right).\end{aligned}$$

For $L^2(\mathbb{R}^d)$, we make 2^d functions by

$$\varphi_1(x_1)\cdot\varphi_2(x_2)\cdots\varphi_d(x_d),\qquad \text{either}\quad \varphi_i=\phi\ \text{ or }\ \varphi_i=\psi. \tag{3.6}$$

Among them, we denote $\phi(x_1)\cdots\phi(x_d)$ by Φ, and remaining 2^d-1 functions by $\psi^{(i)}$ with $i=1,2,\cdots,2^d-1$. With $\widetilde{\phi}$ and $\widetilde{\psi}$, we can have $\widetilde{\Phi}$ and $\widetilde{\psi}^{(i)}$ for $i=1,2,\cdots,2^d-1$ with the same method used in (3.6). We define $\Phi_{k,j}=2^{kd/2}\Phi(2^k\cdot-j)$, and similarly for $\psi^{(i)}_{k,j}$, $\widetilde{\Phi}_{k,j}$, and $\widetilde{\psi}^{(i)}_{k,j}$. Then any $f\in L^2(\mathbb{R}^d)$ can be written as

$$f=\sum_{k\ge k_0}\sum_{j\in\mathbb{Z}^d}\sum_{i=1,\cdots,2^d-1}\langle f,\widetilde{\psi}^{(i)}_{k,j}\rangle\psi^{(i)}_{k,j}+\sum_{l\in\mathbb{Z}^d}\langle f,\widetilde{\Phi}_{k_0,l}\rangle\Phi_{k_0,l}. \tag{3.7}$$

Moreover, we have

$$\|f\|^2_{L^2}\asymp\sum_{k\ge k_0}\sum_{j\in\mathbb{Z}^d}\sum_{i=1,\cdots,2^d-1}\left|\langle f,\widetilde{\psi}^{(i)}_{k,j}\rangle\right|^2+\sum_{l\in\mathbb{Z}^d}\left|\langle f,\widetilde{\Phi}_{k_0,l}\rangle\right|^2. \tag{3.8}$$

When one is concerned with a bounded domain, for example, the unit cube Ω in $\mathbb{R}^d$, then one does not consider all shifts $j \in \mathbb{Z}^d$, but only those shifts for which $\widetilde{\psi}^{(i)}_{k,j}$ intersects Ω nontrivially. Moreover, one must adapt wavelets that overlap the boundary of Ω to preserve L^2-stability on the domain. For details, see, e.g., [6] and [30]. To ignore all complication of this sort, we shall use indices without precisely specifying the domains of the indices of the sums whenever this abbreviation does not cause any confusion.

Remark 3.1.5. If f is a compactly supported function, then the number of j for which $\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle \neq 0$ is less than $C2^{kd}$ for fixed k and i, and the number of l for which $\langle f, \widetilde{\Phi}_{k_0,l}\rangle \neq 0$ is less than $C2^{k_0 d}$ in (3.7) for a constant C, since all $\widetilde{\psi}^{(i)}$ and $\widetilde{\Phi}$ are compactly supported.

3.2 Besov Spaces

For any $h \in \mathbb{R}^d$ and $f : \mathbb{R}^d \to \mathbb{R}$, we define $\triangle^0_h f(x) = f(x)$ and

$$\triangle^{k+1}_h f(x) = \triangle^k_h f(x+h) - \triangle^k_h f(x)$$

for $k = 0, 1, \cdots$.

Let $\mathcal{X}$ be a quasi-normed, complete, and translation-invariant space of real-valued functions on $\mathbb{R}^d$. For a positive integer r, the r-th modulus of smoothness of a function $f \in \mathcal{X}$ is defined by

$$\omega_r(f,t)_{\mathcal{X}} = \sup_{|h|\leq t} \|\triangle^r_h f(\cdot)\|_{\mathcal{X}}.$$

DEFINITION 3.2.1. *Let $\mathcal{X}$ be a quasi-normed, complete, and translation-invariant space of real-valued functions on $\mathbb{R}^d$. Let also $0 < \beta < \infty$, $0 < q \leq \infty$, and r be a positive integer such that $r-1 \leq \beta < r$. The Besov space $B^\beta_q(\mathcal{X})$ is the set of functions f such that their Besov space norm*

$$\|f\|_{B^\beta_q(\mathcal{X})} = \|f\|_{\mathcal{X}} + |f|_{B^\beta_q(\mathcal{X})} < \infty,$$

where the Besov space semi-norm is defined by

$$|f|_{B_q^\beta(\mathcal{X})} = \left(\int_0^\infty (t^{-\beta}\omega_r(f,t)_{\mathcal{X}})^q \frac{dt}{t} \right)^{1/q},$$

with an essential supremum when $q = \infty$.

In this thesis we are interested in Besov spaces $B_q^\beta(L^p(\mathbb{R}^d))$ for $1 < p \leq \infty$ and $B_q^\beta(H^p(\mathbb{R}^d))$ for $0 < p \leq 1$, where $H^p(\mathbb{R}^d)$ is the real Hardy space (for the definition, see, e.g., [43]). To simplify our presentation, throughout this thesis we shall use the following convention:

$$B_{q,p}^\beta(\mathbb{R}^d) = \begin{cases} B_q^\beta(L^p(\mathbb{R}^d)), & \text{if } 1 < p \leq \infty, \\ B_q^\beta(H^p(\mathbb{R}^d)), & \text{if } 0 < p \leq 1. \end{cases}$$

With this convention, we can get more familiar spaces by changing parameters. When $p = q = 2$, $B_{2,2}^\beta(\mathbb{R}^d)$ is the Sobolev space $W^\beta(\mathbb{R}^d)$, and when $0 < \beta < 1$ and $1 \leq p \leq \infty$, $B_{\infty,p}^\beta(\mathbb{R}^d)$ is the Lipschitz space $\mathrm{Lip}(\beta, L^p(\mathbb{R}^d))$.

When $0 < p < 1$ or $0 < q < 1$, then $B_{q,p}^\beta(\mathbb{R}^d)$ are no longer Banach spaces. However, they are always quasi-Banach spaces: The triangle inequality does not hold, but for each space $B_{q,p}^\beta(\mathbb{R}^d)$ there exists a constant C such that for all f and g in $B_{q,p}^\beta(\mathbb{R}^d)$,

$$\|f+g\|_{B_{q,p}^\beta} \leq C\Big(\|f\|_{B_{q,p}^\beta} + \|g\|_{B_{q,p}^\beta}\Big).$$

With a certain abuse of terminology, we shall continue to call these quasi-norms norms.

We can determine whether f is in $B_{q,p}^\beta(\mathbb{R}^d)$ simply by examining its wavelet coefficients. To have such smoothness characterization via the inhomogeneous wavelet decomposition of f, the scaling function ϕ and the dual scaling function $\widetilde{\phi}$ need to satisfy certain conditions. See Corollary 5.2.23 and Corollary 6.7.5 in [32]. All the conditions therein, except regularity condition on ϕ and *Strang-Fix condition* on $\widetilde{\phi}$, are automatically satisfied for ϕ and $\widetilde{\phi}$ in Definition 3.1.2. We now summarize smoothness characterization via the inhomogeneous wavelet decomposition in the following theorem.

Theorem 3.2.2. *Let $\psi, \widetilde{\psi}, \phi$, and $\widetilde{\phi}$ be univariate functions described in Section 3.1. Suppose ϕ has R-th continuous derivative and $\widetilde{\phi}$ satisfies Strang-Fix condition of order $M+1$, i.e.,*

$$\begin{aligned} &\widehat{\widetilde{\phi}}(0) \neq 0 && \textit{and} \\ &\frac{d^\nu \widehat{\widetilde{\phi}}}{d\xi^\nu}(n) = 0 && \textit{for all integers } \nu \textit{ and } n \textit{ such that } 1 \leq \nu < M+1 \textit{ and } n \neq 0. \end{aligned}$$

We assume that $\psi^{(i)}, \widetilde{\psi}^{(i)}, \Phi$, and $\widetilde{\Phi}$ are functions described in Section 3.1. Then as long as $\beta < \min(R, M)$, for any k_0,

$$|f|_{B^\beta_{q,p}} \asymp \left(\sum_{k \geq k_0} 2^{k(\beta + d(1/2 - 1/p))q} \left[\sum_{j,i} |\langle f, \widetilde{\psi}^{(i)}_{k,j} \rangle|^p \right]^{q/p} \right)^{1/q} \tag{3.9}$$

and

$$\|f\|_{B^\beta_{q,p}} \asymp |f|_{B^\beta_{q,p}} + 2^{k_0 d(1/2 - 1/p)} \left(\sum_l |\langle f, \widetilde{\Phi}_{k_0,l} \rangle|^p \right)^{1/p}. \tag{3.10}$$

The proof of Theorem 3.2.2 can be found in Chapter 5 and 6 of [32] for orthogonal wavelets. For biorthogonal wavelets, one can follow the same argument in Chapter 5 and 6 in [32] to prove Theorem 3.2.2. While doing so, one can also replace Strang-Fix condition on $\widetilde{\phi}$ of order $M+1$ by vanishing moments of $\widetilde{\psi}$ of order M, i.e.,

$$\int_{\mathbb{R}} x^m \widetilde{\psi}(x)\, dx = 0, \qquad \text{for } m = 0, 1, \cdots, M,$$

in Theorem 3.2.2.

Remark 3.2.3. Throughout this thesis we use the right hand side of (3.9) as the definition of $|f|_{B^\beta_{q,p}}$ and the right hand side of (3.10) as the definition of $\|f\|_{B^\beta_{q,p}}$ for a fixed integer k_0. We also assume that k_0 is a fixed nonnegative small integer throughout this thesis to avoid any possible confusion.

CHAPTER 4
EMBEDDING, INTERPOLATION, AND DUALITY BETWEEN BESOV SPACES

In this chapter we study embedding, interpolation, and duality between Besov spaces $B^{\beta}_{q,p}(\mathbb{R}^d)$ based on the wavelet sequence norm in Chapter 3.

Throughout this chapter we implicitly assume that ψ, $\widetilde{\psi}$, ϕ, and $\widetilde{\phi}$ satisfy all conditions in Theorem 3.2.2 so that the Besov space $B^{\beta}_{q,p}(\mathbb{R}^d)$ norm is equivalent to the associated sequence norm.

4.1 Embedding

THEOREM 4.1.1.

(1) *If $q' > q$, then*

$$B^{\beta}_{q,p}(\mathbb{R}^d) \hookrightarrow B^{\beta}_{q',p}(\mathbb{R}^d).$$

(2) *If $\beta' < \beta$ and $(\beta - \beta')/d = 1/p - 1/p'$, then*

$$B^{\beta}_{q,p}(\mathbb{R}^d) \hookrightarrow B^{\beta'}_{q,p'}(\mathbb{R}^d)$$

(3) *For any q_1 and q_2, if $\beta' < \beta$, then*

$$B^{\beta}_{q_1,p}(\mathbb{R}^d) \hookrightarrow B^{\beta'}_{q_2,p}(\mathbb{R}^d).$$

Proof. Recall that by definition of Besov norm in this thesis,

$$|f|_{B^{\beta}_{q,p}} = \left(\sum_{k \geq k_0} 2^{k(\beta + d(1/2 - 1/p))q} \left[\sum_{j,i} |\langle f, \widetilde{\psi}^{(i)}_{k,j} \rangle|^p \right]^{q/p} \right)^{1/q}$$

and

$$\|f\|_{B^\beta_{q,p}} = |f|_{B^\beta_{q,p}} + 2^{k_0 d(1/2-1/p)} \left(\sum_l |\langle f, \widetilde{\Phi}_{k_0,l} \rangle|^p \right)^{1/p}.$$

Notice that the second part of $\|f\|_{B^\beta_{q,p}}$ only depends on the parameter p. Since k_0 is a fixed integer, by (2.13) one can find a constant C such that for all f,

$$\text{(a)} \quad 2^{k_0 d(1/2-1/p')} \left(\sum_l |\langle f, \widetilde{\Phi}_{k_0,l} \rangle|^{p'} \right)^{1/p'} \leq C 2^{k_0 d(1/2-1/p)} \left(\sum_l |\langle f, \widetilde{\Phi}_{k_0,l} \rangle|^p \right)^{1/p},$$

if $0 < p \leq p' \leq \infty$.

(1): Let

$$a_k = 2^{k(\beta + d(1/2-1/p))} \left(\sum_{j,i} |\langle f, \widetilde{\psi}^{(i)}_{k,j} \rangle|^p \right)^{1/p}.$$

By (2.13), $\|(a_k)\|_{\ell^{q'}} \leq \|(a_k)\|_{\ell^q}$, since $q' > q$. Notice that by definition of (a_k), $\|(a_k)\|_{\ell^{q'}} = |f|_{B^\beta_{q',p}}$ and $\|(a_k)\|_{\ell^q} = |f|_{B^\beta_{q,p}}$. Combining these results and (a), we have $\|f\|_{B^\beta_{q',p}} \leq C\|f\|_{B^\beta_{q,p}}$. This proves (1).

(2): By hypothesis, $p < p'$ and $\beta + d(1/2 - 1/p) = \beta' + d(1/2 - 1/p')$. Hence by (2.13), we have

$$\text{(b)} \quad \begin{aligned} |f|^q_{B^{\beta'}_{q,p'}} &= \sum_{k \geq k_0} 2^{k(\beta' + d(1/2-1/p'))q} \left(\sum_{j,i} |\langle f, \widetilde{\psi}^{(i)}_{k,j} \rangle|^{p'} \right)^{q/p'} \\ &\leq \sum_{k \geq k_0} 2^{k(\beta + d(1/2-1/p))q} \left(\sum_{j,i} |\langle f, \widetilde{\psi}^{(i)}_{k,j} \rangle|^{p} \right)^{q/p} \\ &= |f|^q_{B^\beta_{q,p}}. \end{aligned}$$

Combining (a) and (b), we have $\|f\|_{B^{\beta'}_{q,p'}} \leq C\|f\|_{B^\beta_{q,p}}$. This proves (2).

(3): Recall that $\beta' < \beta$. It is not difficult to show that

$$B^\beta_{q_1,p}(\mathbb{R}^d) \hookrightarrow B^{\beta'}_{q_2,p}(\mathbb{R}^d) \quad \text{for } q_1 = q_2.$$

For $q_2 > q_1$, we use (1) of this theorem, and have

$$B^\beta_{q_1,p}(\mathbb{R}^d) \hookrightarrow B^\beta_{q_2,p}(\mathbb{R}^d) \hookrightarrow B^{\beta'}_{q_2,p}(\mathbb{R}^d).$$

For $q_2 < q_1$, we write

$$|f|^{q_2}_{B^{\beta'}_{q_2,p}} = \sum_{k\geq k_0} 2^{k(\beta'+d(1/2-1/p))q_2} \Big(\sum_{j,i} |\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle|^p\Big)^{q_2/p}$$
$$= \sum_{k\geq k_0} 2^{k(\beta'-\beta)q_2} \cdot 2^{k(\beta+d(1/2-1/p))q_2} \Big(\sum_{j,i} |\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle|^p\Big)^{q_2/p}.$$

We now use (2.11) on the sum over k with $q_1/(q_1-q_2)$ and q_1/q_2. Then

$$|f|^{q_2}_{B^{\beta'}_{q_2,p}} \leq \Big(\sum_{k\geq k_0} 2^{k(\beta'-\beta)q_2 q_1/(q_1-q_2)}\Big)^{(q_1-q_2)/q_1}$$
$$\times \Big(\sum_{k\geq k_0} 2^{k(\beta+d(1/2-1/p))q_1} \Big[\sum_{j,i} |\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle|^p\Big]^{q_1/p}\Big)^{q_2/q_1}.$$

Since $(\beta'-\beta)q_2q_1/(q_1-q_2) < 0$ and the second term in the right hand side of above equation is $|f|^{q_2}_{B^{\beta}_{q_1,p}}$, we finally have

$$\text{(d)} \qquad |f|^{q_2}_{B^{\beta'}_{q_2,p}} \leq C|f|^{q_2}_{B^{\beta}_{q_1,p}}$$

for a constant C. Combining (a) and (d), we have $\|f\|_{B^{\beta'}_{q_2,p}} \leq C\|f\|_{B^{\beta}_{q_1,p}}$. Thus we completed the proof of (3). $\square$

This theorem shows that the parameter q in $B^{\beta}_{q,p}(\mathbb{R}^d)$ is not as important as p or β. From now on, we consider only Besov spaces of type $B^{\beta}_{p,p}(\mathbb{R}^d)$.

It is well known (see, e.g., [15] and [20]) that

$$\text{(4.1)} \quad B^{\beta}_{p,p}(\mathbb{R}^d) \hookrightarrow L^2(\mathbb{R}^d) \quad \text{if and only if} \quad \beta + d(1/2-1/p) \geq 0 \ \text{ and } \ \beta \geq 0.$$

For instance, $B^{\beta}_{\tau,\tau}(\mathbb{R}^d)$ with $1/\tau = \beta/d + 1/2$ and $\beta \geq 0$ is embedded to $L^2(\mathbb{R}^d)$; since $2 \geq \tau$ for $\beta \geq 0$, using (2.13) one has

$$\text{(4.2)} \qquad \Big(\sum_{k\geq k_0}\sum_{j,i} |\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle|^2\Big)^{1/2} \leq \Big(\sum_{k\geq k_0}\sum_{j,i} |\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle|^\tau\Big)^{1/\tau}$$

and

$$\left(\sum_l |\langle f, \widetilde{\Phi}_{k_0,l}\rangle|^2\right)^{1/2} \le \left(\sum_l |\langle f, \widetilde{\Phi}_{k_0,l}\rangle|^\tau\right)^{1/\tau}. \tag{4.3}$$

Collecting (3) of Definition 3.1.2, (4.2), and (4.3), we have $\|f\|_{L^2} \le C\|f\|_{B^\beta_{\tau,\tau}}$ for a constant C. Thus

$$B^\beta_{\tau,\tau}(\mathbb{R}^d) \hookrightarrow L^2(\mathbb{R}^d) \quad \text{if } 1/\tau = \beta/d + 1/2 \text{ and } \beta \ge 0.$$

These Besov spaces $B^\beta_{\tau,\tau}(\mathbb{R}^d)$ have been used in [4], [13], and [17] for image compression, noise removal, etc. The spaces $B^\beta_{\tau,\tau}(\mathbb{R}^d)$ are quite natural when considering approximation in $L^2(\mathbb{R}^d)$, because these spaces have minimal smoothness to be embedded in $L^2(\mathbb{R}^d)$. This means that given a pair β and $1/\tau = \beta/d + 1/2$, there is no $\beta' < \beta$ or $\tau' < \tau$ such $B^{\beta'}_{\tau,\tau}(\mathbb{R}^d)$ or $B^\beta_{\tau',\tau'}(\mathbb{R}^d)$ is embedded into $L^2(\mathbb{R}^d)$. More importantly, with $B^\beta_{\tau,\tau}(\mathbb{R}^d)$, where $1/\tau = \beta/d + 1/2$, the approximation order of functions can be related to their smoothness order β explicitly:

Remark 4.1.2. If f can be approximated to $\mathcal{O}(N^{-\beta/d})$ in $L^2(\mathbb{R}^d)$ by sums with N nonzero wavelet coefficients, then f is necessarily in $B^\beta_{\tau,\tau}(\mathbb{R}^d)$, $1/\tau = \beta/d + 1/2$. (Because of certain technicalities, this statement only approximates the truth; see [13] for precise statements.) Conversely, if f is in $B^\beta_{\tau,\tau}(\mathbb{R}^d)$, $1/\tau = \beta/d + 1/2$, then scalar quantization of the wavelet coefficients yields compression algorithms with convergence rate of $\mathcal{O}(N^{-\beta/d})$ in $L^2(\mathbb{R}^d)$.

4.2 Interpolation

This is a standard argument from interpolation theory. Suppose $f \in L^\infty(\mathbb{R}^d) \cap B^\beta_{p,p}(\mathbb{R}^d)$. Notice that any wavelets described in Chapter 3 satisfy

$$\begin{aligned}\|\widetilde{\psi}^{(i)}_{k,j}\|_{L^1(\mathbb{R}^d)} &= \int_{\mathbb{R}^d} |2^{kd/2}\widetilde{\psi}^{(i)}(2^k x - j)|\,dx \\ &= 2^{-kd/2}\int_{\mathbb{R}^d} |\widetilde{\psi}^{(i)}(x)|\,dx \\ &\le C2^{-kd/2}\end{aligned}$$

for a constant C. Thus we have

$$(4.4)\qquad \begin{aligned} |\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle| &\le \|f\|_{L^\infty(\mathbb{R}^d)} \|\widetilde{\psi}^{(i)}_{k,j}\|_{L^1(\mathbb{R}^d)} \qquad \text{(by (2.12))} \\ &\le C 2^{-kd/2} \|f\|_{L^\infty(\mathbb{R}^d)}. \end{aligned}$$

For any pair (β', p') such that $\beta' < \beta$ and $\beta' p' = \beta p$, we have

$$(4.5)\qquad \begin{aligned} |f|^{p'}_{B^{\beta'}_{p',p'}} &= \sum_{k\ge k_0} 2^{k(\beta'+d(1/2-1/p'))p'} \sum_{j,i} \left|\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle\right|^{p'} \\ &= \sum_{k\ge k_0} 2^{k(\beta'+d(1/2-1/p'))p'} \sum_{j,i} \left|\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle\right|^{p} \cdot \left|\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle\right|^{p'-p}. \end{aligned}$$

We now apply (4.4) to $\left|\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle\right|^{p'-p}$ in (4.5). Then

$$(4.6)\qquad \begin{aligned} |f|^{p'}_{B^{\beta'}_{p',p'}} &\le C' \|f\|^{p'-p}_{L^\infty} \sum_{k\ge k_0} 2^{k(\beta'+d(1/2-1/p'))p'} 2^{-kd/2(p'-p)} \sum_{j,i} \left|\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle\right|^{p} \\ &= C' \|f\|^{p'-p}_{L^\infty} \sum_{k\ge k_0} 2^{k(\beta+d(1/2-1/p))p} \sum_{j,i} \left|\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle\right|^{p} \\ &= C' \|f\|^{p'-p}_{L^\infty} |f|^{p}_{B^{\beta}_{p,p}}. \end{aligned}$$

With a similar argument, we can show that

$$(4.7)\qquad 2^{k_0 d(1/2-1/p')p'} \sum_{l} \left|\langle f, \widetilde{\Phi}_{k_0,l}\rangle\right|^{p'} \le C'' \|f\|^{p'-p}_{L^\infty} 2^{k_0 d(1/2-1/p)p} \sum_{l} \left|\langle f, \widetilde{\Phi}_{k_0,l}\rangle\right|^{p}.$$

Using (4.6) and (4.7), we prove the following theorem:

Theorem 4.2.1. *If $f \in L^\infty(\mathbb{R}^d) \cap B^{\beta}_{p,p}(\mathbb{R}^d)$, then $f \in B^{\beta'}_{p',p'}(\mathbb{R}^d)$, where $\beta' < \beta$ and $\beta' p' = \beta p$. Moreover,*

$$|f|_{B^{\beta'}_{p',p'}} \le C \|f\|^{1-p/p'}_{L^\infty} |f|^{p/p'}_{B^{\beta}_{p,p}}$$

for a constant C.

