Applications of Orthonormal Bases of Wavelets to Deconvolution

by

Norman C. Corbett

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Applied Mathematics

Waterloo, Ontario, Canada, 1997

0-612-21336-6

The University of Waterloo requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

Abstract

Convolution integral equations arise frequently in many areas of science and engineering. If the kernel of such an equation is well behaved, say integrable, then the task of solving a convolution equation is ill-posed. Indeed, if the kernel is integrable, then the Riemann-Lebesgue Lemma implies that the recovery of high frequency information pertaining to the unknown function will be difficult, if not impossible.

Orthonormal wavelet bases are bases generated by translating and dilating a single function, known as the mother wavelet. One key advantage of these bases is that the mother wavelet can be selected to have fast decay in both the time and frequency domains. This property suggests that wavelet bases may be useful when attempting to solve a convolution equation.

In this thesis, we investigate the applicability of orthonormal wavelet bases with regard to solving convolution equations. In particular, we concentrate on the construction of approximations to the unknown function belonging to scaling function subspaces. We also briefly consider regularization algorithms which are based on the multiresolution analysis, a structure defined by the scaling function associated with the mother wavelet.

Acknowledgements

I would like to acknowledge the generous financial support that I received from the Department of Applied Mathematics, my supervisor and the University of Waterloo.

I want to thank Franklin Mendivil and Leslie Shayer for proof reading the thesis and giving me many valuable suggestions concerning content and presentation.

I would also like to thank Ryan and Michael for occupying the office across from the smoking lounge and for listening patiently to my post nicotine prattle.

Many thanks go to my friends Erik, Fernando, Ian, Janet, Jason and Nat for making my stay in Ducktown tolerable. Thanks go to Haimish for his well-timed ejection of protein based spheres.

Special thanks go to my dear friends Gary and Raelene who stood by me and saw me through difficult times. I would also like to express my appreciation to the good folks at Guiness and Lee Valley without whom civilization as we know it would surely decay into ruins.

Finally, I would like to thank my family who have taught me the lessons of honor, courage and honesty and have supported me in all of my endeavors.

Dedication

To my dearest heart. Leslie. with love.

Contents

List of Figures

Chapter 1

Introduction

Consider the convolution integral equation

$$(\mathcal{G}u)(t) = \int_{-\infty}^{\infty} g(t-\tau)u(\tau)\,d\tau = y(t), \qquad t \in \mathbb{R}, \tag{1.1}$$

where the functions g and y are known. In what follows, we shall address the problem of solving equation 1.1 for the unknown function u, in the case where the kernel g is integrable. This problem is known as the problem of *deconvolution* and when $g \in L^1(\mathbb{R})$, it is ill-posed in the sense of Hadamard (see Section 2.2).

One possible approach to the problem of deconvolution is to assume that the unknown function u admits an expansion in terms of some complete system, say $\{\xi_i\}$. That is, if we assume that

$$u = \sum_i a_i \xi_i,$$

then the problem of solving 1.1 for u is equivalent to the problem of solving the semi-discrete equation

$$\sum_i a_i \int_{-\infty}^{\infty} g(t-\tau)\xi_i(\tau)\,d\tau = y(t)$$

for the unknown scalars $\{a_i\}$.

In this thesis, we will consider a particular type of complete system. Specifically, we will suppose that the unknown function can be expanded in terms of an orthonormal wavelet basis. A wavelet basis is generated by translating and dilating a single function $\psi \in L^2(\mathbb{R})$, known as the mother wavelet. That is, a wavelet basis is a doubly indexed set of functions which are of the form

$$\psi^{j,k}(t) = 2^{j/2}\psi(2^j t - k), \tag{1.2}$$

for $j, k \in \mathbb{Z}$. The expansion of u with respect to the basis generated by 1.2 will take the form

$$u = \sum_{j,k} u^j[k]\psi^{j,k} \tag{1.3}$$

and hence, the problem of solving 1.1 for u is equivalent to the problem of solving the equation

$$\sum_{j,k} u^j[k] \int_{-\infty}^{\infty} g(t-\tau)\psi^{j,k}(\tau)\, d\tau = y(t)$$

for the sequence $\{u^j[k] : j, k \in \mathbb{Z}\}$.

If we use the so-called scaling function ϕ associated with ψ, then 1.3 can be written in the form

$$u = \sum_k u_n[k]\phi^{n,k} + \sum_{j\geq n}\sum_k u^j[k]\psi^{j,k},$$

where

$$\phi^{n,k}(t) = 2^{n/2}\phi(2^n t - k)$$

and $n \in \mathbb{Z}$ is assumed to be fixed. It follows that we can also choose to solve the equation

$$\sum_k u_n[k] \int_{-\infty}^{\infty} g(t-\tau)\phi^{n,k}(\tau)\, d\tau + \sum_{j\geq n}\sum_k u^j[k] \int_{-\infty}^{\infty} g(t-\tau)\psi^{j,k}(\tau)\, d\tau = y(t),$$

for the scalars $u_n[k]$ and $u^j[k]$, where $k \in \mathbb{Z}$ and $j \geq n$.

Wavelet analysis has been an active area of research for well over a decade and can be thought of as an alternative to traditional Fourier analysis. Although relatively new, research concerning the application of wavelet analysis to inverse problems seems promising. It is hoped that the special properties of wavelets can be exploited to yield methods which effectively deal with the ill-posedness of said problems.

With regard to the problem of deconvolution, the article [42] explores the possibility of using the continuous wavelet transform to solve 1.1, while in [18], it shown that wavelet bases can often be used to define a mathematical construct which mimics the singular value decomposition.

In work similar to our own (see [43]), the author uses a wavelet expansion to solve a convolution equation arising from a mixture problem for random variables. The proposed method begins with the assumption that the unknown function belongs to the scaling function subspace V_n, given by the closed linear span

$$V_n = \overline{\bigvee \{\phi^{n,k} : k \in \mathbb{Z}\}}. \tag{1.4}$$

In this case the wavelet coefficients $u^j[k]$ vanish whenever $j \geq n$.

Convolution equations, for which the Fourier transform of the kernel g satisfies

$$|\hat{g}(\omega)| > 0$$

are considered and a method for the recovery of the scaling function coefficients $u_n[k]$, $k \in \mathbb{Z}$, is developed in the case where ϕ is a scaling function of Meyer type. Of course, in most cases, the assumption $u \in V_n$ is not in fact satisfied and the recovery of the scalars $u_n[k]$ provides only an approximation

$$u \approx u_n = \sum_k u_n[k]\phi^{n,k}.$$

In Chapter 5, we make some progress towards the generalization of the results presented in [43]. In particular, the assumption $u \in V_n$, leads us to consider the linear operator $\mathcal{G}|_{V_n}$, the restriction of $\mathcal{G}$ to the subspace V_n. We present necessary and sufficient conditions for the strong invertibility of the operator $\mathcal{G}|_{V_n}$. Sufficient conditions for the weak invertibility of $\mathcal{G}|_{V_n}$ are also presented. These results are valid for a large class of scaling functions (including those of Meyer type) and are based upon the less restrictive assumption that g be integrable. Furthermore, the presentation of the conditions concerning strong and weak invertibility makes use of a continuous 2π-periodic function $\hat{G}_n$, which can be considered to be the spectrum of a particular Toeplitz matrix. The behavior of $\hat{G}_n$ as $|n| \to \infty$ is examined and this investigation leads to a convergence result. Particularly, it is shown that under certain conditions, the approximations u_n converge to the solution u of 1.1 as $n \to \infty$. The material presented in Appendix A is then used to formulate a convergence rate estimate in the case where u belongs to the Sobolev space $H^s(\mathbb{R})$.

Chapter 6 begins with the consideration of the problem of computing the projection $P_m u$, onto V_m, under the assumption that $u \in V_n$, $n > m$. Our investigation leads us to the examination of the operator

$$\mathcal{P}(A)(\omega) = \frac{1}{2}\left\{|H(\omega/2)|^2 A(\omega/2) + |H(\omega/2+\pi)|^2 A(\omega/2+\pi)\right\},$$

where H and A are continuous 2π-periodic functions and

$$|H(\theta)|^2 + |H(\theta+\pi)|^2 = 2$$

for all θ. The operator $\mathcal{P}$ arises in the study of orthonormal bases of wavelets and we consider the behavior of the functions $\mathcal{P}^k(A)$ as $k \to \infty$. In particular, we present two results regarding the convergence of the sequence $\{\mathcal{P}^k(A) : k \in \mathbb{N}\}$ and are thus able to comment upon the sensitivity of $P_m u$ to perturbations in y in the case where m is small.

In related work, found in [30], a multiresolution based regularization algorithm is proposed. This algorithm is based upon the multiresolution analysis, the sequence of nested subspaces $\{V_n : n \in \mathbb{Z}\}$, defined by 1.4. Specifically, the author of [30] uses a multiresolution analysis, defined by the Haar scaling function

$$\phi(t) = \begin{cases} 1, & 0 \leq t < 1 \\ 0, & \text{otherwise} \end{cases} .$$

to construct approximate solutions for a distributed parameter estimation problem.

In the last two sections of Chapter 6, we briefly consider certain aspects of the method introduced in [30], where we show that the proposed regularization algorithm is a special case of the method of C-generalized regularization. Let $\mathcal{C}$ be a linear operator. When using the method of C-generalized regularization, one computes an approximate solution to equation 1.1 by finding the minimizer $u_{c,\alpha}$ of the functional

$$F(u) = \|\mathcal{G}u - y\|^2 + \alpha \|\mathcal{C}u\|^2 . \tag{1.5}$$

In the case of multiresolution regularization, the operator $\mathcal{C}$ can be expressed as the weighted sum of projections

$$\mathcal{C} = \sum_j \lambda_j Q_j,$$

where Q_j is the projection onto the wavelet subspace

$$W_j = \overline{\bigvee \{\psi^{j,k} : k \in \mathbb{Z}\}}.$$

It can be shown that, under appropriate conditions on y, there exists a unique function $u_c^\dagger$, called the C-generalized solution of 1.1, such that $u_{c,\alpha} \to u_c^\dagger$ as $\alpha \to 0$. The special case $\mathcal{C} = I$, I the identity operator, corresponds to Tikhonov, or minimum norm regularization. Furthermore, the corresponding minimizer $u_{i,\alpha}$

tends to $u^\dagger$ as $\alpha \to 0$, where $u^\dagger$ is the generalized solution of 1.1. In general, the two functions $u_c^\dagger$ and $u^\dagger$ are distinct. However there are cases for which the two generalized solutions are close.

We conclude Chapter 6 with a comparison of the minimizers $u_{c,\alpha}$ and $u_{i,\alpha}$ as well as the corresponding generalized solutions $u_c^\dagger$ and $u^\dagger$. It is shown that if the operator C is, in some sense, close to the identity I, then the minimizers and the generalized solutions are close.

The discrete equivalent of equation 1.1 is

$$\sum_k g_{j-k} u_k = y_j, \qquad j \in \mathbb{Z}. \tag{1.6}$$

The articles [33] and [8] are concerned with the application of wavelet methods to the problem of solving 1.6 for the sequence $\{u_k : k \in \mathbb{Z}\}$. In [8], a redundant version of the discrete wavelet transform is used to change 1.6 into a system of discrete convolution equations. The technique of Wiener filtering is the used to solve the individual equations, whereupon the inverse discrete wavelet transform is used to construct an estimate of $\{u_k : k \in \mathbb{Z}\}$. In Chapter 3, we demonstrate that the method discussed in [8] can be regarded as a multiscale regularization algorithm, similar to the algorithm proposed in [30].

In [33], the discrete filters used are defined through the use of the blurring sequence $\{g_k : k \in \mathbb{Z}\}$. In Chapter 3, we provide examples to show that, in general, such filters are not discrete wavelet filters. However, the method introduced does define a new kind of regularization. We investigate this method in some detail and provide a proof of the important property of regularity.

Chapter 2

Convolution equations

2.1 Introduction

Convolution integral equations appear frequently in applications and hence, inverse problems, which are based upon these types of equations, are often encountered. To see how such equations can arise, we consider a simple example. The initial value problem

$$\begin{aligned} y''(t) + ay'(t) &+ by(t) = u(t) \\ y'(0) &= y(0) = 0 \end{aligned} \tag{2.1}$$

models a forced, damped harmonic oscillator. If the Laplace transform is applied to 2.1, then we obtain the equation

$$Y(s) = \frac{U(s)}{s^2 + as + b}, \tag{2.2}$$

where Y and U denote the Laplace transforms of y and u respectively. The inverse Laplace transform can now be applied to 2.2 and the result is the convolution equation

$$y(t) = \int_0^t g(t-\tau)u(\tau)\,d\tau, \tag{2.3}$$

where g is the inverse Laplace transform of $(s^2 + as + b)^{-1}$.

Given a forcing function u, we can use 2.3 to compute the displacement y. This is the forward problem. Since it is often easier to observe the evolution of a system rather than the external forces causing this evolution, a more natural problem might be: Given the function y find the function u. This is one possible inverse problem that is based upon equation 2.3. The other involves the determination of the constants a and b from the functions y and u. The latter inverse problem is known as a *system identification problem*. In this thesis, we will be concerned with the first problem, that is, the determination of the function u from the functions g and y.

2.2 Deconvolution

Suppose that the functions g and y are known. The problem of *deconvolution* involves solving the convolution type integral equation

$$\int_{-\infty}^{\infty} g(t-\tau)u(\tau)\,d\tau = y(t), \quad t \in \mathbb{R}, \tag{2.4}$$

for the unknown function u. The equation 2.4 can be regarded as a model of a linear system, the properties of which are determined by the *kernel* g. With this interpretation in mind, we can regard the function u as the input to this linear system, while y can be thought of as the resulting output.

If $g \in L^1(\mathbb{R})$, then the linear operator $\mathcal{G}$, defined by

$$(\mathcal{G}u)(t) = \int_{-\infty}^{\infty} g(t-\tau)u(\tau)\,d\tau, \tag{2.5}$$

is a continuous mapping of $L^2(\mathbb{R})$ into $L^2(\mathbb{R})$. Furthermore, for any $u \in L^2(\mathbb{R})$, there is a unique $y \in L^2(\mathbb{R})$ such that $y = \mathcal{G}u$. In other words, given $g \in L^1(\mathbb{R})$ and any $u \in L^2(\mathbb{R})$, the problem of determining y is well-posed in the sense of Hadamard. On the other hand, given $g \in L^1(\mathbb{R})$ and $y \in L^2(\mathbb{R})$, the problem of determining u is ill-posed. In particular, at least one of the following conditions:

1. A solution u exists for any $y \in L^2(\mathbb{R})$.
2. The solution u is unique. (2.6)
3. The solution u depends continuously on the *data* y.

will be violated. To see how this happens, it is convenient to have a formal expression for the solution of 2.4. The Fourier convolution theorem allows us to write 2.4 in the alternative form

$$\hat{g}(\omega)\hat{u}(\omega) = \hat{y}(\omega), \quad \omega \in \mathbb{R},$$

which implies that

$$u(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t}\frac{\hat{y}(\omega)}{\hat{g}(\omega)}\,d\omega. \tag{2.7}$$

Now, it is easy to see that if $y \in R(\mathcal{G})$, the range of the operator $\mathcal{G}$, then $\hat{y} = \hat{g}\hat{u}$ and a solution $u \in L^2(\mathbb{R})$ will exist. Unfortunately, when $g \in L^1(\mathbb{R})$, it follows from the Open Mapping Theorem that $R(\mathcal{G})$ is never a closed set. This means that $R(\mathcal{G}) \neq L^2(\mathbb{R})$ and a solution may not exist for an arbitrary $y \in L^2(\mathbb{R})$.

The uniqueness of the solution, when it exists, is equivalent to the condition that the operator $\mathcal{G}$ be injective. This is not always the case. For example, suppose that $\hat{g}$ vanishes identically on some interval I and let u be a solution of 2.4 corresponding to $y \in R(\mathcal{G})$. If $f \in L^2(\mathbb{R})$ is any function such that $\text{supp}(\hat{f}) \subset I$, then $u + f$ is also a solution of 2.4 corresponding to $y \in R(\mathcal{G})$. In general, $\mathcal{G}$ will be an injection as long as $\hat{g}$ vanishes only on a set of measure zero (for example, a countable set of points).

We can rectify some of the problems encountered when considering the questions of existence and uniqueness by generalizing our notion of a solution. Specifically, we say that u is a *least-squares solution* (see [21]) of 2.4 if it is a minimizer of the functional

$$F(u) = \|\mathcal{G}u - y\|^2, \tag{2.8}$$

whenever a minimizer exists. Let $\mathcal{G}^*$ be the adjoint of $\mathcal{G}$. It can be shown that the least-squares solutions must satisfy the *normal equation*

$$\mathcal{G}^*\mathcal{G}u = \mathcal{G}^*y \tag{2.9}$$

and such solutions will exist as long as the output y belongs to the dense subset $R(\mathcal{G}) \oplus R(\mathcal{G})^\perp$ of $L^2(\mathbb{R})$. The set S_y of all least-squares solutions, corresponding to $y \in R(\mathcal{G}) \oplus R(\mathcal{G})^\perp$ is a closed and convex set. Hence, we can assign as a solution to 2.4 the unique function $u^\dagger \in S_y$ of minimal norm. The function $u^\dagger$ is called the *generalized solution* of 2.4 and the linear operator

$$\mathcal{G}^\dagger : R(\mathcal{G}) \oplus R(\mathcal{G})^\perp \to L^2(\mathbb{R})$$

defined by

$$u^\dagger = \mathcal{G}^\dagger y,$$

is referred to as the *generalized inverse*. The generalization that we have just considered allows us to assign meaning to the notion of solution for a larger class of functions y. Moreover, this generalization does, in part, deal with the question of uniqueness. However, the third and perhaps most important condition listed in 2.6 has not been addressed.

The importance of continuity stems from the fact that, in most practical situations, knowledge of y is gained via measurement. Hence, y is not known precisely. Consequently, we must try to extract an approximation of $u^\dagger$ through the use of the corrupted data $y + \delta y$, where $\delta y \in L^2(\mathbb{R})$ represents a small but unknown perturbation. Since in general, $\delta y \notin R(\mathcal{G}) \oplus R(\mathcal{G})^\perp$, a generalized solution need not exist for every observed output. Even if $\delta y \in R(\mathcal{G}) \oplus R(\mathcal{G})^\perp$, the discontinuity of the generalized inverse can lead to an approximation $\mathcal{G}^\dagger(y + \delta y)$ which is arbitrarily far from $u^\dagger$. To see this we consider 2.4 in the special case where $|\hat{g}(\omega)| > 0$ for all $\omega \in \mathbb{R}$. In this case, $\mathcal{G} : L^2(\mathbb{R}) \to R(\mathcal{G})$ is a bijection and for each y in the dense subset $R(\mathcal{G})$, $u^\dagger$ is simply the unique solution u of 2.4. Let

$$\delta y(t) = \sqrt{\frac{2}{\pi}} e^{i\beta t} \frac{\sin(t)}{t},$$

then for every $\beta \in \mathbb{R}$, $\delta y \in R(\mathcal{G})$ and $\|\delta y\| = 1$. If we use the corrupted data $y + \delta y$ to form the approximation

$$u_\delta = \mathcal{G}^{-1}(y + \delta y),$$

then, from 2.7, we obtain

$$u_\delta(t) - u(t) = \left(\mathcal{G}^{-1}\delta y\right)(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega t} \frac{\chi_\beta(\omega)}{\hat{g}(\omega)}\, d\omega,$$

which implies that

$$\|u_\delta - u\|^2 = \int_{-\infty}^{\infty} \left| \frac{\chi_\beta(\omega)}{\hat{g}(\omega)} \right|^2 d\omega,$$

where χ_β denotes the characteristic function of the interval $[-1+\beta, 1+\beta]$. Since $g \in L^1(\mathbb{R})$, $\hat{g}$ is continuous and

$$\|u_\delta - u\|^2 = \int_{-1+\beta}^{1+\beta} |\hat{g}(\omega)|^{-2}\, d\omega \geq \left(\max_{\omega \in [-1+\beta, 1+\beta]} \hat{g}(\omega) \right)^{-2}.$$

The Riemann-Lebesgue Lemma ensures that

$$\lim_{|\omega|\to\infty} \hat{g}(\omega) = 0$$

and therefore,

$$\lim_{|\beta|\to\infty} \|u_\delta - u\|^2 = \lim_{|\beta|\to\infty} \left\|\mathcal{G}^{-1}\delta y\right\|^2 = \infty.$$

Consequently, small high frequency perturbations in the observed output can lead to approximate solutions which are arbitrarily far from the solution obtained from the unperturbed output.

One popular way of dealing with the unboundedness of $\mathcal{G}^\dagger$ is known as the method of *inverse filtering* [3, 40]. When using this method, we seek to remove the ill effects of high frequency perturbations, while preserving the accuracy of the approximation produced. Let $\{W_\alpha(\omega) : \alpha > 0\}$ be a family of continuous functions that satisfies:

1. $|W_\alpha(\omega)/\hat{g}(\omega)| \leq A(\alpha) < \infty$, for all $\alpha > 0$ and $\omega \in \mathbb{R}$.
2. $|1 - W_\alpha(\omega)| \leq B$, uniformly for $\alpha > 0$, $\omega \in \mathbb{R}$ and (2.10)
3. $\lim_{\alpha\to 0^+} |1 - W_\alpha(\omega)| = 0$ pointwise almost everywhere.

where $A(\alpha)$ and B are constants. If we form the approximation

$$u_\alpha(t) = \left(\mathcal{G}_\alpha^\dagger u\right)(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t}\frac{W_\alpha(\omega)}{\hat{g}(\omega)}\hat{y}(\omega)\, d\omega, \tag{2.11}$$

then from the first condition of 2.10, we have that $u_\alpha \in L^2(\mathbb{R})$ for any $y \in L^2(\mathbb{R})$ and any $\alpha > 0$. In particular, it can be shown that

$$\|u_\alpha\| \leq \left\|\mathcal{G}_\alpha^\dagger\right\| \|y\| = A(\alpha)\|y\|,$$

which means that u_α depends continuously on the data y. Furthermore, if $y \in D(\mathcal{G}^\dagger) = R(\mathcal{G}) \oplus R(\mathcal{G})^\perp$, then

$$\begin{aligned} \|u^\dagger - u_\alpha\|^2 &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \left|\widehat{u^\dagger}(\omega) - \frac{W_\alpha(\omega)}{\hat{g}(\omega)}\hat{y}(\omega)\right|^2 d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} |1 - W_\alpha(\omega)|^2 \left|\widehat{u^\dagger}(\omega)\right|^2 d\omega \end{aligned}$$

and in view of the second and third conditions of 2.10, the Lebesgue Dominated Convergence Theorem ensures that

$$\lim_{\alpha\to 0^+} \|u^\dagger - u_\alpha\| = 0.$$

In other words, the operator $\mathcal{G}_\alpha^\dagger$ converges strongly to $\mathcal{G}^\dagger$ on $D(\mathcal{G}^\dagger)$.

In many situations, the function W_α is selected so that $\{W_\alpha : \alpha > 0\}$ is a set of low pass filters. That is, for every $\alpha > 0$

$$W_\alpha(\omega) \approx \begin{cases} 1, & |\omega| \le \Omega_\alpha \\ 0, & \text{otherwise.} \end{cases}$$

For example, if $g(t) = \exp(-|t|)$, then

$$\hat{g}(\omega) = \frac{1}{1+\omega^2}$$

and a suitable family $\{W_\alpha : \alpha > 0\}$ is defined by

$$W_\alpha(\omega) = \frac{1}{(1+\alpha\omega^2)^2}.$$

Suppose the $y \in R(\mathcal{G})$, then since $\mathcal{G}$ is a bijection, the function

$$\begin{aligned} u_\alpha(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t}\frac{1+\omega^2}{(1+\alpha\omega^2)^2}\hat{y}(\omega)\,d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t}(1+\alpha\omega^2)^{-2}\hat{u}(\omega)\,d\omega \end{aligned}$$

can be regarded as a low frequency approximation of the unique solution u.

In the next chapter, we present a brief introduction to wavelet analysis. We are particularly interested in wavelet bases and their applications to solving convolution integral equations. Wavelet bases have the ability to localize both time and frequency in the sense that the functions which comprise these bases can be selected to have good decay in both domains. In particular, the scaling function ϕ can be regarded as a low pass filter and one intuitively expects that the stability of the deconvolution problem can be improved by seeking approximate solution in scaling function subspaces.

Chapter 3

Wavelet analysis

3.1 Introduction

One of the most significant shortcomings of the Fourier transform is its inability to deal effectively with non-stationary signals. Let $\chi_{[0,1]}$ be the characteristic function of the interval $[0, 1]$ and consider the function

$$f(t) = \chi_{[0,1]}(t)\sin(6\pi t)$$

The Fourier transform of f is the function

$$\hat{f}(\omega) = 6\pi\frac{e^{-i\omega} - 1}{\omega^2 - 36\pi^2},$$

which has absolute value

$$\left|\hat{f}(\omega)\right| = 12\pi\frac{|\sin(\omega/2)|}{\omega^2 - 36\pi^2}.$$

We see that $\left|\hat{f}\right|$ reflects the overall frequency content of f and does not give us any information about the significant changes in frequency which occur at the times $t = 0$ and $t = 1$. In fact, this information is contained in the phase of $\hat{f}$, but in some cases can be difficult to extract.

In an effort to overcome such difficulties, D. Gabor (See [9, page 50]) introduced the integral transform

$$(S_\alpha f)(\tau, \omega) = \int_{-\infty}^{\infty} e^{-i\omega t} g_\alpha(t-\tau) f(t)\, dt, \tag{3.1}$$

where

$$g_\alpha(t) = \frac{1}{2\sqrt{\pi\alpha}} e^{-t^2/4\alpha},$$

the idea being that the rapid decay of the Gaussian function, g_α, would allow for the study of the local frequency content of the function f. In fact, for a fixed τ, the function $S_\alpha(\tau, \omega)$ is simply the Fourier transform of the function $g_\alpha(t-\tau)f(t)$, which is approximately $1/(2\sqrt{\alpha\pi})f(t)$ for t near τ. Furthermore, since the Fourier transform of g_α is

$$\hat{g}_\alpha(\omega) = e^{-\alpha\omega^2},$$

the Gabor transform allows one to study a function locally in both time and frequency simultaneously. Different choices of the function g_α lead to a family of transformations known as *short time Fourier transforms*.

3.2 Integral wavelet transform

In a similar way, the *integral wavelet transform* (IWT) ([9, page 60]) also provides a means of studying functions in a local way. However, in this instance, the function f is decomposed into its components with respect to the dilations and translations of a single function ψ, called the *mother wavelet*.

Definition 3.1 *If $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ satisfies the admissibility condition*

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|}\, d\omega < \infty. \tag{3.2}$$

then we define the integral wavelet transform on $L^2(\mathbb{R})$ by

$$\left(\mathcal{W}_\psi f\right)(a,b) = |a|^{-1/2} \int_{-\infty}^{\infty} f(t)\overline{\psi\left(\frac{t-b}{a}\right)}\, dt. \tag{3.3}$$

where $a.\ b \in \mathbb{R}$ and $a \neq 0$.

The variable b in (3.2) can be identified with time . while a can be thought of as a multiple of the reciprocal of frequency. If we choose ψ so that is has rapid decay in both time and frequency. then we can regard $\mathcal{W}_\psi f$ as giving information about the content of the function f near b in time and near c/a (c a constant depending on ψ) in frequency.

It is this ability to localize in time and, most importantly in frequency, which makes the IWT and wavelet analysis in general a plausible tool for deconvolution problems. Indeed one might view the IWT as a type of filter, the characteristics of which are determined by the value of the variable a. As an example, suppose we let

$$\psi(t) = (1-t^2)e^{-t^2/2}. \tag{3.4}$$

then the Fourier transform of ψ is

$$\hat{\psi}(\omega) = \sqrt{2\pi}\omega^2 e^{-\omega^2/2}.$$

The function ψ satisfies the admissibility condition with $C_\psi = 2\pi$. If $\langle \cdot, \cdot \rangle$ denotes the usual inner product on $L^2(\mathbb{R})$, then by 3.3, the IWT of an f

$$\begin{aligned} \left(\mathcal{W}_\psi f\right)(a,b) &= |a|^{-1/2} \left\langle f(t), \left(1 - \left[\frac{t-b}{a}\right]^2\right) \exp\left(-\frac{1}{2}\left[\frac{t-b}{a}\right]^2\right) \right\rangle \\ &= \frac{|a|^{1/2}\mathrm{sgn}(a)}{\sqrt{2\pi}} < \hat{f}(\omega), e^{-i\omega b}(a\omega)^2 e^{-(a\omega)^2/2} >, \end{aligned}$$

where the last equality is obtained by an application of Parseval's identity. The function $(a\omega)^2 e^{-(a\omega)^2/2}$ is concentrated around two peaks centered at $\omega = \pm\sqrt{2}/a$ and as a is increased, the peaks become narrower and move towards $\omega = 0$. As a result, when a is large, the product

$$\hat{f}(\omega)e^{i\omega b}(a\omega)^2 e^{-(a\omega)^2/2},$$

depends primarily on $\hat{f}(\omega)$ for small ω. For other choices of ψ, different localization properties can be achieved, this being one of the major advantages of the IWT.

If we are given the IWT of a function. f, then we can reconstruct f through the use of the formula

$$f(t) = \frac{1}{C_\psi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\mathcal{W}_\psi f)(a.b)|a|^{-1/2}\overline{\psi\left(\frac{t-b}{a}\right)}\, db\frac{da}{a^2}. \tag{3.5}$$

In fact. 3.5 is really a consequence of the Parseval's identity

$$\frac{1}{C_\psi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\mathcal{W}_\psi f)(a,b)\overline{(\mathcal{W}_\psi g)(a,b)}\, db\frac{da}{a^2} = < f,g >. \tag{3.6}$$

which holds for any $f,\ g \in L^2(\mathbb{R})$.

Many variations of 3.3 and 3.5 exist, allowing one to customize the IWT to suit specific needs. Of particular interest is a variation which uses one wavelet for decomposition and a second for reconstruction. Specifically, if the wavelets ψ_1 and ψ_2 satisfy the admissibility condition

$$C_{\psi_1\psi_2} = \int_0^\infty \frac{|\hat{\psi}_1(\omega)||\hat{\psi}_2(\omega)|}{\omega}\, d\omega = \int_0^\infty \frac{|\hat{\psi}_1(-\omega)||\hat{\psi}_2(-\omega)|}{\omega}\, d\omega,$$

then the IWT transform pair, based on $a > 0$, is

$$(\mathcal{W}f)(a,b) = |a|^{-1/2}\int_{-\infty}^{\infty} f(t)\overline{\psi_1\left(\frac{t-b}{a}\right)}\, dt \tag{3.7}$$

and

$$f(t) = \frac{|a|^{-1/2}}{C_{\psi_1 \psi_2}} \int_0^{\infty} \int_{-\infty}^{\infty} (\mathcal{W}f)(a, b) \overline{\psi_2\left(\frac{t-b}{a}\right)}\, db \frac{da}{a^2}. \quad (3.8)$$

The localization and regularity properties of ψ_1 and ψ_2 can be quite different. In fact, in [25] Holschneider examines a two-dimensional extension of 3.7 and 3.8 in which ψ_1 is a distribution. This extension allows Holschneider to use the IWT to invert the Radon transform.

In practical situations, there is a desire to be able to reconstruct the function f from the restriction of its IWT to a discrete grid. For instance, if we let $a = a_0^{-j}$ and $b = a_0^{-j} k b_0$, then we seek a function $\tilde{\psi}$ such that

$$f(t) = \sum_{j,k \in Z} (\mathcal{W}_{\psi} f)(a_0^{-j}, a_0^{-j} k b_0) a_0^{j/2} \tilde{\psi}(a_0^j t - k b_0).$$

The above leads to the theory of *frames*[1] and the study of the *discrete wavelet transform*. The reader is referred to [13] and [24] for more details. If we consider the special case where $a_0 = 2$ and $b_0 = 1$, then it is possible to find a function ψ so that an orthonormal basis is obtained.