4.3 Duality

In analogy with the special case of Sobolev spaces $W^\beta(\mathbb{R}^d)$, the Besov space $B^\beta_{p,p}(\mathbb{R}^d)$ with $\beta < 0$ is understood as the dual space of $B^{\beta'}_{p',p'}(\mathbb{R}^d)$, where $\beta' = -\beta$ and $1/p + 1/p' = 1$. (Here we assume that $1 \le p \le \infty$.) We extend the wavelet-dependent sequence norm to Besov spaces $B^\beta_{q,p}(\mathbb{R}^d)$ with negative β. To do so, we assume that both ϕ and $\widetilde{\phi}$ have R-th continuous derivative, and satisfy Strang-Fix condition of order $M+1$. If $|\beta| < \min(R, M)$, then the duality between the distribution $f \in B^\beta_{p,p}(\mathbb{R}^d)$ and the test function $g \in B^{\beta'}_{p',p'}(\mathbb{R}^d)$ is defined by

$$\langle f, g\rangle = \sum_{k\ge k_0}\sum_{j,i}\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle\langle \psi^{(i)}_{k,j}, g\rangle + \sum_l \langle f, \widetilde{\Phi}_{k_0,l}\rangle\langle \Phi_{k_0,l}, g\rangle.$$

Since $\beta + d(1/2 - 1/p) = -\beta' - d(1/2 - 1/p')$, by (2.12) one can have

$$\begin{aligned}\left|\sum_{j,i}\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle\langle \psi^{(i)}_{k,j}, g\rangle\right| &\le \left(\sum_{j,i}|\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle|^p\right)^{1/p}\left(\sum_{j,i}|\langle \psi^{(i)}_{k,j}, g\rangle|^{p'}\right)^{1/p'}\\ &= a_k \cdot b_k,\end{aligned}$$

where

$$a_k = 2^{k(\beta + d(1/2-1/p))}\left(\sum_{j,i}|\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle|^p\right)^{1/p}$$

and

$$b_k = 2^{k(\beta' + d(1/2-1/p'))}\left(\sum_{j,i}|\langle g, \psi^{(i)}_{k,j}\rangle|^{p'}\right)^{1/p'}.$$

Notice that $\|(a_k)\|_{\ell^p} = |f|_{B^\beta_{p,p}}$ and $\|(b_k)\|_{\ell^{p'}} = |g|_{B^{\beta'}_{p',p'}}$, where we have switched roles of $\widetilde{\psi}^{(i)}$ and $\psi^{(i)}$ in obtaining the equivalent wavelet sequence sums to $|g|^{\beta'}_{p',p'}$. Thus we have

$$\begin{aligned}\left|\sum_{k\ge k_0}\sum_{j,i}\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle\langle \psi^{(i)}_{k,j}, g\rangle\right| &\le \sum_{k\ge k_0} a_k \cdot b_k\\ &\le \left(\sum_{k\ge k_0}|a_k|^p\right)^{1/p}\left(\sum_{k\ge k_0}|b_k|^{p'}\right)^{1/p'}\\ &= |f|_{B^\beta_{p,p}}|g|_{B^{\beta'}_{p',p'}}.\end{aligned} \tag{4.8}$$

With a similar argument, we can show that

$$\left|\sum_l \langle f, \widetilde{\Phi}_{k_0,l}\rangle\langle \Phi_{k_0,l}, g\rangle\right| \le 2^{k_0 d(1/2-1/p)}\left(\sum_l |\langle f, \widetilde{\Phi}_{k_0,l}\rangle|^p\right)^{1/p} \times 2^{k_0 d(1/2-1/p')}\left(\sum_l |\langle g, \Phi_{k_0,l}\rangle|^{p'}\right)^{1/p'}, \tag{4.9}$$

since $d(1/2 - 1/p) = -d(1/2 - 1/p')$. Again, we switch the roles of $(\widetilde{\Phi}, \widetilde{\psi}^{(i)})$ and $(\Phi, \psi^{(i)})$ in obtaining the equivalent wavelet sequence sums to $\|g\|_{B^{\beta'}_{p',p'}}$. Then combining (4.8) and (4.9), we have

$$|\langle f, g\rangle| \le \|f\|_{B^{\beta}_{p,p}} \|g\|_{B^{\beta'}_{p',p'}}.$$

This inequality implies that $B^{\beta}_{p,p}(\mathbb{R}^d)$ is indeed the dual space of $B^{\beta'}_{p',p'}(\mathbb{R}^d)$ and extended definitions can characterize $B^{\beta}_{p,p}(\mathbb{R}^d)$ with negative β.

CHAPTER 5

LINEAR HOMOGENEOUS EQUATIONS

In this chapter we examine three examples for A satisfying $\widehat{A^*Af}(\xi) = |\xi|^{-2\alpha}\widehat{f}(\xi)$,

$$\begin{cases} \text{the identity operator on } L^2(\mathbb{R}^d), & \alpha = 0, \\ \text{the Radon transform on } L^2(\mathbb{R}^2), & \alpha = 1/2, \\ \text{the } 2\pi \times \text{ integration operator on } L^2(\mathbb{R}), & \alpha = 1. \end{cases}$$

The classical method for solving (1.2) has been *filtered backprojection* based on a description of A in the frequency domain. We study this in detail in Section 5.2.

5.1 Examples

The identity operator on $L^2(\mathbb{R}^d)$ satisfies (1.1) with $\alpha = 0$. In this case, we have

$$Y = f + Z$$

in (1.2). Recovering f can be viewed as a noise removal problem.

The Radon transform $\mathcal{R} : L^2(\mathbb{R}^2) \longrightarrow L^2([0,\pi), L^2(\mathbb{R}))$ is defined by

$$\mathcal{R}f(\theta,u) = \int_{L_{\theta,u}} f(x,y)\,ds(x,y),$$

where $ds(x,y)$ is the Euclidean measure on the line

$$L_{\theta,u} = \{(x,y) \mid \; x\cos\theta + y\sin\theta = u\}.$$

The inner product $[\cdot,\cdot]$ in $L^2([0,\pi), L^2(\mathbb{R}))$ is assumed as

$$[F,G] = \int_0^\pi \int_{\mathbb{R}} F(\theta,u)\overline{G(\theta,u)}\,du\,d\theta.$$

THEOREM 5.1.1. (*Fourier slice theorem*)

$$\mathcal{R}f(\theta,\cdot)^\wedge(w) = \widehat{f}(w\cos\theta, w\sin\theta), \tag{5.1}$$

where $\mathcal{R}f(\theta,\cdot)^\wedge(w)$ is the one-dimensional Fourier transform of $\mathcal{R}f(\theta,u)$ as a function of u, and $\widehat{f}$ is the two-dimensional Fourier transform of f.

Proof. See [37]. □

Notice that

$$\begin{aligned}
[\mathcal{R}f,\mathcal{R}g] &= \int_0^\pi \int_{\mathbb{R}} \mathcal{R}f(\theta,u)\overline{\mathcal{R}g(\theta,u)}\,du\,d\theta \\
&= \int_0^\pi \int_{\mathbb{R}} \mathcal{R}f(\theta,\cdot)^\wedge(w)\overline{\mathcal{R}g(\theta,\cdot)^\wedge(w)}\,dw\,d\theta && \text{(by (2.3))} \\
&= \int_0^\pi \int_{\mathbb{R}} \widehat{f}(w\cos\theta, w\sin\theta)\overline{\widehat{g}(w\cos\theta, w\sin\theta)}\,dw\,d\theta && \text{(by (5.1))} \\
&= \int_{\mathbb{R}^2} |\xi|^{-1}\widehat{f}(\xi)\overline{\widehat{g}(\xi)}\,d\xi.
\end{aligned} \tag{5.2}$$

Since $\langle \mathcal{R}^*\mathcal{R}f, g\rangle = [\mathcal{R}f,\mathcal{R}g]$, using (2.3) for $\langle \mathcal{R}^*\mathcal{R}f, g\rangle$, we have

$$\int_{\mathbb{R}^2} \widehat{\mathcal{R}^*\mathcal{R}f}(\xi)\overline{\widehat{g}(\xi)}\,d\xi = \int_{\mathbb{R}^2} |\xi|^{-1}\widehat{f}(\xi)\overline{\widehat{g}(\xi)}\,d\xi.$$

Since g can be arbitrary, we can write

$$\widehat{\mathcal{R}^*\mathcal{R}f}(\xi) = |\xi|^{-1}\widehat{f}(\xi).$$

Thus the Radon transform satisfies (1.1) with $\alpha = 1/2$.

By setting $f = g$ in (5.2), one can show that the Radon transform $\mathcal{R}$ is well-defined for all functions f such that

$$\int_{\mathbb{R}^2} |\xi|^{-1}\left|\widehat{f}(\xi)\right|^2 d\xi \; < \; \infty.$$

Thus the domain of $\mathcal{R}$, denoted by $\mathcal{D}(\mathcal{R})$, is $L^2(\mathbb{R}^2) \cap S^{-1/2}(\mathbb{R}^2)$, where

$$S^\beta(\mathbb{R}^d) = \left\{ f \in S'(\mathbb{R}^d) \mid \int_{\mathbb{R}^d} |\xi|^{2\beta}\left|\widehat{f}(\xi)\right|^2 d\xi \; < \; \infty \right\}. \tag{5.3}$$

Here $S'(\mathbb{R}^d)$ denotes the space of tempered distributions on $\mathbb{R}^d$. For the future use, we define

$$|f|^2_{S^\beta} = \int_{\mathbb{R}^d} |\xi|^{2\beta} |\widehat{f}(\xi)|^2 \, d\xi. \tag{5.4}$$

We define $\mathcal{I} : L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R})$ by

$$\mathcal{I}f(x) = 2\pi \int_{-\infty}^{x} f(t)\, dt.$$

Obviously, the operator $\mathcal{I}$ cannot be defined for all functions in $L^2(\mathbb{R})$. To specify $\mathcal{D}(\mathcal{I})$, notice that by (2.6), one can have

$$2\pi \widehat{f}(\xi) = 2\pi i \xi \, \widehat{\mathcal{I}f}(\xi),$$

since $2\pi f(x) = (\mathcal{I}f)'(x)$. Thus

$$\widehat{\mathcal{I}f}(\xi) = -i\xi^{-1} \widehat{f}(\xi) \tag{5.5}$$

and $\mathcal{D}(\mathcal{I}) = L^2(\mathbb{R}) \cap S^{-1}(\mathbb{R})$.

Notice that

$$\begin{aligned}
\langle \mathcal{I}^*\mathcal{I}f \, , \, g\rangle &= \langle \mathcal{I}f \, , \, \mathcal{I}g\rangle \\
&= \int_{\mathbb{R}} \widehat{\mathcal{I}f}(\xi) \, \overline{\widehat{\mathcal{I}g}(\xi)} \, d\xi && \text{(by (2.3))} \\
&= \int_{\mathbb{R}} (-i\xi^{-1}) \widehat{f}(\xi) (i\overline{\xi^{-1}}) \overline{\widehat{g}(\xi)} \, d\xi && \text{(by (5.5))} \\
&= \int_{\mathbb{R}} |\xi|^{-2} \widehat{f}(\xi) \overline{\widehat{g}(\xi)} \, d\xi.
\end{aligned}$$

This implies that

$$\widehat{\mathcal{I}^*\mathcal{I}f}(\xi) = |\xi|^{-2} \widehat{f}(\xi).$$

Thus the operator $\mathcal{I}$ on $L^2(\mathbb{R})$ satisfies (1.1) with $\alpha = 1$.

5.2 Filtered Backprojection Methods

From (1.1), one's first attempt to solve (1.2) might be

$$\widetilde{f}(x) = \int_{\mathbb{R}^d} e^{2\pi i \xi \cdot x} |\xi|^{2\alpha} \widehat{A^*Y}(\xi)\, d\xi. \tag{5.6}$$

For $\alpha > 0$, however, the solution method in (5.6) generates an unacceptable solution for (1.2) due to instability in high frequency information of f. To avoid this phenomenon, one can consider a filtered solution $\widetilde{f}_w$ such that

$$\widetilde{f}_w(x) = \int_{\mathbb{R}^d} e^{2\pi i \xi \cdot x} w(\xi) |\xi|^{2\alpha} \widehat{A^*Y}(\xi)\, d\xi, \tag{5.7}$$

where the *weight filter* $w(\xi)$ satisfies $0 \leq w(\xi) \leq 1$ and $w(\xi) \to 0$ as $|\xi| \to \infty$. For example, the famous *filtered backprojection* (FBP) method for inverting the Radon transform in medical image processing takes the form (5.7) with $\alpha = 1/2$.

Finding a proper weight filter $w(\xi)$ is very important for the performance of $\widetilde{f}_w$. We consider a family of weight filters $\{w_M\}$ such that

$$w_M(\xi) = \begin{cases} 1, & \text{if } |\xi| \leq M, \\ 0, & \text{otherwise.} \end{cases}$$

Then the mean square error of $\widetilde{f}_{w_M}$ is

$$\|f - \widetilde{f}_{w_M}\|_{L^2}^2 = \int_{|\xi| \leq M} |\xi|^{4\alpha} \left|\widehat{A^*Z}(\xi)\right|^2 d\xi + \int_{|\xi| > M} \left|\widehat{f}(\xi)\right|^2 d\xi. \tag{5.8}$$

We assume that

$$Z = \sigma W, \tag{5.9}$$

where W is the white noise process defined on the underlying space of functions in $\mathcal{Y}$. To deal with this noise model, let us consider the following companion problem to (1.2):

$$y = \widetilde{A} f + z,$$

where $\widetilde{A} : L^2(\mathbb{R}^d) \longrightarrow L^2(\mathbb{R}^d)$ satisfies

$$\widehat{\widetilde{A}f}(\xi) = |\xi|^{-\alpha} \widehat{f}(\xi).$$

Let

$$\widetilde{f}^*_{w_M}(x) = \int_{\mathbb{R}^d} e^{2\pi i \xi \cdot x} w_M(\xi) |\xi|^{2\alpha} \widehat{\widetilde{A}^* y}(\xi)\, d\xi$$

and

$$z = \sigma \widetilde{W}, \tag{5.10}$$

where $\widetilde{W}$ is the white noise process defined on $\mathbb{R}^d$. Since $\widetilde{A}^* \widetilde{A} = A^* A$, and $\widetilde{f}_{w_M}$ and $\widetilde{f}^*_{w_M}$ are obtained by using backprojection methods, the statistical performance (in mean square error measurement) of $\widetilde{f}_{w_M}$ with the noise model (5.9) is equal to that of $\widetilde{f}^*_{w_M}$ with the noise model (5.10), i.e.,

$$E\|f - \widetilde{f}_{w_M}\|^2_{L^2} = E\|f - \widetilde{f}^*_{w_M}\|^2_{L^2}. \tag{5.11}$$

Thus we can estimate an upper bound for $E\|f - \widetilde{f}_{w_M}\|^2_{L^2}$ by doing same thing for $E\|f - \widetilde{f}^*_{w_M}\|^2_{L^2}$. While doing that, we determine M that minimizes that upper bound.

We assume that the true solution f is in $W^\beta(\mathbb{R}^d)$, $\beta > 0$. Notice that for $\beta > 0$,

$$W^\beta(\mathbb{R}^d) = L^2(\mathbb{R}^d) \cap S^\beta(\mathbb{R}^d) \qquad \text{and} \qquad |f|_{W^\beta} = |f|_{S^\beta}. \tag{5.12}$$

With a similar argument used to get (5.8), one can have

$$E\|f - \widetilde{f}^*_{w_M}\|^2_{L^2} = E \int_{|\xi| \le M} |\xi|^{4\alpha} |\widehat{\widetilde{A}^* z}(\xi)|^2 d\xi + \int_{|\xi| > M} |\widehat{f}(\xi)|^2 d\xi. \tag{5.13}$$

Since

$$|\widehat{\widetilde{A}^* z}|^2 = |\xi|^{-2\alpha} |\widehat{z}(\xi)|^2,$$

the first term in the right hand side of (5.13) is

$$\sigma^2 \int_{|\xi|\le M} |\xi|^{2\alpha} d\xi = C\sigma^2 M^{2\alpha+d} \tag{5.14}$$

for a constant C.

One can have

$$\begin{aligned}\int_{|\xi|>M} \left|\widehat{f}(\xi)\right|^2 d\xi &\le M^{-2\beta} \int_{|\xi|>M} |\xi|^{2\beta} \left|\widehat{f}(\xi)\right|^2 d\xi \\ &\le M^{-2\beta} \int_{\mathbb{R}^d} |\xi|^{2\beta} \left|\widehat{f}(\xi)\right|^2 d\xi \\ &\le M^{-2\beta} |f|^2_{W^\beta}. \qquad \text{(see (5.4) and (5.12))}\end{aligned} \tag{5.15}$$

Combining (5.11), (5.13), (5.14), and (5.15), we have

$$E\|f - \widetilde{f}_{w_M}\|^2_{L^2} \le C'\left(\sigma^2 M^{2\alpha+d} + M^{-2\beta}|f|^2_{W^\beta}\right).$$

By comparing the two dominant terms $\sigma^2 M^{2\alpha+d}$ and $M^{-2\beta}|f|^2_{W^\beta}$, we can get a simple approximation to the critical M. One can determine M so that

$$\sigma^2 M^{2\alpha+d} = M^{-2\beta}|f|^2_{W^\beta}$$

or

$$M = \left(\frac{|f|_{W^\beta}}{\sigma}\right)^{1/(\beta+d/2+\alpha)}. \tag{5.16}$$

With this M, we have

$$E\|f - \widetilde{f}_{w_M}\|^2_{L^2} \le C|f|^{(2\alpha+d)/(\beta+d/2+\alpha)}_{W^\beta}\sigma^{2r},$$

with rate exponent

$$r = \frac{\beta}{\beta+d/2+\alpha}. \tag{5.17}$$

It is known (see [22] or Theorem 9.1.1) that the rate of convergence of any method for recovering $f \in W^\beta(\mathbb{R}^d)$ is at most r (5.17), thus no filtered backprojection methods can provide an asymptotically better solution than does $\widetilde{f}_{w_M}$ with M idetermined by (5.16).

The filtered backprojection method $\widetilde{f}_{w_M}$ considers any low-frequency structure to be signal, and any high-frequency structure to be noise, no many how large $|\widehat{f}(\xi)|$ might be. This is not acceptable when we wish to recover the function which can be more meaningfully characterized by its discontinuities, for example, an image with several edges and small extent, since such information lies in the high-frequency domain. Even with more general weight filter w, we cannot avoid this kind of degradation in recovering nonsmooth functions, since $w(\xi) \to 0$ as $|\xi| \to \infty$.

CHAPTER 6
WAVELET-VAGUELETTE DECOMPOSITIONS

In this chapter, we study wavelet-vaguelette decomposition. The terminology vaguelette was used by Meyer in [36] to describe a collection of functions which are wavelet-like. In [22] Donoho constructed a wavelet-vaguelette system for solving homogeneous equations described in Section 5.1 and gave a sufficient condition on wavelets to generate a wavelet-vaguelette system.

We present the definition of the wavelet-vaguelette system in the setting of biorthogonal wavelets in Section 6.1. In Section 6.2 we study some properties of biorthogonal Riesz bases. These properties are useful in finding a sufficient condition on regularity and vanishing moment of biorthogonal wavelets to generate a wavelet-vaguelette system in Section 6.3. The proof in Section 6.3 generalizes Daubechies's work [9] in the discrete wavelet transform.

Throughout this chapter we always assume that A is weakly invertible, i.e., $Af = 0$ implies $f = 0$, and $A(L^2(\mathbb{R}^d)) = \mathcal{Y}$.

6.1 Definition and Background

Definition 6.1.1. *A collection of functions* $\{\psi_{k,j}^{(i)}, \widetilde{\psi}_{k,j}^{(i)}, U_{k,j}^{(i)}, \widetilde{U}_{k,j}^{(i)}\}$ *is called a wavelet-vaguelette system of* $A : L^2(\mathbb{R}^d) \longrightarrow \mathcal{Y}$, *where* $\widehat{A^*Af}(\xi) = |\xi|^{-2\alpha}\widehat{f}(\xi)$, *if*

(1) $\{\psi_{k,j}^{(i)}\}$ *and* $\{\widetilde{\psi}_{k,j}^{(i)}\}$ *are biorthogonal wavelets of* $L^2(\mathbb{R}^d)$,

(2) $\{U_{k,j}^{(i)}\}$ *and* $\{\widetilde{U}_{k,j}^{(i)}\}$ *are biorthogonal Riesz bases of* $\mathcal{Y}$,

(3) $A^*U_{k,j}^{(i)} = 2^{-k\alpha}\widetilde{\psi}_{k,j}^{(i)}$ *and* $A\psi_{k,j}^{(i)} = 2^{-k\alpha}\widetilde{U}_{k,j}^{(i)}$ *for all* k, j, i.

Once we have a wavelet-vaguelette system of A, we can determine wavelet coefficients $\langle f, \widetilde{\psi}_{k,j}^{(i)}\rangle$ of f from vaguelette coefficients $[Af, U_{k,j}^{(i)}]$ of Af. Notice that by (3) of Definition 6.1.1,

$$\begin{aligned} \langle f, \widetilde{\psi}_{k,j}^{(i)} \rangle &= 2^{k\alpha} \langle f, A^* U_{k,j}^{(i)} \rangle \\ &= 2^{k\alpha} [Af, U_{k,j}^{(i)}]. \end{aligned} \tag{6.1}$$

Plugging (6.1) into (1) of Definition 3.1.1, we have the following reproducing formula via the wavelet-vaguelette system $\{\psi_{k,j}^{(i)}, \widetilde{\psi}_{k,j}^{(i)}, U_{k,j}^{(i)}, \widetilde{U}_{k,j}^{(i)}\}$ of A:

$$f = \sum_{k,j,i} 2^{k\alpha} [Af, U_{k,j}^{(i)}] \psi_{k,j}^{(i)}. \tag{6.2}$$

We call (6.2) a wavelet-vaguelette decomposition of f.

We now consider how to construct a wavelet-vaguelette system of A. Notice that any operator A satisfying $\widehat{A^*Af}(\xi) = |\xi|^{-2\alpha} \widehat{f}(\xi)$ has following two properties:

Lemma 6.1.2. *A^*A is translation-invariant, i.e.,*

$$A^*AT_h = T_h A^*A.$$

Proof. Notice that for each $f \in L^2(\mathbb{R}^d)$,

$$\begin{aligned} \widehat{A^*AT_h f}(\xi) &= |\xi|^{-2\alpha} \widehat{T_h f}(\xi) \\ &= |\xi|^{-2\alpha} e^{-2\pi i \xi \cdot h} \widehat{f}(\xi) && \text{(by (2.4))} \\ &= e^{-2\pi i \xi \cdot h} |\xi|^{-2\alpha} \widehat{f}(\xi) \\ &= e^{-2\pi i \xi \cdot h} \widehat{A^*Af}(\xi) \\ &= \widehat{T_h A^*A f}(\xi). && \text{(by (2.4))} \end{aligned}$$

Since f is arbitrarily chosen, $A^*AT_h = T_h A^*A$. ∎

Lemma 6.1.3. *A^*A is homogeneous of order 2α, i.e.,*

$$A^*AD_a = a^{2\alpha} D_a A^*A.$$

Proof. Notice that for each $f \in L^2(\mathbb{R}^d)$,

$$\begin{aligned}
\widehat{A^*AD_a f}(\xi) &= |\xi|^{-2\alpha}\widehat{D_a f}(\xi) \\
&= |\xi|^{-2\alpha} a^d \widehat{f}(a\xi) && \text{(by (2.5))} \\
&= a^{2\alpha} a^d |a\xi|^{-2\alpha}\widehat{f}(a\xi) \\
&= a^{2\alpha}\widehat{D_a A^* A f}(\xi). && \text{(by (1.1) and (2.5))}
\end{aligned}$$

Since f is arbitrarily chosen, $A^*AD_a = a^{2\alpha}D_aA^*A$. $\square$

On $\mathbb{R}^d$, all elements of $\{\widetilde{\psi}^{(i)}_{k,j}\}$ are generated by dilation and translation of basic $2^d - 1$ functions $\{\widetilde{\psi}^{(i)}\}$ as

$$\widetilde{\psi}^{(i)}_{k,j} = 2^{kd/2} D_{2^{-k}} T_{j2^{-k}} \widetilde{\psi}^{(i)}.$$

We define $\pi^{(i)}$ by

$$A^*A\pi^{(i)} = \widetilde{\psi}^{(i)} \tag{6.3}$$

or equivalently

$$\widehat{\pi^{(i)}}(\xi) = |\xi|^{2\alpha}\widehat{\widetilde{\psi}^{(i)}}(\xi).$$

Let

$$\pi^{(i)}_{k,j} = 2^{kd/2} D_{2^{-k}} T_{j2^{-k}} \pi^{(i)}.$$

Then we have

$$\begin{aligned}
A^*A\pi^{(i)}_{k,j} &= A^*A2^{kd/2} D_{2^{-k}} T_{j2^{-k}} \pi^{(i)} \\
&= 2^{-2k\alpha} 2^{kd/2} D_{2^{-k}} A^*A T_{j2^{-k}} \pi^{(i)} && \text{(by Lemma 6.1.3)} \\
&= 2^{-2k\alpha} 2^{kd/2} D_{2^{-k}} T_{j2^{-k}} A^*A \pi^{(i)} && \text{(by Lemma 6.1.2)} \\
&= 2^{-2k\alpha} 2^{kd/2} D_{2^{-k}} T_{j2^{-k}} \widetilde{\psi}^{(i)} && \text{(by (6.3))} \\
&= 2^{-2k\alpha}\widetilde{\psi}^{(i)}_{k,j}.
\end{aligned}$$

Hence

$$A^*(2^{k\alpha}A\pi^{(i)}_{k,j}) = 2^{-k\alpha}\widetilde{\psi}^{(i)}_{k,j}. \tag{6.4}$$

Furthermore,

$$\begin{aligned}
[2^{k\alpha}A\pi^{(i)}_{k,j}, \widetilde{U}^{(i')}_{k',j'}] &= 2^{(k+k')\alpha}[A\pi^{(i)}_{k,j}, A\psi^{(i')}_{k',j'}] && \text{(by (3) of Definition 6.1.1)}\\
&= 2^{(k+k')\alpha}\langle A^*A\pi^{(i)}_{k,j}, \psi^{(i')}_{k',j'}\rangle\\
&= 2^{-(k-k')\alpha}\langle\widetilde{\psi}^{(i)}_{k,j}, \psi^{(i')}_{k',j'}\rangle && \text{(by (6.4))}\\
&= \delta_{(k,j,i),(k',j',i')}.
\end{aligned} \tag{6.5}$$

It implies that if $\{2^{k\alpha}A\pi^{(i)}_{k,j}\}$ and $\{\widetilde{U}^{(i)}_{k,j}\}$ form L^2-stable bases of $\mathcal{Y}$ simultaneously, then by setting

$$U^{(i)}_{k,j} = 2^{k\alpha}A\pi^{(i)}_{k,j}, \tag{6.6}$$

we have a wavelet-vaguelette system $\{\psi^{(i)}_{k,j}, \widetilde{\psi}^{(i)}_{k,j}, U^{(i)}_{k,j}, \widetilde{U}^{(i)}_{k,j}\}$ of A. It is not difficult to show that this construction is the only way of having a wavelet-vaguelette system of A for given biorthogonal wavelets $\{\psi^{(i)}_{k,j}\}$ and $\{\widetilde{\psi}^{(i)}_{k,j}\}$. From now on we use the right-hand side of (6.6) as the definition of $U^{(i)}_{k,j}$.

In [22] Donoho gave a sufficient condition for the existence of the wavelet-vaguelette system. We state it with a slight modification to be fitted into our setting.

THEOREM 6.1.4. (*Donoho*) *Let* $\widetilde{\psi}$ *and* ψ *be univariate functions described in Section* 3.1. *We assume that* $\widetilde{\psi}$ *has the* R*-th continuous derivative and* ψ *has vanishing moments of order* M. *If* $\alpha + d + 1 < \min(R, M)$, *then* $\{\psi^{(i)}_{k,j}, \widetilde{\psi}^{(i)}_{k,j}, U^{(i)}_{k,j}, \widetilde{U}^{(i)}_{k,j}\}$ *forms the wavelet-vaguelette system of* A.

This theorem demands high regularity and large vanishing moments for wavelets to be used for the wavelet-vaguelette system. For instance, $R \geq 4$ for the Radon transform in $\mathbb{R}^2$. Notice that none of first ten Daubechies's compactly supported orthogonal wavelets have this amount of regularity. Many symmetric biorthogonal wavelets with relatively small support size are not included in Theorem 6.1.4 to generate a wavelet-vaguelette system for the Radon transform, either. See, e.g., [5], [11] and [12].

There are several disadvantage in using smoother wavelets in approximating a function that is much less smooth than the given wavelets: First, as long as Theorem 3.2.1 holds, there is no gain in approximation order with smoother wavelets. Second, since smoother wavelets have wider support, they are not as good as less smooth wavelets, which have smaller support, in approximating functions with discontinuities.

As an alternative to Theorem 6.1.4, in Section 6.3 we shall give a sufficient condition that demands lower regularity order on $\widetilde{\psi}$ and fewer vanishing moments on ψ to generate a wavelet-vaguelette system of A.

6.2 Biorthogonal Riesz Bases

The collection of functions $\{\varphi_n\}$ in a separable Hilbert space $\mathcal{H}$ is said to be *near-orthogonal* if for any $(a_n) \in \ell^2$,

$$\left\| \sum_n a_n \varphi_n \right\|_{\mathcal{H}}^2 \asymp \sum_n |a_n|^2.$$

The concept of near-orthogonality is closely related to that of L^2-stability in the following way:

LEMMA 6.2.1. *Let $\mathcal{H}$ be a separable Hilbert space. The collections of functions $\{\varphi_n\}$ and $\{\widetilde{\varphi}_n\}$ are biorthogonal Riesz bases of $\mathcal{H}$ if and only if*

(1) $\langle \varphi_m, \widetilde{\varphi}_n \rangle = \delta_{m,n}$,

(2) $\|\varphi\|_{\mathcal{H}}^2 = \sum_n \langle \varphi, \widetilde{\varphi}_n \rangle \langle \varphi_n, \varphi \rangle$ *for all* $\varphi \in \mathcal{H}$,

(3) *both* $\{\varphi_n\}$ *and* $\{\widetilde{\varphi}_n\}$ *are near-orthogonal.*

Proof. Without loss of generality, we can assume that the index set of n in $\{\varphi_n\}$ and $\{\widetilde{\varphi}_n\}$ is the set of positive integers, since $\mathcal{H}$ is a separable Hilbert space.

$(\rightarrow)$: The proof of (1) immediately follows since $\{\varphi_n\}$ and $\{\widetilde{\varphi}_n\}$ are biorthogonal Riesz bases.

Let $\varphi \in \mathcal{H}$. As being a Riesz basis of $\mathcal{H}$, $\{\varphi_n\}$ represents φ with a sequence (a_n) as $\varphi = \sum_n a_n \varphi_n$. Notice that $\langle \varphi, \widetilde{\varphi}_n \rangle = a_n$ by biorthogonality between $\{\varphi_n\}$ and $\{\widetilde{\varphi}_n\}$. Thus we can represent any $\varphi \in \mathcal{H}$ as

$$\varphi = \sum_n \langle \varphi, \widetilde{\varphi}_n \rangle \varphi_n.$$

In particular,

$$\langle \varphi, \varphi \rangle = \sum_n \langle \varphi, \widetilde{\varphi}_n \rangle \langle \varphi_n, \varphi \rangle.$$

Let $(a_n) \in \ell^2$. Notice that by L^2-stability of $\{\widetilde{\varphi}_n\}$, one can have

$$\text{(a)} \qquad \begin{aligned} \Big\| \sum_{n \le N} a_n \varphi_n \Big\|_{\mathcal{H}}^2 &\asymp \sum_{m \le N} \Big| \langle \sum_{n \le N} a_n \varphi_n, \widetilde{\varphi}_m \rangle \Big|^2 \\ &= \sum_{m \le N} |a_m|^2 \end{aligned}$$

for any positive integer N. Since the equivalence relation in (a) does not depend on N, by letting $N \to \infty$, one can show that

$$\Big\| \sum_n a_n \varphi_n \Big\|_{\mathcal{H}}^2 \asymp \sum_n |a_n|^2.$$

Thus $\{\varphi_n\}$ is near-orthogonal. With a similar argument, we can also show that $\{\widetilde{\varphi}_n\}$ is near-orthogonal.

$(\leftarrow)$: Let

$$\mathcal{H}' = \Big\{ \sum_n a_n \varphi_n \mid (a_n) \in \ell^2 \Big\}.$$

Suppose $\varphi^N = \sum_n a_n^N \varphi_n \to \varphi$ in $\mathcal{H}$ as $N \to \infty$, where $a^N = (a_n^N) \in \ell^2$. Then near-orthogonality of $\{\varphi_n\}$ implies that there exists a sequence $a = (a_n) \in \ell^2$ such that $a^N \to a$ in ℓ^2. Let $\varphi^* = \sum_n a_n \varphi_n$. Then for any N,

$$\|\varphi - \varphi^*\|_{\mathcal{H}} \le \|\varphi - \varphi^N\|_{\mathcal{H}} + \|\varphi^N - \varphi^*\|_{\mathcal{H}}.$$

As $N \to \infty$, $\|\varphi - \varphi^N\|_{\mathcal{H}} \to 0$ by given hypothesis, and $\|\varphi^N - \varphi^*\|_{\mathcal{H}} \to 0$, since $\|\varphi^N - \varphi^*\|_{\mathcal{H}}^2 \asymp \sum_n |a_n^N - a_n|^2$ and $a^N \to a$ in ℓ^2. Thus $\varphi = \varphi^* \in \mathcal{H}'$. We have shown that every limit function of convergent functions in $\mathcal{H}'$ belongs to $\mathcal{H}'$. Therefore $\mathcal{H}'$ is closed in $\mathcal{H}$.

Since $\mathcal{H}'$ is closed in $\mathcal{H}$, one can decompose $\mathcal{H}$ by

$$\text{(b)} \qquad \mathcal{H} = \mathcal{H}' \oplus \mathcal{H}'',$$

where $\mathcal{H}''$ is the orthogonal complement of $\mathcal{H}'$ in $\mathcal{H}$. Notice that for any $\varphi'' \in \mathcal{H}''$,

$$\begin{aligned} \|\varphi''\|^2 &= \sum_n \langle \varphi'', \widetilde{\varphi}_n \rangle \langle \varphi_n, \varphi'' \rangle && \text{(by hypothesis)} \\ &= 0. && \text{(since } \langle \varphi'', \varphi_n \rangle = 0) \end{aligned}$$

It implies that $\mathcal{H}'' = \{0\}$, and hence $\{\varphi_n\}$ spans $\mathcal{H}$. With a similar argument, we can also show that $\{\widetilde{\varphi}_n\}$ spans $\mathcal{H}$.

Let $\varphi \in \mathcal{H}$. Since $\{\varphi_n\}$ spans $\mathcal{H}$, one can find a sequence (a_n) such that $\varphi = \sum_n a_n \varphi_n$. Moreover, biorthogonality between $\{\varphi_n\}$ and $\{\widetilde{\varphi}_n\}$ implies that $a_n = \langle \varphi, \widetilde{\varphi}_n \rangle$. Thus we can represent φ as

$$\varphi = \sum_n \langle \varphi, \widetilde{\varphi}_n \rangle \varphi_n.$$

Since $\{\varphi_n\}$ is near orthogonal, one can show that

$$\|\varphi\|_{\mathcal{H}}^2 \asymp \sum_n |\langle \varphi, \widetilde{\varphi}_n \rangle|^2.$$

Thus $\{\widetilde{\varphi}_n\}$ is L^2-stable. With a similar argument, we can also show that $\{\varphi_n\}$ is L^2-stable. Therefore $\{\varphi_n\}$ and $\{\widetilde{\varphi}_n\}$ are biorthogonal Riesz bases of $\mathcal{H}$. ∎

While proving Lemma 6.2.1, we have shown that any $\varphi \in \mathcal{H}$ can be written as

$$\varphi = \sum_n \langle \varphi, \widetilde{\varphi}_n \rangle \varphi_n,$$

if $\{\varphi_n\}$ and $\{\widetilde{\varphi}_n\}$ are biorthogonal Riesz bases of $\mathcal{H}$.

Lemma 6.2.2. *Let* $\{\varrho_n\}$ *and* $\{\widetilde{\varrho}_n\}$ *be collections of functions in a separable Hilbert space* $\mathcal{S}$ *equipped with the inner product* $(\cdot,\cdot)$*. We assume that*

$$(\varrho_m, \widetilde{\varrho}_n) = \delta_{m,n} \qquad \textit{and} \qquad \|\varrho\|_{\mathcal{S}}^2 = \sum_n (\varrho, \widetilde{\varrho}_n)(\varrho_n, \varrho)$$

for all $\varrho \in \mathcal{S}$*. Let collections of functions* $\{\varphi_n\}$ *and* $\{\widetilde{\varphi}_n\}$ *be biorthogonal Riesz bases of a separable Hilbert space* $\mathcal{H}$*. If*

$$(\varrho_m, \varrho_n) = \langle \varphi_m, \varphi_n \rangle \qquad \textit{and} \qquad (\widetilde{\varrho}_m, \widetilde{\varrho}_n) = \langle \widetilde{\varphi}_m, \widetilde{\varphi}_n \rangle$$

for all m, n*, then* $\{\varrho_n\}$ *and* $\{\widetilde{\varrho}_n\}$ *are biorthogonal Riesz bases of* $\mathcal{S}$*.*

Proof. Without loss of generality, we can assume that the index set of n in $\{\varrho_n\}$ and $\{\widetilde{\varrho}_n\}$ is the set of positive integers, since $\mathcal{S}$ is a separable Hilbert space.

Let $(a_n) \in \ell^2$. Notice that for any positive integer N,

$$\text{(a)} \quad \begin{aligned} (\sum_{n\le N} a_n \widetilde{\varrho}_n, \sum_{m\le N} a_m \widetilde{\varrho}_m) &= \sum_{n\le N}\sum_{m\le N} a_n (\widetilde{\varrho}_n, \widetilde{\varrho}_m)\overline{a_m} \\ &= \sum_{n\le N}\sum_{m\le N} a_n \langle \widetilde{\varphi}_n, \widetilde{\varphi}_m\rangle \overline{a_m} \qquad \text{(by hypothesis)} \\ &= \langle \sum_{n\le N} a_n \widetilde{\varphi}_n, \sum_{m\le N} a_m \widetilde{\varphi}_m \rangle \\ &\asymp \sum_{n\le N} |a_n|^2. \qquad \text{(by Lemma 6.2.1)} \end{aligned}$$

Since the equivalence in (a) does not depend on N, by letting $N \to \infty$, one can show that for all $(a_n) \in \ell^2$,

$$\left\| \sum_n a_n \widetilde{\varrho}_n \right\|_{\mathcal{S}}^2 \asymp \sum_n |a_n|^2.$$

Thus $\{\widetilde{\varrho}_n\}$ is near-orthogonal. With a similar argument, we can also show that $\{\varrho_n\}$ is near-orthogonal. Moreover, by hypothesis, $\{\varrho_n\}$ and $\{\widetilde{\varrho}_n\}$ is biorthogonal and $\|\varrho\|_{\mathcal{S}}^2 = \sum_n (\varrho, \widetilde{\varrho}_n)(\varrho_n, \varrho)$. Thus, by Lemma 6.2.1, $\{\varrho_n\}$ and $\{\widetilde{\varrho}_n\}$ are biorthogonal Riesz bases of $\mathcal{S}$. $\square$

Lemma 6.2.3. *Let $\{\varphi_n\}$ and $\{\widetilde{\varphi}_n\}$ be collections of functions in a separable Hilbert space $\mathcal{H}$. We assume that $\|\varphi\|_{\mathcal{H}}^2 = \sum_n \langle \varphi, \widetilde{\varphi}_n \rangle \langle \varphi_n, \varphi \rangle$ for all $\varphi \in \mathcal{H}$. If for all $\varphi \in \mathcal{H}$,*

$$\sum_n |\langle \varphi, \varphi_n \rangle|^2 \leq C_1 \|\varphi\|_{\mathcal{H}}^2$$

and

$$\sum_n |\langle \varphi, \widetilde{\varphi}_n \rangle|^2 \leq C_2 \|\varphi\|_{\mathcal{H}}^2$$

for positive constants C_1 and C_2, then

$$C_2^{-1} \|\varphi\|_{\mathcal{H}}^2 \leq \sum_n |\langle \varphi, \varphi_n \rangle|^2$$

and

$$C_1^{-1} \|\varphi\|_{\mathcal{H}}^2 \leq \sum_n |\langle \varphi, \widetilde{\varphi}_n \rangle|^2$$

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

Proof. Let $\varphi \in \mathcal{H}$. Notice that by hypothesis, we have

$$\begin{aligned} \|\varphi\|_{\mathcal{H}}^2 &= \sum_n \langle \varphi, \widetilde{\varphi}_n \rangle \langle \varphi_n, \varphi \rangle \\ &\leq \left(\sum_n |\langle \varphi, \widetilde{\varphi}_n \rangle|^2 \right)^{1/2} \left(\sum_n |\langle \varphi, \varphi_n \rangle|^2 \right)^{1/2} && \text{(by (2.11))} \\ &\leq C_2^{1/2} \|\varphi\|_{\mathcal{H}} \left(\sum_n |\langle \varphi, \varphi_n \rangle|^2 \right)^{1/2}. && \text{(by hypothesis)} \end{aligned}$$

Thus we have

$$C_2^{-1} \|\varphi\|_{\mathcal{H}}^2 \leq \sum_n |\langle \varphi, \varphi_n \rangle|^2.$$

With a similar argument, we can also prove that

$$C_1^{-1} \|\varphi\|_{\mathcal{H}}^2 \leq \sum_n |\langle \varphi, \widetilde{\varphi}_n \rangle|^2$$

for all $\varphi \in \mathcal{H}$. ∎

6.3 Sufficient Condition

We assume that ψ and $\widetilde{\psi}$ satisfy

$$
\begin{aligned}
|\widehat{\psi}(w)| &\le C|w|^{a_1}(1+|w|^2)^{-(b_1+a_1)/2}, \\
|\widehat{\widetilde{\psi}}(w)| &\le C|w|^{a_2}(1+|w|^2)^{-(b_2+a_2)/2},
\end{aligned}
\tag{6.7}
$$

for some $a_1 > 0$, $a_2 > 0$, $b_1 > 1/2$, and $b_2 > 1/2$. The conditions $b_1 > 1/2$ and $b_2 > 1/2$ insure that $\psi \in L^2(\mathbb{R})$ and $\widetilde{\psi} \in L^2(\mathbb{R})$.

The constants b_1 and b_2 in (6.7) are closely related to regularity of ψ and $\widetilde{\psi}$, respectively. We note that for any integer n such that $0 \le n < b_1 - 1/2$, $(2\pi i w)^n \widehat{\widetilde{\psi}}(w) \in L^2(\mathbb{R})$. Thus one can consider the inverse Fourier transform of $(2\pi i w)^n \widehat{\widetilde{\psi}}(w)$, and by (2.6), it is the n-th derivative $\widetilde{\psi}^{(n)}$ of $\widetilde{\psi}$. Thus larger b_1 and b_2 in (6.7) imply smoother ψ and $\widetilde{\psi}$, respectively.

On the other hand, the constants a_1 and a_2 in (6.7) are closely related to the number of vanishing moments of ψ and $\widetilde{\psi}$, respectively. Notice that by (2.7),

$$
\frac{d^n \widehat{\psi}}{dw^n}(w) = \int_{\mathbb{R}} e^{-2\pi i w x}(-2\pi i x)^n \psi(x)\, dx.
$$

Thus if ψ has vanishing moments of order M, i.e.,

$$
\int_{\mathbb{R}} x^n \psi(x)\, dx = 0, \qquad \text{whenever } n = 0, 1, \cdots, M,
$$

then

$$
\frac{d^n \widehat{\psi}}{dw^n}(0) = 0, \qquad \text{whenever } n = 0, 1, \cdots, M.
$$

It implies that $\widehat{\psi}(w)$ can be written as

$$
\widehat{\psi}(w) = |w|^M \widehat{g}(w), \tag{6.8}
$$

for some $g \in L^2(\mathbb{R})$. Conversely, if ψ satisfies (6.8), then ψ necessarily has vanishing moments of order M. Hence, the conditions $a_1 > 0$ and $a_2 > 0$ in (6.7) insure that

ψ and $\widetilde{\psi}$ have at least vanishing moments of order 0, respectively. Obviously, larger a_1 and a_2 in (6.7) imply ψ and $\widetilde{\psi}$ with higher vanishing moments, respectively.

Since ψ and $\widetilde{\psi}$ are finite linear combinations of $\{\phi_{1,n}\}$ and $\{\widetilde{\phi}_{1,n}\}$, respectively, we naturally assume that

$$(6.9) \qquad \begin{aligned} |\widehat{\phi}(w)| &\le C(1+|w|^2)^{-b_1/2}, \qquad \text{and} \\ |\widehat{\widetilde{\phi}}(w)| &\le C(1+|w|^2)^{-b_2/2}, \end{aligned}$$

with the same b_1 and b_2 in (6.7). Notice that these decay conditions (6.9) also hold for ψ and $\widetilde{\psi}$ accordingly.

As we mentioned earlier, this section is devoted to find a sufficient condition on ψ and $\widetilde{\psi}$, in terms of their regularity and vanishing moments, to generate a wavelet-vaguelette system $\{\psi_{k,j}^{(i)}, \widetilde{\psi}_{k,j}^{(i)}, U_{k,j}^{(i)}, \widetilde{U}_{k,j}^{(i)}\}$ of A. We shall find regions for a_1, a_2, b_1, and b_2 in (6.7) that insure the existence of the wavelet-vaguelette system $\{\psi_{k,j}^{(i)}, \widetilde{\psi}_{k,j}^{(i)}, U_{k,j}^{(i)}, \widetilde{U}_{k,j}^{(i)}\}$ of A, more specifically, L^2-stability of $\{U_{k,j}^{(i)}\}$ and $\{\widetilde{U}_{k,j}^{(i)}\}$, from ψ and $\widetilde{\psi}$ satisfying (6.7).