3.3 Wavelet bases

An orthonormal wavelet basis of $L^2(\mathbb{R})$ is a basis of the form

$$\left\{\psi^{j,k}(t) = 2^{j/2}\psi(2^j t - k) : j, k \in \mathbb{Z}\right\},$$

which satisfies the orthonormality conditions

$$< \psi^{j,k}, \psi^{l,m} >= \delta_{j,l}\delta_{k,m}.$$

[1] a linearly dependent set which spans the space of interest

If we define the subspaces W_j, $j \in \mathbb{Z}$, by

$$W_j = \overline{\bigvee \{\psi^{j,k} : k \in \mathbb{Z}\}}$$

then we obtain an orthogonal decomposition of $L^2(\mathbb{R})$. That is, we can write

$$L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j. \tag{3.9}$$

Now suppose we define the subspaces V_j, $j \in \mathbb{Z}$ by

$$V_j = \bigoplus_{p=-\infty}^{j-1} W_j \tag{3.10}$$

then the V_j are nested

$$\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots$$

and

$$L^2(\mathbb{R}) = \overline{\bigcup_{j \in \mathbb{Z}} V_j}.$$

Notice that W_j is the orthogonal complement of V_j in V_{j+1}. That is,

$$V_{j+1} = V_j \oplus W_j. \tag{3.11}$$

It turns out that there is a function ϕ, called the *scaling function*, which satisfies

$$\langle \phi^{j,k}, \phi^{j,l} \rangle = \delta_{k,l} \tag{3.12}$$

and

$$V_j = \overline{\bigvee \{\phi^{j,k} : k \in \mathbb{Z}\}} \tag{3.13}$$

In view of the above, any function $f \in L^2(\mathbb{R})$ has an expansion of the form

$$\begin{aligned} f(t) &= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} d_k^j \psi^{j,k}(t), \\ &= \sum_{k=-\infty}^{\infty} c_k^n \phi^{n,k}(t) + \sum_{j=n}^{\infty} \sum_{k=-\infty}^{\infty} d_k^j \psi^{j,k}(t), \end{aligned} \tag{3.14}$$

in which the expansion coefficients c_k^n and d_k^j are computed via the inner products

$$c_k^n = < f, \phi^{n,k} >, \qquad d_k^j = < f, \psi^{j,k} >,$$

and $n \in \mathbb{Z}$ is assumed to be fixed.

We point out that the scaling function ϕ is usually the starting point in the construction of an orthonormal wavelet basis. In fact, one demands that ϕ satisfy the *dilation equation*

$$\phi(t) = \sqrt{2} \sum_{k \in \mathbb{Z}} h_k \phi(2t - k), \tag{3.15}$$

where the *two-scale sequence* $\{h_k : k \in \mathbb{Z}\}$ is assumed to be given. The properties of the two-scale sequence determine the properties of the resulting scaling function. For example, to ensure orthonormality 3.12, or equivalently

$$\sum_{l \in \mathbb{Z}} \left| \hat{\phi}(\omega + 2\pi l) \right|^2 = 1, \tag{3.16}$$

the two-scale sequence must satisfy

$$\sum_p h_{p-2k} h_k = \delta_{p,0}.$$

After a suitable scaling function has been found, the wavelet ψ is defined by an equation of the form

$$\psi(t) = \sqrt{2} \sum_{k \in \mathbb{Z}} (-1)^{k-1} h_{-k-1} \phi(2t - k), \tag{3.17}$$

which can be derived by requiring that the integer translates of ψ span W_0, the orthogonal complement of V_0 in V_1.

The set of nested subspaces $\{V_j : j \in \mathbb{Z}\}$ and the orthogonal decomposition that it defines is known as a *multiresolution analysis* (MRA) of $L^2(\mathbb{R})$ and it is the dilation equation 3.17 that makes this structure possible. If we let P_n denote the orthogonal projection onto the subspace V_n and Q_j denote the orthogonal projection onto the subspace W_j, then, in view of 3.14, any $f \in L^2(\mathbb{R})$ can be decomposed into the orthogonal sum

$$f = P_n f \oplus Q_n f \oplus Q_{n+1} f \oplus \cdots.$$

In practice, the full wavelet expansion of a function is not available and one works with the projection $P_n f$,[2] or equivalently, with the scaling function coefficients $\{c_k^n : k \in \mathbb{Z}\}$. When the coefficients $\{c_k^n : k \in \mathbb{Z}\}$ are known, an algorithm, defined by the equations 3.15 and 3.17 can be used to derive the wavelet coefficients at all scales $j \leq n-1$. This algorithm, called the *discrete wavelet transform* (DWT), is defined by the equations

$$c_k^{n-1} = \sum_p h_{p-2k} c_p^n \tag{3.18}$$

and

$$d_k^{n-1} = \sum_p g_{p-2k} c_p^n, \tag{3.19}$$

where $g_k = (-1)^{k-1} h_{-k-1}$. We can regard each of 3.18 and 3.19 as the operation of convolution with the respective discrete filters $\{h_{-k} : k \in \mathbb{Z}\}$ and $\{g_{-k} : k \in \mathbb{Z}\}$, followed by the operation of down-sampling[3].

Suppose that the *discrete Fourier transform* (DFT) of $\{a_k : k \in \mathbb{Z}\} \in \ell^2(\mathbb{Z})$ is given by

$$A(\omega) = \sum_k a_k e^{-ik\omega},$$

then an application of the DFT to 3.18 and 3.19 results in

$$C^{n-1}(\omega) = \frac{1}{2}\left(\overline{H}(\omega/2) C^n(\omega/2) + \overline{H}(\omega/2+\pi) C^n(\omega/2+\pi)\right) \tag{3.20}$$

[2] $P_n f$ is often referred to as an approximation of f at resolution n, while $Q_j f$ is called the detail of f at resolution j.

[3] also known as decimation.

and

$$D^{n-1}(\omega) = \frac{1}{2}\left(\overline{G}(\omega/2)C^n(\omega/2) + \overline{G}(\omega/2+\pi)D^n(\omega/2+\pi)\right), \qquad (3.21)$$

where C^n, D^n, H and G are the discrete Fourier transforms of the sequences $\{c_k^n : k \in \mathbb{Z}\}$, $\{d_k^n : k \in \mathbb{Z}\}$, $\{h_k : k \in \mathbb{Z}\}$ and $\{g_k : k \in \mathbb{Z}\}$ respectively.

The DWT is invertible and, in particular, if we are given $\{c_k^{n-1} : k \in \mathbb{Z}\}$ and $\{d_k^{n-1} : k \in \mathbb{Z}\}$, then $\{c_k^n : k \in \mathbb{Z}\}$ can be computed via the equation

$$c_k^n = \sum_p h_{k-2p} c_p^{n-1} + \sum_p g_{k-2p} d_p^{n-1}, \qquad (3.22)$$

which has the DFT

$$C^n(\omega) = H(\omega)C^{n-1}(2\omega) + G(\omega)D^{n-1}(2\omega). \qquad (3.23)$$

3.4 Daubechies wavelets

Let us now restrict our attention to a notable class of wavelets which were introduced by I. Daubechies (See [12], [14, page 167]). These wavelets[4] are compactly supported and are called the Daubechies wavelets. Some of their most important properties are:

1. The first N moments of the wavelet vanish. That is,

   $$\mathbf{N}_p = \int_{\mathbb{R}} t^p \psi(t)\,dt = 0, \qquad p = 0, 1, \ldots, N-1,$$

   where N is a positive integer.

2. The scaling function is unimodular,

   $$\int_{\mathbb{R}} \phi(t)\,dt = 1.$$

[4]as well as the scaling functions

3. The wavelet series 3.14 converges exponentially fast with respect to the dilation index. In particular, if the function f is at least N times continuously differentiable, then

$$\left\| f(t) - \sum_k c_l^n \phi^{n,k}(t) \right\| \leq A2^{-nN},$$

for some constant A.

4. The support length of both the wavelet and scaling function is $K = 2N - 1$. Furthermore, the differentiability of both functions increases as N increases.

3.5 Meyer wavelets

A second class of wavelets, which is of theoretical importance, is due to Yves Meyer (see [14, page 137]). G. Walter enlarged this class of wavelets in [44] and produced a family of orthonormal wavelet bases with compactly supported Fourier transforms. To define the Meyer wavelets, we begin with a continuous probability density function p satisfying,

$$\mathrm{supp}(p) = [-\varepsilon, \varepsilon],$$

for $0 < \varepsilon \leq \pi/3$. The Fourier transform of the scaling function is subsequently defined by

$$\hat{\phi}_\varepsilon(\omega) = \sqrt{\int_{-\pi}^{\pi} p(\omega - \nu)\, d\nu} \tag{3.24}$$

and is supported in the interval $[-\pi - \varepsilon, \pi + \varepsilon]$. The Fourier transform of the corresponding wavelet is given by

$$\hat{\psi}_\varepsilon(\omega) = e^{-i\omega/2}\sqrt{\hat{\phi}_\varepsilon^2(\omega/2) - \hat{\phi}_\varepsilon^2(\omega)}. \tag{3.25}$$

One could think of the Meyer class as being a generalization of the Shannon MRA, which has the scaling function

$$\mathrm{sinc}(t) = \frac{\sin(\pi t)}{\pi t}.$$

The Fourier transform of the sinc(t) is the characteristic function

$$\chi_\pi(\omega) = \begin{cases} 1, & \omega \in [-\pi, \pi] \\ 0, & \text{otherwise} \end{cases}$$

and it is not too difficult to show that

$$\lim_{\epsilon \to 0^+} \hat{\phi}_\epsilon(\omega) = \chi_\pi(\omega)$$

for almost all ω.

Chapter 4

Literature review

4.1 Introduction

In its present form, wavelet analysis is a relatively new field, the first articles appearing in the early part of the last decade. The application of wavelet analysis to inverse problems is newer. Many interesting articles have begun to appear and tend to indicate that in some situations wavelet analysis is capable of outperforming established inverse problem methods. However, the evidence is far from conclusive and in some cases, is based solely on numerical simulation. There is a definite need for more research before the possible benefits of applying wavelet analysis to inverse problems are well understood.

4.2 Related work

One of the first articles in which the author considers the utility of wavelet analysis for the solution of an inverse problem is [25]. In this article, the author shows how the continuous wavelet transform can be used to invert the Radon transform (see [15]), which can be defined by

$$(\mathcal{R}f)(\mathbf{p}, \alpha) = \int_{\mathbb{R}^2} \delta(\mathbf{p} \cdot \mathbf{x} - \alpha) f(\mathbf{x})\, d\mathbf{x}, \tag{4.1}$$

where δ is the Dirac distribution and $\mathbf{p}$ is a unit vector. In the subsequent articles [43, 16], this problem is further investigated and it is pointed out in [16] that orthonormal wavelet bases could be used to recover local information about the function f from local information about the Radon transform $\mathcal{R}f$. This fact may have important consequences for the field of computer aided tomography and permit the investigation of tumors on a local basis, reducing a subject's exposure to radiation.

In [4] the authors comment on the possible application of wavelet bases to a wide class of inverse problems based upon the integral equation

$$g(x) = (\mathcal{K}f)(x) = \int_a^b K(x, y) f(y)\, dy, \quad a \le x \le b. \tag{4.2}$$

In this and the earlier paper [5] the authors show that many integral operators of the form 4.2 admit a sparse matrix representation in a orthonormal wavelet basis $\{\psi^{j,k} : j, k \in \mathbb{Z}\}$. That is, most of the entries of the matrix

$$\mathbf{K} = \left[\langle \mathcal{K}\psi^{j,k}, \psi^{l,m}\rangle\right], \tag{4.3}$$

will satisfy the inequality

$$\left|\langle \mathcal{K}\psi^{j,k}, \psi^{l,m}\rangle\right| < \epsilon,$$

where ϵ is some small positive number. This sparse representation of the operator $\mathcal{K}$ permits the multiplication of the matrix $\mathbf{K}$ with a vector of length N in $O(N)$ or $O(N \log N)$ operations. With this efficient method, certain iterative algorithms for the computation of the generalized inverse $\mathbf{K}^\dagger$ become feasible. What is not clear from the discussions in [4] is whether or not wavelet bases offer any particular advantages, aside from sparse representations, in dealing with the ill-posedness (or ill-conditioning) of the inverse problem at hand.

In the papers [30, 31], the distributed parameter system identification problem

$$\Phi(a) = u \tag{4.4}$$

is considered. In this case, $\Phi : L^2[0,1] \to L^2[0,1]$ is a non-linear operator defined by the boundary value problem

$$\begin{cases} -\frac{d^2}{dx^2}(a(x)u(x)) = f(x), & 0 < x < 1 \\ u(0) = u'(0) = 0 & . \end{cases} \tag{4.5}$$

The author restricts his attention to the Haar wavelet basis and assumes that the unknown function $a \in V_n$, where

$$V_n = \overline{\bigvee \{\phi^{n,k} : k \in \mathbb{Z}\}}$$

and ϕ is the Haar scaling function, which can be defined by

$$\phi(t) = \chi_{[0,1]}(t).$$

Numerical evidence is presented which suggests that a multiresolution approach to this particular inverse problem can lead to regularization methods comparable to Tikhonov regularization [22, 23]. For example, in the presence of noisy data u^δ, one seeks to approximate a by minimizing the functional

$$J(a) = \left\|\Phi(a) - u^\delta\right\|^2 \tag{4.6}$$

over the function space $L^2[0,1]$. However, this problem is ill-posed and the small fluctuations present in u^δ can lead to large fluctuations in the approximate solution.

To keep these fluctuations under control, one could instead choose to minimize the functional

$$F_1(a) = \left\|\Phi(a) - u^\delta\right\|^2 + \lambda \left\|\frac{da}{dx}\right\|^2. \tag{4.7}$$

This is a particular case of Tikhonov regularization and the effect of the term $\lambda \|a'(x)\|^2$ in F_1 is to prevent the norm of $a'(x)$, the estimator of the solution a, from growing too large.

Now, it is known that the regularity of a function f can be characterized by its wavelet coefficients. For instance, if $H^s(\mathbb{R})$ is the Sobolev space

$$H^s(\mathbb{R}) = \left\{ f : \int_{-\infty}^{\infty} (1+\omega^2)^s \left|\hat{f}(\omega)\right|^2 d\omega < \infty \right\},$$

then it can be shown that (see for instance [12, pages 298–304]) $f \in H^s(\mathbb{R})$ if and only if

$$\sum_{j,k} \left|\langle f, \psi^{j,k}\rangle\right|^2 (1 + 4^{sj}) < \infty.$$

It follows that the functional

$$F_2(a) = \left\|\Phi(a) - u^\delta\right\|^2 + \sum_{j,k} \lambda_j \left|\langle a, \psi^{j,k}\rangle\right|^2, \tag{4.8}$$

with $\lambda_j \sim 4^{sj}$ as $j \to \infty$, can be used to define a type of regularization which generalizes that defined by 4.7. Functional 4.8 suggests the possibility of a multiscale regularization method based upon wavelet bases. A similar method is suggested by the work of Chen and Lin, which is examined in section 4.4.

In [46], the authors consider the linear moment problem

$$\langle f, g_k\rangle = \mu_k, \quad k \in \mathbb{Z}, \tag{4.9}$$

in which the scalars $\{\mu_k : k \in \mathbb{Z}\}$ and the functions $\{g_k : k \in \mathbb{Z}\}$ are known, while the function $f \in L^2(\mathbb{R})$ is to be estimated.

A classical approach to the problem 4.9 is the Backus-Gilbert method. In this method, the function f is approximated by the sum

$$\tilde{f}(t) = \sum_k \mu_k \alpha_k(t),$$

where the functions $\{\alpha_k : k \in \mathbb{Z}\}$ are selected so that the averaging kernel

$$A(s,t) = \sum_k g_k(s)\alpha_k(t)$$

is a reasonable approximation of the delta distribution. For example, if

$$\delta_n(s,t), \quad n = 0, 1, 2, \ldots$$

defines a delta-sequence converging to $\delta(s-t)$, then according to the so-called *D-criterion*, the functions $\{\alpha_k : k \in \mathbb{Z}\}$ are chosen so that the functional

$$D(\alpha) = \int_{-\infty}^{\infty} (A(s,t) - \delta_n(s,t))^2 \, ds$$

is minimized.

The authors demonstrate that when the D-criterion is used, the assumption $f \in V_n$,

$$V_n = \overline{\bigvee \{\phi^{n,k} : k \in \mathbb{Z}\}} \tag{4.10}$$

can lead to a definite improvement in performance. In particular, in the case of sampled signals, it is shown that the *a priori* condition $f \in V_n$ yields a modification of the Backus-Gilbert method which allows for the complete recovery of f. On the other hand, a straightforward application of the Backus-Gilbert method does not in general allow for the recovery of f from its sampled values. It is also shown that, in all cases, the generalized method performs at least as well as the ordinary method. However, in cases other than that of sampled signals, it is not known whether the generalization provides substantial improvement or not.

The modification of the Backus-Gilbert method, proposed in [46], makes use of the reproducing property of scaling function subspaces 4.10. Suppose that the continuous scaling function ϕ satisfies

$$|\phi(t)| \leq \frac{B}{(1+|t|)^{\beta}}$$

for some $\beta > 1/2$, then the series

$$Q(s,t) = \sum_k \phi(s-k)\phi(t-k)$$

defines a continuous, symmetric kernel. In this case, each subspace V_j has the property that if $f \in V_j$, then

$$f(t) = \int_{-\infty}^{\infty} Q_j(s,t) f(s)\, ds,$$

where

$$Q_j(s,t) = 2^j Q(2^j s, 2^j t).$$

The subspace V_j is said to be a *reproducing kernel Hilbert space* (see, for example, [2]) and this property can be used to convert 4.9 to the new moment problem

$$\langle f, g_k^j \rangle = \mu_k, \quad j = n, n+1, \ldots, \; k \in \mathbb{Z}. \tag{4.11}$$

The function f is now approximated by the sum,

$$\tilde{f}(t) = \sum_k \sum_{j \geq n} \mu_k \alpha_k^j(t).$$

If we define the generalized averaging kernel

$$A(s,t) = \sum_k \sum_{j \geq n} g_k^j(s) \alpha_k^j(t),$$

then, when the D-criterion is employed, the functions $\{\alpha_k^j \; : \; j \geq n, k \in \mathbb{Z}\}$ are selected so that the functional

$$D_w(\alpha) = \sum_{j \geq n} w_j \int_{-\infty}^{\infty} \left(A(s,t) - Q_j(s,t)\right)^2 \, ds,$$

with w_j non-negative weights, is minimized.

4.3 Continuous deconvolution

The convolution equation

$$\int_{-\infty}^{\infty} h(t-\tau)x(\tau)\,d\tau = y(t) \quad t \in \mathbb{R} \tag{4.12}$$

is considered in the paper [42]. It is observed that, in many cases, the convolution kernel h is similar to a scaling function ϕ, and based upon this observation the kernel is used to define a mother wavelet ψ. The continuous wavelet transform

$$(\mathcal{W}x)(t,a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi\left(\frac{s-t}{a}\right) x(s)\,ds$$

is then used to transform 4.12 into a equivalent system of continuous convolution equations indexed by the scale variable a. Each equation in this system is solved separately, yielding the wavelet transform $(\mathcal{W}x)(t,a)$. The inverse wavelet transform

$$x(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{0}^{\infty} (\mathcal{W}x)(s,a)\psi\left(\frac{t-s}{a}\right) \frac{da}{a^{5/2}} ds,$$

with

$$C_\psi = \int_0^\infty \frac{\left|\hat{\psi}(\omega)\right|^2}{\omega} d\omega$$

is then used to recover the function x.

This approach is much like that proposed by Liu et al. (see section 4.4) for the analysis of discrete convolution equations and suffers from one basic disadvantage. An appealing aspects of wavelet analysis is the freedom to choose a wavelet ψ suited to the problem at hand. By using the convolution kernel to define ψ, we lose this freedom.

One of the most powerful techniques for dealing with inverse problems involving compact operators (as defined by 4.2) is known as the *singular value decomposition*

(SVD) (see [29, 22]). Let H be a Hilbert space and suppose that $\mathcal{K} : H \to H$ is a compact linear operator. The singular system for the operator $\mathcal{K}$ is the set $\{v_j, u_j; \mu_j\}$, where $\{v_j\}$ is an orthonormal basis for the orthogonal complement of the null-space $N(\mathcal{K})$, while $\{u_j\}$ is an orthonormal basis for the orthogonal complement of the null-space of the adjoint of $\mathcal{K}$, $N(\mathcal{K}^*)$. The scalars $\{\mu_j\}$ are known as the *singular values* of $\mathcal{K}$, and satisfy

$$\lim_{j\to\infty} \mu_j = 0.$$

The singular system of $\mathcal{K}$ can be defined by the equations

$$\mathcal{K}^* u_j = \mu_j v_j \tag{4.13}$$

and

$$\mathcal{K} v_j = \mu_j u_j. \tag{4.14}$$

It follows that the generalized solution, of minimum norm, of the equation

$$\mathcal{K} f = g,$$

can be expressed as

$$(\mathcal{K}^\dagger g) = \sum_j \frac{\langle g, u_j \rangle}{\mu_j} v_j. \tag{4.15}$$

Since $\{v_j\}$ is an orthonormal basis, the generalized solution $\mathcal{K}^\dagger g$ will exist as a function in H if and only if

$$\sum_j \left| \frac{\langle g, u_j \rangle}{\mu_j} \right|^2 < \infty, \tag{4.16}$$

which holds whenever $g \in R(\mathcal{K})$. If we are faced with noisy data $g^\delta = g + \delta$, then a solution will exist if and only if

$$\sum_j \left| \frac{\langle g^\delta, u_j \rangle}{\mu_j} \right|^2 < \infty$$

and since the inner-products $\{\langle \delta, u_j \rangle\}$ may decay slower than the singular values, $\mathcal{K}^\dagger g^\delta$ need not exist as a function in H.

In such cases, one usually employs some type of regularization procedure in combination with the SVD. For example, if we use as an approximate solution of $\mathcal{K}f = g^\delta$ the minimizer of the functional

$$F(f) = \left\|\mathcal{K}f - g^\delta\right\|^2 + \lambda \left\|f\right\|^2, \tag{4.17}$$

then, through the use of the SVD, we can write the minimizer of F as

$$\tilde{f} = \sum_j W(\mu_j)\langle g^\delta, u_j \rangle v_j, \tag{4.18}$$

where the *window* W is given by

$$W(s) = \frac{s}{\lambda + |s|^2}.$$

Since

$$\sum_j |W(\mu_j)\langle g, u_j \rangle|^2 \leq \frac{1}{\lambda} \sum_j \left|\langle g^\delta, u_j \rangle\right|^2 < \infty,$$

we see that 4.18 is a well-defined element of H for all $\lambda > 0$ and $\delta \in H$.

One major drawback of the SVD is that the basis $\{v_j\}$ may not be well-suited to the function f under investigation. For example, suppose that f is piecewise smooth. The local discontinuities present in f can cause the coefficients

$$\langle f, v_j \rangle = \frac{\langle g, u_j \rangle}{\mu_j}$$

to decay very slowly, and accordingly a significant proportion of the norm of

$$\|\{\langle f, v_j \rangle\}\|^2_{\ell^2} = \sum_j |\langle f, v_j \rangle|^2$$

may be contained in coefficients for which μ_j is small. The windowed SVD deals with the presence of errors in the data g by suppressing the inner-products

$$\langle g^\delta, u_j\rangle = \langle g, u_j\rangle + \langle \delta, u_j\rangle$$

when μ_j is small and. in doing so, may discard significant information about f leading to a poor approximation 4.18. Typically, the basis vectors $\{v_j\}$ are non-local. and this loss of information causes errors that are distributed throughout $\tilde{f}$ and not just near the discontinuities of f.

Wavelet bases hold the promise of being able to circumvent these difficulties. Due to the local nature of the basis ψ, discontinuities of a function f affect only the coefficients $\langle f, \psi^{j,k}\rangle$ which correspond to basis functions near the discontinuities. Furthermore, the suppression or removal of these particular coefficients will produce an approximation which is erroneous only in a neighborhood of said discontinuities. In view of these desirable properties, it makes sense to see if the SVD can be generalized as to allow for the incorporation of wavelet bases.

One possible generalization is explored by Donoho in the paper [18]. Here the author proposes a method for solving the operator equation

$$\mathcal{K}f = g, \tag{4.19}$$

where $\mathcal{K} : D(\mathcal{K}) \subset L^2(\mathbb{R}) \to R(\mathcal{K}) \subset L^2(\mathbb{R})$ is a linear operator. This method makes use of the *wavelet-vaguelette decomposition* (WVD) which emulates some of the more important properties of the SVD, while leaving the user greater flexibility in the selection of a suitable basis for f in 4.19.

The WVD of an operator $\mathcal{K}$ is a collection of three sets of functions:

1. an orthonormal wavelet basis $\{\psi^{j,k} : j, k \in \mathbb{Z}\}$ and

2. two nearly normalized Riesz bases $\{u^{j,k} : j, k \in \mathbb{Z}\}$ and $\{v^{j,k} : j, k \in \mathbb{Z}\}$.

along with a sequence of scalars:

3. *quasi-singular values* $\{\sigma_j : j \in \mathbb{Z}\}$.

The WVD of $\mathcal{K}$ can be defined by the *quasi-singular value relations* which are

$$\mathcal{K}\psi^{j,k} = \sigma_j v^{j,k} \tag{4.20}$$

and

$$\mathcal{K}^* u^{j,k} = \sigma_j \psi^{j,k}. \tag{4.21}$$

We notice that, when the WVD of the operator $\mathcal{K}$ exists, 4.20 implies that $D(\mathcal{K})$ must be a dense subset of $L^2(\mathbb{R})$. Similarly, from 4.21 we see that $R(\mathcal{K}^*)$ must be dense in $L^2(\mathbb{R})$ and since

$$N(\mathcal{K}) = R(\mathcal{K}^*)^{\perp} = \{0\},$$

with $N(\mathcal{K})$ the null space of $\mathcal{K}$, $\mathcal{K}$ must be injective. It follows that the inverse operator $\mathcal{K}^{-1}$ is defined on $R(\mathcal{K})$ although, it need not be continuous. Also, the quasi-singular value relations imply that the sets of functions $\{u^{j,k} : j, k \in \mathbb{Z}\}$ and $\{v^{j,k} : j, k \in \mathbb{Z}\}$ are biorthogonal. Indeed, since $\{\psi^{j,k} : j, k \in \mathbb{Z}\}$ is orthonormal, we have

$$\begin{aligned} \delta_{j,l}\delta_{k,m} &= \langle \psi^{j,k}, \psi^{l,m} \rangle = \frac{1}{\sigma_j} \langle \mathcal{K}^* u^{j,k}, \psi^{l,m} \rangle \\ &= \frac{1}{\sigma_j} \langle u^{j,k}, \mathcal{K}\psi^{l,m} \rangle = \frac{\sigma_l}{\sigma_j} \langle u^{j,k}, v^{l,m} \rangle \end{aligned} \tag{4.22}$$

as required.

With regard to the solution of 4.19, the relation 4.20 implies

$$g = \sum_{j,k} \sigma_j \langle f, \psi^{j,k} \rangle v^{j,k}$$

and through the use of the property 4.22. we obtain

$$\langle f, \psi^{j,k} \rangle = \frac{\langle g, u^{j,k} \rangle}{\sigma_j}. \tag{4.23}$$

In light of 4.23. the Riesz representation theorem implies that the coefficient functional $c_{j,k} : L^2(\mathbb{R}) \to \mathbb{R}$, given by

$$c_{j,k}(g) = \frac{\langle g, u^{j,k} \rangle}{\sigma_j},$$

is continuous with

$$\|c_{j,k}\| = \frac{\|u^{j,k}\|}{|\sigma_j|}.$$

Since the functions $u^{j,k}$ are nearly normalized, we find that

$$\|c_{j,k}\| \approx \frac{D}{|\sigma_j|},$$

for some constant D. It follows that the number $|\sigma_j|^{-1}$ gives an indication of the difficulty in recovering that wavelet coefficients at level j from the data g. For many operators. we expect that

$$\lim_{j \to \infty} \sigma_j = 0,$$

which indicates that higher level wavelet coefficients are increasingly difficult to recover.

Even though the $c_{j,k}$ are continuous, the function f, as defined via 4.23. need not be an element of $L^2(\mathbb{R})$. From 4.23, we have

$$f = \sum_{j,k} \frac{\langle g, u^{j,k} \rangle}{\sigma_j} \psi^{j,k}$$

and $f \in L^2(\mathbb{R})$ if and only if

$$\sum_{j,k} \frac{|\langle g, u^{j,k} \rangle|^2}{|\sigma_j|^2} < \infty, \tag{4.24}$$

which is a condition analogous to Picard's existence criterion. If $g \in R(\mathcal{K})$, then since

$$\langle g, u^{j,k} \rangle = \langle f, \mathcal{K}^* u^{j,k} \rangle = \overline{\sigma}_j \langle f, \psi^{j,k} \rangle,$$

condition 4.24 is satisfied and $f \in L^2(\mathbb{R})$. However, since f must be estimated from the noisy data $g^\delta = g + \delta$, with $\delta \in L^2(\mathbb{R})$, it is advantageous to require that 4.24 holds for any $g \in L^2(\mathbb{R})$. Unfortunately, this requirement cannot be met in all cases. This point is not clear in [18], and to make it so, we investigate some of the consequences of the condition 4.24.

Let us use the notation $\eta^{j,k} = u^{j,k}/\sigma_j$. Now by definition, if 4.24 holds for every $g \in L^2(\mathbb{R})$, then $\{\eta^{j,k} : j, k \in \mathbb{Z}\}$ is said to be a *Bessel sequence* of $L^2(\mathbb{R})$. It can be shown that the $\eta^{j,k}$ define a Bessel sequence if and only if

$$\left\| \sum_{j,k} \alpha_{j,k} \eta^{j,k} \right\|^2 \leq B \sum_{j,k} |\alpha_{j,k}|^2 \tag{4.25}$$

for some constant B (the reader is referred to [47, page 155]). By contrast 4.21 implies that $\{\eta^{j,k} : j, k \in \mathbb{Z}\}$ is a *Riesz-Fischer sequence* or equivalently that

$$\left\| \sum_{j,k} \alpha_{j,k} \eta^{j,k} \right\|^2 \geq A \sum_{j,k} |\alpha_{j,k}|^2. \tag{4.26}$$

with $A = \|\mathcal{K}\|^{-2}$. Together, the inequalities 4.25 and 4.26 indicate that the $\eta^{j,k}$ comprise a *Riesz basis* of $L^2(\mathbb{R})$. A sequence $\{h_k\}$ in a Hilbert space H is said to be a Riesz basis if it is equivalent to an orthonormal basis $\{e_k\}$ of H. That is, there exists a bounded linear operator $\mathcal{T}$, with a bounded inverse, such that

$$\mathcal{T} e_k = h_k.$$

Accordingly, there is an orthonormal basis $\{e^{j,k} : j, k \in \mathbb{Z}\}$ of

$$\overline{\bigvee \{\eta^{j,k} : j, k \in \mathbb{Z}\}} \tag{4.27}$$

and a bounded invertible operator $\mathcal{T}$ such that

$$\mathcal{T}e^{j,k} = \eta^{j,k},$$

for all $j, k \in \mathbb{Z}$. If we let $\xi^{j,k} = \sigma_j v^{j,k}$, then since

$$\delta_{j,l}\delta_{k,m} = \langle \xi^{j,k}, \eta^{l,m} \rangle = \langle \mathcal{T}^* P_\eta \xi^{j,k}, e^{l,m} \rangle$$

we have

$$\mathcal{T}^* P_\eta \xi^{j,k} = e^{j,k}, \tag{4.28}$$

where P_η is the orthogonal projection onto 4.27. Now, from 4.28 we obtain

$$\begin{aligned} \sum_{j,k} |\alpha_{j,k}|^2 &= \left\| \sum_{j,k} \alpha_{j,k} \mathcal{T}^* P_\eta \xi^{j,k} \right\|^2 \\ &\leq \|\mathcal{T}\|^2 \left\| \sum_{j,k} \alpha_{j,k} \xi^{j,k} \right\|^2 \end{aligned}$$

and since 4.20 implies that the $\xi^{j,k}$ define a Bessel sequence, we conclude that $\{\xi^{j,k} : j, k \in \mathbb{Z}\}$ must be a Riesz basis for the subspace $R(\mathcal{K})$. Consequently, the inequalities

$$\tilde{A} \sum_{j,k} |\alpha_{j,k}|^2 \leq \left\| \sum_{j,k} \alpha_{j,k} \xi^{j,k} \right\|^2 \leq \tilde{B} \sum_{j,k} |\alpha_{j,k}|^2$$

must hold and hence $R(\mathcal{K})$ must be a closed subspace of $L^2(\mathbb{R})$. In the case of bounded operators $\mathcal{K}$, since $\mathcal{K}$ is a densely defined bijection with a closed range, the Open Mapping Theorem [27, page 286] implies that $\mathcal{K}^{-1}$ must be continuous. As this need not be the case in general, we conclude that $\{\eta^{j,k} : j, k \in \mathbb{Z}\}$ is not necessarily a Riesz basis and that in general, 4.24 does not hold for all $g \in L^2(\mathbb{R})$.

Even though the sets $\{\xi^{j,k} : j, k \in \mathbb{Z}\}$ and $\{\eta^{j,k} : j, k \in \mathbb{Z}\}$ are not necessarily Riesz bases, one of the main ideas behind the WVD is that the quasi-singular values σ_j can be selected so that the sets $\{u^{j,k} : j, k \in \mathbb{Z}\}$ and $\{v^{j,k} : j, k \in \mathbb{Z}\}$ are Riesz bases. This idea leads us to the notion of *vaguelettes*.

Definition 4.1 *Let $w^{j,k}(t) = 2^{j/2}w(2^j t - k)$ and suppose that the function w satisfies:*

1. $|w(t)| \leq \frac{C_1}{(1+|t|)^\alpha}$ *for some* $\alpha > 0$.
2. $\int_{-\infty}^{\infty} w(t)\,dt = 0$ *and*
3. $|w(s) - w(t)| \leq |s-t|^\beta$ *for some* $\beta > 0$.

then $\left\{w^{j,k} : j, k \in \mathbb{Z}\right\}$ *is said to be a collection of vaguelettes.*

The importance of the definition above is elucidated in the following theorem. which is a consequence of Schur's lemma [36, page 270].

Theorem 4.1 *If* $\left\{w^{j,k} : j, k \in \mathbb{Z}\right\}$ *is a collection of vaguelettes, then there exists a constant B such that*

$$\left\|\sum_{j,k} \alpha_{j,k} w^{j,k}\right\|^2 \leq B \sum_{j,k} |\alpha_{j,k}|^2 .$$

That is, a collection of vaguelettes must be a Bessel sequence.

Suppose that the σ_j can be selected so that the functions $u^{j,k}$ and $v^{j,k}$ define collections of vaguelettes. A sequence $\{h_k\}$ in the Hilbert space H is a Bessel sequence if and only if for any orthonormal basis $\{e_k\}$ of H there exists a a bounded linear operator $\mathcal{T}$ such that

$$\mathcal{T} e_k = h_k .$$

With this fact and property 4.22, we can show that the sets $\left\{u^{j,k} : j, k \in \mathbb{Z}\right\}$ and $\left\{v^{j,k} : j, k \in \mathbb{Z}\right\}$ are Riesz bases.

Donoho has shown that for certain homogeneous operators, the functions $u^{j,k}$ and $v^{j,k}$ do indeed lead to collections of vaguelettes for appropriately defined σ_j and wavelets ψ that satisfy:

1. ψ is M times continuously differentiable and
2. $\int_{-\infty}^{\infty} t^n \psi(t)\, dt = 0$ for $n = 0, 1, \ldots, N$,

with M and N sufficiently large. In particular, many of the homogeneous operators studied in [18], such as integration and fractional integration, can be expressed in the form

$$(Kf)(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\Omega(\omega)}{|\omega|^{\alpha}} \hat{f}(\omega)\, d\omega, \tag{4.29}$$

where $\alpha > 0$ and Ω is homogeneous of degree zero ($\Omega(a\omega) = \Omega(\omega)$ for any $a > 0$). In such cases, the choices $\sigma_j = 2^{-\alpha j}$ and $M, N > \alpha + 2$ are sufficient to ensure that the functions $u^{j,k}$ and $v^{j,k}$ define collections of vaguelettes. Moreover, since these collections are also Riesz bases and hence well-behaved, the rate of decay of the σ_j can be regarded as a measure of the ill-posedness of the inverse problem based upon the operator equation 4.19. For example, the case of ordinary integration corresponds to $\alpha = 1$, while the Abel transform corresponds to $\alpha = 1/2$. Therefore, ordinary integration is about twice as ill-posed as is the Abel transform.

We consider the estimation of f from the noisy data $g^{\delta} = g + \epsilon\delta$, in the case where δ is white noise with variance ϵ^2. To combat the ill-posedness of the inverse problem, Donoho proposes a new type of nonlinear windowing known as *thresholding*. Let $(s)_+$ positive part of s. That is, let

$$(s)_+ = \begin{cases} s, & s > 0 \\ 0, & s \leq 0 \end{cases}$$

and define the function η_t by

$$\eta_t(s) = \operatorname{sgn}(s)(|s| - t)_+.$$

A sequence of non-negative *thresholds* $\{t_j : j \in \mathbb{Z}\}$ is selected and f is approximated by the series

$$\tilde{f}(t) = \sum_{j,k} \eta_{t_j}\left(\frac{\langle g^\delta, u^{j,k}\rangle}{\sigma_j}\right)\psi^{j,k}(t). \tag{4.30}$$

Donoho has shown that in certain Besov spaces, the thresholds can be selected to obtain faster rates of convergence than is possible with standard windowing methods 4.18.

The Besov spaces $B^s_{p,q}$ can be regarded as generalizations of the Sobolev spaces $H^s = B^s_{2,1}$ and the Hölder spaces $C^s = B^s_{\infty,\infty}$. It has been shown in [36, page 199] that for rapidly decreasing wavelets ψ of sufficient regularity, $B^s_{p,q}$ can be completely characterized in terms of wavelet coefficients. In particular, it is shown that $f \in B^s_{p,q}$ if and only if:

1. $\left\{\sum_k \left|\langle f, \phi^{0,k}\rangle\right|^p\right\}^{1/p} < \infty$ and
2. $\left\{\sum_k \left|\langle f, \psi^{j,k}\rangle\right|^p\right\}^{1/p} = 2^{-(1/2-1/p)j}2^{-sj}\varepsilon_j$ for some $\{\varepsilon_j\} \in \ell^q(\mathbb{N})$.

Suppose that f is supported in the interval $[-a, a]$ and B be a ball in $B^s_{p,q}$, then the author defines the *minimax-wavelet risk* by

$$R^W(\epsilon, B) = \inf_{\{t_j\}} \sup_{f\in B} E\left\|\tilde{f} - f\right\|_{L^2[-a,a]}, \tag{4.31}$$

with $\tilde{f}$ given by 4.30 and E the expectation operator. The rate of convergence is defined to be the rate at which 4.31 tends to zero as $\epsilon \to 0^+$, and the author proves that, in the case $p < 2$, this rate is optimal.

G. Walter uses a method inspired by the WVD to obtain expansions for wide-sense stationary stochastic processes that emulate the Karhunen-Loevé transformation [48]. In [44], similar ideas are employed to solve the mixture problem for random variables. Let X, Y and Z be random variables with respective probability density functions f, g and h. If the random variables Y and Z are independent and

$$X = Y + Z,$$

then the probability density functions are related by the equation

$$f(t) = \int_{-\infty}^{\infty} g(\tau)h(t-\tau)\, d\tau.$$

The mixture problem entails estimating the random variable Y from the noisy measurements X. That is, we want to solve for the probability density g, given the densities f and h. This is a problem of deconvolving g from h.

Walter begins by assuming that the unknown function belongs to the subspace V_n. In particular, it is assumed that V_n is generated by an orthonormal scaling function of Meyer type. Walter has shown that the Fourier transform of such a scaling function can be expressed in the form

$$\hat{\phi}^2(\omega) = \int_{-\pi}^{\pi} p(\omega - \nu)\, d\nu,$$

where p is a non-negative unimodular function with $\mathrm{supp}(p) \subset [-\pi/3, \pi/3]$. The Fourier transform of the resulting scaling function will be supported in the interval $[-4\pi/3, 4\pi/3]$ and will be m times continuously differentiable whenever p is $m-1$ times continuously differentiable.

Next, one assumes that $\left|\hat{h}(\omega)\right| > 0$ and defines the functions $u^{n,k}$ by

$$u^{n,k}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t}\frac{\widehat{\phi^{n,k}}(\omega)}{\bar{\hat{h}}(\omega)}\, d\omega \tag{4.32}$$

and the functions $v^{n,k}$ by

$$v^{n,k}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t}\widehat{\phi^{n,k}}(\omega)\hat{h}(\omega)\, d\omega. \tag{4.33}$$

Since $\hat{h}$ has no zeroes and $\hat{\phi}$ is compactly supported with $\text{supp}(\hat{\phi}) \subset [-4\pi/3, 4\pi/3]$, the functions $v^{n,k}$ and, in particular, the functions $u^{n,k}$, are well-defined $L^2(\mathbb{R})$ objects for all $n, k \in \mathbb{Z}$.

It is easy to show that the sets of functions $\{u^{n,k} : k \in \mathbb{Z}\}$ and $\{v^{n,k} : k \in \mathbb{Z}\}$ form Riesz bases for their closed linear spans. Furthermore, since

$$\langle \phi^{n,j}, \phi^{n,k} \rangle = \delta_{j,k},$$

it follows from 4.32 and 4.33 that

$$\langle u^{n,j}, v^{n,k} \rangle = \delta_{j,k}. \tag{4.34}$$

Recall that $g \in V_n$. This means that we can expand g in terms of $\{\phi^{n,k} : k \in \mathbb{Z}\}$ to obtain

$$g(t) = \sum_k a_k^n \phi^{n,k}(t)$$

whereupon we obtain, from 4.33 and the convolution equation, the expansion

$$f(t) = \sum_k a_k^n v^{n,k}(t).$$

If we now use the biorthogonality property 4.34, then we find that

$$a_k^n = \langle f(t), u^{n,k}(t) \rangle.$$

However, since

$$\left\|u^{n,k}\right\|^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\left|\hat{\phi}(\omega)\right|^2}{\left|\hat{h}(2^n\omega)\right|^2} d\omega$$

if for some $\alpha, A > 0$

$$|h(\omega)| \leq \frac{A}{(1+|\omega|)^\alpha},$$

then

$$\begin{aligned}\left\|u^{n,k}\right\|^2 &\geq \frac{1}{2\pi A}\int_{-\infty}^{\infty}(1+|2^n\omega|)^{2\alpha}\left|\hat{\phi}(\omega)\right|^2 d\omega \\ &\geq \frac{1}{A}\left(1+\frac{4^{n\alpha}}{2\pi}\int_{-4\pi/3}^{4\pi/3}|\omega|^{2\alpha}\left|\hat{\phi}(\omega)\right|^2 d\omega\right),\end{aligned}$$

which implies that as $n \to \infty$, the problem of recovering f from g and h becomes increasingly ill-conditioned.

We mention several points not examined by the author. First, even though the two sets $\{u^{n,k} : k \in \mathbb{Z}\}$ and $\{v^{n,k} : k \in \mathbb{Z}\}$ are biorthogonal, they are not in general dual Riesz bases. This is due to the fact that their closed linear spans do not in general coincide. To see that this is the case, suppose that $u^{n,0}$ can be expressed in terms of the basis $\{v^{n,k} : k \in \mathbb{Z}\}$. We find that

$$u^{n,0}(t) = \sum_k b_k^n v^{n,k}(t),$$

which implies that

$$\widehat{u^{n,0}}(\omega) = 2^{n/2}\sum_k b_k^n e^{-ik\omega/2^n}\hat{h}(\omega)\hat{\phi}\left(\omega/2^n\right).$$

However, by definition

$$\widehat{u^{n,0}}(\omega) = \frac{2^{n/2}\hat{\phi}\left(\omega/2^n\right)}{\overline{\hat{h}(\omega)}},$$

which implies that

$$\left|\hat{h}(\omega)\right|^2 \sum_k b_k^n e^{-ik\omega/2^n} = 1 \tag{4.35}$$

and since $\hat{h}$ is not periodic, equation 4.35 cannot hold for all $\omega \in \mathbb{R}$.

Secondly, we note that the orthogonal projection of $\{v^{n,k} : k \in \mathbb{Z}\}$ onto

$$U_n = \overline{\bigvee\left\{u^{n,k} : k \in \mathbb{Z}\right\}}$$

is in fact the dual basis of $\{u^{n,k} : k \in \mathbb{Z}\}$. Suppose we let $\{\tilde{u}^{n,k} : k \in \mathbb{Z}\}$ denote the dual basis. Then, for any $f \in L^2(\mathbb{R})$ the orthogonal projection of f onto U_n is given by

$$(P_u f)(t) = \sum_k \langle f, \tilde{u}^{n,k} \rangle u^{n,k}(t).$$

In particular, if we let $f = v^{n,j}$ and use 4.34, then we find that

$$\left(P_u v^{n,j}\right)(t) = \tilde{u}^{n,j}(t).$$

A consequence of the above is that if $f \in U_n$ then

$$\begin{aligned} f(t) &= \sum_k \langle f, \tilde{u}^{n,k} \rangle u^{n,k}(t) \\ &= \sum_k \langle f, v^{n,k} \rangle u^{n,k}(t). \end{aligned} \tag{4.36}$$

However, if the function f has components in the orthogonal compliment $U_n^{\perp}$, then the expansions 4.36 will not be equal. Since the orthogonal projection of f yields the best mean-square approximation of f in the subspace U_n, there may be circumstances in which the orthogonal projection offers some advantages. However, if n is sufficiently large, the differences may in fact be small.

In the next chapter, we shall consider a similar approach for the solution of convolution integral equations. However, our approach will differ from that of Walter's in the sense that we will make use of dual Riesz bases. Furthermore, the results that we will present are valid for general orthonormal scaling functions.

4.4 Discrete deconvolution

We now turn our attention to two papers in which methods for solving the discrete deconvolution problem

$$\sum_{k\in\mathbb{Z}} h_{j-k}x_k = y_j, \quad j \in \mathbb{Z} \tag{4.37}$$

are presented. Specifically, we will consider the approaches of Lu et al. (see [33]) and that of Chen and Lin [8].

In the article [33], the authors assume that equation 4.37 arises via a discretization of the continuous convolution equation

$$\int_{-\infty}^{\infty} h(t-\tau)x(\tau)\,d\tau = y(t).$$

Furthermore, it is supposed that the so-called blurring sequence $\{h_k : k \in \mathbb{Z}\}$ is in fact a two-scale sequence associated with the orthonormal scaling function ϕ. That is,

$$\phi(t) = \sqrt{2}\sum_k h_k\phi(2t-k),$$

with

$$\langle\phi(t-j), \phi(t-k)\rangle = \delta_{j,k}.$$

If in fact $\{h_k : k \in \mathbb{Z}\}$ is a bona-fide two-scale sequence, then its *discrete Fourier transform* (DFT)

$$H(\omega) = \sum_k h_k e^{-ik\omega}$$

must satisfy the equation

$$|H(\omega)|^2 + |H(\omega+\pi)|^2 = 2, \tag{4.38}$$

or equivalently

$$\sum_j h_{2k-j}h_j = \delta_{k,0}.$$

Two particular convolution kernels h are treated numerically in this paper. The first, a Gaussian kernel, arises in many practical situations. Suppose we assume that the sequence $\{h_k : k \in \mathbb{Z}\}$ is defined by sampling the function h on some regular grid. That is,

$$h_k = \frac{1}{\alpha} e^{-k^2/\alpha^2},$$

for some $\alpha > 0$. In this case we have

$$\sum_j h_{2k-j} h_j = \frac{1}{\alpha^2} \sum_j \exp(-(2k-j)^2/\alpha^2) \exp(-j^2/\alpha^2).$$

It is obvious that the quantity above is positive for any $k \in \mathbb{Z}$ and we conclude that no regular sampling of a Gaussian kernel will give rise to a two-scale sequence.

The second kernel to be considered is the so-called Fejér kernel. We assume that h is of the form

$$h(t) = \frac{1}{\sqrt{2}} \left(\frac{\sin \pi t/2}{\pi t/2} \right)^2,$$

where the scaling has been selected to ensure that $H(0) = \sqrt{2}$, which is implied by the normalization condition $\hat{\phi}(0) = 1$. Since h is band-limited, it is not unreasonable to assume that sampling should be at or above the Nyquist rate. With this in mind, we suppose that $\alpha \geq 1$ and set

$$h_k = \frac{1}{\alpha} h(k/\alpha).$$

The Fourier transform of h is

$$\hat{h}(\omega) = \sqrt{2} \chi_\pi(\omega)(1 - |\omega/\pi|)$$

and by Poisson's summation formula,

$$H(\omega) = \sum_{l=-\infty}^{\infty} \hat{h}(\alpha[\omega + 2\pi l]).$$

If we restrict our attention to the point $\omega = \pi/2$, then equation 4.38 becomes

$$|H(\pi/2)|^2 + |H(3\pi/2)|^2 = 2(1-\alpha/2)^2 \left(\chi_\pi(\alpha\pi/2) + \chi_\pi(-\alpha\pi/2)\right) \tag{4.39}$$

and it follows from 4.39 that 4.38 cannot be satisfied for any equally spaced sampling of the Fejér kernel with $\alpha > 1$ (sampling above the Nyquist rate).

In general, one cannot expect that a sampled convolution kernel will produce a two-scale sequence corresponding to an orthonormal scaling function. In this sense, the method developed Liu et al does not really make use of wavelet analysis. However, as we shall see, the proposed method does make use of an algorithm which is similar to a discrete wavelet transform so that the method is wavelet-like. Even so, one of the advantages of wavelet analysis is the freedom to choose discrete filters which are appropriate for a particular application. Since the filters employed by the current method are defined by the convolution kernel, such freedom is absent in this case.

As we have just mentioned, even if $\{h_k : k \in \mathbb{Z}\}$ is not a two-scale sequence, the proposed method can still be implemented and leads to a new kind of regularization. However, it is not entirely clear from the numerical evidence presented that this method offers any particular advantages over more traditional methods. Moreover, important properties such as convergence and regularity have not been considered. We outline the proposed method and make some comparisons with a standard form of regularization.

Let us suppose that the sequence $\{h_k : k \in \mathbb{Z}\}$ is summable and has been scaled so that

$$\sum_k |h_k| < 1.$$

In this case, we can always find another sequence $\{g_k : k \in \mathbb{Z}\}$, with DFT G such that

$$|H(\omega)|^2 + |G(\omega)|^2 = 1. \tag{4.40}$$

In fact, the choice

$$G(\omega) = \sqrt{1 - |H(\omega)|^2}$$

will do. At this point we should note that in the case of wavelet analysis, the filter G is taken as the mirror of the filter H. In particular,

$$G(\omega) = e^{-i\omega}\overline{H}(\omega + \pi). \tag{4.41}$$

However, unless H satisfies a rescaled version of 4.38, the choice 4.41 will not lead to a perfect reconstruction filter bank.

Now, since $H(0) \neq 0$ in many applications, one can view $\{h_k : k \in \mathbb{Z}\}$ as a low-pass filter, while $\{g_k : k \in \mathbb{Z}\}$ can be thought of as a high-pass filter. The authors use these filters to implement a filter bank scheme for digital signals. This scheme can be defined, in the frequency domain, by the formulae:

1. Decomposition.

$$X_j(\omega) = \begin{cases} X(\omega), & j = 0 \\ H(2^{j-1}\omega)X_{j-1}(\omega), & j = 1, 2, \ldots \end{cases} \tag{4.42}$$

and

$$D^j(\omega) = G(2^{j-1}\omega)X_{j-1}(\omega), \quad j = 1, 2, \ldots, . \tag{4.43}$$

2. Reconstruction.

$$X_{j-1}(\omega) = \overline{H}(2^{j-1}\omega)X_j(\omega) + \overline{G}(2^{j-1}\omega)D^j(\omega) \quad j = 1, 2, \ldots, . \tag{4.44}$$

This filtering scheme is analogous to a redundant (un-decimated) discrete wavelet transform (see [41]).

If we take the DFT of equation 4.37, then we obtain

$$H(\omega)X_0(\omega) = Y(\omega) \tag{4.45}$$

and in light of 4.42, we see that

$$X_1(\omega) = Y(\omega).$$

The deconvolution problem can now be thought of as one of recovering the missing information

$$D^1(\omega) = G(\omega)X_0(\omega).$$

Once D^1 has been obtained, the formula 4.44 can be used to recover an estimate of X_0. For the purposes of illustration, let us restrict our attention to a single level of decomposition and reconstruction.

We obtain from 4.42, 4.43 and 4.44, the special cases

$$\begin{aligned} X_1(\omega) &= H(\omega)X_0(\omega), \\ D^1(\omega) &= G(\omega)X_0(\omega) \end{aligned} \tag{4.46}$$

and

$$X_0(\omega) = \overline{H}(\omega)X_1(\omega) + \overline{G}(\omega)D^1(\omega) \tag{4.47}$$

and combining 4.46 and 4.47, we obtain the *reproducing equation*

$$\begin{pmatrix} X_1(\omega) \\ D^1(\omega) \end{pmatrix} = \begin{pmatrix} |H(\omega)|^2 & H(\omega)\overline{G}(\omega) \\ \overline{H}(\omega)G(\omega) & |G(\omega)|^2 \end{pmatrix} \begin{pmatrix} X_1(\omega) \\ D^1(\omega) \end{pmatrix}. \tag{4.48}$$

This system is singular and accordingly we will restrict our attention to the second equation. The equation relating the missing information D^1 and the data $X_1 = Y$ is

$$\left(1 - |G(\omega)|^2\right) D^1(\omega) = \overline{H}(\omega)G(\omega)Y(\omega). \tag{4.49}$$

The function H may have zeroes on $[-\pi, \pi]$ and since noisy data $Y^\delta = Y + \delta$ need not vanish at these points, equation 4.49 may not be well-posed. The authors suggest that one take as an approximation to D^1, the solution of the equation

$$\left(1 - \alpha |G(\omega)|^2\right) D^1_\alpha(\omega) = \overline{H}(\omega)G(\omega)Y^\delta(\omega),$$

where $\alpha < 1$. Upon letting $\alpha = 1 - \lambda$, the use of 4.38 yields

$$\left(|H(\omega)|^2 + \lambda |G(\omega)|^2\right) D^1_\lambda(\omega) = \overline{H}(\omega)G(\omega)Y^\delta(\omega), \tag{4.50}$$

which is the minimizer of the functional

$$F_1(D) = \left\|HD - GY^\delta\right\|^2 + \lambda \left\|GD\right\|^2.$$

If we apply 4.47, then from 4.50 we find that

$$\begin{aligned} X(\omega) \approx X^\delta_\lambda(\omega) &= \overline{H}(\omega)Y^\delta(\omega) + \overline{G}(\omega)D^1_\lambda(\omega) \\ &= \left(1 + \frac{|G(\omega)|^2}{|H(\omega)|^2 + \lambda |G(\omega)|^2}\right) \overline{H}(\omega)Y^\delta(\omega) \\ &= \left(\frac{1 + \lambda |G(\omega)|^2}{|H(\omega)|^2 + \lambda |G(\omega)|^2}\right) \overline{H}(\omega)Y^\delta(\omega). \end{aligned} \tag{4.51}$$

which could be thought of as the minimizer of the functional

$$\tilde{F}_1(X) = \left\|HX - (1 + \lambda |G|^2)Y^\delta\right\|^2 + \lambda \left\|GX\right\|^2. \tag{4.52}$$

In view of the unusual form of 4.52, there is no reason to suspect that the regularization method this functional defines possesses any of the important properties usually required of more familiar methods. To demonstrate that 4.52 does in fact give rise to a viable regularization method, we first show that in the absence of noise δ

$$\lim_{\lambda \to 0^+} \|X - X_\lambda\| = 0,$$

where X is a solution of $HX = Y$.

Let Z_H be the set

$$Z_H = \{\omega \in [-\pi, \pi] : H(\omega) = 0\}$$

and assume for simplicity that the measure of Z_H is zero. This assumption implies $N(H) = \{0\}$, making the solution X unique (a proof for the more general case $N(H) \neq \{0\}$ is analogous). Now, since

$$\begin{aligned} X(\omega) - X_\lambda(\omega) &= \left(1 - \frac{1 + \lambda |G(\omega)|^2}{|H(\omega)|^2 + \lambda |G(\omega)|^2} |H(\omega)|^2\right) X(\omega) \\ &= \frac{\lambda |G(\omega)|^4}{|H(\omega)|^2 + \lambda |G(\omega)|^2} X(\omega), \end{aligned}$$

we find that

$$\|X - X_\lambda\|^2 = \frac{1}{2\pi} \int_{Z_H^c} \left(\frac{\lambda |G(\omega)|^4}{|H(\omega)|^2 + \lambda |G(\omega)|^2}\right)^2 |X(\omega)|^2 \, d\omega,$$

where

$$Z_H^c = [-\pi, \pi] \setminus Z_H.$$

Observe that for every $\omega \in Z_H^c$,

$$\lim_{\lambda \to 0^+} \frac{\lambda |G(\omega)|^4}{|H(\omega)|^2 + \lambda |G(\omega)|^2} = 0$$

and

$$\frac{\lambda |G(\omega)|^4}{|H(\omega)|^2 + \lambda |G(\omega)|^2} \leq |G(\omega)|^2 \leq 1.$$

The desired result now follows from the Lebesgue dominated convergence theorem.

We now turn our attention to the property of regularity. Specifically, since the choice of the regularization parameter λ will depend on δ, we need to show that it is possible to choose $\lambda = \lambda(\delta)$ so that

$$\lim_{\|\delta\| \to 0^+} \left\|X - X^\delta_{\lambda(\delta)}\right\| = 0, \tag{4.53}$$

where we have used notation

$$X_\lambda^\delta(\omega) = \frac{1+\lambda\,|G(\omega)|^2}{|H(\omega)|^2+\lambda\,|G(\omega)|^2}\overline{H}(\omega)Y^\delta(\omega)$$

to denote the approximate solution obtained from the noisy data Y^δ.

We proceed by considering the behavior of $X_\lambda^\delta - X_\lambda$, where X_λ is the approximate solution in the absence of noise. If we use 4.40, then we find that

$$X_\lambda^\delta(\omega) - X_\lambda(\omega) = \frac{1+\lambda-\lambda\,|H(\omega)|^2}{(1-\lambda)\,|H(\omega)|^2+\lambda}\overline{H}(\omega)\delta(\omega)$$

and if we define the function f_λ by

$$f_\lambda(x) = \left(\frac{1+\lambda-\lambda x}{(1-\lambda)x+\lambda}\right)^2 x.$$

for $x \in [0,1]$, then

$$\left\|X_\lambda^\delta - X_\lambda\right\|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_\lambda\left(|H(\omega)|^2\right)|\delta(\omega)|^2\,d\omega.$$

It can be shown that for $x \in [0,1]$, the function f_λ takes on its maximum value at the point

$$x(\lambda) = \frac{2\lambda^2+1-\sqrt{8\lambda^2+1}}{2\lambda(\lambda-1)},$$

with

$$f_\lambda(x(\lambda)) = \frac{\lambda(3-\sqrt{8\lambda^2+1})^2(\sqrt{8\lambda^2+1}-2\lambda^2-1)}{2(1-\sqrt{8\lambda^2+1})^2(1-\lambda)^3}. \tag{4.54}$$

In light of 4.54 we find that

$$\left\|X_\lambda^\delta - X_\lambda\right\|^2 \leq f_\lambda(x(\lambda))\,\|\delta\|^2$$

and since

$$\begin{aligned} \left\|X - X_\lambda^\delta\right\| &\leq \|X - X_\lambda\| + \left\|X_\lambda^\delta - X_\lambda\right\| \\ &\leq \|X - X_\lambda\| + \sqrt{f_\lambda(x(\lambda))}\,\|\delta\|, \end{aligned} \tag{4.55}$$

the desired result 4.53 will follow as long as we choose $\lambda = \lambda(\delta)$ so that:

1. $\lambda \to 0^+$ as $\|\delta\| \to 0^+$ and

2. $f_\lambda(x(\lambda)) \|\delta\|^2 \to 0^+$ as $\|\delta\| \to 0^+$.

Now, since

$$\sqrt{f_\lambda(x(\lambda))} \sim \frac{1}{2\sqrt{\lambda}} + \frac{3\sqrt{\lambda}}{4} + O(\lambda^{3/2}),$$

we find that the choice

$$\lambda(\delta) = O(\|\delta\|^{2-\alpha}),$$

with $0 < \alpha < 2$, provides the estimate

$$\left\|X^\delta_{\lambda(\delta)} - X_{\lambda(\delta)}\right\| \sim \frac{1}{2} \|\delta\|^{\alpha/2} + O(\|\delta\|^{2-\alpha/2}). \tag{4.56}$$

We point out that the leading term of the estimate 4.56 is consistent with results obtained when minimum norm regularization is applied to compact operators (see for example [29]).

In the special case

$$|H(\omega)|^2 \geq A > 0,$$

we find that

$$\|X - X_\lambda\| \leq \frac{\lambda}{A} \|X\|,$$

whereupon the use of 4.55 yields the error bound

$$\left\|X - X^\delta_{\lambda(\delta)}\right\| \leq \frac{\lambda}{A} \|X\| + \sqrt{f_\lambda(x(\lambda))} \|\delta\|. \tag{4.57}$$

If we consider the leading order behavior of 4.57, then we obtain

$$\left\|X - X^\delta_{\lambda(\delta)}\right\| \sim O(\|\delta\|^{2-\alpha}) + O(\|\delta\|^{\alpha/2}), \tag{4.58}$$

which, once again, is consistent with the results found in [29]. In fact, in the case of *C-generalized regularization* as applied to discrete deconvolution, analysis, similar to that above, yields an estimate, of the type 4.57 with the same leading order behavior as 4.58

In view of this similarity, we suspect that there is little difference between the method of Liu et al. and that of C-generalized regularization. The method of C-generalized regularization can be defined by through the use of the functional

$$F_2(X) = \left\|HX - Y^\delta\right\|^2 + \lambda \left\|GX\right\|^2. \tag{4.59}$$

which is minimized by

$$X_\delta^\lambda(\omega) = \frac{\overline{H}(\omega)}{|H(\omega)|^2 + \lambda\,|G(\omega)|^2} Y^\delta(\omega).$$

We define the inverse filters $\mathcal{H}^\lambda$ and $\mathcal{H}_\lambda$ by

$$\mathcal{H}^\lambda(\omega) = \frac{\overline{H}(\omega)}{|H(\omega)|^2 + \lambda\,|G(\omega)|^2}$$

and

$$\mathcal{H}_\lambda(\omega) = \frac{1 + \lambda\,|G(\omega)|^2}{|H(\omega)|^2 + \lambda\,|G(\omega)|^2}\overline{H}(\omega)$$

and observe that they are related by the equation

$$\mathcal{H}_\lambda(\omega) = (1 + \lambda\,|G(\omega)|^2)\mathcal{H}^\lambda(\omega). \tag{4.60}$$

Plots of the filters $\mathcal{H}^\lambda$ and $\mathcal{H}_\lambda$ in the case of a Gaussian convolution kernel appear in figures 4.1, 4.2 and 4.3. We observe that for small λ, there is almost no difference between the plots of $\mathcal{H}^\lambda$ and $\mathcal{H}_\lambda$, while for larger values of λ, the main difference is near the mid-band frequencies $\omega = \pm\pi/2$. These observations are confirmed by the equation 4.60, from which we conclude that if $||\delta||$ is small, then there will be negligible differences between the two regularization methods.