We begin with $\theta^{(i)}$ and $\widetilde{\theta}^{(i)}$ such that

$$\widehat{\theta^{(i)}}(\xi) = |\xi|^{\alpha}\widehat{\widetilde{\psi}^{(i)}}(\xi) \qquad \text{and} \qquad \widehat{\widetilde{\theta}^{(i)}}(\xi) = |\xi|^{-\alpha}\widehat{\psi^{(i)}}(\xi)$$

for $i = 1, \cdots, 2^d - 1$. We define

$$\theta_{k,j}^{(i)} = 2^{kd/2}\theta^{(i)}(2^k \cdot -j),$$

and similarly for $\widetilde{\theta}_{k,j}^{(i)}$. Notice that by (2.8), one can have

$$(6.10) \qquad \begin{aligned} \widehat{\theta_{k,j}^{(i)}}(\xi) &= 2^{-kd/2}e^{-2\pi i\xi\cdot j2^{-k}}\widehat{\theta^{(i)}}(2^{-k}\xi) \\ &= 2^{-k(d/2+\alpha)}e^{-2\pi i\xi\cdot j2^{-k}}|\xi|^{\alpha}\widehat{\widetilde{\psi}^{(i)}}(2^{-k}\xi), \end{aligned}$$

and similarly

$$(6.11) \qquad \widehat{\pi_{k,j}^{(i)}}(\xi) = 2^{-k(d/2+2\alpha)}e^{-2\pi i\xi\cdot j2^{-k}}|\xi|^{2\alpha}\widehat{\widetilde{\psi}^{(i)}}(2^{-k}\xi).$$

Using (6.10) and (6.11), we can compute $\langle \theta^{(i)}_{k,j}, \theta^{(i')}_{k',j'}\rangle$ by

$$\begin{aligned} &2^{-(k+k')(d/2+\alpha)}\int_{\mathbb{R}^d} e^{-2\pi i\xi\cdot(j2^{-k}-j'2^{-k'})}|\xi|^{\alpha}\widehat{\widetilde{\psi}^{(i)}}(2^{-k}\xi)|\xi|^{\alpha}\overline{\widehat{\widetilde{\psi}^{(i')}}(2^{-k'}\xi)}\,d\xi \\ &\quad = 2^{-(k-k')\alpha}\int_{\mathbb{R}^d}\widehat{\widetilde{\psi}^{(i)}_{k,j}}(\xi)\overline{\widehat{\pi^{(i')}_{k',j'}}(\xi)}\,d\xi \\ &\quad = 2^{-(k-k')\alpha}\langle \widetilde{\psi}^{(i)}_{k,j}, \pi^{(i')}_{k',j'}\rangle. \end{aligned}$$

Since $\widetilde{\psi}^{(i)}_{k,j} = 2^{2k\alpha}A^*A\pi^{(i)}_{k,j}$ (see (6.4)), one has

$$\begin{aligned} \langle \theta^{(i)}_{k,j}, \theta^{(i')}_{k',j'}\rangle &= 2^{(k+k')\alpha}\langle A^*A\pi^{(i)}_{k,j}, \pi^{(i')}_{k',j'}\rangle \\ &= [U^{(i)}_{k,j}, U^{(i')}_{k',j'}]. \qquad \text{(by (6.6))} \end{aligned} \tag{6.12}$$

With a similar argument, we have

$$\langle \widetilde{\theta}^{(i)}_{k,j}, \widetilde{\theta}^{(i')}_{k',j'}\rangle = [\widetilde{U}^{(i)}_{k,j}, \widetilde{U}^{(i')}_{k',j'}] \tag{6.13}$$

and

$$\langle \theta^{(i)}_{k,j}, \widetilde{\theta}^{(i')}_{k',j'}\rangle = [U^{(i)}_{k,j}, \widetilde{U}^{(i')}_{k',j'}]. \tag{6.14}$$

Notice that for any $f \in L^2(\mathbb{R}^d)$,

$$\begin{aligned} \sum_{k,j,i}[Af, U^{(i)}_{k,j}][\widetilde{U}^{(i)}_{k,j}, Af] &= \sum_{k,j,i}\langle f, A^*U^{(i)}_{k,j}\rangle[2^{k\alpha}A\psi^{(i)}_{k,j}, Af] \\ &= \sum_{k,j,i}\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle\langle \psi^{(i)}_{k,j}, A^*Af\rangle \qquad \text{(by (6.4) and (6.6))} \\ &= \langle f, A^*Af\rangle = \|Af\|^2_{\mathcal{Y}}. \end{aligned}$$

Since $A(L^2(\mathbb{R}^d)) = \mathcal{Y}$, we have

$$\|Y\|^2_{\mathcal{Y}} = \sum_{k,j,i}[Y, U^{(i)}_{k,j}][\widetilde{U}^{(i)}_{k,j}, Y] \tag{6.15}$$

for any $Y \in \mathcal{Y}$.

We now summarize what we have proved so far: The collections of functions $\{U_{k,j}^{(i)}\}$ and $\{\widetilde{U}_{k,j}^{(i)}\}$ of $\mathcal{Y}$ satisfy

(1) $[U_{k,j}^{(i)}, \widetilde{U}_{k',j'}^{(i')}] = \delta_{(k,j,i),(k',j',i')}$, ((6.5), (6.6))

(2) $\|Y\|_{\mathcal{Y}}^2 = \sum_{k,j,i}[Y, U_{k,j}^{(i)}][\widetilde{U}_{k,j}^{(i)}, Y]$ for all $Y \in \mathcal{Y}$, ((6.15)),

(3) $[U_{k,j}^{(i)}, U_{k',j'}^{(i')}] = \langle \theta_{k,j}^{(i)}, \theta_{k',j'}^{(i')} \rangle$ and $[\widetilde{U}_{k,j}^{(i)}, \widetilde{U}_{k',j'}^{(i')}] = \langle \widetilde{\theta}_{k,j}^{(i)}, \widetilde{\theta}_{k',j'}^{(i')} \rangle$. ((6.12), (6.13))

Therefore if $\{\theta_{k,j}^{(i)}\}$ and $\{\widetilde{\theta}_{k,j}^{(i)}\}$ are biorthogonal Riesz bases of $L^2(\mathbb{R}^d)$, then by Lemma 6.2.2, $\{U_{k,j}^{(i)}\}$ and $\{\widetilde{U}_{k,j}^{(i)}\}$ are biorthogonal Riesz bases of $\mathcal{Y}$. Thus our goal is finding regions for a_1, a_2, b_1, and b_2 in (6.7) that insure that $\{\theta_{k,j}^{(i)}\}$ and $\{\widetilde{\theta}_{k,j}^{(i)}\}$ are L^2-stable, simultaneously.

Let $\xi = (\xi_1, \cdots, \xi_d) \in \mathbb{R}^d$. From our definition of d-dimensional biorthogonal wavelets, for each $\widetilde{\psi}^{(i)}$ there is at least one n_i such that

$$\psi^{(i)}(x_1, \cdots, x_{n_i}, \cdots, x_d) = \varphi_1(x_1) \cdots \psi(x_{n_i}) \cdots \varphi_d(x_d),$$

where φ_n is $\widetilde{\psi}$ or $\widetilde{\phi}$ for $n \neq n_i$. Notice that

$$|\widehat{\varphi_n}(\xi_n)| \leq C(1 + |\xi_n|^2)^{-b_2/2}, \tag{6.16}$$

whether $\varphi_n = \widetilde{\psi}$ or $\varphi_n = \widetilde{\phi}$. Using (6.16) for $n \neq n_i$ and (6.7) for $n = n_i$, we can bound $|\widehat{\theta^{(i)}}(\xi)|$ by

$$C|\xi|^{\alpha}(1 + |\xi_1|^2)^{-b_2/2} \cdots |\xi_{n_i}|^{a_2}(1 + |\xi_{n_i}|^2)^{-(b_2+a_2)/2} \cdots (1 + |\xi_d|^2)^{-b_2/2}.$$

Using the fact that $|\xi| \leq (1 + |\xi_1|^2)^{1/2} \cdots (1 + |\xi_d|^2)^{1/2}$, we have

$$\begin{aligned} &|\widehat{\theta^{(i)}}(\xi)| \\ &\leq C|\xi_{n_i}|^{a_2}(1 + |\xi_1|^2)^{-(b_2-\alpha)/2} \cdots (1 + |\xi_{n_i}|^2)^{-(b_2+a_2-\alpha)/2} \cdots (1 + |\xi_d|^2)^{-(b_2-\alpha)/2} \\ &\leq C|\xi_{n_i}|^{r_2}(1 + |\xi_1|^2)^{-(b_2-\alpha)/2} \cdots (1 + |\xi_{n_i}|^2)^{-(b_2+r_2-\alpha)/2} \cdots (1 + |\xi_d|^2)^{-(b_2-\alpha)/2} \end{aligned}$$

for any r_2 such that $0 \leq r_2 < a_2$. Thus we have

$$|\widehat{\theta^{(i)}}(\xi)| \leq C|\xi|^{r_2}(1 + |\xi_1|^2)^{-(b_2-\alpha)/2} \cdots (1 + |\xi_d|^2)^{-(b_2-\alpha)/2} \tag{6.17}$$

for any r_2 such that $0 \le r_2 < a_2$. In particular, if we take $r_2 = 0$ in (6.17), then we have

$$|\widehat{\theta^{(i)}}(\xi)| \le C\big(1+|\xi_1|^2\big)^{-(b_2-\alpha)/2} \cdots \big(1+|\xi_d|^2\big)^{-(b_2-\alpha)/2}. \tag{6.18}$$

For $\widetilde{\theta}^{(i)}$, we note that

$$\begin{aligned}
|\widehat{\widetilde{\theta}^{(i)}}(\xi)| &\le C|\xi|^{-\alpha}\big(1+|\xi_1|^2\big)^{-b_1/2} \cdots |\xi_{n_i}|^{a_1}\big(1+|\xi_{n_i}|^2\big)^{-(b_1+a_1)/2} \cdots \big(1+|\xi_d|^2\big)^{-b_1/2} \\
&\le C|\xi_{n_i}|^{a_1-\alpha}\big(1+|\xi_1|^2\big)^{-b_1/2} \cdots \big(1+|\xi_{n_i}|^2\big)^{-(b_1+a_1)/2} \cdots \big(1+|\xi_d|^2\big)^{-b_1/2} \\
&\le C|\xi_{n_i}|^{r_1}\big(1+|\xi_1|^2\big)^{-b_1/2} \cdots \big(1+|\xi_{n_i}|^2\big)^{-(b_1+\alpha+r_1)/2} \cdots \big(1+|\xi_d|^2\big)^{-b_1/2}
\end{aligned}$$

for any r_1 such that $0 \le r_1 < a_1 - \alpha$. Thus we have

$$|\widehat{\widetilde{\theta}^{(i)}}(\xi)| \le C|\xi|^{r_1}\big(1+|\xi_1|^2\big)^{-b_1/2} \cdots \big(1+|\xi_d|^2\big)^{-b_1/2} \tag{6.19}$$

for any r_1 such that $0 \le r_1 < a_1 - \alpha$. In particular, if we take $r_1 = 0$ in (6.19), then we have

$$|\widehat{\widetilde{\theta}^{(i)}}(\xi)| \le C\big(1+|\xi_1|^2\big)^{-b_1/2} \cdots \big(1+|\xi_d|^2\big)^{-b_1/2}. \tag{6.20}$$

From (6.18) and (6.20), it is obvious that $\theta^{(i)}$ and $\widetilde{\theta}^{(i)}$ are in $L^2(\mathbb{R}^d)$ if

$$b_1 > 1 \qquad \text{and} \qquad b_2 > \alpha + 1. \tag{6.21}$$

Moreover, if

$$a_1 > \alpha \qquad \text{and} \qquad a_2 > 0, \tag{6.22}$$

then $\theta^{(i)}$ and $\widetilde{\theta}^{(i)}$ have vanishing moments of at least order 0, since we can take positive r_1 and r_2 in (6.19) and (6.17), respectively. Obviously, conditions on b_1 and b_2 are more than enough to show that $\theta^{(i)}$ and $\widetilde{\theta}^{(i)}$ are well-defined. It turns out that stronger conditions (6.21) and (6.22) are sufficient to show that $\{\psi_{k,j}^{(i)}, \widetilde{\psi}_{k,j}^{(i)}, U_{k,j}^{(i)}, \widetilde{U}_{k,j}^{(i)}\}$ is a wavelet-vaguelette system of A. From now to the end of this section, we assume that a_1, a_2, b_1, and b_2 satisfy (6.21) and (6.22).

Notice that for $f \in L^2(\mathbb{R}^d)$,

$$\begin{aligned}\sum_{k,j,i} |\langle f, \theta_{k,j}^{(i)}\rangle|^2 &= \sum_{k,j,i} \left| \int_{\mathbb{R}^d} \widehat{f}(\xi) 2^{-kd/2} e^{2\pi i \xi \cdot j 2^{-k}} \overline{\widehat{\theta^{(i)}}(2^{-k}\xi)}\, d\xi \right|^2 \qquad \text{(by (2.8))}\\ &= \sum_{k,j,i} 2^{-kd} \left| \int_{I_k} e^{2\pi i \xi \cdot j 2^{-k}} \sum_l \widehat{f}(\xi + l2^k) \overline{\widehat{\theta^{(i)}}(2^{-k}\xi + l)}\, d\xi \right|^2,\end{aligned}$$

where $I_k = [0, 2^k]^d$. Since $\sum_l \widehat{f}(\xi + l2^k) \overline{\widehat{\theta^{(i)}}(2^{-k}\xi + l)}$ is periodic in I_k, one has from (2.9)

$$\begin{aligned}&\sum_j 2^{-kd} \left| \int_{I_k} e^{2\pi i \xi \cdot j 2^{-k}} \sum_l \widehat{f}(\xi + l2^k) \overline{\widehat{\theta^{(i)}}(2^{-k}\xi + l)}\, d\xi \right|^2\\ &\qquad = \int_{I_k} \left| \sum_l \widehat{f}(\xi + l2^k) \overline{\widehat{\theta^{(i)}}(2^{-k}\xi + l)} \right|^2 d\xi.\end{aligned}$$

Using this result, we have

$$\begin{aligned}&\sum_{k,j,i} |\langle f, \theta_{k,j}^{(i)}\rangle|^2\\ &\quad = \sum_{k,i} \int_{I_k} \left| \sum_l \widehat{f}(\xi + l2^k) \overline{\widehat{\theta^{(i)}}(2^{-k}\xi + l)} \right|^2 d\xi\\ &\quad = \sum_{k,i} \int_{I_k} \sum_l \widehat{f}(\xi + l2^k) \widehat{\theta^{(i)}}(2^{-k}\xi + l) \sum_n \overline{\widehat{f}(\xi + n2^k) \widehat{\theta^{(i)}}(2^{-k}\xi + n)}\, d\xi\\ &\quad = \sum_{k,n,i} \int_{\mathbb{R}^d} \widehat{f}(\xi) \overline{\widehat{f}(\xi + n2^k)} \widehat{\theta^{(i)}}(2^{-k}\xi) \overline{\widehat{\theta^{(i)}}(2^{-k}\xi + n)}\, d\xi.\end{aligned}$$

Let

$$C_\theta(\xi) = \sum_{k,i} |\theta^{(i)}(2^{-k}\xi)|^2 \tag{6.23}$$

and

$$R(f)=\sum_{n\neq(0,\cdots,0)}\sum_{k,i}\int_{\mathbb{R}^d}\widehat{f}(\xi)\overline{\widehat{f}(\xi+n2^k)\widehat{\theta^{(i)}}(2^{-k}\xi)\widehat{\theta^{(i)}}(2^{-k}\xi+n)}\,d\xi.$$

Then we have

$$\sum_{k,j,i}|\langle f,\theta^{(i)}_{k,j}\rangle|^2=\int_{\mathbb{R}^d}|\widehat{f}(\xi)|^2C_\theta(\xi)\,d\xi+R(f). \tag{6.24}$$

Notice that

$$\begin{aligned}|R(f)|&\le\sum_{n\neq(0,\cdots,0)}\sum_{k,i}\int_{\mathbb{R}^d}\left|\widehat{f}(\xi)\overline{\widehat{f}(\xi+n2^k)\widehat{\theta^{(i)}}(2^{-k}\xi)\widehat{\theta^{(i)}}(2^{-k}\xi+n)}\right|d\xi\\&\le\sum_{n\neq(0,\cdots,0)}\sum_{k,i}\left(\int_{\mathbb{R}^d}|\widehat{f}(\xi)|^2|\widehat{\theta^{(i)}}(2^{-k}\xi)\widehat{\theta^{(i)}}(2^{-k}\xi+n)|\,d\xi\right)^{1/2}\\&\qquad\times\left(\int_{\mathbb{R}^d}|\widehat{f}(\xi')|^2|\widehat{\theta^{(i)}}(2^{-k}\xi')\widehat{\theta^{(i)}}(2^{-k}\xi'-n)|\,d\xi'\right)^{1/2};\end{aligned} \tag{6.25}$$

here we have used (2.12) and change of variables $\xi'=\xi-n2^k$ in the second factor. We now use (2.11) on the sum over k for the right hand side of the last equation of (6.25). Then

$$\begin{aligned}|R(f)|\le\sum_{n\neq(0,\cdots,0)}\sum_{i}&\left(\sum_k\int_{\mathbb{R}^d}|\widehat{f}(\xi)|^2|\widehat{\theta^{(i)}}(2^{-k}\xi)\widehat{\theta^{(i)}}(2^{-k}\xi+n)|\,d\xi\right)^{1/2}\\&\times\left(\sum_{k'}\int_{\mathbb{R}^d}|\widehat{f}(\xi')|^2|\widehat{\theta^{(i)}}(2^{-k'}\xi')\widehat{\theta^{(i)}}(2^{-k'}\xi'-n)|\,d\xi'\right)^{1/2}.\end{aligned} \tag{6.26}$$

Let

$$\nu_i(s)=\sup_\xi\sum_k|\widehat{\theta^{(i)}}(2^{-k}\xi)\widehat{\theta^{(i)}}(2^{-k}\xi+s)|.$$

Then from (6.26), we have

$$\begin{aligned}|R(f)|&\le\sum_{n\neq(0,\cdots,0)}\sum_i\left(\nu_i(n)\int_{\mathbb{R}^d}|\widehat{f}(\xi)|^2d\xi\right)^{1/2}\left(\nu_i(-n)\int_{\mathbb{R}^d}|\widehat{f}(\xi')|^2d\xi'\right)^{1/2}\\&\le\|f\|^2_{L^2}\sum_{n\neq(0,\cdots,0)}\sum_i\nu_i(n)^{1/2}\nu_i(-n)^{1/2}.\end{aligned}$$

Putting above result and (6.24) together, we have

$$\sum_{k,j,i} |\langle f, \theta_{k,j}^{(i)} \rangle|^2 \le \|f\|_{L^2}^2 \left(\sup_{\xi} C_\theta(\xi) + \sum_{n \neq (0,\cdots,0)} \sum_i \nu_i(n)^{1/2} \nu_i(-n)^{1/2} \right). \tag{6.27}$$

We now use the decay condition of $\widehat{\theta^{(i)}}$ in (6.17) to estimate the right hand side of (6.27). Notice that from (6.23), we can have

$$C_\theta(\xi) \le \sup_{1 \le |\xi_r| \le 2} \sum_{k,i} |\widehat{\theta^{(i)}}(2^{-k}\xi)|^2.$$

Using (6.17) for $k < 0$ with $r_2 = 0$ and for $k \ge 0$ with r_2^* such that $0 < r_2^* < a_2$, we have

$$\begin{aligned} C_\theta(\xi) \le C \sup_{1 \le |\xi_r| \le 2} \Bigg(&\sum_{k=-\infty}^{-1} (1 + |2^{-k}\xi_1|^2)^{-(b_2-\alpha)} \cdots (1 + |2^{-k}\xi_d|^2)^{-(b_2-\alpha)} \\ &+ \sum_{k=0}^{\infty} |2^{-k}\xi|^{2r_2^*} (1 + |2^{-k}\xi_1|^2)^{-(b_2-\alpha)} \cdots (1 + |2^{-k}\xi_d|^2)^{-(b_2-\alpha)} \Bigg) \\ \le C' \Bigg(\sum_{k=-\infty}^{-1} 2^{2k(b_2-\alpha)d} + \sum_{k=0}^{\infty} 2^{-2kr_2^*} \Bigg). \end{aligned}$$

Since $b_2 > \alpha + 1$, the first term is finite. Thus we have

$$\sup_{\xi} C_\theta(\xi) \le C(r_2^*) < \infty, \tag{6.28}$$

for any r_2^* such that $0 < r_2^* < a_2$.

For ν_i in (6.27), we note that

$$\nu_i(s) = \sup_{1 \le |\xi_n| \le 2} \sum_k |\widehat{\theta^{(i)}}(2^{-k}\xi) \widehat{\theta^{(i)}}(2^{-k}\xi + s)|.$$

To have an upper bound of $|\widehat{\theta^{(i)}}(2^{-k}\xi)|$, we use (6.17) for $k < 0$ with $r_2 = 0$ and for $k \ge 0$ with same r_2^* used for estimating $C_\theta(\xi)$. For $|\widehat{\theta^{(i)}}(2^{-k}\xi + s)|$, we use (6.17) with $r_2 = 0$ for any k. Then

$$\nu_i(s) \le C \sup_{1 \le |\xi_n| \le 2} \left(\sum_{k=-\infty}^{-1} T_1 \cdot T_2 + \sum_{k=0}^{\infty} T_3 \cdot T_2 \right),$$

where

$$T_1 = (1+|2^{-k}\xi_1|^2)^{-(b_2-\alpha)/2} \cdots (1+|2^{-k}\xi_d|^2)^{-(b_2-\alpha)/2},$$
$$T_2 = (1+|2^{-k}\xi_1+s_1|^2)^{-(b_2-\alpha)/2} \cdots (1+|2^{-k}\xi_d+s_d|^2)^{-(b_2-\alpha)/2},$$
$$T_3 = |2^{-k}\xi|^{r_2^*}(1+|2^{-k}\xi_1|^2)^{-(b_2-\alpha)/2} \cdots (1+|2^{-k}\xi_d|^2)^{-(b_2-\alpha)/2}.$$

Notice that for any ϵ_2 such that $0 < \epsilon_2 < b_2 - \alpha$, we can have

$$T_2 \le (1+|2^{-k}\xi_1+s_1|^2)^{-(b_2-\alpha-\epsilon_2)/2} \cdots (1+|2^{-k}\xi_d+s_d|^2)^{-(b_2-\alpha-\epsilon_2)/2}. \tag{6.29}$$

We now consider the following inequality; for any x and y in $\mathbb{R}$,

$$\frac{1+y^2}{1+(x-y)^2} \le 1+(x+y)^2, \tag{6.30}$$

since for any x and y, either $y^2 \le (x-y)^2$ or $y^2 \le (x+y)^2$.

For each $(1+|2^{-k}\xi_r+s_r|^2)$ in (6.29), we apply (6.30) with $x = 2^{-k}\xi_r + s_r/2$ and $y = s_r/2$. Then

$$(1+s_r^4/4)(1+|2^{-k}\xi_r|^2)^{-1} \le 1+|2^{-k}\xi_r+s_r|^2. \tag{6.31}$$

Using (6.29) and (6.31), we can bound T_2 by T_2^* such that

$$\begin{aligned} T_2^* &= (1+s_1^2/4)^{-(b_2-\alpha-\epsilon_2)/2} \cdots (1+s_d^2/4)^{-(b_2-\alpha-\epsilon_2)/2} \\ &\quad \times (1+|2^{-k}\xi_1|^2)^{(b_2-\alpha-\epsilon_2)/2} \cdots (1+|2^{-k}\xi_d|^2)^{(b_2-\alpha-\epsilon_2)/2}. \end{aligned} \tag{6.32}$$

With (6.32), we have

$$\nu_i(s) \le C \sup_{1\le|\xi_n|\le 2} \left(\sum_{k=-\infty}^{-1} T_1 \cdot T_2^* + \sum_{k=0}^{\infty} T_3 \cdot T_2^* \right). \tag{6.33}$$

Notice that

$$\begin{aligned} T_1 \cdot T_2^* &= (1+s_1^2/4)^{-(b_2-\alpha-\epsilon_2)/2} \cdots (1+s_d^2/4)^{-(b_2-\alpha-\epsilon_2)/2} \\ &\quad \times (1+|2^{-k}\xi_1|^2)^{-\epsilon_2/2} \cdots (1+|2^{-k}\xi_d|^2)^{-\epsilon_2/2} \end{aligned} \tag{6.34}$$

and

$$T_3 \cdot T_2^* = (1+s_1^2/4)^{-(b_2-\alpha-\epsilon_2)/2} \cdots (1+s_d^2/4)^{-(b_2-\alpha-\epsilon_2)/2} \times |2^{-k}\xi|^{r_2^*} \left(1+|2^{-k}\xi_1|^2\right)^{-\epsilon_2/2} \cdots \left(1+|2^{-k}\xi_d|^2\right)^{-\epsilon_2/2}. \tag{6.35}$$

Plugging (6.34) and (6.35) in (6.33), we have

$$\begin{aligned}
\nu_i(s) &\le C(1+|s_1|^2/4)^{-(b_2-\alpha-\epsilon_2)/2} \cdots (1+|s_d|^2/4)^{-(b_2-\alpha-\epsilon_2)/2} \\
&\quad \times \sup_{1\le|\xi_n|\le 2} \Bigg(\sum_{k=-\infty}^{-1} \left(1+|2^{-k}\xi_1|^2\right)^{-\epsilon_2/2} \cdots \left(1+|2^{-k}\xi_1|^2\right)^{-\epsilon_2/2} \\
&\quad + \sum_{k=0}^{\infty} |2^{-k}\xi|^{r_2^*} \left(1+|2^{-k}\xi_1|^2\right)^{-\epsilon_2/2} \cdots \left(1+|2^{-k}\xi_1|^2\right)^{-\epsilon_2/2} \Bigg) \\
&\le C'(1+|s_1|^2/4)^{-(b_2-\alpha-\epsilon_2)/2} \cdots (1+|s_d|^2/4)^{-(b_2-\alpha-\epsilon_2)/2} \\
&\quad \times \left(\sum_{k=-\infty}^{-1} 2^{k\epsilon_2 d} + \sum_{k=0}^{\infty} 2^{-kr_2^*} \right).
\end{aligned}$$

Thus if $0 < \epsilon_2 < b_2 - \alpha$ and $0 < r_2^* < a_2$, then

$$\nu_i(s) \le C(r_2^*, \epsilon_2)(1+|s_1|^2/4)^{-(b_2-\alpha-\epsilon_2)/2} \cdots (1+|s_d|^2/4)^{-(b_2-\alpha-\epsilon_2)/2} \tag{6.36}$$

for a constant $C(r_2^*, \epsilon_2)$.

With (6.36), we have

$$\begin{aligned}
&\sum_{n\neq(0,\cdots,0)} \sum_i \nu_i(n)^{1/2}\nu_i(-n)^{1/2} \\
&\quad \le C'(r_2^*,\epsilon_2)(2^d-1) \sum_{n_1,\cdots,n_d} (1+n_1^2/4)^{-(b_2-\alpha-\epsilon_2)/2} \cdots (1+n_d^2/4)^{-(b_2-\alpha-\epsilon_2)/2} \\
&\quad \le C''(r_2^*,\epsilon_2) \left(\sum_{m=-\infty}^{\infty} (1+m^2/4)^{-(b_2-\alpha-\epsilon_2)/2} \right)^d.
\end{aligned}$$

Therefore if we take ϵ_2^* such that $b_2 - \alpha - \epsilon_2^* > 1$, then we have

$$\sum_{n\neq(0,\cdot,0)} \sum_i \nu_i(n)^{1/2}\nu_i(-n)^{1/2} \le C'''(r_2^*, \epsilon_2^*) < \infty. \tag{6.37}$$

Combining (6.27), (6.28), and (6.37), we have

$$\sum_{k,j,i} \left|\langle f, \theta^{(i)}_{k,j}\rangle\right|^2 \le C^*(r_2^*, \epsilon_2^*)\|f\|^2_{L^2},$$

where $C^*(r_2^*, \epsilon_2^*) < \infty$, if $0 < r_2^* < a_2$ and $0 < \epsilon_2^* < b_2 - \alpha - 1$.

Since in estimating $\sum_{k,j,i} \left|\langle f, \theta^{(i)}_{k,j}\rangle\right|^2$ we have used only the decay condition on $\widehat{\theta^{(i)}}$ in (6.17), we can easily estimate $\sum_{k,j,i} \left|\langle f, \widetilde{\theta}^{(i)}_{k,j}\rangle\right|^2$ simply by replacing $b_2 - \alpha$ by b_1 and a_2 by $a_1 - \alpha$ in each step. To be more specific, we have

$$\sum_{k,j,i} \left|\langle f, \widetilde{\theta}^{(i)}_{k,j}\rangle\right|^2 \le \widetilde{C}^*(r_1^*, \epsilon_1^*)\|f\|^2_{L^2},$$

where $\widetilde{C}^*(r_1^*, \epsilon_1^*) < \infty$, if $0 < r_1^* < a_1 - \alpha$ and $0 < \epsilon_1^* < b_1 - 1$.

Notice that for any $f \in L^2(\mathbb{R}^d)$,

$$\begin{aligned} \|f\|^2_{L^2} &= \int_{\mathbb{R}^d} \widehat{f}(\xi)\overline{\widehat{f}(\xi)}\, d\xi \\ &= \int_{\mathbb{R}^d} |\xi|^{\alpha}\widehat{f}(\xi) \cdot |\xi|^{-\alpha}\overline{\widehat{f}(\xi)}\, d\xi. \end{aligned} \tag{6.38}$$

For any f_1 and f_2, we can compute $\langle f_1, f_2\rangle$ by

$$\begin{aligned} \sum_{k,j,i}\langle f_1, \widetilde{\psi}^{(i)}_{k,j}\rangle\langle \psi^{(i)}_{k,j}, f_2\rangle = \sum_{k,j,i} &\int_{\mathbb{R}^d} \widehat{f_1}(\xi)2^{-kd/2}e^{2\pi i\xi\cdot j2^{-k}}\overline{\widehat{\widetilde{\psi}^{(i)}}(2^{-k}\xi)}\, d\xi \\ &\times \int_{\mathbb{R}^d} 2^{-kd/2}e^{-2\pi i\xi'\cdot j2^{-k}}\widehat{\psi^{(i)}}(2^{-k}\xi')\overline{\widehat{f_2}(\xi')}\, d\xi'. \end{aligned} \tag{6.39}$$

We apply (6.39) with $\widehat{f_1}(\xi) = |\xi|^{\alpha}\widehat{f}(\xi)$ and $\widehat{f_2}(\xi) = |\xi|^{-\alpha}\widehat{f}(\xi)$. Then from (6.38), we have

$$\begin{aligned} \|f\|^2_{L^2} = \sum_{k,j,i} &\int_{\mathbb{R}^d} |\xi|^{\alpha}\widehat{f}(\xi)2^{-kd/2}e^{2\pi i\xi\cdot j2^{-k}}\overline{\widehat{\widetilde{\psi}^{(i)}}(2^{-k}\xi)}\, d\xi \\ &\times \int_{\mathbb{R}^d} 2^{-kd/2}e^{-2\pi i\xi'\cdot j2^{-k}}\widehat{\psi^{(i)}}(2^{-k}\xi')|\xi'|^{-\alpha}\overline{\widehat{f}(\xi')}\, d\xi' \\ = \sum_{k,j,i} &\int_{\mathbb{R}^d} \widehat{f}(\xi)|2^{-k}\xi|^{\alpha}2^{-kd/2}e^{2\pi i\xi\cdot j2^{-k}}\overline{\widehat{\widetilde{\psi}^{(i)}}(2^{-k}\xi)}\, d\xi \\ &\times \int_{\mathbb{R}^d} 2^{-kd/2}e^{-2\pi i\xi'\cdot j2^{-k}}|2^{-k}\xi'|^{-\alpha}\widehat{\psi^{(i)}}(2^{-k}\xi')\overline{\widehat{f}(\xi')}\, d\xi'. \end{aligned}$$

Since

$$\widehat{\theta_{k,j}^{(i)}}(\xi) = 2^{-kd/2} e^{-2\pi i \xi \cdot j 2^{-k}} \left|2^{-k}\xi\right|^{\alpha} \widehat{\psi^{(i)}}(2^{-k}\xi)$$

and

$$\widehat{\widetilde{\theta}_{k,j}^{(i)}}(\xi) = 2^{-kd/2} e^{-2\pi i \xi \cdot j 2^{-k}} \left|2^{-k}\xi\right|^{-\alpha} \widehat{\psi^{(i)}}(2^{-k}\xi)$$

we have

$$\|f\|_{L^2}^2 = \sum_{k,j,i} \langle f, \theta_{k,j}^{(i)} \rangle \langle \widetilde{\theta}_{k,j}^{(i)}, f \rangle.$$

We now summarize what we proved for $\{\theta_{k,j}^{(i)}\}$ and $\{\widetilde{\theta}_{k,j}^{(i)}\}$:

(1) $\langle \theta_{k,j}^{(i)}, \widetilde{\theta}_{k',j'}^{(i')} \rangle = \delta_{(k,j,i),(k',j',i')}$.

(2) For any $f \in L^2(\mathbb{R}^d)$, $\|f\|_{L^2}^2 = \sum_{k,j,i} \langle f, \theta_{k,j}^{(i)} \rangle \langle \widetilde{\theta}_{k,j}^{(i)}, f \rangle$.

(3) If $a_1 > \alpha$, $b_1 > 1$, $a_2 > 0$, and $b_2 > \alpha + 1$, then

$$\sum_{k,j,i} \left|\langle f, \theta_{k,j}^{(i)} \rangle\right|^2 \leq C\|f\|_{L^2}^2 \qquad \text{and} \qquad \sum_{k,j,i} \left|\langle f, \widetilde{\theta}_{k,j}^{(i)} \rangle\right|^2 \leq \widetilde{C}\|f\|_{L^2}^2.$$

Therefore, by Lemma 6.2.3, $\{\theta_{k,j}^{(i)}\}$ and $\{\widetilde{\theta}_{k,j}^{(i)}\}$ are biorthogonal Riesz bases of $L^2(\mathbb{R}^d)$. Finally, we summarize what we proved in this section as follows:

THEOREM 6.3.1. *Let $\psi, \widetilde{\psi}, \phi$, and $\widetilde{\phi}$ be univariate functions described in Section 3.1. Let $\{\psi_{k,j}^{(i)}, \widetilde{\psi}_{k,j}^{(i)}, U_{k,j}^{(i)}, \widetilde{U}_{k,j}^{(i)}\}$ be the collections of functions described in Section 6.1. If*

$$\begin{aligned}
|\widehat{\psi}(w)| &\leq C|w|^{a_1}(1+|w|^2)^{-(b_1+a_1)/2}, && a_1 > \alpha \;, b_1 > 1, \\
|\widehat{\widetilde{\psi}}(w)| &\leq C|w|^{a_2}(1+|w|^2)^{-(b_2+a_2)/2}, && a_2 > 0 \;, b_2 > \alpha + 1, \\
|\widehat{\phi}(w)| &\leq C(1+|w|^2)^{-b_1/2}, && \text{and} \\
|\widehat{\widetilde{\phi}}(w)| &\leq C(1+|w|^2)^{-b_2/2},
\end{aligned}$$

then $\{\psi_{k,j}^{(i)}, \widetilde{\psi}_{k,j}^{(i)}, U_{k,j}^{(i)}, \widetilde{U}_{k,j}^{(i)}\}$ are the wavelet-vaguelette system of A.

Obviously, Theorem 6.3.1 allows more wavelets to be used in wavelet-vaguelette system for a given operator A than Theorem 6.1.4 does. More importantly, sufficient conditions listed in Theorem 6.3.1 do not depend on dimension d.

CHAPTER 7
VAGUELETTE COEFFICIENTS AND VARIATIONAL PROBLEMS

In this chapter we study a recursive algorithm in computing vaguelette coefficients $[Y, U_{k,j}^{(i)}]$ for given $Y \in \mathcal{Y}$. This recursive algorithm is naturally derived when wavelets that are obtained from scaling functions as in (4) of Definition 3.1.2 are used for a wavelet-vaguelette system of A.

We consider a family of variational problems to solve (1.2). The main idea is based on a smoothness characterization of Besov space $B_{1,1}^{\beta_0}(\mathbb{R}^d)$ via wavelets and the L^2-stability of vaguelettes for $\mathcal{Y}$.

From now on, we assume that univariate functions $\psi, \widetilde{\psi}, \phi$, and $\widetilde{\phi}$ satisfy

$$
\begin{aligned}
&|\widehat{\psi}(w)| \le C|w|^{a_1}(1+|w|^2)^{-(b_1+a_1)/2}, \qquad && a_1 > 2\alpha, \;\; b_1 > \max(1, |\beta_0|),\\
&|\widehat{\widetilde{\psi}}(w)| \le C|w|^{a_2}(1+|w|^2)^{-(b_2+a_2)/2}, \qquad && a_2 > \max(0, |\beta_0|), \;\; b_2 > 2\alpha + 1,\\
&|\widehat{\phi}(w)| \le C(1+|w|^2)^{-b_1/2}, && \text{and}\\
&|\widehat{\widetilde{\phi}}(w)| \le C(1+|w|^2)^{-b_2/2}. &&
\end{aligned}
\tag{7.1}
$$

The conditions in (7.1) are more than enough to have a wavelet-vaguelette system of A (see Theorem 6.3.1), but additional assumptions are required to have a smoothness characterization of $B_{1,1}^{\beta_0}(\mathbb{R}^d)$ via wavelets (see Theorem 3.2.2). and stability in computing vaguelette coefficients by the recursive algorithm.

7.1 Vaguelette Coefficients

In this section we consider the case $d = 2$. The multi-dimensional algorithm can follows easily from the 2-dimensional one. Throughout this section we assume that $A(L^2(\mathbb{R}^2)) = \mathcal{Y}$.

We define $\rho \in S'(\mathbb{R}^2)$ by

$$\widehat{\rho}(\xi) = |\xi|^{2\alpha}\widehat{\widetilde{\Phi}}(\xi). \tag{7.2}$$

Let $\rho_{k,j} = 2^k \rho(2^k \cdot -j)$. We define

$$V_{k,j} = 2^{k\alpha} A\rho_{k,j}. \tag{7.3}$$

Then by following the same argument used to show $A^*A\pi_{k,j}^{(i)} = 2^{-2k\alpha}\widetilde{\psi}_{k,j}^{(i)}$ in (6.4), one can show that $A^*A\rho_{k,j} = 2^{-2k\alpha}\widetilde{\Phi}_{k,j}$. Thus

$$\begin{aligned} 2^{k\alpha}[Af, V_{k,j}] &= 2^{2k\alpha}[Af, A\rho_{k,j}] \\ &= 2^{2k\alpha}\langle f, A^*A\rho_{k,j}\rangle \\ &= \langle f, \widetilde{\Phi}_{k,j}\rangle. \end{aligned}$$

Moreover, since $2^{k\alpha}[Af, U_{k,j}^{(i)}] = \langle f, \widetilde{\psi}_{k,j}^{(i)}\rangle$ for $i = 1, 2, 3$, by using the fast wavelet transform associated with $\{\widetilde{\psi}_{k,j}^{(i)}\}$, one can get $\{2^{k\alpha}[Af, U_{k,j}^{(i)}]\}_{\{k_0 \le k < m,\ j,\ i=1,2,3\}}$ and $\{2^{k_0\alpha}[Af, V_{k_0,n}]\}_n$ from $\{2^{m\alpha}[Af, V_{m,l}]\}_l$. Since $A(L^2(\mathbb{R}^2)) = \mathcal{Y}$, we can summarize it as follows:

THEOREM 7.1.1. *For each $Y \in \mathcal{Y}$, we have*

$$\begin{aligned} 2^{(k-1)\alpha}[Y, V_{k-1,j}] &= \sum_{n_1,n_2} \widetilde{h}_{n_1-2j_1}\widetilde{h}_{n_2-2j_2} 2^{k\alpha}[Y, V_{k,j}], \\ 2^{(k-1)\alpha}[Y, U_{k-1,j}^{(1)}] &= \sum_{n_1,n_2} \widetilde{h}_{n_1-2j_1}\widetilde{g}_{n_2-2j_2} 2^{k\alpha}[Y, V_{k,j}], \\ 2^{(k-1)\alpha}[Y, U_{k-1,j}^{(2)}] &= \sum_{n_1,n_2} \widetilde{g}_{n_1-2j_1}\widetilde{h}_{n_2-2j_2} 2^{k\alpha}[Y, V_{k,j}], \\ 2^{(k-1)\alpha}[Y, U_{k-1,j}^{(3)}] &= \sum_{n_1,n_2} \widetilde{g}_{n_1-2j_1}\widetilde{g}_{n_2-2j_2} 2^{k\alpha}[Y, V_{k,j}]. \end{aligned} \tag{7.4}$$

Using (7.4) successively on $\{2^{k\alpha}[Y, V_{k,j}]\}$, we can get $\{2^{k\alpha}[Y, U_{k,j}^{(i)}]\}_{\{k_0 \le k < m,\ j,\ i=1,2,3\}}$ and $\{2^{k_0\alpha}[Y, V_{k_0,n}]\}_n$ from $\{2^{m\alpha}[Y, V_{m,l}]\}_l$.

As analogy with the inhomogeneous wavelet decomposition, we can have an *inhomogeneous wavelet-vaguelette decomposition*:

THEOREM 7.1.2. *Let* $\{\psi_{k,j}^{(i)}, \widetilde{\psi}_{k,j}^{(i)}, U_{k,j}^{(i)}, \widetilde{U}_{k,j}^{(i)}\}$ *be a wavelet-vaguelette system of* A. *Then*

$$f = \sum_{k\geq k_0}\sum_{j,i} 2^{k\alpha}[Af, U_{k,j}^{(i)}]\psi_{k,j}^{(i)} + \sum_{l} 2^{k_0\alpha}[Af, V_{k_0,l}]\Phi_{k_0,l}.$$

7.2 Variational Problems

We consider a family of variational problems, which naturally give rise to a parametrized solution class $\widetilde{f}_{\gamma,\beta_0}$: Given a positive parameter γ and Besov space $B_{1,1}^{\beta_0}(\mathbb{R}^d)$, find a function $\widetilde{f}_{\gamma,\beta_0}$ that minimizes over all possible functions g in $B_{1,1}^{\beta_0}(\mathbb{R}^d)$ the functional

$$\|Y - Ag\|_{\mathcal{Y}}^2 + 2\gamma\|g\|_{B_{1,1}^{\beta_0}}. \tag{7.5}$$

Using L^2-stability of $\{U_{k,j}^{(i)}\}$ and smoothness characterization of wavelets, we have

$$\begin{aligned}
\|Y - Ag\|_{\mathcal{Y}}^2 &\asymp \sum_{k,j,i} \left|[Y - Ag, U_{k,j}^{(i)}]\right|^2 \\
&= \sum_{k,j,i} \left|[Y, U_{k,j}^{(i)}] - [Ag, U_{k,j}^{(i)}]\right|^2 \\
&= \sum_{k,j,i} \left|[Y, U_{k,j}^{(i)}] - 2^{k\alpha}\langle g, A^*A\pi_{k,j}^{(i)}\rangle\right|^2 && \text{(by (6.6))} \\
&= \sum_{k,j,i} \left|[Y, U_{k,j}^{(i)}] - 2^{-k\alpha}\langle g, \widetilde{\psi}_{k,j}^{(i)}\rangle\right|^2 && \text{(by (6.4))}
\end{aligned}$$

and

$$|g|_{B_{1,1}^{\beta_0}} = \sum_{k\geq k_0}\sum_{j,i} 2^{k(\beta_0-d/2)}|\langle g, \widetilde{\psi}_{k,j}^{(i)}\rangle|.$$

Combining these sequence sums, we have following equivalent sequence sums to the functional (7.5):

$$\sum_{k,j,i} 2^{-2k\alpha}\left|2^{k\alpha}[Y, U_{k,j}^{(i)}] - \langle g, \widetilde{\psi}_{k,j}^{(i)}\rangle\right|^2 + 2\gamma\sum_{k\geq k_0}\sum_{j,i} 2^{k(\beta_0-d/2)}|\langle g, \widetilde{\psi}_{k,j}^{(i)}\rangle|, \tag{7.6}$$

which can be minimized by minimizing each term separately. Notice that $a|b-x|^2 + 2c|x|$, where $a > 0$, $c > 0$, is minimized when

$$x = \begin{cases} b - c/a, & \text{if } b - c/a, \\ 0, & \text{if } |b| \leq c/a, \\ b + c/a, & \text{if } b + c/a. \end{cases}$$

With the shrinkage operator S_μ in (2.2), we can denote this minimum point simply by $S_{c/a}(b)$. We apply this to the case where $a = 2^{-2k\alpha}, b = 2^{k\alpha}[Y, U_{k,j}^{(i)}]$, and $c = \gamma 2^{k(\beta_0 - d/2)}$ for $k \geq k_0$, then (7.6) is minimized by choosing the function $\widetilde{f}^*_{\gamma,\beta_0}$ such that

$$\widetilde{f}^*_{\gamma,\beta_0} = \sum_{k,j,i} S_{\mu_k}(2^{k\alpha}[Y, U_{k,j}^{(i)}])\psi_{k,j}^{(i)}, \tag{7.7}$$

where

$$\mu_k = \begin{cases} \gamma 2^{k(\beta_0 - d/2 + 2\alpha)}, & \text{for } k \geq k_0, \\ 0, & \text{for } k < k_0. \end{cases} \tag{7.8}$$

Since $\mu_k = 0$ if $k < k_0$, one can rewrite (7.7) as

$$\widetilde{f}^*_{\gamma,\beta_0} = \sum_{k \geq k_0} \sum_{j,i} S_{\mu_k}(2^{k\alpha}[Y, U_{k,j}^{(i)}])\psi_{k,j}^{(i)} + \sum_l 2^{k_0\alpha}[Y, V_{k_0,l}]\Phi_{k_0,l}. \tag{7.9}$$

With equivalence between function norm and wavelet sequence sums, we now suggest $\widetilde{f}^*_{\gamma,\beta_0}$ as a solution method for solving (1.2).

CHAPTER 8

DISCRETIZATION AND NOISE MODEL

In this chapter we discretize the functional form (1.2). For the perturbation data error in observation procedure, we assume a statistical noise model.

8.1 Discretization

We recall the given problem (1.2) of this thesis:

$$Y = Af + Z. \tag{1.2}$$

In practice we are only able to have finitely many data $\{Y_i\}_{i=0,1,\cdots,N-1}$ for Y. For this data set, we assume that sufficiently many data $\{Y_i\}_{i=0,1,\cdots,N-1}$ are observed to approximate the underlying Y with a negligible error.