Finally, we should mention that, as in [33], we have restricted our consideration to a single application of 4.42 and 4.43 to 4.45. As is demonstrated by the authors,

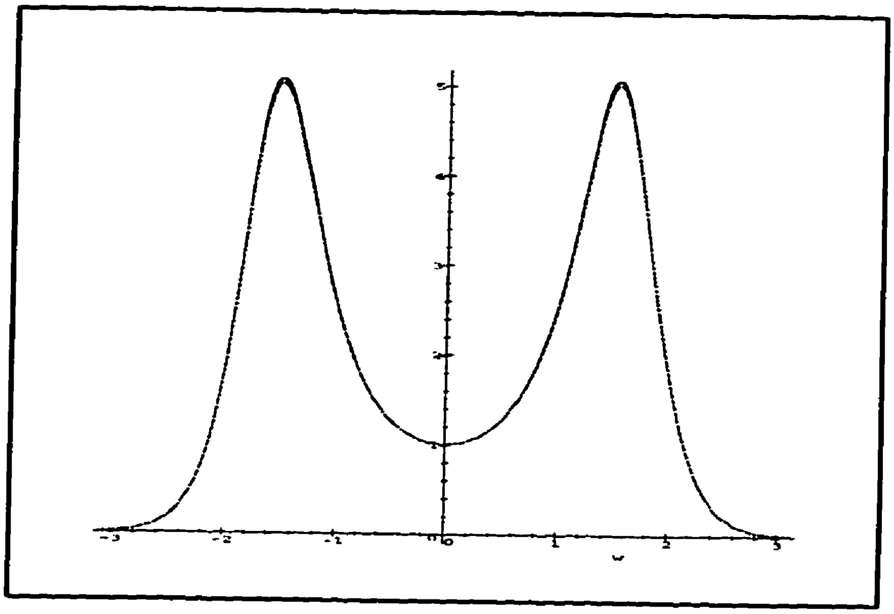

Figure 4.1: $\mathcal{H}^{\lambda}$, $\mathcal{H}_{\lambda}$, $\lambda = .01$

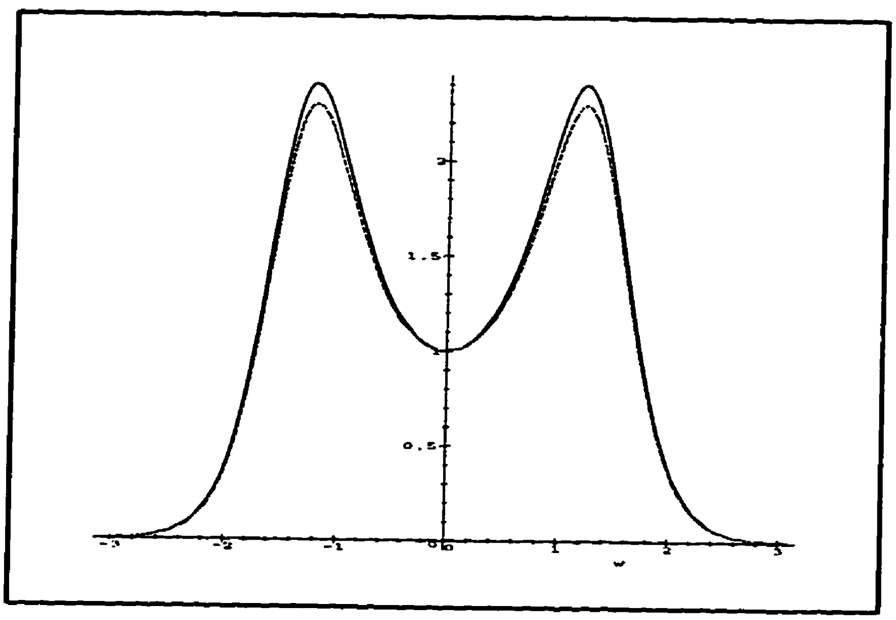

Figure 4.2: $\mathcal{H}^{\lambda}$, $\mathcal{H}_{\lambda}$, $\lambda = .05$

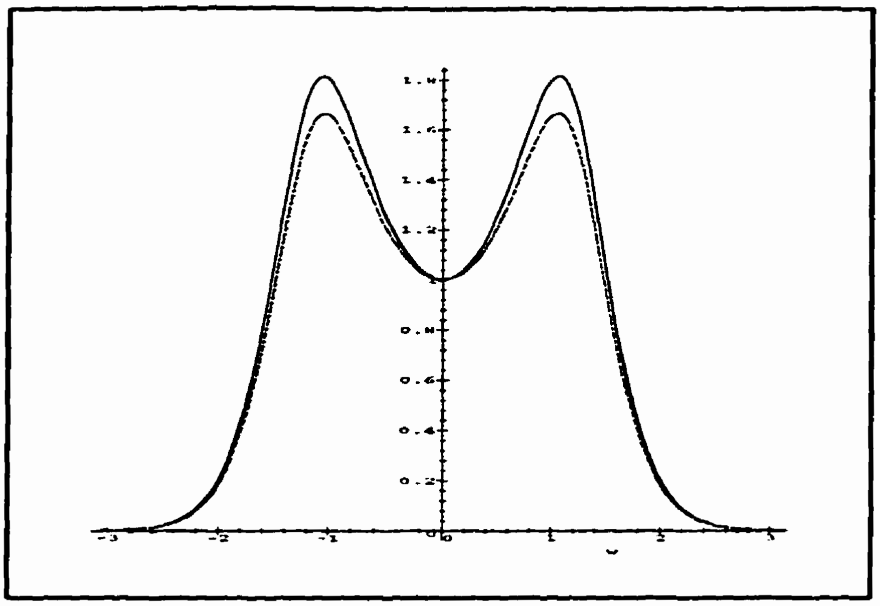

Figure 4.3: $\mathcal{H}^\lambda$, $\mathcal{H}_\lambda$, $\lambda = .1$

it is possible to derive more general reproducing equations in analogy with 4.48. This leads to the possibility of a regularization method, with more than one regularization parameter, which may improve the flexibility of the proposed method. However, we shall leave consideration of this generalization for future research.

We now consider the second article in which a discrete deconvolution problem is addressed. The main focus of the paper [8] is the equation

$$\sum_{k=-\infty}^{\infty} b[j-k]x[k] = y[j], \quad j \in \mathbb{Z}, \tag{4.61}$$

where the sequence $\{x[k] : k \in \mathbb{Z}\}$ is assumed to be obtained by sampling a $1/f$ fractal process $x(t)$. In general, a stochastic process x is said to be a $1/f$ process if its empirical *power spectral density* $\hat{R}$ satisfies

$$\hat{R}(\omega) \sim \frac{\sigma^2}{|\omega|^\gamma}, \tag{4.62}$$

for some γ such that $0 < \gamma < 2$. The authors consider the problem of estimating the sequence $\{x[k] : k \in \mathbb{Z}\}$ from the noisy data

$$y^{\delta}[k] = y[k] + \delta[k], \quad k \in \mathbb{Z},$$

with $\{\delta[k] : k \in \mathbb{Z}\}$ a sampling of a white noise process $\delta(t)$.

Due to the non-stationary nature of x (i.e. self-similarity and slow decay of correlation) standard statistical filtering technique, such as Wiener filtering, prove to be inadequate. The authors propose a multiscale filtering technique which combines Wiener filtering and an un-decimated wavelet transform. This method relies on the ability of the wavelet transform to remove, or at least reduce, non-stationary effects making the subsequent Wiener filtering more effective.

The ability of wavelet transforms to reduce non-stationary behavior in $1/f$ processes is not completely unjustified. In the special case of fractional Brownian motion, Flandrin (see [19]) has shown that the continuous wavelet transform of x, given by

$$(\mathcal{W}x)(t, a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(s)\psi\left(\frac{s-t}{a}\right) ds,$$

is a wide-sense stationary process in t with power spectral density

$$\hat{S}_a(\omega) = a\sigma^2 \frac{\left|\hat{\psi}(a\omega)\right|^2}{|\omega|^{2H+1}}.$$

In the related work [35], Masry shows that for each fixed j, the wavelet coefficients

$$x^j[k] = \langle x, \psi^{j,k} \rangle$$

of a stochastic process x, with wide-sense stationary increments, forms a wide-sense stationary sequence $\{x^j[k] : k \in \mathbb{Z}\}$.

Admittedly, these results are far from being complete and the utility of wavelet transforms for the analysis of general $1/f$ processes is still under investigation. However, it is noted in [45] that there is strong empirical evidence to suggest that wavelet transforms serve as whitening filters for all $1/f$ processes. That is to say, if x is a zero mean $1/f$ process with power spectral density 4.62, the wavelet coefficients $x^j[k]$ are weakly correlated along and across scales and satisfy

$$E\left\{\left|x^j[k]\right|^2\right\} \approx \sigma^2 2^{-\gamma j}.$$

It is the properties above that make the proposed method plausible.

Let $\{h_k : k \in \mathbb{Z}\}$ and $\{g_k : k \in \mathbb{Z}\}$ be the two-scale sequences corresponding to the orthonormal wavelet basis $\{\psi^{j,k} : j, k \in \mathbb{Z}\}$. In particular, we suppose that these sequences satisfy the dilation equations

$$\phi(t) = \sqrt{2}\sum_k h_k \phi(2t - k)$$

and

$$\psi(t) = \sqrt{2}\sum_k g_k \phi(2t - k)$$

with $\{g_k : k \in \mathbb{Z}\}$ the mirror of $\{h_k : k \in \mathbb{Z}\}$, given by $g_k = (-1)^{k+1} h_{k-1}$. One can define the un-decimated discrete wavelet transform (analogous to 4.42 and 4.43) of the sequence $\{x[k] : k \in \mathbb{Z}\}$ via the equations

$$X_j(\omega) = \overline{H}(2^{j-1}\omega) X_{j-1}(\omega), \quad j = 1, 2, \ldots \tag{4.63}$$

and

$$X^j(\omega) = \overline{G}(2^{j-1}\omega) X_{j-1}(\omega), \quad j = 1, 2, \ldots, \tag{4.64}$$

where $X_0 = X$ is the the DFT of $\{x[k] : k \in \mathbb{Z}\}$. If we take the DFT of equation 4.61, the we obtain

$$B(\omega) X(\omega) = Y(\omega) \tag{4.65}$$

and by repeatedly applying equations 4.63 and 4.64, we can derive a hierarchy of equations equivalent to 4.65. For example, a single iteration of this process produces the system of equations

$$\begin{aligned} B(\omega)X^1(\omega) &= Y^1(\omega), \\ B(\omega)X_1(\omega) &= Y_1(\omega), \end{aligned} \tag{4.66}$$

where $Y^1 = \overline{G}Y$ and $Y_1 = \overline{H}Y_1$. Since X^1 and X_1 must be estimated from the noisy data

$${}^{\delta}Y^1(\omega) = \overline{G}(\omega)Y^{\delta}(\omega)$$

and

$${}^{\delta}Y_1(\omega) = \overline{H}(\omega)Y^{\delta}(\omega),$$

the authors propose the method of Wiener filtering for the approximation of X^1 and X_1. Once suitable estimates of X^1 and X_1 have been obtained, X_0 can be estimated by applying a down sampled inverse wavelet transform.

Wiener filtering is a least-squares technique for the estimation of a stochastic process from noisy or imprecise measurements (see [39]). For instance, if this technique is applied to 4.61, then $\{x[k] : k \in \mathbb{Z}\}$ is approximated by the sequence

$$\tilde{x}[k] = \sum_j c[k-j]y^{\delta}[j], \quad k \in \mathbb{Z}.$$

The filter $\{c[k] : k \in \mathbb{Z}\}$ is to be selected so that the expected mean-square error

$$E\left\{|e[k]|^2\right\} = E\left\{|x[k] - \tilde{x}[k]|^2\right\}$$

is minimized. If one assumes that $\{x[k] : k \in \mathbb{Z}\}$ and $\{\delta[k] : k \in \mathbb{Z}\}$ are wide-sense stationary sequences with zero cross-correlation, then

$$E\left\{|e[k]|^2\right\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} |1 - C(\omega)B(\omega)|^2\, \hat{R}_x(\omega)\, d\omega + \frac{1}{2\pi}\int_{-\pi}^{\pi} |C(\omega)|^2\, \hat{R}_{\delta}(\omega)\, d\omega. \tag{4.67}$$

where $\hat{R}_x$ and $\hat{R}_\delta$ are the DFTs of the autocorrelation sequences

$$R_x[k] = E\{x[j]x[j+k]\}$$

and

$$R_\delta[k] = E\{\delta[j]\delta[j+k]\}$$

If it is now assumed that $\{x[k] : k \in \mathbb{Z}\}$ and $\{\delta[k] : k \in \mathbb{Z}\}$ are white noise sequences with the respective variances σ_x^2 and σ_δ^2, then 4.67 can be written as

$$E\left\{|e[k]|^2\right\} = \frac{\sigma_x^2}{2\pi}\int_{-\pi}^{\pi} |1 - C(\omega)B(\omega)|^2\, d\omega + \frac{\sigma_\delta^2}{2\pi}\int_{-\pi}^{\pi} |C(\omega)|^2\, d\omega. \qquad (4.68)$$

Equation 4.68 defines a functional $J(C) = E\left\{|e[k]|^2\right\}$ and if we minimize this functional on the Hilbert space $L^2[-\pi, \pi]$, then we obtain the result

$$C(\omega) = \frac{\sigma_x^2 \overline{B}(\omega)}{\sigma_\delta^2 + \sigma_x^2 |B(\omega)|^2}. \qquad (4.69)$$

We emphasize that the derivation of 4.69 depends heavily on the assumption that $\{x[k] : k \in \mathbb{Z}\}$ and $\{\delta[k] : k \in \mathbb{Z}\}$ are wide-sense stationary white noise sequences. With regard to the solution of system 4.66, it is the role of the wavelet transform as a whitening filter for $1/f$ processes that makes this assumption reasonable. Before further consideration of the system 4.66, we list several points regarding the filter $\{c[k] : k \in \mathbb{Z}\}$:

1. Stability. The filter $\{c[k] : k \in \mathbb{Z}\}$ can be computed from its DFT 4.69 through the use of the equation

$$c[k] = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ik\omega} C(\omega)\, d\omega$$

and if we assume that $\{b[k] : k \in \mathbb{Z}\}$ is a sequence in $\ell^1(\mathbb{Z})$, then it is easy to show that

$$\sum_k |c[k]|^2 < \infty.$$

However it is not entirely clear that $\{c[k] : k \in \mathbb{Z}\}$ is a stable filter in the sense that

$$\sum_k |c[k]| < \infty.$$

In this case however, a theorem, due to Wiener (see for example [26]), ensures that $\{c[k] : k \in \mathbb{Z}\} \in \ell^1(\mathbb{Z})$ whenever $\{b[k] : k \in \mathbb{Z}\} \in \ell^1(\mathbb{Z})$.

2. Causality. If $\{b[k] : k \in \mathbb{Z}\}$ is causal, that is $b_k = 0$ for all $k = -1, -2, \ldots$ then $\{c[k] : k \in \mathbb{Z}\}$ will not be causal. The Wiener filtering technique can be adapted to produce a causal filter, which may be desirable for certain applications. This complication in considered in some detail in [8] and will not be pursued further here.

3. Regularization. The DFT of the filter $\{c[k] : k \in \mathbb{Z}\}$ is identical in form to that obtained by a minimum norm regularization scheme in which the regularization parameter λ is predetermined by the variances σ_x^2 and σ_δ^2. This observation suggests the possibility of a redundant multiscale regularization method for deterministic convolution equations.

Let us now return to 4.66 and suppose that the assumptions needed for Wiener filtering hold. That is, let us suppose that $\{x^1[k] : k \in \mathbb{Z}\}$, $\{x_1[k] : k \in \mathbb{Z}\}$, $\{\delta^1[k] : k \in \mathbb{Z}\}$ and $\{\delta_1[k] : k \in \mathbb{Z}\}$ are white noise sequences with respective variances $\sigma_{x^1}^2$, $\sigma_{x_1}^2$, $\sigma_{\delta^1}^2$ and $\sigma_{\delta_1}^2$. We should point out that in view of 4.62 and the fact that $\hat{\phi}(0) = 1$, the sequence $\{x_1[k] : k \in \mathbb{Z}\}$ may not be well-defined. In such cases, 4.64 should be used to replace 4.66 with a system of equation for the detail sequences $\{x^j[k] : j \geq 1, k \in \mathbb{Z}\}$ only.

If we apply Wiener filtering to each equation in 4.66, in the presence of noisy data, we obtain the approximations

$$X_\lambda^1(\omega) = \frac{\overline{B}(\omega)}{\lambda + |B(\omega)|^2} {}^\delta Y^1(\omega) \tag{4.70}$$

and

$$X_1^\mu(\omega) = \frac{\overline{B}(\omega)}{\mu + |B(\omega)|^2} {}^\delta Y_1(\omega), \tag{4.71}$$

where

$$\lambda = \frac{\sigma_{x^1}^2}{\sigma_{\delta^1}^2}$$

and

$$\mu = \frac{\sigma_{x_1}^2}{\sigma_{\delta_1}^2}.$$

We recall that, in the frequency domain, the operation of down-sampling is given by the mapping

$$X(\omega) \longmapsto (\mathcal{D}X)(\omega) = \frac{1}{2}\left(X(\omega/2) + X(\omega/2 + \pi)\right). \tag{4.72}$$

If 4.70 and 4.71 are down-sampled and the inverse discrete wavelet transform (DWT) applied to the resulting functions, then the approximation

$$\begin{aligned} X(\omega) \approx X_\lambda^\mu(\omega) &= \frac{1}{2}\left(|H(\omega)|^2 \mathcal{B}_\mu(\omega) + |G(\omega)|^2 \mathcal{B}_\lambda(\omega)\right) Y^\delta(\omega) \\ &+ \frac{1}{2}\left(H(\omega)\overline{H}(\omega+\pi)\mathcal{B}_\mu(\omega+\pi)\right. \\ &+ \left. G(\omega)\overline{G}(\omega+\pi)\mathcal{B}_\lambda(\omega+\pi)\right) Y^\delta(\omega+\pi), \end{aligned} \tag{4.73}$$

with

$$\mathcal{B}_\gamma(\omega) = \frac{\overline{B}(\omega)}{\gamma + |B(\omega)|^2},$$

is obtained. Alternatively, if the undecimated inverse DWT is applied directly to 4.70 and 4.71, then the approximation

$$X(\omega) \approx \tilde{X}_\lambda^\mu(\omega) = \frac{1}{2}\left(|H(\omega)|^2 \mathcal{B}_\mu(\omega) + |G(\omega)|^2 \mathcal{B}_\lambda(\omega)\right) Y^\delta(\omega) \tag{4.74}$$

is produced. Now, since the filters H and G possess the properties:

1. $|H(\omega)|^2 + |G(\omega)|^2 = 2$ and

2. $H(\omega)\overline{H}(\omega+\pi) + G(\omega)\overline{G}(\omega+\pi) = 0$,

we see that, when $\lambda = \mu$,

$$X(\omega) \approx X_\lambda^\lambda(\omega) = \tilde{X}_\lambda^\lambda(\omega) = B_\lambda(\omega)Y^\delta(\omega),$$

which is the approximation one expects when ordinary Wiener filtering is applied directly to 4.65.

We emphasize that the method of Wiener filtering is dependent upon the knowledge of the power spectral densities of the sequences $\{x[k] : k \in \mathbb{Z}\}$ and $\{\delta[k] : k \in \mathbb{Z}\}$ and, in many cases, such specific knowledge may not be available. However, as we have mentioned, 4.73 and 4.74 suggest the possibility of a multiscale regularization method, which is a natural generalization of Tikhonov's regularization 4.17.

Chapter 5

Methods based on scaling functions

5.1 Introduction

In this chapter, we will examine some of the consequences of the assumption $u \in V_n$ with regard to equation 2.4. In particular, we are interested in cases for which this assumption leads to a modified problem which is well-posed. We show that there are conditions, involving the spectrum of a real, symmetric Toeplitz matrix, which ensure that the operator[1] $\mathcal{G}|_{V_n}$ is either weakly or strongly invertible. Subsequently, some of the properties of this matrix are investigated and then used to establish a convergence result as well as comment upon the condition of the modified problem.

[1]The restriction of $\mathcal{G}$ to V_n.

5.2 Deconvolution of functions in scaling function subspaces

Let us assume that the unknown function u belongs to the scaling function subspace

$$V_n = \overline{\bigvee \{\phi^{n,k} : k \in \mathbb{Z}\}},$$

where ϕ is an orthonormal scaling function. If we let $u_n[k] = \langle u, \phi^{n,k} \rangle$, then u admits an expansion of the form

$$u(t) = \sum_k u_n[k]\phi^{n,k}(t), \tag{5.1}$$

which allows us to rewrite 2.4 as

$$\sum_k u_n[k] \int_{-\infty}^{\infty} g(t-\tau)\phi^{n,k}(\tau)\,d\tau = y(t). \tag{5.2}$$

Now, since

$$\begin{aligned}
\left(\mathcal{G}\phi^{n,k}\right)(t) &= \int_{-\infty}^{\infty} g(t-\tau)2^{n/2}\phi(2^n\tau - k)\,d\tau \\
&= \int_{-\infty}^{\infty} g(t - k/2^n - \tau)2^{n/2}\phi(2^n\tau)\,d\tau \\
&= \left(\mathcal{G}\phi^{n,0}\right)(t - k/2^n),
\end{aligned}$$

if we use the notation $\xi_n(t) = \left(\mathcal{G}\phi^{n,0}\right)(t)$, then 5.2 becomes

$$\sum_k u_n[k]\xi_n(t - k/2^n) = y(t). \tag{5.3}$$

We see that under the assumption $u \in V_n$, the problem of solving 2.4 for the unknown function u is equivalent to the problem of solving 5.3 for the unknown sequence $\{u_n[k] : k \in \mathbb{Z}\}$. In terms of operators, we have simply replaced the equation

$$\mathcal{G}u = y \tag{5.4}$$

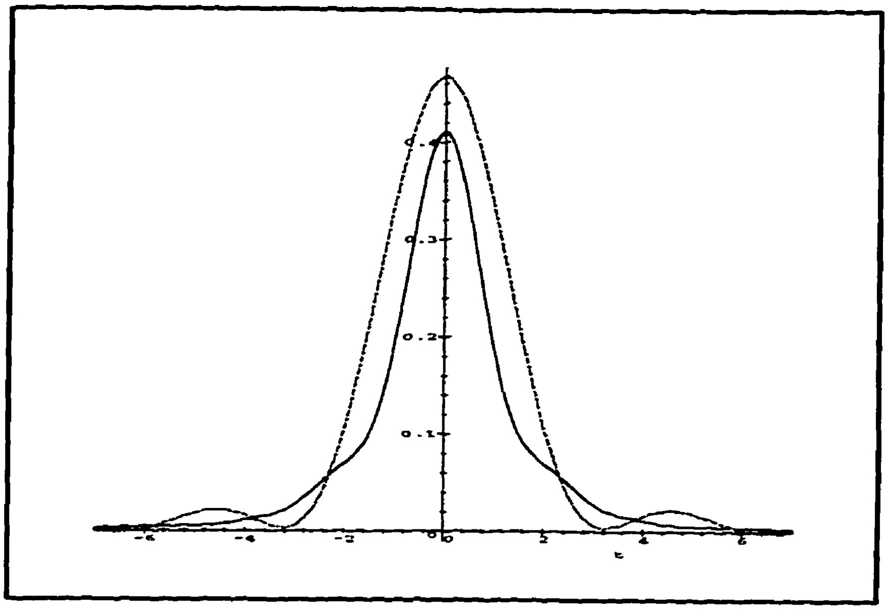

Figure 5.1: The functions ξ_{-1} and ξ_0

with the new equation

$$\mathcal{G}|_{V_n} u = y, \tag{5.5}$$

where $\mathcal{G}|_{V_n}$ denotes the restriction of $\mathcal{G}$ to the subspace V_n.

The expansion 5.3 is similar to the scaling function expansion 5.1 in the sense that the relevant functions are obtained by translating a single function by the amount $k/2^n$. However, unlike 5.1, the basic function ξ_n depends on n, and as the resolution is changed we obtain expansions with respect to different functions. This is due to the fact that the operations of translation and dilation do not commute. Figure 5.1 shows typical examples of the function ξ_n in the case where $g(t) = e^{-|t|}$.

In section 2.2, it was shown that the problem defined by 2.4 is ill-posed. In the coming sections, we shall consider circumstances under which the new problem, defined by 5.3, is well-posed. This consideration will entail an investigation of the sequence of functions $\{\xi_n(t - k/2^n) : k \in \mathbb{Z}\}$. In particular, we are interested in the properties of $\{\xi_n(t - k/2^n) : k \in \mathbb{Z}\}$ which will enable us to recover $\{u_n[k] : k \in \mathbb{Z}\}$ in a unique and continuous way from the specified data. In the next section, we begin this investigation by considering the simpler problem of recovering a single scaling function coefficient $u_n[k]$.

5.3 Coefficient functionals and weak invertibility

Consider the equation 5.4 and let z be some element of $L^2(\mathbb{R})$. Suppose that, instead of solving 5.4 for u, we seek to recover the *moment* $\langle u, z\rangle$ from y. If we define the linear functional $c : L^2(\mathbb{R}) \to \mathbb{R}$ by

$$c(y) = \langle u, z\rangle$$

then is can be shown that c is continuous if and only if $z \in R(\mathcal{G}^*)$, where $\mathcal{G}^*$ is the adjoint of $\mathcal{G}$ (see [1, 18]).

We are particularly interested in the functionals $c_{n,k} : \overline{\mathcal{G}(V_n)} \to \mathbb{R}$, which are defined by

$$c_{n,k}(y) = \langle u, \phi^{n,k}\rangle. \tag{5.6}$$

Specifically, we want to know when the functionals 5.6 are continuous. Since we have assumed that $u \in V_n$, we are working with the operator equation 5.5 and accordingly, the functionals 5.6 will be continuous as long as $\phi^{n,k}$ belongs to the range of the adjoint of the operator $\mathcal{G}|_{V_n}$ for all $k \in \mathbb{Z}$.

Definition 5.1 *If the coefficient functionals $c_{n,k}$ are continuous for all $k \in \mathbb{Z}$, then we will say that the operator $\mathcal{G}|_{V_n}$ is weakly invertible.*

We begin by considering the adjoint of $\mathcal{G}|_{V_n}$. Let H_1 and H_2 be Hilbert spaces with respective inner-products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$. Suppose that $\mathcal{K} : H_1 \to H_2$ is a bounded linear operator. The adjoint $\mathcal{K}^* : H_2 \to H_1$ is the bounded linear operator defined by the equation

$$\langle \mathcal{K}x, y \rangle_2 = \langle x, \mathcal{K}^* y \rangle_1.$$

The existence of the operator $\mathcal{K}^*$ is guaranteed by the Riesz Representation Theorem.

We can regard $\mathcal{G}|_{V_n}$ as a linear operator mapping the Hilbert space V_n into the Hilbert space $\overline{\mathcal{G}(V_n)} \equiv \overline{R(\mathcal{G}|_{V_n})}$. The adjoint of $\mathcal{G}|_{V_n}$ is defined by

$$\langle \mathcal{G}|_{V_n} f, h \rangle = \langle f, (\mathcal{G}|_{V_n})^* h \rangle,$$

where $f \in V_n$ and $h \in \overline{\mathcal{G}(V_n)}$. We notice that for $f \in V_n$

$$\langle \mathcal{G}|_{V_n} f, h \rangle = \langle \mathcal{G} f, h \rangle = \langle f, \mathcal{G}^* h \rangle$$

and since $P_n f = f$, we find that

$$\langle \mathcal{G}|_{V_n} f, h \rangle = \langle f, P_n \mathcal{G}^* h \rangle.$$

It now follows that

$$(\mathcal{G}|_{V_n})^* = P_n \mathcal{G}^* |_{\overline{\mathcal{G}(V_n)}}. \tag{5.7}$$

and hence the linear functionals $c_{n,k}$ are continuous if and only if

$$\phi^{n,k} \in R\left((\mathcal{G}|_{V_n})^*\right) = P_n \mathcal{G}^* \left(\overline{\mathcal{G}(V_n)}\right).$$

We immediately notice that if $\left\{\phi^{n,k} : k \in \mathbb{Z}\right\} \subset R\left((\mathcal{G}|_{V_n})^*\right)$, we must have

$$\overline{R\left((\mathcal{G}|_{V_n})^*\right)} = V_n$$

and since, for any bounded linear operator $\mathcal{K}$, $\overline{R(\mathcal{K}^*)} = N(\mathcal{K})^\perp$, the above implies

$$N(\mathcal{G}|_{V_n}) = \{0\}.$$

That is, if the $c_{n,k}$ are continuous, $\mathcal{G}|_{V_n}$ must be a bijection onto its range. Furthermore, we observe that

$$\langle u(t), \phi^{n,k}(t)\rangle = \langle u(t + k/2^n), \phi^{n,0}(t)\rangle.$$

Let $u^k(t) = u(t + k/2^n)$, then if $y = \mathcal{G}u$, we find that

$$y(t + k/2^n) = \left(\mathcal{G}u^k\right)(t)$$

and hence

$$\begin{aligned} c_{n,k}(y) &= \langle u(t + k/2^n), \phi^{n,0}(t)\rangle \\ &= c_{n,0}(y(t + k/2^n)). \end{aligned}$$

In other words, since V_n is closed under translations which are integer multiples of 2^{-n}, the $c_{n,k}$ are continuous if and only if $c_{n,0}$ is continuous. If we want to recover all of the $u_n[k]$ from the function $y = \mathcal{G}u$, it is enough to ensure that $\phi^{n,0} \in R((\mathcal{G}|_{V_n})^*)$.

Before considering conditions which ensure the weak invertibility of $\mathcal{G}|_{V_n}$, we examine some of the immediate consequences of

$$\phi^{n,0} \in R((\mathcal{G}|_{V_n})^*), \tag{5.8}$$

which will be of use in the work to come. Condition 5.8 implies that there exists some $\tilde{\xi}_n \in \overline{\mathcal{G}(V_n)}$ such that

$$\phi^{n,0} = P_n\mathcal{G}^*\tilde{\xi}_n \tag{5.9}$$

and since, for any $u \in V_n$,

$$u_n[0] = \langle u, \phi^{n,0}\rangle = \langle \mathcal{G}u, \tilde{\xi}_n\rangle,$$

we conclude that

$$c_{n,k}(y) = \langle y(t), \tilde{\xi}_n(t - k/2^n)\rangle. \tag{5.10}$$

Also, if 5.8 holds, then the $c_{n,k}$ are continuous and the Riesz representation theorem ensures the existence of a unique function $\nu_n \in \overline{\mathcal{G}(V_n)}$ such that

$$c_{n,0}(y) = \langle y, \nu_n\rangle.$$

Accordingly, the function $\tilde{\xi}_n = \nu_n$, defined by 5.9 is unique.

The sequences

$$X_n = \{\xi_n(t - k/2^n) : k \in \mathbb{Z}\}$$

and

$$\tilde{X}_n = \left\{\tilde{\xi}_n(t - k/2^n) : k \in \mathbb{Z}\right\}$$

are biorthogonal. This follows easily from the orthonormality of $\left\{\phi^{n,k} : k \in \mathbb{Z}\right\}$, which implies that

$$\begin{aligned} \delta_{j,k} &= \langle \phi^{n,j}(t), \phi^{n,k}(t)\rangle \\ &= \langle \phi^{n,j}(t), \left(P_n\mathcal{G}^*\tilde{\xi}_n(\cdot - k/2^n)\right)(t)\rangle \\ &= \langle \xi_n(t - j/2^n), \tilde{\xi}_n(t - k/2^n)\rangle. \end{aligned} \tag{5.11}$$

The set, $\tilde{X}_n$ is a Riesz-Fischer sequence with

$$\left\|\sum_k \alpha_k \tilde{\xi}_n(t - k/2^n)\right\|^2 \geq \|\mathcal{G}\|^{-2} \sum_k |\alpha_k|^2, \tag{5.12}$$

which follows from 5.9 and the Parseval's relation

$$\sum_k |\alpha_k|^2 = \left\|\sum_k \alpha_k \phi^{n,k}\right\|.$$

Indeed, since

$$\begin{aligned}\left\|\sum_k \alpha_k \phi^{n,k}\right\|^2 &= \left\|\sum_k \alpha_k P_n \mathcal{G}^* \tilde{\xi}_n(t-k/2^n)\right\|^2 \\ &\leq \|\mathcal{G}^*\|^2 \left\|\sum_k \tilde{\xi}_n(t-k/2^n)\right\|^2\end{aligned}$$

and $\|\mathcal{G}\| = \|\mathcal{G}^*\|$, we see that 5.12 holds. Furthermore, the sequence $\tilde{X}_n$ is complete in $\overline{\mathcal{G}(V_n)}$. To see this, we suppose that $\nu \in \overline{\mathcal{G}(V_n)}$ is such that

$$\langle \nu(t), \tilde{\xi}_n(t-k/2^n)\rangle = 0, \tag{5.13}$$

for all $k \in \mathbb{Z}$. We want to show that 5.13 implies $\nu = 0$. We know that there exists $\nu^j \in \mathcal{G}(V_n)$, $j \in \mathbb{N}$ such that

$$\lim_{j\to\infty} \|\nu - \nu^j\| = 0$$

and since strong convergence [2] implies weak convergence [3], 5.13 yields

$$\lim_{j\to\infty} \langle \nu^j(t), \tilde{\xi}_n(t-k/2^n)\rangle = \langle \nu(t), \tilde{\xi}_n(t-k/2^n)\rangle = 0,$$

for all $k \in \mathbb{Z}$. We can always find $\sigma^j \in V_n$ such that $\mathcal{G}\sigma^j = \nu^j$ and through the use of 5.9, we obtain

$$\langle \nu^j(t), \tilde{\xi}_n(t-k/2^n)\rangle = \langle \sigma^j(t), \phi^{n,k}(t)\rangle.$$

The set $\{\phi^{n,k} : k \in \mathbb{Z}\}$ is complete in V_n and accordingly, the sequence $\{\sigma^j : j \in \mathbb{N}\}$ converges weakly to zero. The operator $\mathcal{G}$ is continuous and hence $\{\nu^j : j \in \mathbb{N}\}$ also converges weakly to zero. However, we already know that $\{\nu^j : j \in \mathbb{N}\}$ converges strongly to ν and since the strong and weak limits must agree, we conclude that $\nu = 0$ and therefore $\tilde{X}_n$ is complete.

---

[2]Convergence in the norm.

[3]Convergence with respect to all continuous linear functionals.

We mention that X_n is a Bessel sequence with

$$\left\| \sum_k \alpha_k \xi(t - k/2^n) \right\|^2 \leq \|\mathcal{G}\|^2 \sum_k |\alpha_k|^2 \tag{5.14}$$

and a proof similar to that above shows that X_n is also complete in $\overline{\mathcal{G}(V_n)}$.

We return our attention to the weak invertibility of the operator $\mathcal{G}|_{V_n}$. We provide a characterization of this property which makes use of the *Gram matrix* of the sequence X_n. The Gram matrix of X_n, given by

$$\mathbf{G}_n = [\langle \xi_n(t - k/2^n), \xi_n(t - j/2^n) \rangle], \tag{5.15}$$

is a symmetric *Toeplitz* matrix and, in light of 5.14, is a bounded linear operator mapping $\ell^2(\mathbb{Z})$ into $\ell^2(\mathbb{Z})$. An explicit expression for the entries of 5.15 can be obtained by considering the Fourier coefficients of an appropriate 2π-periodic function. Moreover, this function can be defined in terms of the functions $\hat{\phi}$ and $\hat{g}$.

To find this function, we note that

$$\begin{aligned}
\langle \xi_n(t), \xi_n(t - k/2^n) \rangle &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik\omega/2^n} \left| \hat{\xi}_n(\omega) \right|^2 d\omega \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik\omega} \left| \hat{g}(2^n\omega)\hat{\phi}(\omega) \right|^2 d\omega \\
&= \frac{1}{2\pi} \lim_{L\to\infty} \int_{-\pi}^{\pi} e^{ik\omega} \sum_{|l|\leq L} \left| \hat{g}(2^n[\omega + 2\pi l])\hat{\phi}(\omega + 2\pi l) \right|^2 d\omega.
\end{aligned}$$

Also note that

$$\begin{aligned}
\lim_{L\to\infty} \sum_{|l|\leq L} & \left| \hat{g}[2^n(\omega + 2\pi l)]\hat{\phi}(\omega + 2\pi l) \right|^2 \\
&\leq \sup_{\omega\in\mathbb{R}} |\hat{g}(\omega)|^2 \sum_l \left| \hat{\phi}(\omega + 2\pi l) \right|^2 = \|\mathcal{G}\|^2.
\end{aligned}$$

Hence, if we use the notation

$$\hat{G}_n(\omega) = \sum_l \left| \hat{g}(2^n[\omega + 2\pi l])\hat{\phi}(\omega + 2\pi l) \right|^2, \tag{5.16}$$

then the Lebesgue Dominated Convergence Theorem (see, for example [38]) allows us to write

$$g_n[k] = \langle \xi_n(t), \xi_n(t - k/2^n) \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik\omega} \hat{G}_n(\omega)\, d\omega. \tag{5.17}$$

That is, the entries of 5.15 are in fact the Fourier coefficients of the function 5.16 and in particular,

$$\mathbf{G}_n = [g_n[j-k]].$$

We are now in a position to prove the following theorem, which provides a characterization of the weak invertibility of $\mathcal{G}|_{V_n}$ in terms of $\hat{G}_n$.

Theorem 5.1 *Let $\hat{G}_n$ be defined as in 5.16. If there exists a sequence $\{A^j : j \in \mathbb{N}\} \subset L^2[-\pi, \pi]$ and a constant M such that*

$$\lim_{j\to\infty} \left\| A^j \hat{G}_n - 1 \right\|^2 = \lim_{j\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| A^j(\omega) \hat{G}_n(\omega) - 1 \right|^2 d\omega = 0 \tag{5.18}$$

and $\|A^j\| \leq M$, then $\phi^{n,0} \in R\left((\mathcal{G}|_{V_n})^*\right)$.

Proof. Let

$$A^j(\omega) = \sum_k \alpha_k^j e^{-ik\omega}$$

and define the functions $\nu^j \in \mathcal{G}(V_n)$ by

$$\nu^j(t) = \sum_k \alpha_k^j \xi_n(t - k/2^n).$$

If we now let

$$\beta_p^j = \langle \nu^j(t), \xi_n(t - k/2^n) \rangle = \sum_k \alpha_k^j g_n[p-k],$$

then, by 5.18

$$\lim_{j\to\infty} \sum_p \left| \beta_p^j - \delta_{p,0} \right|^2 = 0. \tag{5.19}$$

It now follows that, since $\{\xi_n(t-k/2^n) : k \in \mathbb{Z}\}$ is complete in $\mathcal{G}(V_n)$ and $\|A^j\| \leq M$, the sequence $\{\nu^j : j \in \mathbb{N}\}$ converges weakly to some $\nu \in \overline{\mathcal{G}(V_n)}$.

Since $P_n\mathcal{G}^*$ is continuous, $P_n\mathcal{G}^*\nu^j$ converges weakly to $P_n\mathcal{G}^*\nu$. Now

$$\left(P_n\mathcal{G}^*\nu^j\right)(t) = \sum_k \beta_k^j \phi^{n,k}(t)$$

and, in view of 5.19, we have

$$\lim_{j\to\infty} \left\|P_n\mathcal{G}^*\nu^j - \phi^{n,0}\right\| = 0.$$

Since the strong and weak limits must coincide, we conclude that

$$P_n\mathcal{G}^*\nu = \phi^{n,0}$$

and therefore 5.8 holds. ∎

Observe that if $1/\hat{G}_n \in L^2[-\pi,\pi]$, then $\mathcal{G}|_{V_n}$ will be weakly invertible. We also observe that if $\mathcal{G}$ is weakly invertible, then so is $\mathcal{G}|_{V_n}$ for any n. However, the converse does not hold. For example, suppose that the Fourier transform of the convolution kernel is

$$\hat{g}(\omega) = \begin{cases} \left|\omega(\omega^2-\pi^2)\right|^{1/4}, & -\pi < \omega < \pi \\ 0, & \text{otherwise} \end{cases}$$

and let ϕ be the Shannon scaling function. That is

$$\hat{\phi}(\omega) = \begin{cases} 1, & -\pi < \omega < \pi \\ 0, & \text{otherwise} \end{cases}.$$

The operator $\mathcal{G}$ is not injective and hence, is not weakly invertible. On the other hand, for $\omega \in (-\pi,\pi)$,

$$\hat{G}_0(\omega) = \hat{g}(\omega)$$

and since $1/\hat{G}_0 \in L^2[-\pi,\pi]$, the operator $\mathcal{G}|_{V_0}$ is weakly invertible.

Finally, we note that even if the functionals $c_{n,k}$ are continuous, the inverse of the operator $\mathcal{G}|_{V_n}$ need not be continuous. Since V_n and $\ell^2(\mathbb{Z})$ are isometric, the last statement is equivalent to saying that

$$\sum_k |c_{n,k}(y)|^2 < \infty, \tag{5.20}$$

need not hold for all $y \in \overline{\mathcal{G}(V_n)}$. We recall that $\{\nu_k : k \in \mathbb{Z}\}$ is said to be a Bessel sequence as long as

$$\left\| \sum_k \alpha_k \nu_k \right\|^2 \leq B \sum_k |\alpha_k|^2 . \tag{5.21}$$

According to [47, page 154], 5.21 is equivalent to

$$\sum_k |\langle \sigma, \nu_k \rangle|^2 \leq B \|\sigma\|^2 \tag{5.22}$$

for any σ. In view of 5.22 and 5.10, we see that 5.20 will hold if and only if $\tilde{X}_n$ is a Bessel sequence.

5.4 Strong invertibility

Suppose that $\mathcal{G}|_{V_n}$ is weakly invertible. As we have seen in the previous section, this means that $\mathcal{G}|_{V_n}$ is a bijection onto its range. However, $(\mathcal{G}|_{V_n})^{-1}$ need not be continuous. Our present objective is to derive verifiable conditions which ensure that $(\mathcal{G}|_{V_n})^{-1}$ is continuous. In particular, we are interested in finding the properties of the functions g and ϕ which will ensure that the inverse problem defined by 5.5 is well-posed.

Definition 5.2 *If $(\mathcal{G}|_{V_n})^{-1}$ is a continuous linear operator mapping $\overline{\mathcal{G}(V_n)}$ onto V_n, then we say that $\mathcal{G}|_{V_n}$ is strongly invertible.*

Recall that under the assumption $u \in V_n$, equation 2.4 can be written as

$$\mathcal{G}|_{V_n} u = y,$$

or equivalently, as

$$\sum_k u_n[k]\xi_n(t - k/2^n) = y(t).$$

Since

$$\left\|\sum_k a_k \phi^{n,k}\right\|^2 = \sum_k |a_k|^2,$$

the subspace V_n and $\ell^2(\mathbb{Z})$ are isometric and therefore $\mathcal{G}|_{V_n}$ will have a continuous inverse if and only if the linear operator defined by

$$y(t) \mapsto \{u_n[k] : k \in \mathbb{Z}\} \tag{5.23}$$

is continuous. In turn, the linear operator defined by 5.23 will be continuous if and only if there exists a constant C such that

$$\sum_k |u_n[k]|^2 = \sum_k |c_{n,k}(y)|^2 \le C\|y\|^2$$

and since

$$c_{n,k}(y) = \langle y(t), \tilde{\xi}_n(t - k/2^n)\rangle,$$

we see that $(\mathcal{G}|_{V_n})^{-1}$ will be continuous if and only if

$$\sum_k \left|\langle y(t), \tilde{\xi}_n(t - k/2^n)\rangle\right|^2 \le C\|y\|^2 \tag{5.24}$$

for any $y \in \overline{\mathcal{G}(V_n)}$. Since 5.24 is equivalent to

$$\left\|\sum_k a_k \tilde{\xi}_n(t - k/2^n)\right\|^2 \le C \sum_k |a_k|^2, \tag{5.25}$$

the operator $\mathcal{G}|_{V_n}$ will be strongly invertible if and only if the sequence of functions

$$\tilde{X}_n = \left\{\tilde{\xi}_n(t - k/2^n) : k \in \mathbb{Z}\right\}$$

forms a Bessel sequence of $\overline{\mathcal{G}(V_n)}$. We have already pointed out that whenever the sequence $\tilde{X}_n$ exists, it must be a Riesz-Fischer sequence. Hence, if $\tilde{X}_n$ is a Bessel sequence, then it must be a Riesz basis of $\overline{\mathcal{G}(V_n)}$.

Inequality 5.24, or equivalently 5.25, can be difficult to establish when one works with the sequence $\tilde{X}_n$. However, it can be shown that $\tilde{X}_n$ is a Bessel sequence if and only if the set

$$X_n = \{\xi_n(t - k/2^n) : k \in \mathbb{Z}\}$$

is a Riesz-Fischer sequence so that we may work with the sequence X instead. Since the sequence X_n is defined explicitly by

$$\xi_n(t) = \left(\mathcal{G}\phi^{n,0}\right)(t),$$

this approach turns out to be easier.

A proof of the aforementioned equivalence makes use of the relevant definitions and the biorthogonality of the sequences X_n and $\tilde{X}_n$. Assume that $\tilde{X}_n$ forms a Bessel sequence of $\overline{\mathcal{G}(V_n)}$ and let

$$y(t) = \sum_k b_k \xi_n(t - k/2^n).$$

The sequence X_n is a Bessel sequence and hence the function y is a well defined element of $\overline{\mathcal{G}(V_n)}$ for any $\{b_k : k \in \mathbb{Z}\} \in \ell^2(\mathbb{Z})$. The biorthogonality of X_n and $\tilde{X}_n$ yields

$$b_k = \langle y(t), \tilde{\xi}_n(t - k/2^n)\rangle$$

and through the use of 5.24 we obtain

$$\sum_k |b_k|^2 \leq C \left\| \sum_k b_k \xi_n(t - k/2^n) \right\|^2,$$

which implies that X_n is a Riesz-Fischer sequence.

Let us now suppose that X_n is a Riesz-Fischer sequence with

$$\left\|\sum_k a_k \xi_n(t-k/2^n)\right\|^2 \geq C \sum_k |a_k|^2.$$

We use Parseval's relation for the orthonormal basis $\{\phi^{n,k} : k \in \mathbb{Z}\}$ to obtain the inequality

$$\left\|\sum_k a_k \xi_n(t-k/2^n)\right\|^2 \geq C \left\|\sum_k a_k \phi^{n,k}(t)\right\|^2. \tag{5.26}$$

Recall that X_n is complete in $\overline{\mathcal{G}(V_n)}$ and therefore 5.26 implies that there exists a continuous linear operator $\mathcal{T} : \overline{\mathcal{G}(V_n)} \to V_n$ such that

$$\phi^{n,k}(t) = (\mathcal{T}\xi_n(\cdot - k/2^n))(t).$$

The orthonormality of $\{\phi^{n,k} : k \in \mathbb{Z}\}$ is used once again to obtain

$$\delta_{j,k} = \langle \mathcal{T}^*\mathcal{T}\xi_n(t-j/2^n), \xi_n(t-k/2^n)\rangle$$

and, since $\tilde{X}_n$ is the unique sequence in $\overline{\mathcal{G}(V_n)}$ biorthogonal to X, we conclude that

$$\begin{aligned}\tilde{\xi}_n(t-k/2^n) &= (\mathcal{T}^*\mathcal{T}\xi_n(\cdot - k/2^n))(t) \\ &= \left(\mathcal{T}^*\mathcal{T}\mathcal{G}\phi^{n,k}\right)(t).\end{aligned}$$

Finally, the operator $\mathcal{T}^*\mathcal{T} : \overline{\mathcal{G}(V_n)} \to \overline{\mathcal{G}(V_n)}$ is continuous and therefore $\tilde{X}_n$ is a Bessel sequence with

$$\left\|\sum_k b_k \tilde{\xi}_n(t-k/2^n)\right\|^2 \leq \|\mathcal{T}\|^4 \|\mathcal{G}\|^2 \sum_k |b_k|^2.$$

We now focus on establishing conditions which ensure that X_n is a Riesz-Fischer sequence of $\overline{\mathcal{G}(V_n)}$. In view of the fact that X_n is a Bessel sequence, such conditions will also guarantee that X_n is a Riesz basis.

The sequence X_n is said to be a Riesz basis of $\overline{\mathcal{G}(V_n)}$ if it is complete and if there exist positive constants A and B such that

$$A \sum_k |d_k|^2 \leq \left\| \sum_k d_k \xi_n(t - k/2^n) \right\|^2 \leq B \sum_k |d_k|^2 . \tag{5.27}$$

Inequality 5.14 ensures that a suitable choice for a *upper Riesz bound* B satisfies

$$B \leq \|\mathcal{G}\|^2 .$$

It is the existence of a *lower Riesz bound* A that is, in general, difficult to verify. However, since the sequence X_n is generated by translating a single function ξ_n, we can use Parseval's relation for the Fourier transform to simplify the estimation of A.

We have

$$\begin{aligned} \left\| \sum_k d_k \xi_n(t - k/2^n) \right\|^2 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \sum_k d_k e^{-ik\omega/2^n} \hat{\xi}_n(\omega) \right|^2 d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \sum_k d_k e^{-ik\omega} \right|^2 \left| \hat{g}(2^n\omega)\hat{\phi}(\omega) \right|^2 d\omega \\ &= \frac{1}{2\pi} \lim_{L\to\infty} \sum_{|l|\leq L} \int_{(2l-1)\pi}^{(2l+1)\pi} \left| \sum_k d_k e^{-ik\omega} \right|^2 \left| \hat{g}(2^n\omega)\hat{\phi}(\omega) \right|^2 d\omega. \end{aligned}$$

whereupon the use of the notation

$$D(\omega) = \sum_k d_k e^{-ik\omega}$$

and the change of variable $\omega = \nu + 2\pi l$ yields the equation

$$\left\| \sum_k d_k \xi_n(t - k/2^n) \right\|^2 = \frac{1}{2\pi} \lim_{L\to\infty} \int_{-\pi}^{\pi} |D(\omega)|^2 \sum_{|l|\leq L} \left| \hat{g}(2^n[\omega + 2\pi l])\hat{\phi}(\omega + 2\pi l) \right|^2 d\omega.$$

We have already seen that

$$\sum_{|l|\leq L} \left|\hat{g}(2^n[\omega + 2\pi l])\hat{\phi}(\omega + 2\pi l)\right|^2 \leq \|\mathcal{G}\|^2,$$

for almost all ω and hence the Lebesgue Dominated Convergence Theorem implies that

$$\left\|\sum_k d_k \xi_n(t - k/2^n)\right\|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |D(\omega)|^2 \hat{G}_n(\omega)\, d\omega, \tag{5.28}$$

where $\hat{G}_n(\omega)$ is as in 5.16. namely,

$$\hat{G}_n(\omega) = \sum_l \left|\hat{g}(2^n[\omega + 2\pi l])\hat{\phi}(\omega + 2\pi l)\right|^2.$$

In view of 5.28. we see that 5.27 and

$$A \leq \int_{-\pi}^{\pi} \tilde{D}(\omega)\hat{G}_n(\omega)\, d\omega \leq B, \tag{5.29}$$

with

$$\tilde{D}(\omega) = \frac{|D(\omega)|^2}{2\pi \|D\|^2}, \tag{5.30}$$

are equivalent.

Suppose that $\hat{G}_n$ satisfies the inequalities

$$A \leq \hat{G}_n(\omega) \leq B, \tag{5.31}$$

for almost all ω, then since the function 5.30 is unimodular, if 5.31 holds, then 5.29 (and hence 5.27) must hold. In other words, if we can show that there exist positive constants A and B such that the inequalities 5.31 hold for almost all ω, then X_n must be a Riesz basis for its closed linear span.

For our present purposes, what we have just shown is enough. However, it is interesting to note that in most practical cases 5.31 and 5.29 are, in fact, equivalent.

For instance if $\hat{G}_n$ is continuous and we select $\tilde{D}$ from some appropriate δ-sequence. then it can be shown that 5.29 implies 5.31.

The assumption that $\hat{G}_n$ is continuous is not overly restrictive. For example, we know that g, $\phi \in L^1(\mathbb{R})$ and therefore, the functions $\hat{g}$ and $\hat{\phi}$ will be continuous. Furthermore, since

$$\hat{G}_n(\omega) \leq \|\mathcal{G}\|^2 \sum_l \left|\hat{\phi}(\omega + 2\pi l)\right|^2 = \|\mathcal{G}\|^2, \tag{5.32}$$

we see that if the series

$$\sum_l \left|\hat{\phi}(\omega + 2\pi l)\right|^2$$

converges uniformly, then the series 5.16 converges uniformly and $\hat{G}_n$ will be continuous.

It is not difficult to find conditions sufficient to ensure the uniform convergence of $\sum_l \left|\hat{\phi}(\omega + 2\pi l)\right|^2$. Henceforth we will assume that the scaling functions we use satisfy

$$\left|\hat{\phi}(\omega)\right| \leq \frac{C}{1 + |\omega|^\alpha}, \tag{5.33}$$

for some $\alpha > 1/2$. It follows that

$$\sum_l \left|\hat{\phi}(\omega + 2\pi l)\right|^2 \leq C^2 \sum_l \frac{1}{1 + |\omega + 2\pi l|^{2\alpha}}.$$

and since we can restrict our attention to $\omega \in [-\pi, \pi]$, we find that

$$\begin{aligned} \sum_l \left|\hat{\phi}(\omega + 2\pi l)\right|^2 &\leq C^2 \left(1 + \sum_{l=1}^{\infty} \frac{1}{1 + ([2l-1]\pi)^{2\alpha}} + \sum_{l=-1}^{-\infty} \frac{1}{1 + ([2l+1]\pi)^{2\alpha}}\right) \\ &= C^2 \left(1 + 2\sum_{l=1}^{\infty} \frac{1}{1 + ([2l-1]\pi)^{2\alpha}}\right). \end{aligned}$$

In light of the inequality,

$$\sum_{l=1}^{\infty} \frac{1}{1 + ([2l-1]\pi)^{2\alpha}} \leq \frac{1}{\pi^{2\alpha}} \sum_{l=1}^{\infty} \frac{1}{l^{2\alpha}} < \infty,$$

we conclude that $\sum_l |\phi(\omega + 2\pi l)|^2$ converges uniformly on compact subsets of $\mathbb{R}$.

Many orthonormal scaling functions will satisfy the inequality 5.33 for some $\alpha > 1/2$. For example, if ϕ is a Meyer scaling function, then $\hat{\phi}$ has compact support and 5.33 is satisfied for any choice of α.

If ϕ is a Daubechies scaling function, then ϕ is compactly supported in some interval $I \subset \mathbb{R}$. We can choose ϕ from the Daubechies family so that it is N-times continuously differentiable. It follows that integration by parts leads to the inequality

$$\begin{aligned} (1 + |\omega|^N) \left|\hat{\phi}(\omega)\right| &= \left|\int_I e^{-i\omega t}\phi^{(N)}(t)\,dt\right| + \left|\int_I e^{-i\omega t}\phi(t)\,dt\right| \\ &\leq C. \end{aligned}$$

When ϕ belongs to the Battle-Lemarié family, we have

$$\hat{\phi}(\omega) = M(\omega)\left(\frac{\sin(\omega/2)}{\omega/2}\right)^N,$$

where M is a continuous 2π-periodic function and N is a positive integer. The inequality 5.33 is satisfied with $\alpha = N$.

When $\hat{G}_n$ is continuous, we can define the so-called *optimal Riesz bounds* for the sequence X_n. In particular, we let

$$A_n = \min_{\omega\in[-\pi,\pi]} \hat{G}_n(\omega) \tag{5.34}$$

and

$$B_n = \max_{\omega\in[-\pi,\pi]} \hat{G}_n(\omega). \tag{5.35}$$

We reiterate that the upper bound B_n will always exist and is bounded above by $\|\mathcal{G}\|^2$. Furthermore, it is the existence of the lower bound A_n which can be difficult to establish. With regard to the inverse problem at hand, it is the lower bound which is most important. In fact, we have shown that $(\mathcal{G}|_{V_n})^{-1}$ will be continuous if and only if $A_n > 0$. By borrowing ideas used by Cohen (see [14, page 182]) in the study of orthonormal wavelet bases, we can find necessary and sufficient conditions for the existence of a positive A_n. We will need the following definition:

Definition 5.3 *A compact set K is said to be congruent to $[-\pi, \pi]$ modulo 2π (we will use the notation $K \equiv [-\pi, \pi] \bmod 2\pi$) if:*

1. *$|K| = 2\pi$ and*
2. *for every $\omega \in [-\pi, \pi]$, there is an integer $l \in \mathbb{Z}$ such that $\omega + 2\pi l \in K$.*

We now state and prove the following theorem, which gives necessary and sufficient conditions for the existence of A_n in terms of the functions $\hat{g}$ and $\hat{\phi}$.

Theorem 5.2 *Suppose that $\hat{\phi}$ satisfies 5.33, then there exists an $A > 0$ such that*

$$\hat{G}_n(\omega) \geq A \tag{5.36}$$

for all ω if and only if there exists some constant $C > 0$ and a compact set $K \equiv [-\pi, \pi] \bmod 2\pi$ such that

$$\left|\hat{g}(2^n\omega)\hat{\phi}(\omega)\right| \geq C \tag{5.37}$$

for all $\omega \in K$.

Proof. Suppose that 5.37 holds for all $\omega \in K \equiv [-\pi, \pi] \bmod 2\pi$. We want to show that this assumption implies 5.36 holds for some $A > 0$. To do this, we assume the contrary. That is, there exists at least one $\omega^* \in [-\pi, \pi]$ such that the continuous function $\hat{G}_n$ satisfies

$$\hat{G}_n(\omega^*) = 0,$$

which implies

$$\left|\hat{g}(2^n[\omega^* + 2\pi l])\hat{\phi}(\omega^* + 2\pi l)\right| = 0$$

for all $l \in \mathbb{Z}$.

By definition, for any $\omega \in [-\pi, \pi]$ there exists $m \in \mathbb{Z}$ such that $\omega + 2\pi m \in K$. It follows that we can choose m so that $\omega^* + 2\pi m \in K$, which contradicts the original assumption.

We now assume that 5.36 holds and show that 5.37 must hold. Suppose that $0 < \epsilon < A/2$. Since

$$\sum_l \left|\hat{g}(2^n[\omega + 2\pi l])\hat{\phi}(\omega + 2\pi l)\right|^2$$

converges uniformly for $\omega \in [-\pi, \pi]$, there exists a positive integer L such that

$$\left|\hat{G}_n(\omega) - \sum_{|l| \le L} \left|\hat{g}(2^n[\omega + 2\pi l])\hat{\phi}(\omega + 2\pi l)\right|^2\right| \le \epsilon$$

and hence

$$\sum_{|l| \le L} \left|\hat{g}(2^n[\omega + 2\pi l])\hat{\phi}(\omega + 2\pi l)\right|^2 \ge \hat{G}_n(\omega) - \epsilon \ge A/2 \tag{5.38}$$

for all $\omega \in [-\pi, \pi]$.

Now, in view of inequality 5.38, for each $\omega \in [-\pi, \pi]$, there exists l_ω such that $|l_\omega| \le L$ and

$$\left|\hat{g}(2^n[\omega + 2\pi l_\omega])\hat{\phi}(\omega + 2\pi l_\omega)\right|^2 \ge \frac{A}{2(2L+1)}. \tag{5.39}$$

Since $\hat{g}$ and $\hat{\phi}$ are continuous, there exists a neighborhood M_ω of ω such that

$$\left|\hat{g}(2^n[\omega + 2\pi l_\omega])\hat{\phi}(\omega + 2\pi l_\omega)\right|^2 \ge \frac{A}{4(2L+1)}$$

for all $\omega \in M_\omega$. If we define the sets R_l by

$$R_l = \left\{\omega \in [-\pi, \pi] : \left|\hat{g}(2^n[\omega + 2\pi l])\hat{\phi}(\omega + 2\pi l)\right|^2 \ge \frac{A}{4(2L+1)}\right\},$$

for $l = -L, -L+1, \ldots L$, then, each R_l is a subset of $[-\pi, \pi]$ such that

$$M_\omega \subset R_{l_\omega}.$$

Since $\{M_\omega : \omega \in [-\pi, \pi]\}$ is a cover of $[-\pi, \pi]$, $\{R_l : l = -L, \ldots, L\}$ is a sequence of subsets of $[-\pi, \pi]$ such that

$$[-\pi, \pi] = \bigcup_{l=-L}^{L} R_l.$$

If we now define the sets S_l by,

$$S_{-L} = R_{-L}$$

and

$$S_l = R_l \setminus \bigcup_{p=-L}^{l-1} S_p,$$

then the S_l, $l = -L, -L+1, \ldots L$ form a sequence of disjoint subsets of $[-\pi, \pi]$ satisfying

$$[-\pi, \pi] = \bigcup_{l=-L}^{L} \overline{S_l}.$$

Let us use the notation $S_l + 2\pi l$ to signify the sets

$$S_l + 2\pi l = \{\omega : \omega - 2\pi l \in S_l\}$$

and subsequently define the compact set K by

$$K = \bigcup_{l=-L}^{L} \overline{S_l + 2\pi l}.$$

Since the sets S_l are disjoint,

$$|K| = \left| \bigcup_{l=-L}^{L} \overline{S_l + 2\pi l} \right| = \sum_{l=-L}^{L} |S_l| = 2\pi$$

hence $K \equiv [-\pi, \pi] \bmod 2\pi$ and

$$\left|\hat{g}(2^n\omega)\hat{\phi}(\omega)\right|^2 \geq C^2 = \frac{A}{4(2L+1)}$$

for all $\omega \in K$. ∎

Suppose that the hypothesis of 5.2 is satisfied, then X_n is a Riesz basis of the subspace $\overline{\mathcal{G}(V_n)}$ satisfying the inequalities 5.27 with $A = A_n$ and $B = B_n$. According to Young [47, page 32], there exists a Riesz basis $\{\nu_k : k \in \mathbb{Z}\}$ of $\overline{\mathcal{G}(V_n)}$, biorthogonal to X_n and satisfying the inequalities

$$\frac{1}{B_n}\sum_k |a_k|^2 \leq \left\|\sum_k a_k\nu_k(t)\right\|^2 \leq \frac{1}{A_n}\sum_k |a_k|^2. \tag{5.40}$$

The Riesz basis $\{\nu_k : k \in \mathbb{Z}\}$ is known as the *dual basis* and, since $\tilde{X}_n$ is the unique sequence in $\overline{\mathcal{G}(V_n)}$ biorthogonal to X_n, we must have

$$\nu_k(t) = \tilde{\xi}_n(t - k/2^n).$$

It is not too difficult to show that the Fourier transform of $\tilde{\xi}_n$ is given explicitly by

$$\hat{\tilde{\xi}}_n(\omega) = \frac{\hat{\xi}_n(\omega)}{\hat{G}_n(\omega/2^n)}, \tag{5.41}$$

which implies that

$$\tilde{\xi}_n(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t}\frac{\hat{\xi}_n(\omega)}{\hat{G}_n(\omega/2^n)}\,d\omega. \tag{5.42}$$

To justify the representation 5.42, we simply need to show that the sequence $\tilde{X}_n$, generated by 5.42, is biorthogonal to X_n and is contained in the subspace $\overline{\mathcal{G}(V_n)}$.

If $\tilde{\xi}_n$ is given by 5.42, then

$$\langle\tilde{\xi}_n(t - k/2^n), \xi_n(t - k/2^n)\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ip\omega/2^n}\frac{\left|\hat{\xi}_n(\omega)\right|^2}{\hat{G}_n(\omega/2^n)}\,d\omega,$$

where $p = k - j$. We use the equation

$$\hat{\xi}_n(\omega) = 2^{-n/2}\hat{g}(\omega)\hat{\phi}(\omega/2^n)$$

to obtain

$$\begin{aligned}\langle\tilde{\xi}_n(t-k/2^n), \xi_n(t-k/2^n)\rangle &= \frac{2^{-n}}{2\pi}\int_{-\infty}^{\infty} e^{ip\omega/2^n}\frac{\left|\hat{g}(\omega)\hat{\phi}(\omega/2^n)\right|^2}{\hat{G}_n(\omega/2^n)}\,d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ip\omega}\frac{\left|\hat{g}(2^n\omega)\hat{\phi}(\omega)\right|^2}{\hat{G}_n(\omega)}\,d\omega \\ &= \frac{1}{2\pi}\lim_{L\to\infty}\int_{-\pi}^{\pi} e^{ip\omega}\frac{\sum_{|l|\le L}\left|\hat{g}(2^n[\omega+2\pi l])\hat{\phi}(\omega+2\pi l)\right|^2}{\hat{G}_n(\omega)}\,d\omega.\end{aligned}$$

whereupon the Lebesgue Dominated Convergence theorem implies that

$$\langle\tilde{\xi}_n(t-k/2^n), \xi_n(t-k/2^n)\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ip\omega}\,d\omega = \delta_{p,0},$$

as required.

We now show that the function $\tilde{\xi}_n$, defined by 5.42, belongs to the subspace $\overline{\mathcal{G}(V_n)}$. If we define the sequence $\{\tilde{g}_n[k] : k \in \mathbb{Z}\}$ by

$$\tilde{g}_n[k] = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{e^{ik\omega}}{\hat{G}_n(\omega)}\,d\omega \tag{5.43}$$

then, since $\hat{G}_n$ is continuous with

$$\hat{G}_n(\omega) \ge A_n > 0,$$

we have

$$\sum_k |\tilde{g}_n[k]|^2 < \infty.$$

Equation 5.41 can be used to obtain

$$\hat{\tilde{\xi}}_n(\omega) = \sum_k \tilde{g}_n[k]e^{-ik\omega/2^n}\hat{\xi}_n(\omega)$$

and hence, the inverse Fourier transform yields

$$\tilde{\xi}_n(t) = \sum_k \tilde{g}_n[k]\xi_n(t - k/2^n). \tag{5.44}$$

Since X_n is a Bessel sequence, we see that 5.44 is a well defined function in $\overline{\mathcal{G}(V_n)}$. The subspace $\overline{\mathcal{G}(V_n)}$ is closed under translations by integer multiples of 2^{-n} and therefore the result follows.

Every function $f \in \overline{\mathcal{G}(V_n)}$ can be expanded in terms of either X_n or $\tilde{X}_n$. The biorthogonality of X_n and $\tilde{X}_n$ implies that these expansions are of the form

$$\begin{aligned} f(t) &= \sum_k \langle f(t), \xi_n(t - k/2^n)\rangle \tilde{\xi}_n(t - k/2^n) \\ &= \sum_k \langle f(t), \tilde{\xi}_n(t - k/2^n)\rangle \xi_n(t - k/2^n). \end{aligned} \tag{5.45}$$

We note that, for a general $f \in L^2(\mathbb{R})$, the series in 5.45 represent the orthogonal projection of f onto $\overline{\mathcal{G}(V_n)}$. Now, when $y \in \overline{\mathcal{G}(V_n)}$, the solution of 5.5

$$\mathcal{G}|_{V_n} u = y$$

has the expansion

$$u(t) = \sum_k \langle y(t), \tilde{\xi}_n(t - k/2^n)\rangle \phi^{n,k}(t). \tag{5.46}$$

If we let $u_n[k] = \langle u, \phi^{n,k}\rangle$ and $y_n[k] = \langle y(t), \xi_n(t - k/2^n)\rangle$, then equations 5.44 and 5.46 can be combine to yield

$$u_n[k] = \sum_p \tilde{g}_n[k-p] y_n[p].$$

It follows that the coefficients $u_n[k]$ can be computed by convolving the sequence $\{y_n[k] : k \in \mathbb{Z}\}$ with the inverse filter $\{\tilde{g}_n[k] : k \in \mathbb{Z}\}$. Since $\hat{G}_n(\omega) \geq A_n > 0$, $\{\tilde{g}_n[k] : k \in \mathbb{Z}\} \in \ell^1(\mathbb{Z})$ whenever $\{g_n[k] : k \in \mathbb{Z}\}$. That is, the inverse filter will be stable whenever the forward filter is stable.