We wish to approximate the true solution f by

$$\widetilde{f}^*_{\gamma,\beta_0,m} = \sum_{k\geq k_0}^{m-1} \sum_{j,i} S_{\mu_k}(2^{k\alpha}[Y, U^{(i)}_{k,j}])\psi^{(i)}_{k,j} + \sum_{l} 2^{k_0\alpha}[Y, V_{k_0,l}]\Phi_{k_0,l}. \tag{8.1}$$

In this approach, we have three different types of errors. First, in approximating Y from $\{Y_i\}_{\{i=0,1,\cdots,N-1\}}$ and computing $[Y, U^{(i)}_{k,j}]$ and $[Y, V_{k_0,l}]$ from the approximation to Y, we cannot avoid a certain error. However, we shall ignore this type of error in this thesis. Second, an error is introduced in having finite m in $\widetilde{f}^*_{\gamma,\beta_0,m}$ (8.1). Notice that the smallest error we can have with $\widetilde{f}^*_{\gamma,\beta_0,m}$ (8.1) in the mean square measurement is equivalent to the sequence sums

$$\sum_{k\geq m} \sum_{j,i} |\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle|^2.$$

The integer m in (8.1) is closely related to N, the number of data. Increasing m for fixed N does not necessarily give a better solution since it may generate another error in computing $[Y, U_{k,j}^{(i)}]$ and $[Y, V_{k_0,l}]$. We assume that we can take a positive integer m in (8.1) such that

$$2^{md} = N.$$

Again, we ignore this type of error in this thesis. Third, the solution method $\widetilde{f}^*_{\gamma,\beta_0,m}$ (8.1) itself makes an error in approximating

$$f_m = \sum_{k \geq k_0}^{m-1} \sum_{j,i} \langle f, \widetilde{\psi}_{k,j}^{(i)} \rangle \psi_{k,j}^{(i)} + \sum_{l} \langle f, \widetilde{\Phi}_{k_0,l} \rangle \Phi_{k_0,l}. \tag{8.2}$$

Since

$$\begin{aligned} \|f_m - \widetilde{f}^*_{\gamma,\beta_0,m}\|_{L^2}^2 \asymp \sum_{k=k_0}^{m-1} \sum_{j,i} |\langle f, \widetilde{\psi}_{k,j}^{(i)} \rangle - S_{\mu_k}(2^{k\alpha}[Y, U_{k,j}^{(i)}])|^2 \\ + \sum_{l} |\langle f, \widetilde{\Phi}_{k_0,l} \rangle - 2^{k_0\alpha}[Y, V_{k_0,l}]|^2, \end{aligned} \tag{8.3}$$

it suffices to consider the right hand side of (8.3) to control the error between f_m and $\widetilde{f}^*_{\gamma,\beta_0,m}$. In this thesis we are only interested in this type of error.

8.2 Noise Model

We assume that a white noise model for the observation error Z in (1.2). Let

$$Z = \sigma W \tag{8.4}$$

for a constant $\sigma > 0$, where W is the white noise process defined on the underlying space of functions in $\mathcal{Y}$. This assumption naturally impose a noise model for the discrete data $\{Y_i\}$ such that

$$Y_i = (Af)_i + Z_i,$$

where Z_i are independent and identically distributed as $N(0, \sigma_0^2)$ for some $\sigma_0 > 0$. Since the number of data is N, it is reasonable to assume that

$$\begin{aligned} \sigma_0^2 &= N\sigma^2 \\ &= 2^{md}\sigma^2. \end{aligned} \tag{8.5}$$

We now examine the effect from this noise model (8.4) in $[Y, U_{k,j}^{(i)}]$ and $[Y, V_{k_0,l}]$ in $\widetilde{f}^*_{\gamma,\beta_0,m}$. Notice that for the white noise process W, $[W, U_{k,j}^{(i)}]$ are mean zero Gaussian random variables for all k, j, i and

$$\mathrm{Var}[W, U_{k,j}^{(i)}] = \|U_{k,j}^{(i)}\|_{\mathcal{Y}}^2.$$

Since

$$\begin{aligned}
\|U_{k,j}^{(i)}\|_{\mathcal{Y}}^2 &= [2^{k\alpha} A\pi_{k,j}^{(i)}, 2^{k\alpha} A\pi_{k,j}^{(i)}] && \text{(by (6.6))} \\
&= 2^{2k\alpha} \langle \pi_{k,j}^{(i)}, A^* A \pi_{k,j}^{(i)} \rangle \\
&= \langle \pi_{k,j}^{(i)}, \widetilde{\psi}_{k,j}^{(i)} \rangle && \text{(by (6.4))} \\
&= \int_{\mathbb{R}^d} \widehat{\pi_{k,j}^{(i)}}(\xi) \overline{\widehat{\widetilde{\psi}_{k,j}^{(i)}}(\xi)}\, d\xi && \text{(by (2.3))} \\
&= 2^{-kd} \int_{\mathbb{R}^d} \widehat{\pi^{(i)}}(2^{-k}\xi) \overline{\widehat{\widetilde{\psi}^{(i)}}(2^{-k}\xi)}\, d\xi && \text{(by (2.8))} \\
&= 2^{-kd} \int_{\mathbb{R}^d} |2^{-k}\xi|^{2\alpha} |\widehat{\widetilde{\psi}^{(i)}}(2^{-k}\xi)|^2\, d\xi \\
&= \int_{\mathbb{R}^d} |\xi'|^{2\alpha} |\widehat{\widetilde{\psi}}(\xi')|^2\, d\xi',
\end{aligned}$$

one can show that $[Z, U_{k,j}^{(i)}]$ are mean zero Gaussian random variables with

$$\mathrm{Var}[Z, U_{k,j}^{(i)}] = \sigma^2 \int_{\mathbb{R}^d} |\xi|^{2\alpha} |\widehat{\widetilde{\psi}^{(i)}}(\xi)|^2\, d\xi.$$

Similarly, we have mean zero Gaussian random variables $[Z, V_{k_0,l}]$ such that

$$\mathrm{Var}[Z, V_{k_0,l}] = \sigma^2 \int_{\mathbb{R}^d} |\xi|^{2\alpha} |\widehat{\widetilde{\Phi}}(\xi)|^2\, d\xi. \tag{8.6}$$

Notice that

$$
\begin{aligned}
2^{k\alpha}[Y, U_{k,j}^{(i)}] &= 2^{k\alpha}[Af + Z, U_{k,j}^{(i)}] && \text{(by (1.2))} \\
&= 2^{k\alpha}\langle f, A^* U_{k,j}^{(i)}\rangle + 2^{k\alpha}[Z, U_{k,j}^{(i)}] \\
&= 2^{k\alpha}\langle f, A^* 2^{k\alpha} A\pi_{k,j}^{(i)}\rangle + 2^{k\alpha}[Z, U_{k,j}^{(i)}] && \text{(by (6.6))} \\
&= \langle f, \widetilde{\psi}_{k,j}^{(i)}\rangle + 2^{k\alpha}[Z, U_{k,j}^{(i)}]. && \text{(by (6.4))}
\end{aligned}
$$

With a similar argument, we have

$$2^{k_0\alpha}[Y, V_{k_0,l}] = \langle f, \widetilde{\Phi}_{k_0,l}\rangle + 2^{k_0\alpha}[Z, V_{k_0,l}].$$

Thus the solution method $\widetilde{f}^*_{\gamma,\beta_0,m}$ (8.1) can be rewritten as

$$
\begin{aligned}
\widetilde{f}^*_{\gamma,\beta_0,m} = &\sum_{k_0 \le k < m} \sum_{j,i} S_{\mu_k}(\langle f, \widetilde{\psi}_{k,j}^{(i)}\rangle + 2^{k\alpha}[Z, U_{k,j}^{(i)}])\psi_{k,j}^{(i)} \\
&+ \sum_l (\langle f, \widetilde{\Phi}_{k_0,l}\rangle + 2^{k_0\alpha}[Z, V_{k_0,l}])\Phi_{k_0,l}.
\end{aligned} \tag{8.7}
$$

Before we close this section, we define some notations for the future use. Let

$$c_i = \int_{\mathbb{R}^d} |\xi|^{2\alpha} \left|\widehat{\widetilde{\psi}^{(i)}}(\xi)\right|^2 d\xi \tag{8.8}$$

for $i = 1, \cdots, 2^d - 1$,

$$c_0 = \int_{\mathbb{R}^d} |\xi|^{2\alpha} \left|\widehat{\widetilde{\Phi}}(\xi)\right|^2 d\xi, \tag{8.9}$$

$$c_* = \min(c_1, \cdots, c_{2^d-1}), \tag{8.10}$$

and

$$c^* = \max(c_1, \cdots, c_{2^d-1}). \tag{8.11}$$

CHAPTER 9
ERROR ESTIMATES

In this chapter we determine γ and β_0 that minimize an upper bound of $E\|f_m - \widetilde{f}^*_{\gamma,\beta_0,m}\|^2_{L^2}$, where the expectation operator E is needed to deal with the statistical noise model proposed in Chapter 8.

9.1 Background

In [22] Donoho assumed the same white noise model as in this thesis, and suggested

$$\widetilde{f}_{(a_{k,j,i})} = \sum_{k\geq k_0}\sum_{j,i} S_{a_{k,j,i}}(2^{k\alpha}[Y, U^{(i)}_{k,j}])\psi^{(i)}_{k,j} + \sum_{l} 2^{k_0\alpha}[Y, V_{k_0,l}] \tag{9.1}$$

as a solution method for solving (1.2). This is exactly same as $\widetilde{f}^*_{\gamma,\beta_0}$ (7.9) except for possibly different shrinkage parameters. For the performance of $\widetilde{f}_{(a_{k,j,i})}$, he proved the following theorem:

THEOREM 9.1.1. (*Donoho*) *If the true solution f is known to lie in a ball $\mathcal{F}(C) = \{f \mid \|f\|_{B^\beta_{q,p}} \leq C\}$ of the Besov space $B^\beta_{q,p}(\mathbb{R}^d)$ with*

$$\beta > (2\alpha + d)(1/p - 1/2),$$

then

$$\inf_{(a_{k,j,i})} \sup_{f\in\mathcal{F}(C)} E\|f - \widetilde{f}_{(a_{k,j,i})}\|^2_{L^2} \asymp \sigma^{2r_M}$$

as $\sigma \to 0$, with rate exponent

$$r_M = \frac{\beta}{\beta + d/2 + \alpha}.$$

Moreover, this wavelet shrinkage method attains the optimal rate of convergence:

$$\inf_{(a_{k,j,i})} \sup_{f\in\mathcal{F}(C)} E\|f-\widetilde{f}_{(a_{k,j,i})}\|_{L_2}^2 \le \mathit{Constant}\cdot \inf_{\widetilde{f}} \sup_{f\in\mathcal{F}(C)} E\|f-\widetilde{f}\|_{L_2}^2,$$

where $\widetilde{f}$ ranges over all possible methods. This is faster than the rate of convergence of any linear methods:

$$\inf_{\widetilde{f}^*} \sup_{f\in\mathcal{F}(C)} E\|f-\widetilde{f}^*\|_{L_2}^2 \asymp \sigma^{2r_L}$$

as $\sigma \to 0$, where $\widetilde{f}^*$ ranges over all possible linear methods, with rate exponent

$$r_L = \frac{\beta + d(1/2 - 1/p^*)}{\beta + d(1-1/p^*) + \alpha},$$

where $p^* = \min(2, p)$, which is smaller than r_M in case $p < 2$.

This theorem implies that if one uses the optimal shrinkage parameter, then the corresponding wavelet shrinkage method is the best *estimator* in *minimax* sense. (For definitions of these statistical terminologies, see, e.g., [1].) However, Donoho in [22] did not give a method for finding that parameter.

In [31] Kolaczyk used the shrinkage parameter

$$a_{k,j,i} = \sqrt{2\log(2^{2k})}\,2^{k/2}\sigma c_i^{1/2}, \tag{9.2}$$

in (9.1) for the tomographic reconstruction. For the definition of c_i, see (8.8). This choice of shrinkage amount is motivated by the VisuShrink method of Donoho and Johnstone [23]. For details, see, e.g., [23] and [31].

One may ask what is the optimal shrinkage parameter to attain the optimal rate of convergence in Theorem 9.1.1 and what is the rate of convergence through (9.1) in solving (1.2) when the true solution f is known to lie in $B^{\beta}_{p,p}(\mathbb{R}^d)$ with

$$d(1/p - 1/2) \le \beta \le (2\alpha + d)(1/p - 1/2).$$

For these questions, we present an error estimation of $\widetilde{f}^{*}_{\gamma,\beta_0,m}$ (8.1) for $f \in B^{\beta}_{p,p}(\mathbb{R}^d)$, $d(1/p-1/2) \le \beta \le (2\alpha+d)(1/p-1/2)$ with an explicit shrinkage parameter in the next section.

9.2 Error Estimates of $\widetilde{f}^{*}_{\gamma,\beta_0,m}$

In this section we estimate an upper bound of $E\|f_m - \widetilde{f}^{*}_{\gamma,\beta_0,m}\|^2_{L^2}$. While doing that, we determine β_0, which gives the main effect on the wavelet shrinkage procedure, and then we choose γ, which corresponds to a subtle part of the algorithm.

Throughout this section, we assume that f is compactly supported in the unit cube of $\mathbb{R}^d$. Thus we need only j and l for which $\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle \ne 0$ and $\langle f, \widetilde{\Phi}_{k_0,l}\rangle \ne 0$ in (8.1). Moreover, the number of such $j \le C2^{kd}$ for fixed k and i, and that of such $l \le C2^{k_0 d}$ for a constant C. Having these results in mind, we start to estimate an upper bound of $E\|f_m - \widetilde{f}^{*}_{\gamma,\beta_0,m}\|^2_{L^2}$.

From (8.2) and (8.7), we note that $f_m - \widetilde{f}^{*}_{\gamma,\beta_0,m}$ is

$$\sum_{k=k_0}^{m-1} \sum_{\langle f,\widetilde{\psi}^{(i)}_{k,j}\rangle \ne 0} \Big(\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle - S_{\mu_k}(\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle + 2^{k\alpha}[Z, U^{(i)}_{k,j}]) \Big) \psi^{(i)}_{k,j} + \sum_{\langle f,\widetilde{\Phi}_{k_0,l}\rangle \ne 0} 2^{k_0\alpha}[Z, V_{k_0,l}]\Phi_{k_0,l}.$$

Hence using (3.8) one can bound $E\|f_m - \widetilde{f}^{*}_{\gamma,\beta_0,m}\|^2_{L^2}$ by a constant multiple of

$$\sum_{k=k_0}^{m-1} \sum_{\langle f,\widetilde{\psi}^{(i)}_{k,j}\rangle \ne 0} E\big|\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle - S_{\mu_k}(\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle + 2^{k\alpha}[Z, U^{(i)}_{k,j}])\big|^2 + \sum_{\langle f,\widetilde{\Phi}_{k_0,l}\rangle \ne 0} 2^{2k_0\alpha}\mathrm{Var}[Z, V_{k_0,l}].$$

We note that by (8.6) and (8.9), $\mathrm{Var}[Z, V_{k_0,l}] = c_0\sigma^2$ for all l. Thus we can bound $E\|f_m - \widetilde{f}^{*}_{\gamma,\beta_0,m}\|^2_{L^2}$ by a constant multiple of

$$\sum_{k=k_0}^{m-1} \sum_{\langle f,\widetilde{\psi}^{(i)}_{k,j}\rangle \ne 0} E\big|\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle - S_{\mu_k}(\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle + 2^{k\alpha}[Z, U^{(i)}_{k,j}])\big|^2 + c_0 2^{k_0(2\alpha+d)}\sigma^2,$$

where we have used the fact that the number of l for which $\langle f, \widetilde{\Phi}_{k_0,l}\rangle \neq 0$ is less than $C2^{k_0 d}$ for a constant C.

LEMMA 9.2.1. *If* $X \sim N(0, \tau^2)$, *then*

$$E|t - S_\mu(t+X)|^2 \leq \begin{cases} \mu^2 + \tau^2, & \text{if } |t| > \mu, \\ t^2 + E|S_\mu(X)|^2, & \text{if } |t| \leq \mu, \end{cases}$$

where

$$E|S_\mu(X)|^2 = 2\tau^2 \int_{y > \frac{\mu}{\tau}} (y - \frac{\mu}{\tau})^2 P(y)\, dy. \qquad (P(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2})$$

Proof. See [4] and [24]. □

Let

$$\Lambda_k = \{(j,i) | |\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle| > \mu_k\}$$

and

$$\widetilde{\Lambda}_k = \{(j,i) | 0 < |\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle| \leq \mu_k\}.$$

We apply Lemma 9.2.1 to the case where $t = \langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle$, $X = 2^{k\alpha}[Z, U^{(i)}_{k,j}]$, $\tau^2 = 2^{2k\alpha} c_i \sigma^2$, and $\mu = \mu_k$ (see (7.8)). Then we have

$$E\|f_m - \widetilde{f}^*_{\gamma,\beta_0,m}\|^2_{L^2} \leq C\Big(S_0(f) + S_1(f) + S_2(f) + S_3(f) + S_4(f)\Big), \tag{9.3}$$

where

$$S_0(f) = c_0 2^{k_0(2\alpha+d)}\sigma^2,$$
$$S_1(f) = \sum_{k=k_0}^{m-1} \sum_{(j,i)\in\Lambda_k} \mu_k^2,$$
$$S_2(f) = \sum_{k=k_0}^{m-1} \sum_{(j,i)\in\Lambda_k} c_i 2^{2k\alpha}\sigma^2,$$
$$S_3(f) = \sum_{k=k_0}^{m-1} \sum_{(j,i)\in\widetilde{\Lambda}_k} |\langle f, \widetilde{\psi}_{k,j}^{(i)}\rangle|^2, \qquad \text{and}$$
$$S_4(f) = \sum_{k=k_0}^{m-1} \sum_{(j,i)\in\widetilde{\Lambda}_k} 2^{2k\alpha+1} c_i \sigma^2 \int_{t>t_{k,i}} (t-t_{k,i})^2 P(t)\, dt,$$

where $t_{k,i} = c_i^{-1/2} 2^{k(\beta_0-d/2+\alpha)} \frac{\gamma}{\sigma}$. For the definition of c_i, see (8.8).

Let $f \in B^{\beta}_{p_0,p_0}(\mathbb{R}^d)$, where

$$p_0 = \frac{2\alpha+d}{\beta+d/2+\alpha}. \tag{9.4}$$

For $S_1(f)$, notice that

$$\begin{aligned} S_1(f) &= \sum_{k=k_0}^{m-1} \sum_{(j,i)\in\Lambda_k} \mu_k^{2-p_0}\cdot\mu_k^{p_0} \\ &= \sum_{k=k_0}^{m-1} 2^{-k(\beta+d(1/2-1/p_0))p_0} \mu_k^{2-p_0} 2^{k(\beta+d(1/2-1/p_0))p_0} \sum_{(j,i)\in\Lambda_k} \mu_k^{p_0}. \end{aligned}$$

Let

$$a = \frac{\gamma}{\sigma}. \tag{9.5}$$

Since $\mu_k = a\sigma 2^{k(\beta_0-d/2+2\alpha)}$ (see (7.8) and (9.5)), one has

$$S_1(f) \le a^{2-p_0}\sigma^{2-p_0} T(\beta_0,\beta,p_0) \sum_{k=k_0}^{m-1} 2^{k(\beta+d(1/2-1/p_0))p_0} \sum_{(j,i)\in\Lambda_k} \mu_k^{p_0}, \tag{9.6}$$

where

$$T(\beta_0,\beta,p_0)=\max_{k_0\le k<m}\left\{2^{k((\beta_0-d/2+2\alpha)(2-p_0)-(\beta+d(1/2-1/p_0))p_0)}\right\}.$$

By arranging the exponent of 2, we have

$$T(\beta_0,\beta,p_0)=\max_{k_0\le k<m}\left\{2^{k(\beta_0(2-p_0)+4\alpha-2\alpha p_0-\beta p_0)}\right\}.$$

Thus we have

$$S_1(f)\le |f|^{p_0}_{B^{\beta}_{p_0,p_0}}a^{2-p_0}\sigma^{2-p_0}T(\beta_0,\beta,p_0), \tag{9.7}$$

because in (9.6),

$$\begin{aligned}\sum_{k=k_0}^{m-1}2^{k(\beta+d(1/2-1/p_0))p_0}\sum_{(j,i)\in\Lambda_k}\mu_k^{p_0}&\le\sum_{k=k_0}^{m-1}2^{k(\beta+d(1/2-1/p_0))p_0}\sum_{j,i}|\langle f,\widetilde{\psi}^{(i)}_{k,j}\rangle|^{p_0}\\&\le |f|^{p_0}_{B^{\beta}_{p_0,p_0}}.\end{aligned}$$

For $S_2(f)$, since $\sigma^2=a^{-2}2^{-2k(\beta_0-d/2+2\alpha)}\mu_k^2$ (see (7.8) and (9.5)), one has

$$S_2(f)\le c^*\sum_{k_0=k}^{m-1}\sum_{(j,i)\in\Lambda_k}2^{2k\alpha}a^{-2}2^{-k(2\beta_0-d+4\alpha)}\mu_k^2.$$

We use the same argument that is used for $S_1(f)$. Then we have

$$\begin{aligned}S_2(f)\le c^*\sum_{k_0=k}^{m-1}&a^{-2}2^{-k((\beta+d(1/2-1/p_0))p_0+2\beta_0-d+2\alpha)}\mu_k^{2-p_0}\\&\times 2^{k(\beta+d(1/2-1/p_0))p_0}\sum_{(j,i)\in\Lambda_k}\mu_k^{p_0}.\end{aligned}$$

Moreover, since $a^{-2}\mu_k^{2-p_0}=a^{-p_0}\sigma^{2-p_0}2^{k(\beta_0-d/2+2\alpha)(2-p_0)}$, we have

$$S_2(f)\le c^*a^{-p_0}\sigma^{2-p_0}T^*(\beta_0,\beta,p_0)\sum_{k_0=k}^{m-1}2^{k(\beta+d(1/2-1/p_0))p_0}\sum_{(j,i)\in\Lambda_k}\mu_k^{p_0},$$

where

$$T^*(\beta_0, \beta, p_0) = \max_{k_0 \le k < m} \left\{ 2^{-k((\beta + d(1/2 - 1/p_0))p_0 + 2\beta_0 - d + 2\alpha - (\beta_0 - d/2 + 2\alpha)(2 - p_0))} \right\}.$$

By arranging the exponent of 2, we can write

$$T^*(\beta_0, \beta, p_0) = \max_{k_0 \le k < m} \left\{ 2^{-k(\beta_0 p_0 + \beta p_0 - 2\alpha - d + 2\alpha p_0)} \right\}.$$

With a similar argument used for $S_1(f)$, we have

$$\sum_{k=k_0}^{m-1} 2^{k(\beta + d(1/2 - 1/p_0))p_0} \sum_{(j,i) \in \Lambda_k} \mu_k^{p_0} \le |f|^{p_0}_{B^\beta_{p_0,p_0}}.$$

Thus

$$S_2(f) \le c^* |f|^{p_0}_{B^\beta_{p_0,p_0}} a^{-p_0} \sigma^{2-p_0} T^*(\beta_0, \beta, p_0). \tag{9.8}$$

For $S_3(f)$, we can bound them as follows:

$$\begin{aligned} S_3(f) &= \sum_{k_0=k}^{m-1} \sum_{(j,i) \in \widetilde{\Lambda}_k} |\langle f, \widetilde{\psi}^{(i)}_{k,j} \rangle|^2 \\ &\le \sum_{k_0=k}^{m-1} \sum_{(j,i) \in \widetilde{\Lambda}_k} |\langle f, \widetilde{\psi}^{(i)}_{k,j} \rangle|^{p_0} \cdot \mu_k^{2-p_0} \\ &= \sum_{k_0=k}^{m-1} 2^{-k(\beta + d(1/2 - 1/p_0))p_0} \mu_k^{2-p_0} 2^{k(\beta + d(1/2 - 1/p_0))p_0} \sum_{(j,i) \in \widetilde{\Lambda}_k} |\langle f, \widetilde{\psi}^{(i)}_{k,j} \rangle|^{p_0}. \end{aligned}$$

We now use the same argument used for $S_1(f)$. Then we have

$$S_3(f) \le |f|^{p_0}_{B^\beta_{p_0,p_0}} a^{2-p_0} \sigma^{2-p_0} T(\beta_0, \beta, p_0). \tag{9.9}$$

Combining (9.7), (9.8), and (9.9), for $f \in B^\beta_{p_0,p_0}(\mathbb{R}^d)$, we can bound $S_1(f) + S_2(f) + S_3(f)$ by

$$|f|^{p_0}_{B^\beta_{p_0,p_0}} \sigma^{2-p_0} \left(2a^{2-p_0} T(\beta_0, \beta, p_0) + c^* a^{-p_0} T^*(\beta_0, \beta, p_0) \right). \tag{9.10}$$

We now consider the case when $f \in B^{\beta}_{p,p}(\mathbb{R}^d)$, where

$$d(1/p - 1/2) \le \beta < (2\alpha + d)(1/p - 1/2). \tag{9.11}$$

Notice that if $\beta = (2\alpha + d)(1/p - 1/2)$, then $B^{\beta}_{p,p}(\mathbb{R}^d)$ is the same as $B^{\beta}_{p_0,p_0}(\mathbb{R}^d)$ with p_0 in (9.4). We define

$$\beta^* = \begin{cases} \frac{\alpha+d/2}{\alpha}(\beta + d(1/2 - 1/p)), & \text{if } \alpha > 0, \\ \beta, & \text{if } \alpha = 0, \end{cases} \tag{9.12}$$

and

$$p_0^* = \frac{2\alpha + d}{\beta^* + d/2 + \alpha}. \tag{9.13}$$

Notice that for any pair (β, p) satisfying (9.11), $\beta^* < \beta$ and $(\beta - \beta^*)/d = 1/p - 1/p_0^*$ if $\alpha > 0$. Thus by (2) of Theorem 4.1.1, we have

$$B^{\beta}_{p,p}(\mathbb{R}^d) \hookrightarrow B^{\beta^*}_{p_0^*,p_0^*}(\mathbb{R}^d) \quad \text{and} \quad |f|_{B^{\beta^*}_{p_0^*,p_0^*}} \le |f|_{B^{\beta}_{p,p}}. \tag{9.14}$$

If $\alpha = 0$, we bound $S_1(f)+S_2(f)+S_3(f)$ by (9.10) by replacing p_0 with $d/(\beta+d/2)$.