In light of 5.46, the operator $(\mathcal{G}|_{V_n})^{-1} : \overline{\mathcal{G}(V_n)} \to V_n$ is such that

$$\left\|(\mathcal{G}|_{V_n})^{-1} y\right\|^2 = \sum_k \left|\langle y(t), \tilde{\xi}_n(t - k/2^n)\rangle\right|^2. \tag{5.47}$$

If we use the equivalence defined by 5.24 and 5.25, then, definition 5.41, can be used to show that $\tilde{X}_n$ is a Bessel sequence with optimal bound A_n^{-1}. Therefore, equation 5.47 implies

$$\left\|(\mathcal{G}|_{V_n})^{-1} y\right\|^2 \leq \frac{\|y\|^2}{A_n} \tag{5.48}$$

and since A_n^{-1} is as small as is possible, we find that

$$\left\|(\mathcal{G}|_{V_n})^{-1}\right\| = \frac{1}{\sqrt{A_n}}. \tag{5.49}$$

Suppose that we have the corrupted observation $y + \delta y$, $\delta y \in L^2(\mathbb{R})$, of the true data y. We use a series of the form 5.46 to form the approximation u_δ of the exact solution u. The error in this approximation is

$$u_\delta(t) - u(t) = \sum_k \langle \delta y(t), \tilde{\xi}_n(t - k/2^n)\rangle \phi^{n,k}(t).$$

and hence

$$\|u_\delta - u\|^2 = \sum_k \left|\langle \delta y(t), \tilde{\xi}_n(t - k/2^n)\rangle\right|^2 \leq \frac{\|\delta y\|^2}{A_n},$$

from which we see that the number $A_n^{-1/2}$ gives an upper bound on the relative error in the approximation u_δ.

Theorem 5.2 can be used to provide an interesting contrast between the operators $\mathcal{G}$ and $\mathcal{G}|_{V_n}$ when the Fourier transform of the kernel g is positive. Many convolution kernels, such as the Gaussian

$$g(t) = e^{-t^2},$$

have Fourier transforms satisfying

$$|\hat{g}(\omega)| > 0, \tag{5.50}$$

for all $\omega \in \mathbb{R}$. When 5.50 holds. $\mathcal{G}$ is a bijection onto its range and $\mathcal{G}^{-1}$ exists. However. since $g \in L^1(\mathbb{R})$, $\mathcal{G}^{-1}$ cannot be continuous.

On the other hand,

$$\sum_l \left|\hat{\phi}(\omega + 2\pi l)\right|^2 = 1 > 0$$

and Theorem 5.2 imply the existence of some compact set $K \equiv [-\pi, \pi] \bmod 2\pi$ upon which

$$\left|\hat{\phi}(\omega)\right| \geq D > 0.$$

Since K is compact. inequality 5.50 ensures that there is a constant C such that

$$\left|\hat{g}(2^n\omega)\hat{\phi}(\omega)\right| \geq C > 0,$$

for all $\omega \in K$ and all $n \in \mathbb{Z}$. Therefore, if $\hat{g}$ satisfies 5.50. the operator $\mathcal{G}|_{V_n}$ is strongly invertible for all $n \in \mathbb{Z}$. However, as we shall see,

$$\lim_{n\to\infty} A_n = 0$$

and, as a consequence, the condition number

$$\begin{aligned} \kappa(\mathcal{G}|_{V_n}) &= \|\mathcal{G}|_{V_n}\| \left\|(\mathcal{G}|_{V_n})^{-1}\right\| \\ &= \sqrt{\frac{B_n}{A_n}} \end{aligned} \tag{5.51}$$

can be large for large n. Accordingly, the inverse problem defined by 5.5

$$\mathcal{G}|_{V_n} u = y$$

can be very ill-conditioned even when, technically, it is well-posed.

Finally, one might argue that 5.2 is of more theoretical than practical value. In fact. condition 5.37 could be difficult to verify for arbitrary functions $\hat{\phi}$ and $\hat{g}$. However, for many examples, ϕ is a low pass filter with a Fourier transform that is concentrated on the interval $[-\pi, \pi]$. This fact can be used to derive a simple sufficient condition for the existence of a positive A_n. For instance, suppose that ϕ is a Meyer scaling function, then $\hat{\phi}(\omega)$ is positive for all $\omega \in (-\pi - \epsilon, \pi + \epsilon)$ and $\text{supp}(\hat{\phi}) = [-\pi - \epsilon, \pi + \epsilon]$ for some $0 < \epsilon \leq \pi/3$. It follows that $\mathcal{G}|_{V_n}$ is strongly invertible if and only if $|\hat{g}(2^n\omega)| \geq D$ for some compact set $K \equiv [-\pi, \pi] \bmod 2\pi \subset (-\pi - \epsilon, \pi + \epsilon)$. Hence, a sufficient condition for the strong invertibility of $\mathcal{G}|_{V_n}$ is that $|\hat{g}(2^n\omega)|$ be bounded below on $[-\pi, \pi]$

The same condition is valid when ϕ is a Daubechies scaling function. In particular, it is shown in Appendix A, that if ϕ is a Daubechies scaling function, then $\left|\hat{\phi}(\omega)\right|$ must be bounded below for $\omega \in [-\pi, \pi]$. This means that $\left|\hat{g}(2^n\omega)\hat{\phi}(\omega)\right|$ will be bounded below on $[-\pi, \pi]$ if and only if $|\hat{g}(2^n\omega)|$ is bounded below on $[-\pi, \pi]$.

5.5 The function $\hat{G}_n(\omega)$

In this section, we begin with an investigation of the behavior of the function $\hat{G}_n$ for large $|n|$. In doing so, we will be in a position to prove that the function

$$u_n(t) = \sum_k \langle y(t), \tilde{\xi}_n(t - k/2^n)\rangle \phi^{n,k}(t) \tag{5.52}$$

converges to the solution u of 5.4 in the special case where $y \in R(\mathcal{G})$, $|\hat{g}(\omega)| > 0$, for $\omega \in \mathbb{R}$ and

$$|\hat{g}(\omega_2)| \leq |\hat{g}(\omega_1)|$$

for all $\omega_2 \geq \omega_1 \geq \Omega > 0$. This examination of $\hat{G}_n$ will also facilitate comments concerning the Riesz bounds A_n and B_n. In particular, we will be able to provide a justification of

$$\lim_{n \to \infty} \kappa(\mathcal{G}|_{V_n}) = \infty \tag{5.53}$$

and

$$\lim_{n\to-\infty} \kappa(\mathcal{G}|_{V_n}) = 1, \tag{5.54}$$

where $\kappa(\mathcal{G}|_{V_n})$ is the condition number, defined by 5.51.

Let us consider $\hat{G}_n$ for large positive n. In particular, we will show that

$$\lim_{n\to\infty} \hat{G}_n(\omega/2^n) = |\hat{g}(\omega)|^2 \tag{5.55}$$

uniformly on compact subsets of $\mathbb{R}$, whenever

$$\sum_l \left|\hat{\phi}(\omega + 2\pi l)\right|^2 = 1 \tag{5.56}$$

converges uniformly on compact subsets of $\mathbb{R}$.

First of all, we use 5.56 to write

$$\begin{aligned}\left|\hat{G}_n(\omega/2^n) - |\hat{g}(\omega)|^2\right| &= \left|\sum_l \left(\left|\hat{g}(\omega + 2^{n+1}\pi l)\right|^2 - |\hat{g}(\omega)|^2\right)\left|\hat{\phi}(\omega/2^n + 2\pi l)\right|^2\right| \\ &\leq \sum_l \left|\left|\hat{g}(\omega + 2^{n+1}\pi l)\right|^2 - |\hat{g}(\omega)|^2\right|\left|\hat{\phi}(\omega/2^n + 2\pi l)\right|^2 \end{aligned} \tag{5.57}$$

and note that, since

$$\left|\left|\hat{g}(\omega + 2^{n+1}\pi l)\right|^2 - |\hat{g}(\omega)|^2\right| \leq 2\|\mathcal{G}\|^2,$$

the last series in 5.57 converges uniformly.

Without loss of generality, we assume that $\omega \in I = [-\Omega, \Omega]$ and $n \in \mathbb{N}$. Since 5.56 converges uniformly, for any $\epsilon > 0$, we can choose $L \in \mathbb{N}$, independent of n, such that

$$\sum_{|l|>L} \left|\hat{\phi}(\omega + 2\pi l)\right|^2 \leq \frac{\epsilon}{2\|\mathcal{G}\|^2}$$

for all $\omega \in I$. It is easy to see that

$$\sum_{|l|>L} \left|\hat{\phi}(\omega/2^n + 2\pi l)\right|^2 \leq \frac{\epsilon}{2\|\mathcal{G}\|^2}$$

for all $\omega \in [-2^n\Omega, 2^n\Omega] \supset I$ and we can now use 5.57 to obtain the estimate

$$\left|\hat{G}_n(\omega/2^n) - |\hat{g}(\omega)|^2\right| \leq \sum_{|l|\leq L} \left||\hat{g}(\omega + 2^{n+1}\pi l)|^2 - |\hat{g}(\omega)|^2\right| \left|\hat{\phi}(\omega/2^n + 2\pi l)\right|^2 + \epsilon. \tag{5.58}$$

which holds for all $\omega \in I$

If we let

$$D_l^n(\omega) = \left||\hat{g}(\omega + 2^{n+1}\pi l)|^2 - |\hat{g}(\omega)|^2\right| \left|\hat{\phi}(\omega/2^n + 2\pi l)\right|^2,$$

then $D_0^n \equiv 0$. Suppose that $0 < |l| \leq L$. Since $\hat{\phi}$ satisfies 5.56 and $\hat{\phi}(0) = 1$, we have

$$\hat{\phi}(2\pi l) = 0,$$

for all $l \neq 0$ and hence

$$\lim_{n\to\infty} D_l^n(\omega) = 0, \tag{5.59}$$

for all $\omega \in I$ and $0 \leq |l| \leq L$. The functions $\hat{g}$ and $\hat{\phi}$ are continuous so the convergence of the limit 5.59 is uniform for $\omega \in I$. If 5.59 and 5.58 are combined, then we find that

$$\lim_{n\to\infty} \max_{\omega\in I} \left|\hat{G}_n(\omega/2^n) - |\hat{g}(\omega)|^2\right| = 0,$$

for any $\Omega > 0$.

We now turn our attention to the behavior of $\hat{G}_n$ as $n \to -\infty$. In this case, we assert that

$$\lim_{n\to-\infty} \hat{G}_n(\omega) = |\hat{g}(0)|^2 \tag{5.60}$$

uniformly. For any $\epsilon > 0$, there exists an $L \in \mathbb{N}$ such that

$$\left|\hat{G}_n(\omega) - |\hat{g}(0)|^2\right| \leq \sum_{|l|\leq L} \left||\hat{g}(2^n[\omega + 2\pi l])|^2 - |\hat{g}(0)|^2\right| \left|\hat{\phi}(\omega + 2\pi l)\right|^2 + \epsilon,$$

for all $\omega \in I$ and since

$$\lim_{n\to-\infty} \left| |\hat{g}(2^n[\omega + 2\pi l])|^2 - |\hat{g}(0)|^2 \right| = 0,$$

a nearly identical argument yields the desired result 5.60.

Consider 5.60 and recall the definitions 5.34 and 5.35 of A_n and B_n. It follows that,

$$\lim_{n\to-\infty} A_n = \lim_{n\to-\infty} B_n = |\hat{g}(0)|^2$$

and, as long as $\hat{g}(0) \neq 0$, we deduce that 5.54 holds. This means that, for small enough n, the inverse problem, defined by equation 5.5, will be well-posed whenever $\hat{g}(0) \neq 0$. On the other hand, 5.59 implies

$$\lim_{n\to\infty} A_n = 0$$

and

$$\lim_{n\to\infty} B_n = \|\mathcal{G}\|^2 .$$

Accordingly, even when $\kappa(\mathcal{G}|_{V_n})$ is finite for all $n \in \mathbb{Z}$, the problem posed by 5.5 becomes increasingly ill-conditioned as $n \to \infty$.

The behavior described above is not unexpected. Since $g \in L^1(\mathbb{R})$, we know that $\hat{g}(\omega)$ tends to zero as $|\omega| \to \infty$. This means that the high frequency components of u become increasingly difficult to recover. In light of the fact that $\phi^{n,k}$ is a low pass filter with a Fourier transform that is essentially supported in $[-2^n\pi, 2^n\pi]$, one expects that $\kappa(\mathcal{G}|_{V_n})$ will increase as the support width of $\widehat{\phi^{n,k}}$ increases.

In some instances, the rate at which A_n tends to zero as $n \to \infty$ can be estimated. Suppose that

$$|\hat{g}(\omega)| \leq \frac{C}{1+|\omega|},$$

then

$$\begin{aligned}\hat{G}_n(\omega) &\leq C^2 \sum_l \frac{1}{1+4^n(\omega+2\pi l)^2} \\ &= C^2 \frac{\sinh(2^{-n})}{2^{n+1}(\cosh(2^{-n})-\cos(\omega))}.\end{aligned}$$

Since

$$\frac{\sinh(x)}{1+\cosh(x)} \leq 1.$$

for all $x \geq 0$, it follows that

$$A_n \leq C^2 2^{-n-1} \tag{5.61}$$

and hence, the lower Riesz bounds A_n decay exponentially fast as $n \to \infty$.

We now turn our attention to convergence of 5.52 in the aforementioned special case.

Theorem 5.3 *Let $y \in R(\mathcal{G})$, with $|\hat{g}(\omega)| > 0$ for all $\omega \in \mathbb{R}$ and suppose that there exists a $\Omega > 0$ such that*

$$|\hat{g}(\omega_2)| \leq |\hat{g}(\omega_1)|, \tag{5.62}$$

whenever $\omega_2 \geq \omega_1 \geq \Omega$. If ϕ is an orthonormal scaling function which satisfies:

1. $\left|\hat{\phi}(\omega)\right| \geq D > 0$, *for all* $\omega \in [-\pi, \pi]$ *and*
2. $\left|\hat{\phi}(\omega)\right| \leq \frac{C}{1+|\omega|^\alpha}$, *for some* $\alpha > 1/2$.

then, for any n, the function

$$u_n(t) = \sum_k \langle y(t), \tilde{\xi}_n(t-k/2^n)\rangle \phi^{n,k}(t)$$

is a well defined element of V_n such that

$$\lim_{n\to\infty} \|u - u_n\| = 0, \tag{5.63}$$

where u is the unique solution of $\mathcal{G}u = y$.

Proof. Let P_n denote the orthogonal projection onto the subspace V_n and Q_j denote the orthogonal projection onto W_j, then

$$\|u - u_n\|^2 = \|P_n u - u_n\|^2 + \sum_{j \geq n} \|Q_j u\|^2 .$$

For any $f \in L^2(\mathbb{R})$ we have

$$\lim_{n \to \infty} \sum_{j \geq n} \|Q_j f\|^2 = 0$$

and hence we need only show that

$$\lim_{n \to \infty} \|P_n u - u_n\| = 0. \tag{5.64}$$

Since the function $y \in R(\mathcal{G})$ and $|\hat{g}(\omega)| > 0$, it follows that there is a unique $u \in L^2(\mathbb{R})$ such that

$$\hat{g}(\omega)\hat{u}(\omega) = \hat{y}(\omega)$$

and we can now use 5.41 to obtain

$$\begin{aligned} \langle y(t), \tilde{\xi}_n(t - k/2^n) \rangle &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\overline{\hat{g}(\omega)}}{\hat{G}_n(\omega/2^n)} \hat{y}(\omega) e^{ik\omega/2^n} 2^{-n/2} \overline{\hat{\phi}(\omega/2^n)}\, d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{|\hat{g}(\omega)|^2}{\hat{G}_n(\omega/2^n)} \hat{u}(\omega) e^{ik\omega/2^n} 2^{-n/2} \overline{\hat{\phi}(\omega/2^n)}\, d\omega. \end{aligned} \tag{5.65}$$

If we define the linear operator $\tilde{\mathcal{M}}_n$ by

$$\left(\tilde{\mathcal{M}}_n f\right)(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} \frac{\hat{f}(\omega)}{\hat{G}_n(\omega/2^n)}\, d\omega, \tag{5.66}$$

then, $\tilde{\mathcal{M}}_n$ is continuous with

$$\left\|\tilde{\mathcal{M}}_n\right\| = \frac{1}{A_n} < \infty$$

and, in view of 5.65, we can write u_n in the form

$$u_n = P_n \mathcal{G}^* \mathcal{G} \tilde{\mathcal{M}}_n u, \tag{5.67}$$

from which we see that, for each n, u_n is a well-defined element of V_n.

In view of 5.67, we can write

$$\begin{aligned} \|P_n u - u_n\| &= \left\| P_n(\mathcal{I} - \mathcal{G}^* \mathcal{G} \tilde{\mathcal{M}}_n)u \right\| \\ &\leq \left\| (\mathcal{I} - \mathcal{G}^* \mathcal{G} \tilde{\mathcal{M}}_n)u \right\| \end{aligned}$$

and consequently, if we can show that the sequence of operators $\left\{ \mathcal{G}^* \mathcal{G} \tilde{\mathcal{M}}_n \right\}$ converges strongly to the identity, then we will have established the desired result. Towards this end, we use 5.66 to obtain

$$\begin{aligned} \left\| (\mathcal{I} - \mathcal{G}^* \mathcal{G} \tilde{\mathcal{M}}_n)u \right\|^2 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \frac{|\hat{g}(\omega)|^2}{\hat{G}_n(\omega/2^n)} - 1 \right|^2 |\hat{u}(\omega)|^2 \, d\omega \qquad (5.68) \\ &= \frac{1}{2\pi} \int_{I'} \left| \frac{|\hat{g}(\omega)|^2}{\hat{G}_n(\omega/2^n)} - 1 \right|^2 |\hat{u}(\omega)|^2 \, d\omega \\ &+ \frac{1}{2\pi} \int_{\mathbb{R} \setminus I'} \left| \frac{|\hat{g}(\omega)|^2}{\hat{G}_n(\omega/2^n)} - 1 \right|^2 |\hat{u}(\omega)|^2 \, d\omega. \end{aligned}$$

where $I' = [-\Omega', \Omega']$ and $\Omega' > 0$ arbitrary. Since $\hat{G}_n(\omega/2^n)$ converges to $|\hat{g}(\omega)|^2$ uniformly on compact subsets of $\mathbb{R}$ and $u \in L^2(\mathbb{R})$, it is enough to show that

$$h_n(\omega) = \frac{|\hat{g}(\omega)|^2}{\hat{G}_n(\omega/2^n)} \leq C', \qquad (5.69)$$

with C' independent of n.

Let

$$I_q^n = [2^n \pi(2q-1), 2^n \pi(2q+1)]$$

and

$$M_q^n = \max_{\omega \in I_q^n} h_n(\omega).$$

Since h_n is an even function, we can restrict our attention to values of ω such that $\omega \geq 0$ and accordingly, we can assume that $q \in \mathbb{N}$. In addition, we suppose that

$n \geq N_1$, where $N_1 \in \mathbb{N}$ is such that

$$N_1 \geq \log_2\left(\frac{\Omega}{\pi}\right)$$

and consequently $I_q^n \subset [\Omega, \infty)$, for all $q \geq 1$. It follows from 5.62 and the fact that $\hat{G}_n$ is periodic that, if $\omega \geq \Omega$, then

$$h_n(\omega + 2^{n+1}\pi q) \leq h_n(\omega),$$

for all $q \geq 1$, and hence

$$M_1^n \geq M_2^n \geq M_3^n \geq \cdots,$$

from which we obtain

$$h_n(\omega) \leq \max(M_0^n, M_1^n). \tag{5.70}$$

Suppose that $\omega \in I_0^n$, then

$$\hat{G}_n(\omega/2^n) \geq |\hat{g}(\omega)\phi(\omega/2^n)|^2 \geq D^2 |\hat{g}(\omega)|^2,$$

which implies

$$M_0^n \leq \frac{1}{D^2}. \tag{5.71}$$

On the other hand, if $\omega \in I_1^n$, then

$$\hat{G}_n(\omega/2^n) \geq D^2 \left|\hat{g}(\omega - 2^{n+1}\pi)\right|^2$$

and therefore ,

$$h_n(\omega) \leq \left|\frac{\hat{g}(\omega)}{D\hat{g}(\omega - 2^{n+1}\pi)}\right|^2, \tag{5.72}$$

for $\omega \in [-2^n\pi, 2^n\pi]$. We know that

$$\lim_{|\omega|\to\infty} \hat{g}(\omega) = 0$$

and, in view of 5.62, there exists $N_2 \geq N_1$ such that for all $n \geq N_2$ for all $n \geq N_2$

$$\min_{\omega \in I_1^n} \left|\hat{g}(\omega - 2^{n+1}\pi)\right| = |\hat{g}(2^n\pi)|$$

and by similar reasoning,

$$\max_{\omega \in I_1^n} |\hat{g}(\omega)| = |\hat{g}(2^n\pi)| .$$

Inequality 5.72 now implies

$$M_1^n \leq \frac{1}{D^2}$$

and, from 5.70 we conclude that 5.69 holds with $C' = 1/D^2$.

It now follows from 5.69, that for any $\epsilon > 0$, we can choose $\Omega' > 0$ so that

$$\left\|\left(\mathcal{I} - \mathcal{G}^*\mathcal{G}\tilde{\mathcal{M}}_n\right) u\right\|^2 \leq \frac{1}{2\pi} \int_{I'} \left|\frac{|\hat{g}(\omega)|^2}{\hat{G}_n(\omega/2^n)} - 1\right|^2 |\hat{u}(\omega)|^2 \, d\omega + \epsilon$$

and therefore ,

$$\lim_{n\to\infty} \left\|\left(\mathcal{I} - \mathcal{G}^*\mathcal{G}\tilde{\mathcal{M}}_n\right) u\right\| = 0,$$

as required. ∎

In certain instances, rate estimates for 5.64 are simple to derive. Suppose that $|\hat{g}(\omega)| > 0$ for $\omega \in \mathbb{R}$ and that

$$|\hat{g}(\omega_2)| \leq |\hat{g}(\omega_1)|$$

for all $\omega_2 \geq \omega_1 \geq 0$. We begin by considering the quantity

$$\left|\frac{|\hat{g}(\omega)|^2}{\hat{G}_n(\omega/2^n)} - 1\right| = \frac{1}{\hat{G}_n(\omega/2^n)} \left|\hat{G}_n(\omega/2^n) - |\hat{g}(\omega)|^2\right| , \tag{5.73}$$

for $\omega \in I_n = [-2^n\pi, 2^n\pi]$. The monotonicity of $\hat{g}$ yields the inequality

$$\begin{aligned} \left|\hat{G}_n(\omega/2^n) - |\hat{g}(\omega)|^2\right| &\leq \left|1 - \left|\hat{\phi}(\omega/2^n)\right|^2\right| \{|\hat{g}(\omega)|^2 + \sum_{l\neq 0} \left|\hat{\phi}(\omega/2^n + 2\pi l)\right|^2 |\hat{g}(\omega)|^2\} \\ &= 2\left|1 - \left|\hat{\phi}(\omega/2^n)\right|^2\right| |\hat{g}(\omega)|^2 , \end{aligned} \tag{5.74}$$

where we have used 5.56, which is

$$\sum_l \left|\hat{\phi}(\omega + 2\pi l)\right|^2 = 1.$$

We combine inequalities 5.73 and 5.74 to obtain

$$\left|\frac{|\hat{g}(\omega)|^2}{\hat{G}_n(\omega/2^n)} - 1\right| \leq 2\frac{|\hat{g}(\omega)|^2}{\hat{G}_n(\omega/2^n)}\left|1 - \left|\hat{\phi}(\omega/2^n)\right|^2\right|$$

and, through the use of

$$\hat{G}_n(\omega/2^n) \geq D^2 |\hat{g}(\omega)|^2,$$

we find that

$$\left|\frac{|\hat{g}(\omega)|^2}{\hat{G}_n(\omega/2^n)} - 1\right| \leq \frac{2}{D^2}\left|1 - \left|\hat{\phi}(\omega/2^n)\right|^2\right|, \tag{5.75}$$

valid for all $\omega \in I_n$. It is interesting to note that the bound 5.75 does not depend on the decay of $\hat{g}$ as $|\omega| \to \infty$. Rather, it is the behavior of $\hat{\phi}$ near $\omega = 0$ which governs 5.75.

Suppose that $1 < \beta < 2$ and let $J_n = [-\beta^n\pi, \beta^n\pi]$, then $J_n \subset I_n$ for all n and, from 5.75, we have

$$\begin{aligned}
\|\mathcal{I} - \mathcal{G}^*\mathcal{G}\mathcal{M}_n u\|^2 &\leq \frac{2}{D^4\pi}\int_{J_n}\left|1 - \left|\hat{\phi}(\omega/2^n)\right|^2\right|^2 |\hat{u}(\omega)|^2\, d\omega \\
&\quad + \frac{1}{2\pi}\int_{\mathbb{R}\setminus J_n}\left|\frac{|\hat{g}(\omega)|^2}{\hat{G}_n(\omega/2^n)} - 1\right|^2 |\hat{u}(\omega)|^2\, d\omega \\
&\leq \frac{4}{D^4}\max_{\omega\in J_n}\left|1 - \left|\hat{\phi}(\omega/2^n)\right|^2\right|^2 \|u\|^2 \\
&\quad + \frac{D^2+1}{2\pi D^2}\int_{\mathbb{R}\setminus J_n} |\hat{u}(\omega)|^2\, d\omega,
\end{aligned} \tag{5.76}$$

and hence, the rate at which 5.64 converges is essentially determined by the behavior of $\hat{\phi}$ for ω near 0 and the decay of $\hat{u}$ as $|\omega| \to \infty$.

Let us assume that ϕ is a Daubechies scaling function and that u belongs to the Sobolev space $H^s(\mathbb{R})$. then

$$\|u\|_s^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty}(1+\omega^2)^s\,|\hat{u}(\omega)|^2\,d\omega < \infty,$$

which implies that

$$\begin{aligned}\frac{1}{2\pi}\int_{\mathbb{R}\setminus J_n}|\hat{u}|^2\,d\omega &= \frac{1}{2\pi}\int_{\mathbb{R}\setminus J_n}\frac{(1+\omega^2)^s}{(1+\omega^2)^s}\,|\hat{u}(\omega)|^2\,d\omega \\ &\le \frac{1}{(1+\pi^2\beta^{2n})^s}\,\|u\|_s^2 \le \frac{1}{\pi^2\beta^{2sn}}\,\|u\|_s^2. \end{aligned} \tag{5.77}$$

Furthermore. it is shown in Appendix A that

$$1-\left|\hat{\phi}(\omega)\right|^2 = \mu_N\omega^{2N} + O(\omega^{2N+2})$$

and hence, there exists a constant $M_N > 0$ such that

$$\left|1-\left|\hat{\phi}(\omega/2^n)\right|^2\right| \le M_N(\beta/2)^{2Nn}. \tag{5.78}$$

for all $\omega \in J_n$. As a result of 5.77 and 5.78, we have the inequality

$$\|(\mathcal{I}-\mathcal{G}^*\mathcal{G}\mathcal{M}_n)\,u\|^2 \le \frac{4M_N^2}{D^4}(\beta/2)^{4Nn}\,\|u\|^2 + \frac{D^2+1}{\pi^2D^2}\beta^{-2sn}\,\|u\|_s^2. \tag{5.79}$$

To obtain a rate estimate for 5.63, we must examine the rate at which

$$\sum_{j\ge n}\left\|Q_ju^\dagger\right\|^2 \to 0$$

as $n\to\infty$. Under the assumption that $u \in H^s(\mathbb{R})$, it can be shown that (see [14, page 299])

$$\sum_{j\ge n}(1+4^{sj})\,\|Q_ju\|^2 < \infty$$

and hence, there exists a constant $C_3 > 0$ such that

$$\begin{aligned}\sum_{j\ge n}\|Q_ju\|^2 &= \sum_{j\ge n}\frac{(1+4^{sj})}{(1+4^{sj})}\,\|Q_ju\|^2 \\ &\le C_3 4^{-sn},\end{aligned}$$

which, in combination with 5.79, yields the result

$$\|u - u_n\|^2 \leq C_1(\beta/2)^{4Nn} + C_2\beta^{-2sn} + C_3 4^{-sn}, \tag{5.80}$$

for all $\beta \in (1,2)$. Finally, we point out that, if $N \geq s/2$, then the optimal choice of β leads us to the estimate

$$\|u - u_n\|^2 \leq C 2^{-4nNs/(2N+s)},$$

for some constant C.

Chapter 6

Multiresolution based methods

6.1 Introduction

Up to this point, our work has not made use of the multiresolution analysis (MRA) generated by the scaling function ϕ. In this chapter, we begin an investigation of two methods which utilize the MRA in an effort to solve the operator equation

$$\mathcal{G}_n u = y, \tag{6.1}$$

where we have used the notation $\mathcal{G}_n \equiv \mathcal{G}|_{V_n}$.

The first of these methods is based upon the wavelet expansion,

$$\begin{aligned} u &= P_m u + Q_m u + \cdots Q_{n-1} u \\ &= \sum_k \langle u, \phi^{m,k} \rangle \phi^{m,k} + \sum_{j=m}^{n-1} \sum_k \langle u, \psi^{j,k} \rangle \psi^{j,k} \end{aligned} \tag{6.2}$$

of $u \in V_n$. With this method, we seek to improve the condition of the inverse problem 6.1 by selectively recovering the orthogonal projections $P_m u$ and $Q_j u$, $j = m, \dots, n-1$, onto the subspaces V_m and W_j respectively.

The second method is a multiresolution regularization algorithm which is due to J. Liu. In the paper [30], the author presents empirical evidence which suggests that MRA based regularization may be a useful tool for the solution of certain inverse problems. In particular, the author employs a multiresolution regularization method, based on the Haar MRA, to solve a distributed parameter estimation problem. We will present a preliminary investigation of certain theoretical aspects of MRA based regularization, and demonstrate that this method can be regarded as a special case of C-generalized regularization. Moreover, we shall examine circumstances under which the generalized and C-generalized solutions are, in some sense, close.

6.2 MRA decomposition techniques

Consider the inverse problem posed by 6.1, and suppose that $\mathcal{G}_n$ is invertible. As we have seen, we can write the solution of 6.1 in the form

$$\begin{aligned} u &= \mathcal{G}_n^{-1} y \\ &= \sum_k \langle y(t), \tilde{\xi}_n(t-k/2^n)\rangle \phi^{n,k}(t). \end{aligned} \tag{6.3}$$

Since $\mathcal{G}_n$ is invertible, we know that the set of functions $\left\{\tilde{\xi}_n(t-k/2^n) : k \in \mathbb{Z}\right\}$ is a Riesz basis with Riesz bounds $0 < B_n^{-1} \le A_n^{-1} < \infty$. It follows that the condition number

$$\kappa_n \equiv \kappa(\mathcal{G}_n) = \|\mathcal{G}_n\| \left\|\mathcal{G}_n^{-1}\right\| = \sqrt{\frac{B_n}{A_n}} \tag{6.4}$$

is finite. However, even though $\kappa_n < \infty$, if n is large, then A_n can be close to zero and hence, κ_n can be quite large for large n. This means that the inverse problem 6.1 can be extremely ill-conditioned even when it is well-posed. Consequently, the presence of small errors in the data can give rise to approximate solutions which deviate substantially (in the norm on $L^2(\mathbb{R})$) from the exact solution.[1]

We now turn our attention to the problem of recovering the functions $P_m u$ and $Q_j u$, $j = m, \ldots, n-1$, from the data y. Since $\{\phi^{m,k}, \psi^{j,k} : j = m, \ldots, n-1, k \in \mathbb{Z}\}$ is an orthonormal basis for V_n, this is equivalent to the recovery of the sequences

$$\{u_m[k] : k \in \mathbb{Z}\} \text{ and } \{u^j[k] : j = m, \ldots, n-1, k \in \mathbb{Z}\}, \tag{6.5}$$

where we have used the notation $u_m[k] = \langle u, \phi^{m,k} \rangle$ and $u^j[k] = \langle u, \psi^{j,k} \rangle$. We know that $u_n[k] = \langle y(t), \tilde{\xi}_n(t - k/2^n) \rangle$ and therefore we can apply the discrete wavelet transform to the sequence $\{u_n[k] : k \in \mathbb{Z}\}$ to obtain the sequences 6.5. In particular, we have the decomposition formulae

$$u_{j-1}[k] = \sum_l h_{l-2k} u_j[k] \tag{6.6}$$

and

$$u^{j-1}[k] = \sum_l g_{l-2k} u_j[k], \tag{6.7}$$

where $j = m+1, m = 2, \ldots, n$.