From the relations (9.12), (9.13) between β^* and p_0^*, it is obvious that by following the same routine as we did for $f \in B^{\beta}_{p_0,p_0}(\mathbb{R}^d)$, we can have

$$S_1(f)+S_2(f)+S_3(f) \le |f|^{p_0^*}_{B^{\beta^*}_{p_0^*,p_0^*}} \sigma^{2-p_0^*} \left(2a^{2-p_0^*} T(\beta_0, \beta^*, p_0^*) + c^* a^{-p_0^*} T^*(\beta_0, \beta^*, p_0^*) \right)$$

for $f \in B^{\beta^*}_{p_0^*,p_0^*}(\mathbb{R}^d)$. Furthermore, using (9.14), we can bound $S_1(f)+S_2(f)+S_3(f)$ by

$$|f|^{p_0^*}_{B^{\beta}_{p,p}} \sigma^{2-p_0^*} \left(2a^{2-p_0^*} T(\beta_0, \beta^*, p_0^*) + c^* a^{-p_0^*} T^*(\beta_0, \beta^*, p_0^*) \right) \tag{9.15}$$

for $f \in B^{\beta}_{p,p}(\mathbb{R}^d)$.

We now determine β_0, which minimizes $S_1(f)+S_2(f)+S_3(f)$ for a fixed a. To do so, we examine the exponents of 2 in $T(\beta_0, \beta^*, p_0^*)$ and $T^*(\beta_0, \beta^*, p_0^*)$. Obviously, to reduce the contributions from $S_1(f)+S_2(f)+S_3(f)$, it is desirable if our choice of β_0 satisfies

$$\beta_0(2-p_0^*)+4\alpha-2\alpha p_0^*-\beta^* p_0^* \le 0 \qquad \text{and} \qquad \beta_0 p_0^*+\beta^* p_0^*-2\alpha-d+2\alpha p_0^* \ge 0, \tag{9.16}$$

so that $T(\beta_0,\beta^*,p_0^*)$ and $T^*(\beta_0,\beta^*,p_0^*)$ are bounded independently of m. Notice that

$$\begin{aligned}
&\beta_0(2-p_0^*)+4\alpha-2\alpha p_0^*-\beta^* p_0^* \\
&\qquad = \frac{2\beta_0\beta^*}{\beta^*+d/2+\alpha}+\frac{4\alpha\beta^*}{\beta^*+d/2+\alpha}-\frac{(2\alpha+d)\beta^*}{\beta^*+d/2+\alpha} \\
&\qquad = (\beta_0-d/2+\alpha)\frac{2\beta^*}{\beta^*+d/2+\alpha}
\end{aligned} \tag{9.17}$$

and

$$\begin{aligned}
\beta_0 p_0^*+\beta^* p_0^*-2\alpha-d+2\alpha p_0^* &= (\beta_0+\beta^*+2\alpha)\frac{2\alpha+d}{\beta^*+d/2+\alpha}-2\alpha-d \\
&= (\beta_0-d/2+\alpha)\frac{2\alpha+d}{\beta^*+d/2+\alpha}.
\end{aligned} \tag{9.18}$$

From (9.16), (9.17), and (9.18), it is obvious that we must take

$$\beta_0=\beta_0^*=d/2-\alpha.$$

With this β_0^*, we have

$$T(\beta_0^*,\beta^*,p_0^*)=T^*(\beta_0^*,\beta^*,p_0^*)=1.$$

Therefore, from (9.15), we have

$$S_1(f)+S_2(f)+S_3(f)\le |f|_{B^\beta_{p,p}}^{p_0^*}\,\sigma^{2-p_0^*}(2a^{2-p_0^*}+c^*a^{-p_0^*}) \tag{9.19}$$

for $f\in B^\beta_{p,p}(\mathbb{R}^d)$.

For $S_4(f)$, we note that $t_{k,i} = c_i^{-1/2} 2^{k(\beta_0^* - d/2+\alpha)} a \leq c_*^{-1/2} a$ with our choice for $\beta_0^* = d/2 - \alpha$. Notice that as a function of x, $\int_{t>x} (t-x)^2 P(t)\, dt$ is strictly decreasing for $x > 0$. Thus by bounding $|\widetilde{\Lambda}_k|$ by $(2^d - 1)2^{kd}$, we have

$$S_4(f) \leq 2c^*(2^d - 1) \sum_{k=k_0}^{m-1} 2^{k(2\alpha+d)} \sigma^2 \int_{t>c_*^{-1/2}a} \left(t - c_*^{-1/2} a\right)^2 P(t)\, dt \tag{9.20}$$

Finally, combining (9.3), (9.19), and (9.20), we can bound $E\|f - \widetilde{f}^*_{\gamma,\beta_0^*,m}\|^2_{L^2}$ by a constant multiple of

$$\begin{aligned} &|f|^{p_0^*}_{B^\beta_{p,p}} \sigma^{2-p_0^*} (2a^{2-p_0^*} + c^* a^{-p_0^*}) \\ &\qquad + \frac{2c^*(2^d-1)\cdot 2^{m(2\alpha+d)}\sigma^2}{2^{2(d/2+\alpha)} - 1} \int_{t>c_*^{-1/2}a} \left(t - c_*^{-1/2}a\right)^2 P(t) dt, \end{aligned} \tag{9.21}$$

for $f \in B^\beta_{p,p}(\mathbb{R}^d)$. Obviously, we ignored the contribution from $S_0(f)$, which is fixed and small. We now find the parameter $\gamma = a\sigma$ in $\widetilde{f}^*_{\gamma,\beta_0^*,m}$ by choosing a that minimizes (9.21).

The equation (9.21) is our main upper bound of $E\|f_m - \widetilde{f}^*_{\gamma,\beta_0^*,m}\|^2_{L^2}$. We emphasize that given only two parameters characterizing the smoothness of f (β and $|f|_{B^\beta_{p,p}}$), the known parameter α that determines the ill-posedness of the reconstruction procedure, the known constants c^* and c_*, and an estimate of the standard deviation σ of the noise in the observation procedure, one can numerically compute the minimum point a^* of the equation (9.21) as a function of a and use $\gamma^* = a^*\sigma$ to determine $\widetilde{f}^*_{\gamma^*,\beta_0^*,m}$ that minimizes our upper bound on the error.

To have an asymptotic result for (9.21), we use following inequality; for any $x \geq 0$,

$$\int_{t>x} (t-x)^2 P(t)\, dt \leq \frac{\sqrt{2\pi}}{2} P(x). \tag{9.22}$$

This is rather rough estimation since one can replace the right hand side of (9.22) by $2x^{-3}P(x)$ for $x \geq 1$ (see, e.g., [4]).

Using (9.22), we can bound $2 \cdot 2^{m(2\alpha+d)}\sigma^2 \int_{t>c_*^{-1/2}a} (t - c_*^{-1/2}a)^2 P(t)\, dt$ in (9.21) by

$$2^{m(2\alpha+d)}\sigma^2 e^{-c_*^{-1}a^2/2}.$$

We now compare two dominant terms $|f|^{p_0^*}_{B^\beta_{p,p}}\sigma^{2-p_0^*}$ and $2^{m(2\alpha+d)}\sigma^2 e^{-c_*^{-1}a^2/2}$ to get a simple approximation to the critical a. If

$$\frac{|f|_{B^\beta_{p,p}}}{\sigma} \le 2^{md(2\alpha+d)/p_0^*},$$

then by setting

$$|f|^{p_0^*}_{B^\beta_{p,p}}\sigma^{2-p_0^*} = 2^{m(2\alpha+d)}\sigma^2 e^{-c_*^{-1}a^2/2},$$

we can get the critical point a_1 such that

$$a_1 = c_*^{1/2}\sqrt{(\frac{4\alpha}{d} + 2 - p_0^*)\log 2^{md} - 2p_0^* \log \frac{|f|_{B^\beta_{p,p}}}{\sigma_0}},$$

where $\sigma_0^2 = 2^{md}\sigma^2$. With this a_1, we have

$$E\|f_m - \widetilde{f}^*_{\gamma,\beta_0^*,m}\|^2_{L^2} \le C|f|^{p_0^*}_{B^\beta_{p,p}}\sigma^{2-p_0^*}\left(\left[(4\alpha/d + 2 - p_0^*)\ln 2^{md}\right]^R + 1\right), \tag{9.23}$$

where

$$R = \begin{cases} (2-p_0^*)/2, & \text{if } a_1 \ge 1, \\ p_0^*/2, & \text{if } a_1 < 1. \end{cases}$$

On the other hand, if

$$\frac{|f|_{B^\beta_{p,p}}}{\sigma} > 2^{m(2\alpha+d)/p_0^*},$$

then the first term in (9.21) dominates the whole sum. In this case we can take $a = 1$. Thus we can have

$$E\|f_m - \widetilde{f}^*_{\gamma,\beta_0^*,m}\|^2_{L^2} \le C'|f|^{p_0^*}_{B^\beta_{p,p}}\sigma^{2-p_0^*}. \tag{9.24}$$

As we mentioned earlier, we find a near-optimal a^* by minimizing (9.21). However, with (9.23) and (9.24), we can have following asymptotic result by ignoring the logarithm term in (9.23):

THEOREM 9.2.1. *Let f be a compactly supported function in $B^{\beta}_{p,p}(\mathbb{R}^d)$, where*

$$d(1/p - 1/2) \leq \beta \leq (2\alpha + d)(1/p - 1/2).$$

Then with $\beta_0^* = d/2 - \alpha$ and $\gamma^* = a^*\sigma$, where a^* is the minimum point of (9.21),

$$E\left\|f_m - \widetilde{f}^*_{\gamma^*,\beta_0^*,m}\right\|^2_{L^2} \asymp \sigma^{2r},$$

as $\sigma \to 0$, with the exponent

$$r = \frac{\beta + d(1/2 - 1/p)}{\beta + d(1/2 - 1/p) + \alpha}, \qquad \text{if } \alpha > 0,$$

or

$$r = \frac{\beta}{\beta + d/2}, \qquad \text{if } \alpha = 0.$$

Proof. From (9.23) and (9.24), it is obvious that the rate of convergence is $1 - p_0^*/2$. If $\alpha > 0$, then

$$1 - p_0^*/2 = \frac{\beta^*}{\beta^* + d/2 + \alpha}.$$

Since $\beta^* = \frac{\alpha + d/2}{\alpha}(\beta + d(1/2 - 1/p))$ for $\alpha > 0$,

$$r = \frac{\beta + d(1/2 - 1/p)}{\beta + d(1/2 - 1/p) + \alpha},$$

if $\alpha > 0$. For $\alpha = 0$, $p_0^* = d/(\beta + d/2)$. Thus we have

$$r = \frac{\beta}{\beta + d/2},$$

if $\alpha = 0$. ∎

We now assume that the true solution f is not only compactly supported but also bounded. Suppose $f \in B^{\beta}_{p,p}(\mathbb{R}^d)$, where $d(1/p-1/2) \le \beta < (2\alpha+d)(1/p-1/2)$. Using (2) of Theorem 4.2.1, we can show that

$$f \in B^{\beta'}_{p',p'}(\mathbb{R}^d) \quad \text{and} \quad |f|_{B^{\beta'}_{p',p'}} \le \|f\|_{L^\infty}^{1-p/p'} |f|^{p/p'}_{B^{\beta}_{p,p}} \tag{9.25}$$

for any pair (β', p') satisfying $\beta' < \beta$, $\beta' p' = \beta p$, and

$$p' = \frac{2\alpha + d}{\beta' + d/2 + \alpha}. \tag{9.26}$$

From (9.26), by replacing β^* by β' and p_0^* by p' we can follow the same routine as we did to get (9.21) for $f \in B^{\beta^*}_{p_0^*,p_0^*}(\mathbb{R})$, where $p_0^* = \frac{2\alpha+d}{\beta^*+d/2+\alpha}$. Thus we can bound $E\|f - \widetilde{f}^*_{\gamma,\beta_0^*,m}\|^2_{L^2}$ by a constant multiple of

$$\begin{aligned} &|f|^{p'}_{B^{\beta'}_{p',p'}} \sigma^{2-p'}(2a^{2-p'} + c^* a^{-p'}) \\ &\quad + \frac{2c^*(2^d-1)\cdot 2^{m(2\alpha+d)}\sigma^2}{2^{2(d/2+\alpha)}-1} \int_{t > c_*^{-1/2}a} (t - c_*^{-1/2}a)^2 P(t)\,dt. \end{aligned} \tag{9.27}$$

THEOREM 9.2.2. *Let f be a compactly supported function on $\mathbb{R}^d$. We assume that $f \in L^\infty(\mathbb{R}^d) \cap B^{\beta}_{p,p}(\mathbb{R}^d)$, where*

$$d(1/p - 1/2) \le \beta \le (2\alpha + d)(1/p - 1/2).$$

Then with $\beta_0^* = d/2 - \alpha$ and $\gamma^* = a^*\sigma$, where a^* is the minimum point of (9.27),

$$E\|f_m - \widetilde{f}^*_{\gamma^*,\beta_0^*,m}\|^2_{L^2} \asymp \sigma^{2r_*}, \tag{9.28}$$

as $\sigma \to 0$, with the exponent

$$r = \frac{\beta p}{2\alpha + d}.$$

In particular, for $f \in B^{\beta}_{\tau,\tau}(\mathbb{R}^d)$, where $1/\tau = \beta/d + 1/2$, we have

$$r = \frac{d}{2\alpha + d} \cdot \frac{\beta}{\beta + d/2}.$$

Proof. As we did to get the asymptotic result for (9.21), with a rough estimation for a near optimal a, we can have

$$E\|f_m - \widetilde{f}_{\gamma,\beta_0,m}\|_{L^2}^2 \asymp \sigma^{2-p'}.$$

Since

$$\begin{aligned} 1 - p'/2 &= 1 - \frac{\alpha + d/2}{\beta' + d/2 + \alpha} \\ &= \frac{\beta'}{\beta' + d/2 + \alpha} \\ &= \frac{\beta' p'}{2\alpha + d} \\ &= \frac{\beta p}{2\alpha + d}, \end{aligned}$$

we have

$$r = \frac{\beta p}{2\alpha + d}$$

as the rate of convergence for $f \in L^\infty(\mathbb{R}^d) \cap B^\beta_{p,p}(\mathbb{R}^d)$.

For $f \in L^\infty(\mathbb{R}^d) \cap B^\beta_{\tau,\tau}(\mathbb{R}^d)$, since

$$r = \frac{\beta \tau}{2\alpha + d},$$

we have

$$r = \frac{d}{2\alpha + d} \cdot \frac{\beta}{\beta + d/2},$$

since $1/\tau = \beta/d + 1/2$. □

CHAPTER 10
COMPUTATIONS: TOMOGRAPHIC RECONSTRUCTION

We conducted tomographic reconstruction ($\alpha = 1/2$) experiments using $\widetilde{f}^*_{\gamma^*,\beta_0^*,m}$ in (8.1) with $\beta_0^* = 1/2$ and $\gamma^* = a^*\sigma$, where a^* is chosen as the (numerical) minimum point of (9.21). Our main conclusion is that this shrinkage parameter leads to smaller error and better reconstruction than using the parameter suggested by Kolaczyk [31]. The wavelet shrinkage method often exhibits certain artifacts, which do not appear in reconstructed images from the filtered backprojection method. Those artifacts in the wavelet shrinkage method are largely due to the fact that wavelet bases are not rotationally invariant. However, "rotational averaging technique" associated with the wavelet shrinkage method reduces those artifacts dramatically, and outperforms not only the standard wavelet shrinkage method but also the traditional filtered backprojection method in the mean square error measurement.

Our main computations are applied to the phantom image f_8 of 256×256 pixels in Figure 1. With the assumption that the true intensity field f of the digital image f_8 is bounded, we scaled f_8 to have 256 as the brightest pixel and 1 as the darkest one. We computed Radon projection data, $\mathcal{R}f$, at uniformly spaced 256 angles, and at uniformly spaced 256 points for each angle. Thus, the number of data, $N = 65536$, and $m = 8$ in $\widetilde{f}^*_{\gamma^*,\beta_0^*,m}$.

We added independent and identically distributed Gaussian noise with standard deviations $\sigma_0^{(25)} = 4.0734$, $\sigma_0^{(20)} = 7.2434$, and $\sigma_0^{(15)} = 12.8813$ to $\mathcal{R}f$. Here the superscripts in standard deviations denote *signal-to-noise ratio* (SNR) defined by

$$\mathrm{SNR} = 10\log_{10}\left(\frac{\sum_{i=0}^{255}\sum_{j=0}^{255}|\mathcal{R}f(\theta_i,u_j)|^2}{65536\times\sigma_0^2}\right).$$

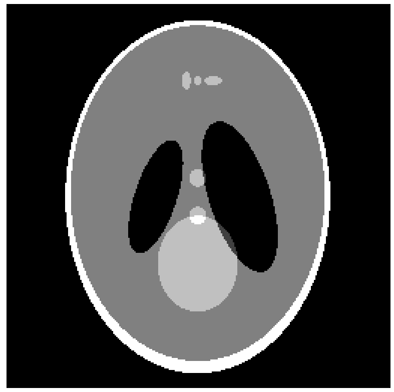

Fig. 1. Original image.

We use biorthogonal wavelets ${}_3\widetilde{\phi}$, ${}_{3,9}\widetilde{\psi}$, ${}_{3,9}\phi$, and ${}_{3,9}\psi$ illustrated on page 276 of [10]. These functions satisfy Theorem 6.3.1 for the Radon transform in two dimension. i We modified biorthogonal wavelets at the boundary in a way equivalent to assuming that the phantom image is periodic. We also choose $k_0 = 4$ in (8.1).

We assume that $f \in B^{\beta}_{p,p}(\mathbb{R}^2)$ for $\beta > 0$, where

$$p = \frac{3}{\beta + 3/2}. \tag{10.1}$$

To estimate the smoothness order β of f, we note that

$$\begin{aligned}
\sum_{k\geq 4} \sum_{|\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle| < 2^{k/2}\gamma} |\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle|^2 &= \sum_{k\geq 4} \sum_{|\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle| < 2^{k/2}\gamma} |\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle|^{2-p} \cdot |\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle|^p \\
&\leq \gamma^{2-p} \sum_{k\geq 4} 2^{k(2-p)/2} \sum_{j,i} |\langle f, \widetilde{\psi}^{(i)}_{k,j}\rangle|^p \\
&= \gamma^{2-p} |f|^p_{B^{\beta}_{p,p}},
\end{aligned} \tag{10.2}$$

where we have used the fact

$$\begin{aligned}
(\beta + 2(1/2 - 1/p))p &= \beta p + p - 2 \\
&= (3/p - 3/2)p + p - 2 \qquad \text{(by (10.1))} \\
&= (2-p)/2
\end{aligned} \tag{10.3}$$

in (10.2). Let

$$E(\gamma) = \left(\sum_{k \geq 4} \sum_{|\langle f, \widetilde{\psi}_{k,j}^{(i)} \rangle| < 2^{k/2}\gamma} |\langle f, \widetilde{\psi}_{k,j}^{(i)} \rangle|^2 \right)^{1/2}.$$

Then from (10.2), we have

$$E(\gamma)^2 \leq \gamma^{2-p} |f|_{B_{p,p}^{\beta}}^p. \tag{10.4}$$

On the other hand, if we denote the number of (j, i) for which $|\langle f, \widetilde{\psi}_{k,j}^{(i)} \rangle| \geq 2^{k/2}\gamma$ by $N_k(\gamma)$, then we have

$$\begin{aligned} &\sum_{k \geq 4} N_k(\gamma)(2^{k/2}\gamma)^p 2^{k(\beta+2(1/2-1/p))p} \\ &\qquad \leq \sum_{k \geq 4} 2^{k(\beta+2(1/2-1/p))p} \sum_{j,i} |\langle f, \widetilde{\psi}_{k,j}^{(i)} \rangle|^p \\ &\qquad = |f|_{B_{p,p}^{\beta}}^p. \end{aligned} \tag{10.5}$$

Notice that $2^{k(\beta+2(1/2-1/p)+1/2)p} = 2^k$ by (10.3). We define

$$N(\gamma) = \sum_{k \geq 4} N_k(\gamma) 2^k.$$

Then from (10.5), we have

$$\gamma^p \leq N(\gamma)^{-1} |f|_{B_{p,p}^{\beta}}^p. \tag{10.6}$$

Combining (10.1), (10.4), and (10.6), we have

$$E(\gamma) \leq N(\gamma)^{-\beta/3} |f|_{B_{p,p}^{\beta}}. \tag{10.7}$$

It is remarked in [7] that (10.7) is invertible, i.e., if we observe

$$E(\gamma) \leq C N(\gamma)^{-\beta/3} \tag{10.8}$$

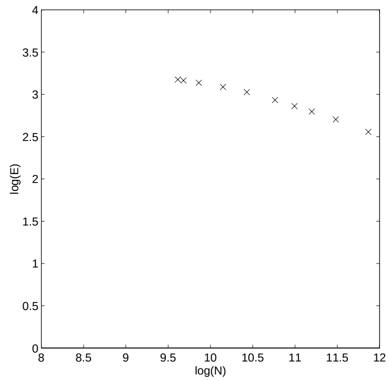

Fig. 2. $\log N(\gamma)$ vs $\log E(\gamma)$.

for some β and C, then one can conclude that $f \in B^{\beta}_{p,p}(\mathbb{R}^2)$, $p = \frac{3}{\beta+3/2}$, and one can define an equivalent semi-norm on $B^{\beta}_{p,p}(\mathbb{R}^2)$ such that in this semi-norm $|f|_{B^{\beta}_{p,p}} = C$. (This statement is not correct, but is close enough to the truth to be used in practice; see [7] for the precise statement.) We use (10.7) in estimating the smoothness order β of f. We compute $E(\gamma)$ and $N(\gamma)$ for several γ, and estimate β and $|f|_{B^{\beta}_{p,p}}$ from $\log E(\gamma)$ and $\log N(\gamma)$ graph at Figure 2. With this approach, we estimated

$$\beta \approx 0.8040$$

and

$$|f|_{B^{\beta}_{p,p}} \approx 337.6403$$

with $p = 1.3021$.

We computed c_1 in (8.8) numerically by

$$\begin{aligned} c_1 &= \int_{\mathbb{R}^2} \sqrt{\xi_1^2 + \xi_2^2} \left|\widehat{\widetilde{\psi}}(\xi_1)\right|^2 \left|\widehat{\widetilde{\phi}}(\xi_2)\right|^2 d\xi_1 d\xi_2 \\ &\approx 0.2525, \end{aligned}$$

and similarly we have $c_2 \approx 0.2525$ and $c_3 \approx 0.4400$. Thus we take 0.2525 as the approximated value for c_* and 0.4400 for c^*.

We compute the minimum point of (9.21) as a function of a, with $p_0^* = p = 1.3021$, $\alpha = 1/2$, $|f|_{B^\beta_{p,p}}^{p_0^*}$ (≈ 1960.2), c^* (≈ 0.4400), and c_* (≈ 0.2525), by the bisection method for each $\sigma = 2^{-8}\sigma_0$. We denote the resulting algorithm $\widetilde{f}^*_{\gamma^*,\beta_0^*,m}$ by f_C, i.e.,

$$f_C = \sum_{k\geq 4}^{7} \sum_{j,i} S_{a\sigma 2^{k/2}}(2^{k/2}[Y, U_{k,j}^{(i)}])\psi_{k,j}^{(i)} + \sum_l 2^2[Y, V_{k_0,l}]\Phi_{4,l},$$

where a is chosen as the numerical minimum point of (9.21) for given $\sigma = 2^{-8}\sigma_0$, and $E\|f_8 - f_C\|_{L^2}^2$ by E_C.

We also consider

$$f_C^{(R)} = \frac{1}{r}\sum_{r=0}^{R-1}\Bigg(\sum_{k\geq 4}^{7}\sum_{j,i} S_{a\sigma 2^{k/2}}(2^{k/2}[Y(\theta - \pi r/2R, \cdot), U_{k,j}^{(i)}])\psi_{k,j}^{(i)} \\ + \sum_l 2^2[Y(\theta - \pi r/2R, \cdot), V_{k_0,l}]\Phi_{4,l}\Bigg),$$

where a is, again, chosen as the numerical minimum point of (9.21) for each σ with same estimated values (α, $p_0^*, p, |f|_{B^\beta_{p,p}}^{p_0^*}, c^*$, and c_*) used in f_C. Thus $f_C^{(R)}$ is the average of R images reconstructed from

$$Y(\theta, \cdot),\ \ Y(\theta - \pi/2R, \cdot),\ \ \cdots,\ \ Y(\theta - \pi(R-1)/2R, \cdot)$$

applying f_C to each rotated data set. We conducted experiments using $f_C^{(2)}$. The mean square error $E\|f_8 - f_C^{(2)}\|_{L^2}^2$ is denoted by $E_C^{(2)}$.

In [31] Kolaczyk used the shrinkage parameter

$$a_{k,j,i} = \sqrt{2\log(2^{2k})}2^{k/2}\sigma c_i^{1/2} \tag{10.8}$$

in (9.1) for the tomographic reconstruction. This choice of shrinkage parameter is motivated by the VisuShrink method of Donoho and Johnstone [23]. For the comparison purpose, we define

$$f_V = \sum_{k=6}^{7}\sum_{j,i} S_{a_{k,j,i}}(2^{k/2}[Y, U_{k,j}^{(i)}])\psi_{k,j}^{(i)} + \sum_l 2^3[Y, V_{6,l}]\Phi_{6,l},$$

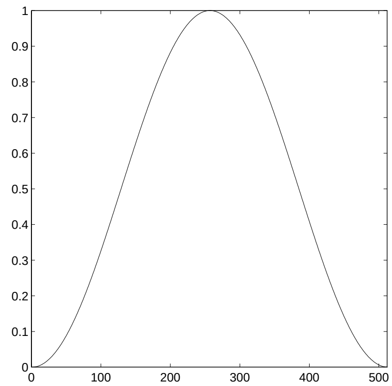

Fig. 3. Hamming weight filter.

Notice that f_V only shrinks the first two highest level wavelet coefficients. This is what Kolaczyk [31] suggested for tomographic reconstruction. (In [31] Meyer's wavelets are used for experiments.) The mean square error $\|f_8 - f_V\|_{L^2}^2$ is denoted by E_V.

We also consider the filtered backprojection method using the Hamming weight filter w_H (see, e.g., [40]) with cutoff 0.5. We plotted w_H for 512 points in Figure 3. We define f_F by

$$f_F = \int_{\mathbb{R}^2} e^{2\pi i \xi \cdot x} w_H(\xi)|\xi| \widehat{\mathcal{R}^* Y}(\xi) d\xi.$$

We denote $E\|f_8 - f_F\|_{L^2}^2$ by E_F.

In f_C and f_V, we first compute

$$\begin{aligned} 2^4[Y, V_{8,l}] &= 2^4[Y, 2^4 \mathcal{R}\rho_{8,l}] \\ &= 2^8 \langle \mathcal{R}^* Y, \rho_{8,l} \rangle \\ &= \int_{\mathbb{R}^2} \widehat{\mathcal{R}^* Y}(\xi) e^{2\pi i \xi \cdot j 2^{-8}} \overline{\widehat{\rho}}(2^{-8}\xi)\, d\xi, \\ &= 2^{-8} \int_{\mathbb{R}^2} |\xi| \widehat{\mathcal{R}^* Y}(\xi) e^{2\pi i \xi \cdot j 2^{-8}} \overline{\widehat{\widehat{\Phi}}}(2^{-8}\xi)\, d\xi \qquad \text{(see (7.2))} \end{aligned}$$

and use the recursive algorithm described in Section 7.1 to get $\{2^{k/2}[Y, U_{k,j}^{(i)}]\}_{\{4 \le k < 7, j\}}$ and $\{2^2[Y, V_{4,l}]\}_l$.

For $\mathcal{R}^*Y$, we note that

$$\begin{aligned}\widehat{\mathcal{R}^*\mathcal{R}f}(w\cos\theta, w\sin\theta) &= |w|^{-1}\widehat{f}(w\cos\theta, w\sin\theta)\\ &= |w|^{-1}\mathcal{R}f(\theta,\cdot)^\wedge(w). \qquad \text{(by Theorem 5.1.1)}\end{aligned}$$

Thus we can compute $\mathcal{R}^*Y$ by

$$\widehat{\mathcal{R}^*Y}(w\cos\theta, w\sin\theta) = |w|^{-1}Y(\theta,\cdot)^\wedge(w).$$

In numerical computations involving the Fourier transform, we used the *fast Fourier transform* (see, e.g., [29]). To reduce the artifacts from the fast Fourier transform, we *zero-padded* the projection data of 256×256 to 256×512 (see, e.g., [40]).

Table 1

SNR	E	E_C	$E_C^{(2)}$	E_V	E_F
25	2205.8	690.3	652.2	699.2	658.8
20	3731.5	753.4	680.1	845.8	1195.2
15	5584.0	932.6	791.1	1214.2	2413.6

Table 1 contains the results of our tests. We calculated mean square errors E_C, $E_C^{(2)}$, E_V, and E_F with three different SNR 25, 20, and 15. The second column in Table 1 show mean square errors of the direct backprojection method (without filtering or wavelet shrinkage) for each noise level. We also report that $\|f_8\|_{L^2}^2 = 10100$.

Figures 4-(a), 4-(b), and 4-(c) show 256×256 reconstructions based on f_C for data with SNR 25, 20, and 15, respectively. The algorithm f_C has a certain drawback in reconstruction. We can see "square blocks"'near edges in Figure 4. This is largely due to the fact that wavelet bases are not rotationally invariant.

The algorithm $f_C^{(2)}$ is designed to reduce this phenomenon. Figure 5-(a), 5-(b), and 5-(c) show 256×256 reconstruction based on $f_C^{(2)}$ for data with SNR 25, 20, and 15, respectively. Most of the artifacts in the reconstructed images of Figure 4 are disappeared in those of Figure 5. Moreover, $f_C^{(2)}$ overally outperforms f_C, f_V, and f_F.

The algorithm f_V has less artifacts, but it is outperformed by f_C and $f_C^{(2)}$. Since f_V employs the wavelet shrinkage method for the first two highest wavelet coefficients, it can be viewed a combined method of the filtered backprojection described in Section 5.2 and wavelet shrinkage. Figures 6-(a), 6-(b), and 6-(c) show 256×256 reconstructions based on f_V for data with SNR 25, 20, and 15, respectively.

Figures 7-(a), 7-(b), and 7-(c) show 256×256 reconstructions based on f_F for data with SNR 25, 20, and 15, respectively. For high noise levels, f_F is outperformed by the other three methods with significant differences in the mean square error measurement.

We applied $f_C^{(2)}$ to real *positron emission tomography* (PET) data set. Four images in Figure 8 show reconstruction using $f_C^{(2)}$ with four different shrinkage parameters. We also present images reconstructed from the same PET data in Figure 9. The images in Figure 9 are reconstructed by using the filtered backprojection method, but we do not know exactly what weight filter is employed. We leave it to the reader to compare the overall visual quality of reconstructions in Figure 8 and Figure 9.

We believe that minimizing our bound on the error (9.27) leads to near-optimal shrinkage parameters for tomographic reconstruction with wavelet shrinkage. Moreover, our technique for estimating the smoothness of images leads to accurate estimates of the true smoothness of images. We also can predict accurately the performance of the wavelet tomographic reconstruction algorithm using only two smoothness parameters β and $|f|_{B^{\beta}_{p,p}}$. We also believe that rotational averaging techniques in $f_C^{(2)}$ remove most of the artifacts of wavelet shrinkage methods such as in f_C and f_V, while preserving a great amount of noise removal.

Fig. 4-(a). f_C with SNR = 25.

Fig. 4-(b). f_C with SNR = 20.

Fig. 4-(c). f_C with SNR = 15.

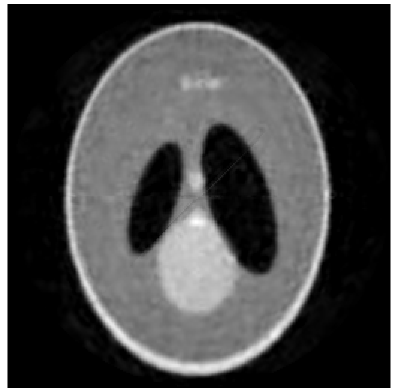

Fig. 5-(a). $f_C^{(2)}$ with SNR = 25.

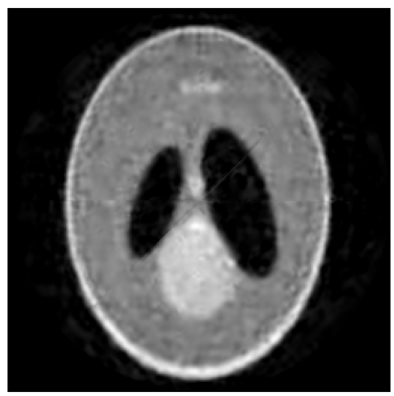

Fig. 5-(b). $f_C^{(2)}$ with SNR = 20.

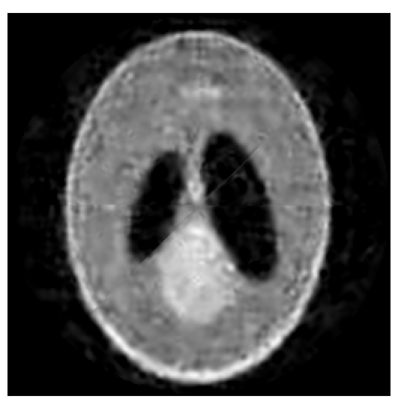

Fig. 5-(c). $f_C^{(2)}$ with SNR = 15.

Fig. 6-(a). f_V with SNR = 25.

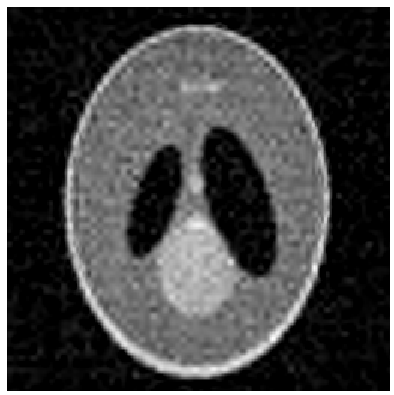

Fig. 6-(b). f_V with SNR = 20.

Fig. 6-(c). f_V with SNR = 15.

Fig. 7-(a). f_F with SNR = 25.

Fig. 7-(b). f_F with SNR = 20.

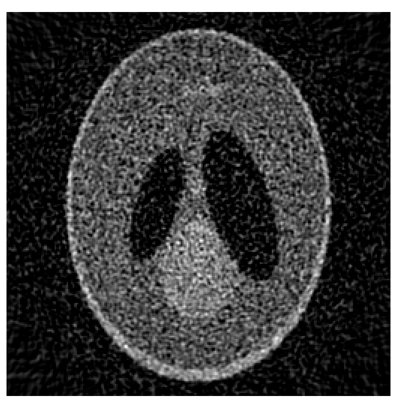

Fig. 7-(c). f_F with SNR = 15.

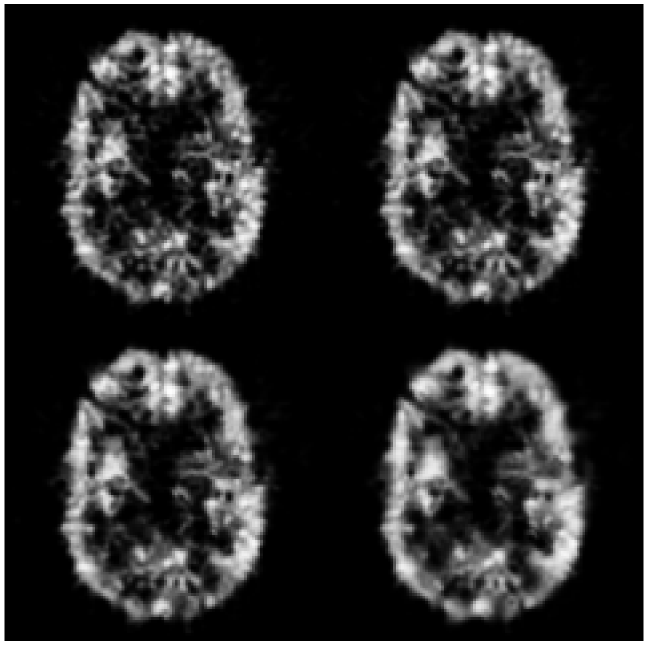

Fig. 8. reconstructed images with $f_C^{(2)}$.

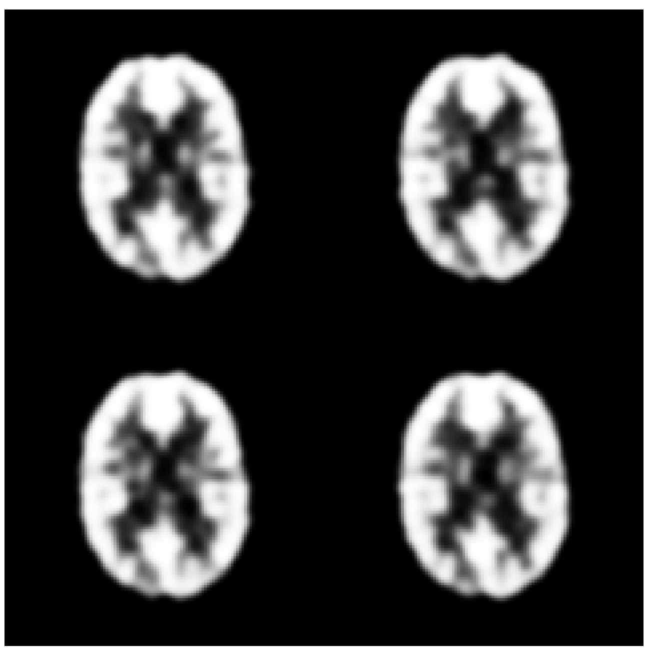

Fig. 9. reconstructed images with FBP.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] J. O. Berger, *Statistical decision theory and Bayesian analysis*, Springer-Verlag, 1985.

[2] C. Bouman and K. Sauer, *Bayesian estimation of transmission tomograms using segmentation based optimization*, IEEE Trans. Nuclear Science, 39, 4, pp. 1144–1152.

[3] T. Cai, *Minimax wavelet estimations via block thresholding*, Technical Report #96-41, Purdue University.

[4] A. Chambolle, R. A. DeVore, N.-Y. Lee, and B. J. Lucier, *Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage*, IEEE Transactions on Image Processing to appear.

[5] A. Cohen, I. Daubechies, and J.-C. Feauveau, *Biorthogonal bases of compactly supported wavelets*, Comm. Pure Appl. Math., 45 (1992), pp. 485–560.

[6] A. Cohen, I. Daubechies, and P. Vial, *Wavelets on the interval and fast wavelet transforms*, Appl. Comput. Harmonic Analysis, 1 (1993), pp. 54–81.

[7] A. Cohen, R. A. DeVore, and R. Hochmuth, *Restricted nonlinear approximation*, in preparation.

[8] I. Daubechies, *Orthonormal basis of compactly supported wavelets*, Comp. Pure Appl. Math., 46 (1988), pp. 909–996.

[9] I. Daubechies, *The wavelet transform, time-frequency localization and signal analysis*, IEEE Trans. Inform. Theory, 36 (1990), pp. 961–1005.

[10] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference series in Applied Mathematics 91, SIAM, Philadelphia, 1992.

[11] I. Daubechies and J. Lagarias, *Two-scale difference equations I. Existence and global regularity of solutions*, SIAM J. Math. Anal., 22 (1991), pp. 1388–1410.

[12] I. Daubechies and J. Lagarias, *Two-scale difference equations II. Local regularity, infinite products of matrices and fractals*, SIAM J. Math. Anal., 23 (1992), pp. 1031–1079.

[13] R. DeVore, B. Jawerth, and B. Lucier, *Image Compression through wavelet transform coding*, IEEE Trans. Information Theory, 38, 2 (1992), pp. 719–746; Special issue on Wavelet Transforms and Multiresolution Analysis.

[14] R. DeVore, B. Jawerth, and B. Lucier, *Surface compression*, Computer Aided Geom. Design, 9 (1992), pp. 219–239.

[15] R. DeVore, B. Jawerth, and V. Popov, *Compression of wavelet decompositions*, Amer. J. Math., 114 (1992), pp. 737–785.

[16] R. DeVore and G. Lorentz, *Constructive Approximation*, Springer–Verlag, New York, 1993.

[17] R. A. DeVore and B. J. Lucier, *Fast wavelet techniques for near-optimal image processing*, in IEEE Military Communications Conference Record, San Diego, October 11-14, 1992, IEEE Press, Piscataway, NJ, 1992, pp. 1129–1135.

[18] R. A. DeVore and B. J. Lucier, *Classifying the smoothness of images: Theory and applications to wavelet image processing*, in ICIP–94: Proceedings of the 1994 IEEE International Conference on Image Processing, Austin, TX, November 13–16 II, IEEE Press, Los Alamitos, CA, 1994, pp. 6–10.

[19] D. A. DeVore, P. Petrushev, and X. M. Yu, *Nonlinear wavelet approximations in the space $C(\mathbb{R}^d)$*, in Progress in Approximation Theory, Proceedings of the US/USSR Conference on Approximation, Tampa, 1990, Springer-Verlag, New York, 1992, pp. 261–283.

[20] R. A. DeVore and V. Popov, *Interpolation of Besov spaces*, Trans. Amer. Math. Soc., 305 (1988), pp. 397–414.

[21] D. Donoho, *De-noising by soft-thresholding*, IEEE Trans. Information Theory, 41 (1995), pp. 613–627.

[22] D. Donoho, *Nonlinear solution of linear inverse problem by wavelet-vaguelette decomposition*, Appl. Comput. Harmon. Anal., 2 (1995), pp. 101–126.

[23] D. Donoho and I. Johnstone, *Adapting to unknown smoothness via wavelet shrinkage*, J. Amer. Statist. Assoc., 90 (1995), pp. 1200–1224.

[24] D. Donoho and I. Johnstone, *Neo-classical minimax problems, thresholding and adaptive function estimations*, Bernoulli, 2 (1996), pp. 39–62.

[25] D. Donoho and I. Johnstone, *Ideal spatial adaptation by wavelet shrinkage*, Biometrika, 81 (1994), pp. 425–455.

[26] D. Donoho, I. Johnstone, G. Kerkyacharian, and D. Picard, *Wave-*

let shrinkage: Asymptopia ?, J. Roy. Statist. Soc. Ser. B, 57 (1995), pp. 301–369.

[27] M. Frazier and B. Jawerth, *A discrete transform and decompositions of distribution spaces*, J. of Functional Anal., 93 (1990), pp. 34-170.

[28] C. Herley and M. Vetterli, *Biorthogonal bases of symmetric compactly supported wavelets*, in in Wavelets, Fractals, and Fourier Transforms, Oxford Univ. Press, New York, 1993, pp. 991–1008.

[29] A. C. Kak and M. Slaney, *Principles of computerized tomographic imaging*, IEEE Press, 1988.

[30] D.-G. Kim, *Wavelet decompositions and function spaces on the unit cube*, Ph.D. Thesis, Purdue University, 1994.

[31] E. D. Kolaczyk, *A wavelet shrinkage approach to tomographic image reconstruction*, J. Amer. Statist. Assoc., 91 (1996), pp. 1079–1090.

[32] G. C. Kyriazis, *Wavelet decompositions and spaces of functions*, Ph.D. Thesis, University of South Carolina, 1992.

[33] B. J. Lucier, M. Kalergi, W. Qian, R. A. DeVore, R. A. Clark, E. B. Saff, and L. P. Clarke, *Wavelet compression and segmentation of mammographic images*, J. of Digital Imaging, 7 (1994), pp. 27–38.

[34] P. Maass, *Wavelet-projection methods for inverse problems*, preprint.

[35] Y. Meyer, *Ondelettes et opérateurs I: Ondelettes*, Hermann, Paris, 1990; English transl. by D. H. Salinger, *Wavelets and operators*, Cambridge Univ. Press, Cambridge, 1992.

[36] Y. Meyer, *Ondelettes et opérateurs II: Opérateurs de Calderón-Zygmund*, Hermann, Paris, 1990.

[37] F. Natterer, *The mathematics of computerized tomography*, John Wiley and Sons, New York, 1986.

[38] A. Rosenfeld and A. C. Kak, *Digital picture processing I*, Academic Press, 1982.

[39] Y. Royden, *Real Analysis*, Macmillan Publishing, New York, 1968.

[40] J. L. C. Sanz, E. B. Hinkle, and A. K. Jain, *Radon and projection transform-based computer vision*, Springer-Verlag, 1987.

[41] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Univ. Press, Princeton, New Jersey, 1970.

[42] A. N. Tikhonov, *Solution of incorrectly formulated problems and regularization method*, Soviet Math. Dok., 4 (1963), pp. 1035.

[43] H. TRIBEL, *Theory of function spaces II*, Monographs in Mathematics, vol. 78, Birkhauser Verlag, Basel, 1983.

[44] R. M. YOUNG, *An Introduction to Nonharmonic Fourier Series*, Academic Press, New York, 1976.

VITA

VITA

Namyong Lee was born on Aug 1, 1967 in Seoul, Korea. He began his collegiate years at Seoul National University in Seoul in 1986, and completed the Bachelor of Science degree in Mathematics in 1990 at the same university. He pursued his graduate studies in the Department of Mathematics at Purdue University, West Lafayette, Indiana in 1992 and received the Doctor of Philosophy degree in Mathematics in 1997.