If we define the functions ν_j and ν^j by

$$\nu_{j-1}(t) = \sum_l h_l \nu_j(t - l/2^j) \tag{6.8}$$

[1]Consider the inverse problem defined by the operator equation $\mathcal{A}f = g$. Suppose that the data g is contaminated with error δg and let the error induced in the solution be δf. It can be shown that the relative errors in the solution and the data are related via the inequality

$$\frac{\|\delta f\|}{\|f\|} \le \kappa(\mathcal{A}) \frac{\|\delta g\|}{\|g\|}.$$

Consequently, when the condition number $\kappa(\mathcal{A})$ is large, the relative error in the solution can be large even when the relative error in the data is small. When the condition number is large, the inverse problem is said to be *ill-conditioned*.

and

$$\nu^{j-1}(t) = \sum_l g_l \nu_j(t - l/2^j), \tag{6.9}$$

with $j = m+1, \ldots, n$ and $\nu_n \equiv \tilde{\xi}_n$, then, through the use of formulae 6.6 and 6.7, we can show that

$$u_j[k] = \langle y(t), \nu_j(t - k/2^j)\rangle \tag{6.10}$$

and

$$u^j[k] = \langle y(t), \nu^j(t - k/2^j)\rangle. \tag{6.11}$$

We now use the functionals 6.10 and 6.11 to write expansion 6.2 in the form

$$u = \sum_k \langle y(t), \nu_m(t - k/2^m)\rangle \phi^{m,k} + \sum_{j=m}^{n-1} \sum_k \langle y(t), \nu^j(t - k/2^j)\rangle \psi^{j,k}.$$

We mention that, since $\{h_k\}, \{g_k\} \in \ell^1(\mathbb{Z})$, if $\tilde{\xi}_n \in L^p(\mathbb{R})$, then the functions ν_j and ν^j are well-defined elements of $L^p(\mathbb{R})$ for all $j \leq n$.

Recall that, if n is large, then the condition number 6.4 can be large. Since

$$\|\mathcal{G}_n\|^2 = B_n \leq \|\mathcal{G}\|^2,$$

the magnitude of κ_n depends primarily on the quantity

$$\|\mathcal{G}_n^{-1}\|^2 = \frac{1}{A_n}. \tag{6.12}$$

Now, the decomposition

$$\|\mathcal{G}_n^{-1} y\|^2 = \|P_m \mathcal{G}_n^{-1} y\|^2 + \cdots + \|Q_{n-1} \mathcal{G}_n^{-1} y\|^2,$$

leads us to the inequality

$$\|\mathcal{G}_n^{-1}\|^2 \leq \|P_n \mathcal{G}_n^{-1}\|^2 + \cdots + \|Q_{n-1} \mathcal{G}_n^{-1}\|^2,$$

which implies that

$$\kappa_n \le \|\mathcal{G}\| \left(\left\| P_m \mathcal{G}_n^{-1} \right\|^2 + \cdots \left\| Q_{n-1} \mathcal{G}_n^{-1} \right\|^2 \right)^{1/2}. \tag{6.13}$$

Accordingly, the size of the condition number κ_n is dependent upon the magnitudes of the quantities

$$\left\| P_m \mathcal{G}_m^{-1} \right\| \text{ and } \left\| Q_j \mathcal{G}_n^{-1} \right\|, \tag{6.14}$$

for $j = m, \ldots, n$. Moreover, the magnitude of the norms 6.14 can be used as an indication of which functions $P_m u$ and $Q_j u$ are the most difficult to construct in the presence of noisy data. For instance, suppose that the observed data is of the form $y + \delta y$, where δy represents small but unknown error. The magnitude of the error of the approximate solution

$$u_\delta = \mathcal{G}_n^{-1}(y + \delta y)$$

is given by

$$\left\| \mathcal{G}_n^{-1} \delta y \right\| = \left(\left\| P_m \mathcal{G}_n^{-1} \delta y \right\|^2 + \cdots + \left\| Q_{n-1} \mathcal{G}_n^{-1} \delta y \right\|^2 \right)^{1/2}.$$

If, in particular, $\|Q_j \mathcal{G}_n^{-1}\|$ is large, then the magnitude of the error in $Q_j u_\delta$, given by

$$\left\| Q_j \mathcal{G}_n^{-1} \delta y \right\| \le \left\| Q_j \mathcal{G}_n^{-1} \right\| \|\delta y\|,$$

can be large even when $\|\delta y\|$ is small.

To obtain estimates for the norms 6.14, we appeal to Parseval's relation which yields

$$\left\| P_m \mathcal{G}_n^{-1} y \right\|^2 = \sum_k |\langle y(t), \nu_m(t - k/2^m) \rangle|^2$$

and

$$\left\| Q_j \mathcal{G}_n^{-1} y \right\|^2 = \sum_k \left| \langle y(t), \nu^j(t - k/2^j) \rangle \right|^2.$$

We can now use the equivalent definitions of a Bessel sequence 5.21 and 5.22 to infer that

$$\left\|P_m\mathcal{G}_n^{-1}y\right\|^2 \leq D_m \left\|y\right\|^2$$

if and only if

$$\left\|\sum_k \alpha_k \nu_m(t-k/2^m)\right\|^2 \leq D_m \sum_k |\alpha_k|^2, \tag{6.15}$$

while

$$\left\|Q_j\mathcal{G}_n^{-1}y\right\|^2 \leq D^j \left\|y\right\|^2$$

if and only if

$$\left\|\sum_k \alpha_k \nu^j(t-k/2^j)\right\|^2 \leq D^j \sum_k |\alpha_k|^2. \tag{6.16}$$

In other words, we can obtain estimates for the norms 6.14 by finding the bounds of the appropriate Bessel sequences.

Let us restrict our attention to the problem of estimating D_m. The left hand side of inequality 6.15 can be written as

$$\begin{aligned}\left\|\sum_k \alpha_k \nu_m(t-k/2^m)\right\|^2 &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \left|A(\omega/2^m)\hat{\nu}_m(\omega)\right|^2 d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |A(\omega)|^2 N_m(\omega)\, d\omega.\end{aligned}$$

where

$$N_m(\omega) = 2^m \sum_l |\hat{\nu}_m(2^m[\omega + 2\pi l])|^2. \tag{6.17}$$

If we assume that N_m is continuous, then we have the estimate

$$\left\|P_m\mathcal{G}_n^{-1}\right\|^2 = D_m = \max_{\omega\in[-\pi,\pi]} N_m(\omega). \tag{6.18}$$

Now, through the use of 6.8, we can derive an expression relating the functions N_m and $\hat{G}_n$. In particular, we can combine the Fourier transform of equation 6.8 with equation 6.17 to obtain

$$N_m(\omega) = \mathcal{P}\left(N_{m+1}\right)(\omega),$$

where $\mathcal{P} : C^0[-\pi, \pi] \mapsto C^0[-\pi, \pi]$ is the map

$$\mathcal{P}(A)(\omega) = \frac{1}{2}\left\{|H(\omega/2)|^2 A(\omega/2) + |H(\omega/2+\pi)|^2 A(\omega/2+\pi)\right\}. \tag{6.19}$$

If we repeat this process $n-m-1$ more times, then we find that

$$N_m(\omega) = \mathcal{P}^{n-m}\left(N_n\right)(\omega),$$

where

$$N_n(\omega) = 2^n \sum_l \left|\hat{\tilde{\xi}}_n(2^n[\omega + 2\pi l])\right|^2 = \frac{1}{\hat{G}_n(\omega)},$$

which means that, for all $m \leq n$,

$$\left\|P_m \mathcal{G}_n^{-1}\right\|^2 = \max_{\omega\in[-\pi,\pi]} \mathcal{P}^{n-m}\left(1/\hat{G}_n\right)(\omega). \tag{6.20}$$

If we define the operator $\mathcal{Q}$ by

$$\mathcal{Q}(A)(\omega) = \frac{1}{2}\left\{|G(\omega/2)|^2 A(\omega/2) + |G(\omega/2+\pi)|^2 A(\omega/2+\pi)\right\}. \tag{6.21}$$

then a similar derivation yields the result

$$\left\|Q_j \mathcal{G}_n^{-1}\right\|^2 = \max_{\omega\in[-\pi,\pi]} \mathcal{Q}\left(\mathcal{P}^{n-m-1}\left(1/\hat{G}_n\right)\right)(\omega). \tag{6.22}$$

It should be noted that the Fourier coefficients of $\mathcal{P}^{n-m}(1/\hat{G}_n)$ are the entries of the Toeplitz matrix

$$[\langle \nu_m(t-k/2^m), \nu_m(t-l/2^m)\rangle],$$

$k, l \in \mathbb{Z}$, which is the Gram matrix of $\{\nu_m(t-k/2^m) : k \in \mathbb{Z}\}$. In particular, the entries of this Gram matrix are related to the Fourier coefficients of $\mathcal{P}^{n-m}(1/\hat{G}_n)$ via

$$\langle \nu_m(t-k/2^m), \nu_m(t-l/2^m)\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i(k-l)}\mathcal{P}^{n-m}(1/\hat{G}_n)(\omega)\, d\omega.$$

We contrast this to the case examined in the previous chapter. Specifically, the entries of the Gram matrix of the set $\left\{\tilde{\xi}(t-k/2^m) : k \in \mathbb{Z}\right\}$ are generated, in a similar fashion, by the function $1/\hat{G}_m$.

In general,

$$1/\hat{G}_m(\omega) \neq \mathcal{P}^{n-m}\left(1/\hat{G}_n\right)(\omega). \tag{6.23}$$

Expression 6.23 highlights one difference between two possible approaches to the construction of an approximation to u in the subspace V_m. In the first, we assume $u \in V_m$. The functions $\tilde{\xi}_m(t-k/2^m)$ are then used to produce

$$\begin{aligned} u(t) \approx u_1(t) &= \sum_k \langle y(t), \tilde{\xi}_m(t-k/2^m)\rangle \phi^{m,k}(t) \\ &= \left(\mathcal{G}_m^{-1} y\right)(t). \end{aligned} \tag{6.24}$$

The second approach begins with the assumption $u \in V_n$, for some $n > m$, whereupon the functions $\nu_m(t-k/2^m)$ give rise to the approximation

$$\begin{aligned} u(t) \approx u_2(t) &= \sum_k \langle y(t), \nu_m(t-k/2^m)\rangle \phi^{m,k}(t) \\ &= \left(P_m \mathcal{G}_n^{-1} y\right)(t). \end{aligned} \tag{6.25}$$

Approximations 6.24 and 6.25 are usually distinct and, as we have just shown

$$\left\|P_m \mathcal{G}_n^{-1}\right\|^2 = \max_{\omega \in [-\pi,\pi]} \mathcal{P}^{n-m}\left(1/\hat{G}_n\right)(\omega),$$

while

$$\left\|\mathcal{G}_m^{-1}\right\|^2 = \max_{\omega \in [-\pi,\pi]} 1/\hat{G}_m(\omega).$$

In view of 6.23, we will generally have

$$\left\|\mathcal{G}_m^{-1}\right\| \neq \left\|P_m \mathcal{G}_n^{-1}\right\|, \tag{6.26}$$

which means that the sensitivity of 6.24 and 6.25 to any noise in the data y will differ according to the magnitudes of the norms in expression 6.26. However, in the previous chapter we proved that

$$\lim_{m \to -\infty} \hat{G}_m(\omega) = |\hat{g}(0)|^2$$

and, as we shall show in the next section,

$$\lim_{k \to \infty} \mathcal{P}^k (A)(\omega) = A(0).$$

It follows that

$$\lim_{m \to -\infty} \|\mathcal{G}_m^{-1}\| = \lim_{m \to -\infty} \|P_m \mathcal{G}_n^{-1}\| = |\hat{g}(0)|$$

and therefore, when we consider sensitivity to noisy data and restrict our attention to the construction of low resolution approximation, then we expect little difference in the two approaches.

In some cases, the norms 6.20 and 6.22 can be difficult to estimate. In the next section, an examination of the operator $\mathcal{P}$ enables us to make some general statements concerning the limiting behavior of 6.20 and 6.22.

6.3 The operator $\mathcal{P}$

The operator $\mathcal{P}$ arises in the study of orthonormal wavelet bases (see [14, page 190]). Here, an examination of the fixed points of $\mathcal{P}$ provides necessary and sufficient conditions for the characterization of *two-scale symbols* H, which give rise to orthonormal wavelet bases.

Many of the properties of $\mathcal{P}$ arise directly from the conditions

$$|H(\theta)|^2 + |H(\theta + \pi)|^2 = 2 \tag{6.27}$$

and

$$|H(0)|^2 = 2. \tag{6.28}$$

In fact, with conditions 6.27 and 6.28 in mind, it is easy to show that:

1. $\mathcal{P}^N(A)(0) = A(0)$,
2. $\mathcal{P}^N(A)(\pi) = A(\pi/2^n)$, and (6.29)
3. if C is a constant, then $\mathcal{P}(C) = C$,

where A is a 2π-periodic function and $N \in \mathbb{N}$. A comment concerning condition 3 above is in order; it can be shown that if H is a two-scale symbol, satisfying 6.27 and 6.28, then H will give rise to an orthonormal wavelet basis if and only if the only fixed points of $\mathcal{P}$ are constants.

We have restricted our attention to two-scale sequences $\{h_k\} \in \ell^1(\mathbb{Z})$. The function H is therefore continuous, and we can regard $\mathcal{P}$ as an operator mapping continuous 2π-periodic functions to continuous 2π-periodic functions. In view of the conditions 6.29, it is not unreasonable to suspect that

$$\lim_{N\to\infty} \left\|\mathcal{P}^N(A)(\omega) - A(0)\right\|_\infty = 0, \tag{6.30}$$

where $\|\cdot\|_\infty$ is the usual norm on $C^0[-\pi,\pi]$. We will prove that 6.30 does indeed hold under the appropriate conditions. However, we first establish the weaker result

$$\lim_{N\to\infty} \left\|\mathcal{P}^N(A) - A(0)\right\|_1 = 0,$$

where $\|\cdot\|_1$ is the norm on $L^1[-\pi,\pi]$. To do this, it is convenient to have an explicit representation of $\mathcal{P}^N(A)$, which is given in the following:

Lemma 6.1 *If the operator $\mathcal{P}$ is as in 6.19, then for all $N \in \mathbb{N}$*

$$\mathcal{P}^N(A)(\omega) = 2^{-N} \sum_{k=0}^{2^N-1} \left\{ \prod_{p=1}^{N} |H(2^{-p}[\omega + 2\pi k])|^2 \right\} A(2^{-N}[\omega + 2\pi k]). \tag{6.31}$$

Proof. Suppose that $N = 1$, then 6.31 becomes

$$\begin{aligned} \mathcal{P}(A)(\omega) &= \frac{1}{2} \sum_{k=0}^{1} \prod_{p=1}^{1} |H(2^{-p}[\omega + 2\pi k])|^2 A(1/2[\omega + 2\pi k]) \\ &= \frac{1}{2} \left\{ |H(\omega/2)|^2 A(\omega/2) + |H(\omega/2 + \pi)|^2 A(\omega/2 + \pi) \right\}. \end{aligned}$$

as required. Assume that 6.31 holds for $N = M$ and consider $\mathcal{P}^{M+1}(A)$. We have

$$\mathcal{P}^{M+1}(A)(\omega) = \frac{1}{2} \left\{ |H(\omega/2)|^2 \mathcal{P}^M(A)(\omega/2) + |H(\omega/2 + \pi)|^2 \mathcal{P}^M(\omega/2 + \pi) \right\}. \tag{6.32}$$

which, after some algebra, can be written as

$$\begin{aligned} \mathcal{P}^{M+1}(A)(\omega) &= 2^{-M-1} \sum_{k=0}^{2^M-1} \prod_{p=1}^{M+1} |H(2^{-p}[\omega + 4\pi l])|^2 A(2^{-M-1}[\omega + 4\pi k]) \\ &+ 2^{-M-1} \sum_{k=0}^{2^M-1} \prod_{p=1}^{M+1} |H(2^{-p}[\omega + 2\pi(2k+1)])|^2 A(2^{-M-1}[\omega + 2\pi(2k+1)]) \\ &= 2^{-M-1} \sum_{k=0}^{2^{M+1}-1} \prod_{p=1}^{M+1} |H(2^{-p}[\omega + 2\pi k])|^2 A(2^{-M-1}[\omega + 2\pi k]) \end{aligned}$$

and therefore the desired result follows by induction. ∎

Since $\mathcal{P}^N(1) = 1$ for all N, Lemma 6.1 immediately implies that

$$2^{-N} \sum_{k=0}^{2^N-1} \left\{ \prod_{p=1}^{N} |H(2^{-p}[\omega + 2\pi k])|^2 \right\} = 1 \tag{6.33}$$

for all $N \in \mathbb{N}$. We can now prove the following theorem.

Theorem 6.1 *Let ϕ be an orthonormal scaling function such that $|\phi| \geq C$ for all $\omega \in [-\pi, \pi]$. If $A \in L^1[-\pi, \pi]$ is a 2π-periodic function, continuous near $\omega = 0$, then*

$$\lim_{N\to\infty} \left\|\mathcal{P}^N(A) - A(0)\right\|_1 = 0.$$

Proof. In view of 6.31 and 6.33, we can write

$$\left|\mathcal{P}^N(A) - A(0)\right| \leq \sum_{k=0}^{2^N-1} \prod_{p=1}^{N} \left|\frac{1}{\sqrt{2}}H(2^{-p}[\omega + 2\pi k])\right|^2 \left|A(2^{-N}[\omega + 2\pi k]) - A(0)\right|.$$

which implies

$$\|\mathcal{P}(A) - A(0)\|_1 \leq \frac{1}{2\pi}\int_0^{2\pi} \sum_{k=0}^{2^N-1} \prod_{p=1}^{N} \left|H(2^{-p}[\omega + 2\pi k])\right|^2 \left|A(2^{-N}[\omega + 2\pi k]) - A(0)\right| d\omega.$$

If we make the change of variable $\theta = \omega + 2\pi k$ in the integral above, then we obtain the inequality

$$\left\|\mathcal{P}^N(A) - A(0)\right\|_1 \leq \frac{1}{2\pi}\int_0^{2^{N+1}\pi} \prod_{p=1}^{N} \left|\frac{1}{\sqrt{2}}H(2^{-p}\theta)\right|^2 \left|A(2^{-N}\theta) - A(0)\right| d\theta.$$

whereupon the change of variable $\theta = 2^N\omega$ yields the result

$$\|\mathcal{P}(A) - A(0)\|_1 \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} 2^N \prod_{p=1}^{N} \left|\frac{1}{\sqrt{2}}H(2^{N-p}\omega)\right|^2 |A(\omega) - A(0)|\, d\omega. \qquad (6.34)$$

Now, from the identity

$$\hat{\phi}(\omega) = \frac{1}{\sqrt{2}}H(\omega/2)\hat{\phi}(\omega/2),$$

one can derive the equation

$$\left|\hat{\phi}(\omega)\right|^2 = \prod_{p=1}^{N} \left|\frac{1}{\sqrt{2}}H(2^{-p}\omega)\right|^2 \left|\hat{\phi}(\omega/2^N)\right|^2 \qquad (6.35)$$

and hence. 6.34 is equivalent to

$$\begin{aligned}\|\mathcal{P}(A) - A(0)\|_1 &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} 2^N \frac{\left|\hat{\phi}(2^N\omega)\right|^2}{\left|\hat{\phi}(\omega)\right|^2} |A(\omega) - A(0)|\, d\omega \\ &\leq \frac{C}{2\pi}\int_{-\pi}^{\pi} 2^N \left|\hat{\phi}(2^N\omega)\right|^2 |A(\omega) - A(0)|\, d\omega.\end{aligned}$$

The non-negative function $\left|\hat{\phi}(\omega)\right|^2$ is continuous and satisfies

$$\int_{-\infty}^{\infty} \left|\hat{\phi}(\omega)\right|^2 d\omega = 2\pi.$$

Hence. the set of functions $\{2^N/2\pi \left|\hat{\phi}(2^N\omega)\right|^2 : N \in \mathbb{N}\}$ forms a δ-sequence. We conclude that $\mathcal{P}^N(A) \to A(0)$, as $N \to \infty$ in the norm on $L^1[-\pi,\pi]$. ▮

Under certain circumstances, the results of the previous theorem can be extended to include the spaces $L^p[-\pi,\pi]$, $1 \leq p < \infty$. Suppose that $A \in L^\infty[-\pi,\pi]$, then $A \in L^p[-\pi,\pi]$, for $1 \leq p < \infty$. Furthermore, since

$$\mathcal{P}(A)(\omega) \leq \sup_{\omega\in[-\pi,\pi]} |A(\omega)|,$$

it is easy to show that

$$\mathcal{P}^N(A)(\omega) \leq \sup_{\omega\in[-\pi,\pi]} |A(\omega)|$$

for all $N \in \mathbb{N}$. It now follows that

$$\left|\mathcal{P}^N(A) - A(0)\right|^p \leq \left(\sup_{\omega\in[-\pi,\pi]} |A(\omega)| + |A(0)|\right)^{p-1} \left|\mathcal{P}^N(A) - A(0)\right|,$$

which implies that

$$\left\|\mathcal{P}^N(A) - A(0)\right\|_p \leq D \left\|\mathcal{P}^N(A) - A(0)\right\|_1,$$

for some constant D.

Although of some theoretical interest, in view of 6.20 and 6.22, the results of Theorem 6.1 are of little practical value. Fortunately, with some additional hypotheses, it can be shown that the sequence $\{\mathcal{P}^N(A) : N \in \mathbb{N}\}$ converges uniformly. We have the following:

Theorem 6.2 *Let $A \in C^0[0, 2\pi]$ and suppose that ϕ is an orthonormal scaling function such that $\left|\hat{\phi}(\omega)\right| \geq C$ for all $\omega \in [-\pi, \pi]$. If the series*

$$\sum_l \left|\hat{\phi}(\omega + 2\pi l)\right|^2 \tag{6.36}$$

converges uniformly, then

$$\lim_{N\to\infty} \max_{\omega\in[0,2\pi]} \left|\mathcal{P}^N(A)(\omega) - A(0)\right| = 0.$$

Proof. If we use identity 6.33 to write

$$\mathcal{P}^N(A)(\omega) - A(0) = \sum_{k=0}^{2^N-1} \prod_{p=1}^{N} \left|\frac{1}{\sqrt{2}} H(2^{-p}[\omega + 2\pi k])\right|^2 \left(A(2^{-N}[\omega + 2\pi k]) - A(0)\right).$$

then we can use the fact that $\mathcal{P}^N(A) - A(0)$ is 2π-periodic to obtain

$$\mathcal{P}^N(A)(\omega) - A(0) = \sum_{k=-2^{N-1}}^{2^{N-1}-1} \prod_{p=1}^{N} \left|\frac{1}{\sqrt{2}} H(2^{-p}[\omega + 2\pi k])\right|^2 \left(A(2^{-N}[\omega + 2\pi k]) - A(0)\right).$$

which implies that

$$\begin{aligned} \left|\mathcal{P}^N(A)(\omega) - A(0)\right| &\leq \sum_{k=-2^{N-1}}^{2^{N-1}-1} \prod_{p=1}^{N} \left|\frac{1}{\sqrt{2}} H(2^{-p}[\omega + 2\pi k])\right|^2 \left|A(2^{-N}[\omega + 2\pi k]) - A(0)\right| \\ &= \sum_{k=-2^{N-1}}^{2^{N-1}-1} \frac{\left|\hat{\phi}(\omega + 2\pi k)\right|^2}{\left|\hat{\phi}(2^{-N}[\omega + 2\pi k])\right|^2} \left|A(2^{-N}[\omega + 2\pi k]) - A(0)\right|, \end{aligned} \tag{6.37}$$

where we have used 6.35 to obtain the last line above.

Consider the functions $\left|\hat{\phi}(2^{-N}[\omega + 2\pi k])\right|$. If $\omega \in [0, 2\pi]$, then

$$2^{-N}[\omega + 2\pi k] \in \left[k\pi/2^{N-1}, (k+1)\pi/2^{N-1}\right]$$

and

$$\bigcup_{k=-2^{N-1}}^{2^{N-1}-1} \left[k\pi/2^{N-1}, (k+1)\pi/2^{N-1}\right] = [-\pi, \pi].$$

We can now use the fact the $\left|\hat{\phi}\right|$ is bounded below on $[-\pi, \pi]$ to obtain, from 6.37, the inequality

$$\begin{aligned} \left|\mathcal{P}^N(A)(\omega) - A(0)\right| &\leq C \sum_{k=-2^{N-1}}^{2^{N-1}-1} \left|\hat{\phi}(\omega + 2\pi k)\right|^2 \left|A(2^{-N}[\omega + 2\pi k]) - A(0)\right| \\ &\leq C \sum_{|k|\leq q} \left|\hat{\phi}(\omega + 2\pi k)\right|^2 \left|A(2^{-N}[\omega + 2\pi k]) - A(0)\right| \\ &+ C \sum_{|k|>q} \left|\hat{\phi}(\omega + 2\pi k)\right|^2 \left|A(2^{-N}[\omega + 2\pi k]) - A(0)\right|, \quad (6.38) \end{aligned}$$

for some $q \in \mathbb{N}$.

Consider the third series in 6.38. Since

$$\left|A(2^{-N}[\omega + 2\pi k]) - A(0)\right| \leq 2 \|A\|_\infty,$$

the uniform convergence of the series 6.36 implies that we can choose $q \in \mathbb{N}$, independently of N, such that for any $\epsilon > 0$,

$$\begin{aligned} C \sum_{|k|>q} &\left|\hat{\phi}(\omega + 2\pi k)\right|^2 \left|A(2^{-N}[\omega + 2\pi k]) - A(0)\right| \\ &\leq 2C \|A\|_\infty \sum_{|k|>q} \left|\hat{\phi}(\omega + 2\pi k)\right|^2 \leq \epsilon/2. \end{aligned} \tag{6.39}$$

With regard to the second series, the continuity of the function A ensures that for any $\epsilon > 0$, there exists a $\delta > 0$ such that

$$C\left|A(\theta) - A(0)\right| \leq \epsilon/2,$$

whenever $|\theta| \leq \delta$. Now. for $|k| \leq q$ and $\omega \in [0, 2\pi]$, we have

$$\left|2^{-N}[\omega + 2\pi k]\right| \leq 2^{-N+1}\pi(q+1)$$

and hence, we can choose $M \in \mathbb{N}$, independent of k, such that

$$\left|2^{-N}[\omega + 2\pi k]\right| \leq \delta$$

for all $N \geq M$. Since

$$\sum_{|k|\leq q} \left|\hat{\phi}(\omega + 2\pi k)\right|^2 \leq 1,$$

it follow that

$$\begin{aligned} C \sum_{|k|\leq q} \left|\hat{\phi}(\omega + 2\pi k)\right|^2 \left|A(2^{-N}[\omega + 2\pi k]) - A(0)\right| \\ \leq \sum_{|k|\leq q} \left|\hat{\phi}(\omega + 2\pi k)\right|^2 \leq \epsilon/2, \end{aligned} \tag{6.40}$$

whenever $N \geq M$.

If we now return our attention to 6.38, we see that inequalities 6.39 and 6.40 imply that

$$\left\|\mathcal{P}^N(A) - A(0)\right\|_\infty \leq \epsilon$$

for all $N \geq M$. Hence, the desired result follows. ∎

In light of the results of Theorem 6.2, equation 6.20 immediately yields

$$\lim_{m\to-\infty} \left\|P_m \mathcal{G}_n^{-1}\right\|^2 = \frac{1}{\hat{G}_n(0)} \tag{6.41}$$

and since $\hat{\phi}(2\pi k) = \delta_{k,0}$, we have

$$\lim_{m\to-\infty} \left\|P_m \mathcal{G}_n^{-1}\right\| = \frac{1}{|\hat{g}(0)|}.$$

Let us investigate some of the implications of 6.41. Suppose that:

1. $u \in V_n$.

2. The operator $\mathcal{G}_n^{-1}$ is continuous.

3. The observed data is of the form $y + \delta y$, where $y = \mathcal{G}_n u$ and δy represents a small, but unknown error.

If the function

$$u_m^\delta = P_m \mathcal{G}_n^{-1}(y + \delta y), \tag{6.42}$$

in the subspace V_m , $m < n$, is used as an approximation to u, then the squared error in this approximation is given by

$$\begin{aligned} \left\|u - u_m^\delta\right\|^2 &= \|u - P_m u\|^2 + \left\|P_m \mathcal{G}_n^{-1} \delta y\right\|^2 \\ &= \sum_{j=m}^{n-1} \left\|Q_j \mathcal{G}_n^{-1} y\right\|^2 + \left\|P_m \mathcal{G}_n^{-1} \delta y\right\|^2. \end{aligned}$$

Since $\|y\| \leq \|\mathcal{G}_n\| \, \|u\|$, the relative error must satisfy

$$\begin{aligned} \frac{\left\|u - u_m^\delta\right\|^2}{\|u\|^2} &\leq \frac{\sum_{j=m}^{n-1} \|Q_j \mathcal{G}_n^{-1} y\|^2}{\|\mathcal{G}_n^{-1} y\|^2} + \kappa_{n,m}^2 \frac{\|\delta y\|^2}{\|y\|^2}, \\ &= \frac{\left\|P_m^\perp \mathcal{G}_n^{-1} y\right\|^2}{\|P_m^\perp \mathcal{G}_n^{-1} y\|^2 + \|P_m \mathcal{G}_n^{-1} y\|^2} + \kappa_{n,m}^2 \frac{\|\delta y\|^2}{\|y\|^2}. \end{aligned} \tag{6.43}$$

where

$$\kappa_{n,m} = \|\mathcal{G}_n\| \left\|P_m \mathcal{G}_n^{-1}\right\|$$

and $P_m^\perp$ is the orthogonal projection onto $V_m^\perp$. Inequality 6.43 demonstrates that the relative error in the approximation 6.42 depends on two distinct sources. The first source is resolution error which results from projecting onto the subspace V_m. The second source is the data error. Furthermore, the sensitivity of the approximation 6.42 to the error δy is governed by the magnitude of the scalar $\kappa_{n,m}$.

In many instances, the convolution kernel is such that

$$\|\mathcal{G}\| = \max_{\omega \in \mathbb{R}} |\hat{g}(\omega)| = |\hat{g}(0)|,$$

then, since $\|\mathcal{G}_n\| \leq \|\mathcal{G}\|$ and $\hat{G}_n(0) = |\hat{g}(0)|^2$, we have

$$\|\mathcal{G}_n\| = |\hat{g}(0)|.$$

It follows that

$$\lim_{m \to -\infty} \kappa_{n,m} = 1 \tag{6.44}$$

and hence the sensitivity to error in the data decreases as $m \to -\infty$. However.

$$\lim_{m \to -\infty} \frac{\left\|P_m^{\perp} \mathcal{G}_n^{-1} y\right\|^2}{\left\|P_m^{\perp} \mathcal{G}_n^{-1} y\right\|^2 + \left\|P_m \mathcal{G}_n^{-1} y\right\|^2} = 1 \tag{6.45}$$

and hence the error due to resolution increases to its maximum value as $m \to -\infty$. This is a common feature of many inverse problems.

6.4 Multiresolution regularization

In this section, we turn out attention to the MRA based regularization algorithm proposed by Liu in [30]. This algorithm is based upon the functional

$$F_1(u) = \|\mathcal{G}u - y\|^2 + \sum_j \lambda_j \|Q_j u\|^2, \tag{6.46}$$

where Q_j is the orthogonal projection onto the wavelet subspace W_j and $\lambda_j \geq 0$. The minimizing function u_λ of 6.46 can be regarded as an approximate solution of the equation

$$\mathcal{G}u = y, \tag{6.47}$$

the properties of which are determined by the scalars λ_j. Moreover, this algorithm can be thought of as a generalization of the method of Tikhonov regularization, which is based upon the functional

$$F_2(u) = \|\mathcal{G}u - y\|^2 + \lambda \|u\|^2 .$$

In fact, since

$$\|u\|^2 = \sum_j \|Q_j u\|^2 ,$$

it is easy to see that if,

$$\lambda = \cdots = \lambda_{-1} = \lambda_0 = \lambda_1 = \cdots ,$$

then the functionals F_1 and F_2 are identical.

The presence of the term $\lambda_j \|Q_j u\|^2$ in the functional 6.46 serves to prevent the norm of $Q_j u_\lambda$ from being too large. As the scalar λ_j is made larger, the norm $\|Q_j u_\lambda\|^2$ is made smaller.

We now introduce a related functional which allows us to view multiresolution regularization as a special case of C-generalized regularization [3, pages 52–99]. Suppose that the scaling function subspace V_0 represents the coarsest scale of interest, and consider the modified functional

$$F(u) = \|\mathcal{G}u - y\|^2 + \alpha \left(\|P_0 u\|^2 + \sum_{j \geq 0} \lambda_j^2 \|Q_j u\|^2 \right) , \tag{6.48}$$

where $\alpha > 0$. In light of the equation

$$\|P_0 u\|^2 + \sum_{j \geq 0} \lambda_j^2 \|Q_j u\|^2 = \left\| P_0 u + \sum_{j \geq 0} \lambda_j Q_j u \right\|^2 ,$$

we see that if we define the linear operator $C_\lambda : L^2(\mathbb{R}) \to L^2(\mathbb{R})$, by

$$C_\lambda f = P_0 f + \sum_{j \geq 0} \lambda_j Q_j f , \tag{6.49}$$

then the functional 6.48 can be rewritten in the form

$$F(u) = \|\mathcal{G}u - y\|^2 + \alpha \|\mathcal{C}_\lambda u\|^2 . \tag{6.50}$$

Hence, we can consider the multiresolution regularization scheme, defined by 6.48 (or equivalently, by 6.50), to be a special case of C-generalized regularization, where the *smoothing operator* $\mathcal{C}_\lambda$ is the weighted sum of projection 6.49.

The standard theory of C-generalized regularization can now be applied. First of all, we must ensure that the operator $\mathcal{C}_\lambda$ has the following properties:

1. The null spaces of $\mathcal{G}$ and $\mathcal{C}_\lambda$ must satisfy

$$N(\mathcal{G}) \cap N(\mathcal{C}_\lambda) = \{0\}.$$

2. The operator $\mathcal{C}_\lambda$ must be a closed linear operator with a dense domain. That is

$$\overline{D(\mathcal{C}_\lambda)} = L^2(\mathbb{R}).$$

   Furthermore, the range of $\mathcal{C}_\lambda$ must be all of $L^2(\mathbb{R})$, or equivalently

$$R(\mathcal{C}_\lambda) = L^2(\mathbb{R}).$$

3. The set

$$\mathcal{G}(N(\mathcal{C}_\lambda)) = \left\{ f \in L^2(\mathbb{R}) : f = \mathcal{G}h,\ h \in N(\mathcal{C}_\lambda) \right\}$$

   must be closed.

We will say that an operator, $\mathcal{C}_\lambda$, satisfying the conditions above, is an *admissible smoothing operator*.

Since $\mathcal{C}_\lambda$ must be onto $L^2(\mathbb{R})$, none of the scalars λ_j can vanish. Indeed, if $\lambda_j = 0$, then $R(\mathcal{C}_\lambda)$ will be orthogonal to the wavelet subspace W_j, in violation of condition 3 above. Consequently,

$$\|\mathcal{C}_\lambda f\|^2 = \|P_0 f\|^2 + \sum_{j\geq 0} \lambda_j^2 \|Q_j f\|^2 = 0$$

if and only if $f = 0$. Accordingly, $N(\mathcal{C}_\lambda) = \{0\}$ and condition 1 above is satisfied. The domain of $\mathcal{C}_\lambda$ is the set

$$D(\mathcal{C}_\lambda) = \left\{f \in L^2(\mathbb{R}) : \|\mathcal{C}_\lambda f\| < \infty\right\} \tag{6.51}$$

and since, for any $f \in V_n$

$$\|\mathcal{C}_\lambda f\|^2 = \|P_0 f\|^2 + \sum_{j=0}^{n-1} \lambda_j^2 \|Q_j f\|^2 < \infty,$$

we have $V_n \subset D(\mathcal{C}_\lambda)$ for any n. We conclude that $D(\mathcal{C}_\lambda)$ is a dense subset as long as $\lambda_j > 0$.

Now, we must ensure that $\mathcal{C}_\lambda$ is a closed operator. Since $\mathcal{C}_\lambda$ is a bijection onto $L^2(\mathbb{R})$, the Closed Graph Theorem [27, page 292] implies that the inverse operator $\mathcal{C}_\lambda^{-1}$ is continuous. The operator in question is given by

$$\mathcal{C}_\lambda^{-1} f = P_0 f + \sum_{j\geq 0} \lambda_j^{-1} Q_j f, \tag{6.52}$$

which implies that

$$\left\|\mathcal{C}_\lambda^{-1} f\right\|^2 = \|P_0 f\|^2 + \sum_{j\geq 0} \lambda_j^{-2} \|Q_j f\|^2,$$

from which we conclude that the sequence of scalars λ_j^{-1} must be bounded, or equivalently

$$\lambda_j \geq \gamma > 0 \tag{6.53}$$

for some constant γ.

If, in fact, the scalars λ_j are bounded above as well, then the operator $\mathcal{C}_\lambda$ is also continuous. However, in most situations, we will be more interested in the case where

$$\lim_{j\to\infty} \lambda_j = \infty. \tag{6.54}$$

As we shall see, the minimizer u_α^λ of functional 6.50 belongs to the subspace $D(\mathcal{C}_\lambda^2) \subset D(\mathcal{C}_\lambda)$. If the asymptotic behavior of the λ_j is selected appropriately, then *a priori* assumptions about the smoothness of the minimizer u_α^λ can be addressed. For example, if the λ_j satisfy

$$\lambda_j \sim 2^{sj},$$

as $j \to \infty$, then it can be shown that u_α^λ belongs to the Sobolev space

$$H^s(\mathbb{R}) = \left\{ f : \int_{-\infty}^{\infty} (1 + |\omega|^2)^s \left| \hat{f}(\omega) \right|^2 d\omega < \infty \right\}.$$

We point out that, in certain instances, asymptotic conditions for the λ_j, sufficient to ensure that

$$\int_{-\infty}^{\infty} W(\omega) \left| \hat{u}_\alpha^\lambda(\omega) \right|^2 d\omega < \infty, \tag{6.55}$$

are readily established. For instance, suppose that W is a continuous, even, real valued and non-decreasing weight function, and let ψ be an orthonormal wavelet of Meyer type. The expansion of any $f \in L^2(\mathbb{R})$ with respect to this wavelet basis can be written in the form

$$f(t) = \sum_k f_0[k]\phi^{0,k}(t) + \sum_{j\geq 0}\sum_k f^j[k]\psi^{j,k}(t),$$

so that the Fourier transform of f has the expansion

$$\hat{f}(\omega) = F_0(\omega)\hat{\phi}(\omega) + \sum_{j\geq 0} 2^{-j/2} F(\omega/2^j)\hat{\psi}(\omega/2^j).$$

Now, recall that $\text{supp}(\hat{\phi}) = [-\pi-\varepsilon, \pi+\varepsilon]$ and $\text{supp}(\hat{\psi}(\omega)) = [-2(\pi+\varepsilon), -\pi-\varepsilon] \cup [\pi+\varepsilon, 2(\pi+\varepsilon)]$, and accordingly,

$$\begin{aligned}\left|\hat{f}(\omega)\right|^2 &\leq \left|F_0(\omega)\hat{\phi}(\omega)\right|^2 + \sum_{j\geq 0}\left|2^{-j/2}F^j(\omega/2^j)\hat{\psi}(\omega/2^j)\right|^2 \\ &+ 2\left|F_0(\omega)\hat{\phi}(\omega)\right|\left|F^0(\omega)\hat{\psi(\omega)}\right| \\ &+ 2\sum_{j\geq 0}\left|2^{-j/2}F^j(\omega/2^j)\hat{\psi}(\omega/2^j\right|\left|2^{-j-1/2}F^{j+1}(\omega/2^{j+1})\hat{\psi}(\omega/2^{j+1}\right|,\end{aligned}$$

whereupon an application of the Cauchy-Schwartz inequality for sums yields

$$\begin{aligned}\left|\hat{f}(\omega)\right|^2 &\leq \left|F_0(\omega)\hat{\phi}(\omega)\right|^2 + \sum_{j\geq 0}\left|2^{-j/2}F^j(\omega/2^j)\hat{\psi}(\omega/2^j)\right|^2 \\ &+ 2\sqrt{\left|F_0(\omega)\hat{\phi}(\omega)\right|^2 + \sum_{j\geq 0}\left|2^{-j/2}F^j(\omega/2^j)\hat{\psi}(\omega/2^j)\right|^2}\sqrt{\sum_{j\geq 0}\left|2^{-j/2}F^j(\omega/2^j)\hat{\psi}(\omega/2^j)\right|^2} \\ &\leq 3\left(\left|F_0(\omega)\hat{\phi}(\omega)\right|^2 + \sum_{j\geq 0}\left|2^{-j/2}F^j(\omega/2^j)\hat{\psi}(\omega/2^j)\right|^2\right) \qquad (6.56)\end{aligned}$$

We can now use inequality 6.56 to obtain

$$\begin{aligned}\frac{1}{2\pi}\int_{-\infty}^{\infty} W(\omega)\left|\hat{f}(\omega)\right|^2 d\omega &\leq \frac{3}{2\pi}\int_{-\infty}^{\infty} W(\omega)\left|F_0(\omega)\hat{\phi}(\omega)\right| d\omega \\ &+ \frac{3}{2\pi}\sum_{j\geq 0}\int_{-\infty}^{\infty} W(2^j\omega)\left|F^j(\omega)\hat{\psi}(\omega)\right|^2 d\omega \\ &\leq 3W(\pi+\varepsilon)\left\|P_0 f\right\|^2 + 3\sum_{j\geq 0} W(2^{j+1}(\pi+\varepsilon))\left\|Q_j f\right\|^2,\end{aligned}$$

which implies

$$\begin{aligned}\frac{1}{2\pi}\int_{-\infty}^{\infty} W(\omega)\left|\hat{f}(\omega)\right|^2 d\omega &\leq 3W(\pi+\varepsilon)\left\|P_0 f\right\|^2 + 3\sup_{j\geq 0}\left(\frac{W(2^{j+1}(\pi+\varepsilon))}{\lambda_j^2}\right)\sum_{j\geq 0}\lambda_j^2\left\|Q_j f\right\|^2 \\ &\leq 3\max\left(W(\pi+\varepsilon), \sup_{j\geq 0}\left(\frac{W(2^{j+1}(\pi+\varepsilon))}{\lambda_j^2}\right)\right)\left\|C_\lambda f\right\|^2\end{aligned}$$

Since $\left\|C_\lambda u_\alpha^\lambda\right\| < \infty$, it follows that, if W is such that

$$\sup_{j\geq 0}\frac{W(2^{j+1}(\pi+\varepsilon))}{\lambda_j^2} < \infty,$$

then 6.55 holds.

The scalars λ_j are real and bounded below, so that the operator $\mathcal{C}_\lambda$ is self-adjoint and positive definite. The Euler equation of functional 6.50 is

$$\left(\mathcal{G}^*\mathcal{G} + \alpha\mathcal{C}_\lambda^2\right) u_\alpha^\lambda = \mathcal{G}^* y \tag{6.57}$$

and, in light of inequality 6.53, we find that

$$\left\langle \left(\mathcal{G}^*\mathcal{G} + \alpha\mathcal{C}_\lambda^2\right) f, f\right\rangle \geq \left\langle \alpha\mathcal{C}_\lambda^2 f, f\right\rangle \geq \frac{\alpha}{\gamma} \|f\|^2 .$$

It follows that a solution to 6.57 exists for any $y \in L^2(\mathbb{R})$ and can be written as

$$u_\alpha^\lambda = \left(\mathcal{G}^*\mathcal{G} + \alpha\mathcal{C}_\lambda^2\right)^{-1} \mathcal{G}^* y. \tag{6.58}$$

Since

$$\left(\mathcal{A}^*\mathcal{A} + \alpha I\right)^{-1} \mathcal{A}^* = \mathcal{A}^* \left(\mathcal{A}\mathcal{A}^* + \alpha I\right)^{-1}$$

and

$$\mathcal{G}\mathcal{G}^* + \alpha\mathcal{C}_\lambda^2 = \mathcal{C}_\lambda \left\{ \left(\mathcal{G}\mathcal{C}_\lambda^{-1}\right)^* \left(\mathcal{G}\mathcal{C}_\lambda^{-1}\right) + \alpha I \right\} \mathcal{C}_\lambda .$$

equation 6.58 can be rewritten as

$$u_\alpha^\lambda = \mathcal{C}_\lambda^{-2}\mathcal{G}^* \left\{ \left(\mathcal{G}\mathcal{C}_\lambda^{-1}\right) \left(\mathcal{G}\mathcal{C}_\lambda^{-1}\right)^* + \alpha I \right\}^{-1} y \tag{6.59}$$

and hence, $u_\alpha^\lambda \in D\left(\mathcal{C}_\lambda^2\right)$.

6.5 Comparison of C-generalized and generalized solutions

We now turn our attention to questions which concern the relationship between the C-generalized and generalized solution of 6.1. We shall demonstrate that, in appropriate circumstances. there is little difference between the C-generalized and generalized solutions.

Let P be the projection onto the closure of the subspace $R(\mathcal{G})$. It is well known that if $Py \in \mathcal{G}(D(\mathcal{C}_\lambda))$, then there exists a unique function $u_c^\dagger$ such that

$$\lim_{\alpha \to 0^+} \left\| u_\alpha^\lambda - u_c^\dagger \right\| = 0.$$

The function $u_c^\dagger$ is called the *C-generalized solution* of equation 6.1.

Suppose that S_y is the set of all least-squares solutions of 6.1 corresponding to y, that is

$$S_y = \{u : \mathcal{G}u = Py\}, \tag{6.60}$$

then $u_c^\dagger \in S_y \cap D(\mathcal{C}_\lambda)$ is the unique least-squares solution which minimizes the functional

$$\nu(f) = \|\mathcal{C}_\lambda f\|^2. \tag{6.61}$$

In contrast, the generalized solution $u^\dagger \in S_y$ is the unique minimizer of the functional

$$\rho(f) = \|f\|^2 \tag{6.62}$$

and since $\mathcal{G}(D(\mathcal{C}_\lambda)) \subset R(\mathcal{G})$, $u^\dagger$ will exist whenever $u_c^\dagger$ exists. However, $D(\mathcal{C}_\lambda) \neq L^2(\mathbb{R})$ and accordingly, there will be cases where $u^\dagger$ exists, but $u_c^\dagger$ does not.

In general, the C-generalized and generalized solutions will be distinct. However, there are conditions under which $u_c^\dagger$ and $u^\dagger$ will be the same, or nearly so. Obviously, if $\mathcal{G}$ is injective and $y \in R(\mathcal{G})$, then there exists a unique solution of 6.1 and therefore $u_c^\dagger = u^\dagger$. More generally, if the operators $\mathcal{G}$ and $\mathcal{C}_\lambda$ commute, then $u_c^\dagger = u^\dagger$. A justification of this simple, but seemingly unknown fact follows immediately from 6.59, which now implies that $u_\alpha^\lambda \in R(\mathcal{G}^*) \subset N(\mathcal{G})^\perp$ and, since $N(\mathcal{G})^\perp$ is closed, $u_c^\dagger \in N(\mathcal{G})^\perp$. It is well known that the set S_y is the affine subspace

$$S_y = u^\dagger \oplus N(\mathcal{G}),$$

where $u^\dagger$ is the unique least-squares solution in $N(\mathcal{G})^\perp$. We conclude that $u_c^\dagger = u^\dagger$ whenever $\mathcal{G}$ and $\mathcal{C}_\lambda$ commute.

The projection operators P_0 and Q_j do not usually commute with the convolution operator $\mathcal{G}$. Consequently, the operators $\mathcal{C}_\lambda^{-1}$ and $\mathcal{G}$ do not, in general, commute. There is, however, one notable exception. The projections, P_0 and Q_j, corresponding to the Shannon scaling function and wavelet can be expressed in the form

$$(P_0 f)(t) = \int_{-\infty}^{\infty} \mathrm{sinc}(t-\tau) f(\tau)\, d\tau$$

and

$$(Q_j f)(t) = \int_{-\infty}^{\infty} \left\{2^{j+1}\mathrm{sinc}(2^{j+1}(t-\tau)) - 2^j \mathrm{sinc}(2^j(t-\tau))\right\} f(\tau)\, d\tau,$$

both of which commute with the convolution operator $\mathcal{G}$. Therefore, in the Shannon case, the approximation u_α^λ converges to $u_c^\dagger = u^\dagger$ as $\alpha \to 0^+$.

Before we consider further comparisons of the function $u_c^\dagger$ and $u^\dagger$, we consider a case where minimizers u_α^λ and u_α are close in the sense of the norm. In practice, we seek to approximate the generalized solutions $u^\dagger$ and $u_c^\dagger$ through the use of the corrupted data $y + \delta y$. In some cases, $u^\dagger$ and $u_c^\dagger$ will not be well-defined, as it may happen that

$$P\delta y \notin \mathcal{G}\left(D(\mathcal{C}_\lambda)\right).$$

Even if $P\delta y \in \mathcal{G}\left(D(\mathcal{C}_\lambda)\right)$, the approximations formed from $y + \delta y$ can differ substantially from the generalized solutions $u^\dagger$ and $u_c^\dagger$. In such cases, the minimizers u_α and u_α^λ can be used as approximations of $u^\dagger$ and $u_c^\dagger$ respectively. Intuitively, one expects that if the smoothing operator $\mathcal{C}_\lambda$ is close to the identity I, then u_α^λ will be close to u_α. The next theorem illustrates one case in which this happens.

Theorem 6.3 *Suppose that the self-adjoint operators $\mathcal{C}_N$ are admissible smoothing operators and that*

$$\lim_{N\to\infty} \left\|\mathcal{C}_N^{-1} - I\right\| = 0.$$

Let u_α^N be the solution of the Euler equation

$$\left(\mathcal{G}^*\mathcal{G} + \alpha\mathcal{C}_N^2\right) u_\alpha^N = \mathcal{G}^* y, \tag{6.63}$$

while u_α denotes the unique solution of

$$\left(\mathcal{G}^*\mathcal{G} + \alpha I\right) u_\alpha = \mathcal{G}^* y, \tag{6.64}$$

then

$$\lim_{N\to\infty} \left\|u_\alpha^N - u_\alpha\right\| = 0$$

for any $\alpha > 0$

Proof. Equations 6.63 and 6.64 yield the equation

$$u_\alpha^N - u_\alpha = \left\{\left(\mathcal{G}^*\mathcal{G} + \alpha\mathcal{C}_N^2\right)^{-1} - \left(\mathcal{G}^*\mathcal{G} + \alpha I\right)^{-1}\right\} \mathcal{G}^* y. \tag{6.65}$$

Let us use the notation

$$\mathcal{G}_\alpha = \mathcal{G}^*\mathcal{G} + \alpha I$$

and

$$\mathcal{G}_N = \mathcal{G}\mathcal{C}_N^{-1},$$

then we can write

$$\begin{aligned} \mathcal{G}^*\mathcal{G} + \alpha\mathcal{C}_N^2 &= \mathcal{C}_N\left(\mathcal{G}_N^*\mathcal{G}_N + \alpha I\right)\mathcal{C}_N \\ &= \mathcal{C}_N\left(\left(\mathcal{G}_N^*\mathcal{G}_N - \mathcal{G}^*\mathcal{G}\right)\mathcal{G}_\alpha^{-1} + I\right)\mathcal{G}_\alpha\mathcal{C}_N. \end{aligned}$$

If we let

$$\gamma_N = \|\mathcal{G}_N^*\mathcal{G}_N - \mathcal{G}^*\mathcal{G}\|,$$

then, since $\mathcal{C}_N^{-1}$ converges uniformly to I, $\gamma_N \to 0$ as $N \to \infty$. Furthermore, for any fixed $\alpha > 0$, we can choose $N_1 \in \mathbb{N}$ so that

$$\|(\mathcal{G}_N^*\mathcal{G}_N - \mathcal{G}^*\mathcal{G})\,\mathcal{G}_\alpha\| \le \gamma_N/\alpha < 1. \tag{6.66}$$

for all $N \ge N_1$. It follows that, for all $N \ge N_1$, the operator

$$\mathcal{A}_N = (\mathcal{G}_N^*\mathcal{G}_N - \mathcal{G}^*\mathcal{G})\,\mathcal{G}_\alpha^{-1} + I$$

is invertible and accordingly, 6.65 can be written in the form

$$u_\alpha^N - u_\alpha = \left(\mathcal{C}_N^{-1}\mathcal{G}_\alpha^{-1}\mathcal{A}_N^{-1}\mathcal{C}_N^{-1} - \mathcal{G}_\alpha^{-1}\right)\mathcal{G}^*y \tag{6.67}$$

In view of equation 6.67 and the fact that

$$\lim_{N\to\infty} \|\mathcal{B}_N\mathcal{D}_N - \mathcal{B}\mathcal{D}\| = 0$$

whenever

$$\lim_{N\to\infty} \|\mathcal{B}_N - \mathcal{B}\| = \lim_{N\to\infty} \|\mathcal{D}_N - \mathcal{D}\| = 0,$$

we need only show that

$$\lim_{N\to\infty} \left\|\mathcal{A}_N^{-1} - I\right\| = 0. \tag{6.68}$$

Recall that $N \ge N_1$ and hence inequality 6.66 holds. It follows that $\mathcal{A}_N^{-1}$ admits the Neumann series

$$\mathcal{A}_N^{-1} = \sum_{p=0}^{\infty}(-1)^p\left((\mathcal{G}_N^*\mathcal{G}_N - \mathcal{G}^*\mathcal{G})\,\mathcal{G}_\alpha^{-1}\right)^p,$$

which implies that

$$\begin{aligned} \left\|\mathcal{A}_N^{-1} - I\right\| &\leq \sum_{p=1}^{\infty}(-1)^p \left\|\left(\mathcal{G}_N^*\mathcal{G}_N - \mathcal{G}^*\mathcal{G}\right)\mathcal{G}_\alpha^{-1}\right\|^p \\ &\leq \gamma_N/\alpha \sum_{p=0}^{\infty}(\gamma_N/\alpha)^p = \frac{\gamma_N}{\alpha - \gamma_N} \end{aligned}$$

and since $\gamma_N \to 0$ as $N \to \infty$, we see that $\mathcal{A}_N^{-1}$ converges uniformly to I as $N \to \infty$.

Finally, from 6.67, we obtain

$$\lim_{N\to\infty} \left\|u_\alpha^N - u_\alpha\right\| \leq \left\|\mathcal{G}^*y\right\| \lim_{N\to\infty} \left\|\mathcal{C}_N^{-1}\mathcal{G}_\alpha^{-1}\mathcal{A}_N^{-1}\mathcal{C}_N^{-1} - \mathcal{G}_\alpha^{-1}\right\| = 0.$$

as required. ▮

The condition $\mathcal{C}_N^{-1} \to I$ uniformly is quite restrictive. Unfortunately, it is not clear that the hypothesis of the previous theorem can be weakened. Furthermore, if $\mathcal{C}_N$ is a weighted sum of projections of the type defined below, then we can not exhibit examples that satisfy the required conditions of the theorem for all $f \in L^2(\mathbb{R})$.

However, there are sets of functions for which a weighted sum of projections $\mathcal{C}_N^{-1}$ will in fact converge uniformly to the identity. Suppose that

$$\mathcal{C}_N = P_0 + \sum_{j\geq 0} \lambda_j^N Q_j,$$

where

$$\lim_{N\to\infty} \lambda_j^N = 1$$

and that $f \in V_n$. It follows that

$$\begin{aligned} \left\|\left(\mathcal{C}_N^{-1} - I\right) f\right\|^2 &= \sum_{j=0}^{n-1} \left(\frac{1-\lambda_j^N}{\lambda_j^N}\right)^2 \left\|Q_j f\right\|^2 \\ &\leq \max_{0\leq j\leq n-1} \left(\frac{1-\lambda_j^N}{\lambda_j^N}\right)^2 \left\|f\right\|^2 \end{aligned}$$

and therefore the C_N^{-1} converges to I uniformly for all f in the scaling function subspace V_n.

Let us return our attention to a comparison of the functions $u^\dagger$ and $u_c^\dagger$. In the next theorem, we examine a particular situation for which the generalized solutions are close in a weak sense.

Theorem 6.4 *Suppose that the operator C_N is given by*

$$C_N = P_N + \sum_{j \geq N} \lambda_j Q_j,$$

$N \in \mathbb{N}$, where $\{\lambda_j : j \in \mathbb{N}\}$ is a non-decreasing sequence of real numbers, such that $\lambda_j \geq 1$. Assume that $Py \in \mathcal{G}(D(C_N))$ and let $u_N^\dagger \in S = u^\dagger \oplus N(\mathcal{G})^\perp$ be the unique least-squares solution which minimizes the functional

$$\nu_N(f) = \|C_N f\|^2,$$

then, for any $h \in L^2(\mathbb{R})$, there exists a subsequence

$$\left\{u_{N_k}^\dagger : k \in \mathbb{N}\right\}$$

such that

$$\lim_{k \to \infty} \langle u_{N_k}^\dagger - u^\dagger, h \rangle = 0. \tag{6.69}$$

Proof. First, we note that the domain of C_N does not depend on N. Indeed, since

$$\sum_{j=N}^{N+M-1} \lambda_j^2 \|Q_j f\|^2 < \infty,$$

we see that, for any $M, N \in \mathbb{N}$

$$D \equiv D(C_N) = D(C_{N+M}). \tag{6.70}$$

Now, for any $f \in D$, we have

$$\begin{aligned} \|C_N f\|^2 &= \|P_N f\|^2 + \sum_{j \geq N} \lambda_j^2 \|Q_j f\|^2 \\ &\leq \|P_N f\|^2 + \|Q_N f\|^2 + \sum_{j \geq N+1} \lambda_j^2 \|Q_j f\|^2 = \|C_{N+1} f\|^2 \end{aligned}$$

and since $u_N^\dagger \in S \cap D$ is the unique function minimizing the functional ν_N, we have

$$\left\|C_N u_N^\dagger\right\| \geq \left\|C_{N+1} u_N^\dagger\right\| > \left\|C_{N+1} u_{N+1}^\dagger\right\| \geq \cdots, \tag{6.71}$$

which implies that the numbers $\left\|C_N u_N^\dagger\right\|$ form a decreasing sequence. Consequently

$$\left\|C_N u_N^\dagger\right\| \leq \left\|C_0 u_0^\dagger\right\| \tag{6.72}$$

and therefore, there exists a subsequence $\{C_{N_k} u_{N_k}^\dagger : k \in \mathbb{N}\}$ which converges weakly to some $u \in L^2(\mathbb{R})$.

We now show that the subsequence $\{u_{N_k}^\dagger : k \in \mathbb{N}\}$ also converges weakly and has the same limit u. The sequence $\{u_N^\dagger : N \in \mathbb{N}\}$ is bounded. Indeed, since $\lambda_j \geq 1$, for any $f \in D$

$$\|C_N f\| \geq \|f\|$$

from which we obtain the inequality

$$\left\|u_N^\dagger\right\| \leq \left\|C_N u_N^\dagger\right\| \leq \left\|C_0 u_0^\dagger\right\|. \tag{6.73}$$

Assume, without loss of generality, that $M \leq N$, then $P_M C_N u_N^\dagger = P_M u_N^\dagger$. Since P_M is continuous, we find that, for all $h \in L^2(\mathbb{R})$

$$\lim_{k \to \infty} \langle P_M(u_{N_k}^\dagger - u), h \rangle = 0.$$

Furthermore, since

$$\begin{aligned} \left|\langle u_{N_k}^\dagger - u, h \rangle\right| &\leq \left|\langle P_M(u_{N_k}^\dagger - u), h \rangle\right| + \left|\langle u_{N_k}^\dagger - u, P_M h - h \rangle\right| \\ &\leq \left|\langle P_M(u_{N_k}^\dagger - u), h \rangle\right| + \left(\left\|C_0 u_0^\dagger\right\| + \|u\|\right) \|P_M h - h\| \end{aligned}$$

and

$$\lim_{M\to\infty} \|P_M h - h\| = \sum_{j\geq M} \|Q_j h\|^2 = 0.$$

we see that

$$\lim_{k\to\infty} \left|\langle u^{\dagger}_{N_k} - u, h\rangle\right| = 0. \tag{6.74}$$

The sequence $\{u^{\dagger}_N : N \in \mathbb{N}\}$ is contained in the set S of all least-squares solutions. As we have already seen, S is a closed affine subspace and is therefore weakly closed. It now follows that the weak limit u, of the subsequence $\{u^{\dagger}_{N_k} : N \in \mathbb{N}\}$, is a least-squares solution.

To complete the proof, we need to show that the function u is in fact the generalized solution $u^{\dagger}$. We begin by showing that $f^{\dagger}_N = \mathcal{C}_N u^{\dagger}_N$ is the generalized solution of the equation

$$\mathcal{G}\mathcal{C}_N^{-1} f = y. \tag{6.75}$$

Denote by S_N the set of all least-squares solutions of 6.75. If $f \in S_N$, then f must satisfy the Euler equation

$$\left(\mathcal{G}\mathcal{C}_N^{-1}\right)^* \left(\mathcal{G}\mathcal{C}_N^{-1}\right) f = \left(\mathcal{G}\mathcal{C}_N^{-1}\right)^* y,$$

or equivalently

$$\mathcal{C}_N^{-1}\left(\mathcal{G}^*\mathcal{G}\mathcal{C}_N^{-1} f - \mathcal{G}^* y\right) = 0$$

and since $N(\mathcal{C}_N^{-1}) = \{0\}$, every $f \in S_N$ must satisfy

$$\mathcal{G}^*\mathcal{G}\mathcal{C}_N^{-1} f = \mathcal{G}^* y. \tag{6.76}$$

In view of 6.76, for any $f \in S_N$, the function $\mathcal{C}_N^{-1} f$ is a least-squares solution in $S \cap D$ and, since $\mathcal{C}_N$ is well-defined on $S \cap D$, for any $u \in S \cap D$, $\mathcal{C}_N u$ is a least-squares solution of 6.75. Furthermore, since for any $u \in S \cap D$,

$$\|f\| = \|\mathcal{C}_N u\|$$

we conclude that $f_N^\dagger$ is the generalized solution of 6.75.

The function $f_N^\dagger$ is the unique element of S_N that belongs to the subspace $N(\mathcal{G}\mathcal{C}_N^{-1})^\perp$. Furthermore, the subspace $R(\mathcal{C}_N^{-1}\mathcal{G}^*)$ is dense in $N(\mathcal{G}\mathcal{C}_N^{-1})^\perp$, which implies that for any sequence $\{\epsilon_N : N \in \mathbb{N}\}$ of positive numbers, there exists a function $v_N \in R(\mathcal{G}^*)$ such that

$$\left\| f_N^\dagger - \mathcal{C}_N^{-1} v_N \right\| \leq \epsilon_N.$$

We choose the sequence $\{\epsilon_N : N \in \mathbb{N}\}$ so that

$$\lim_{N\to\infty} \epsilon_N = 0$$

and show that $\{\mathcal{C}_{N_k}^{-1} h_{N_k} : k \in \mathbb{N}\}$ converges weakly to u, the weak limit of $\{f_{N_k}^\dagger : k \in \mathbb{N}\}$. Since

$$\begin{aligned} \left|\langle \mathcal{C}_{N_k}^{-1} h_{N_k} - u, v\rangle\right| &\leq \left|\langle \mathcal{C}_{N_k}^{-1} f_{N_k}^\dagger - u, v\rangle\right| + \left|\langle f_{N_k}^\dagger - u, v\rangle\right| \\ &\leq \|h\|\, \epsilon_{N_k} + \left|\langle f_{N_k}^\dagger - u, v\rangle\right|, \end{aligned}$$

we see immediately that

$$\lim_{k\to\infty} \left|\langle \mathcal{C}_{N_k}^{-1} h_{N_k}, v\rangle\right|$$

for any $v \in L^2(\mathbb{R})$.

Notice that

$$P_M \mathcal{C}_N^{-1} h_N = P_M h_N$$

and once again, it can be shown that $\{h_{N_k} : k \in \mathbb{N}\}$ converges weakly to u. Finally, since $h_N \in N(\mathcal{G})^\perp$, we conclude that u must be the generalized solution $u^\dagger$. ▮

We mention that is not immediately obvious whether or not the previous theorem can be extended to address strong convergence.

Chapter 7

Conclusions

In this thesis, we have considered certain aspects of the applications of wavelet analysis to the problem of deconvolution. In particular, we have addressed some of the basic theoretical considerations of the problem deconvolution with wavelet bases. This is merely a beginning. Although the properties of wavelet bases are attractive and the empirical results found in the literature are encouraging, a fair evaluation of wavelet analysis, with regard to inverse problems, is ongoing.

The results we have presented raise many questions and, if some of these questions are answered, the perhaps the aforementioned evaluation will be more complete. In Chapter 5, the properties of the function $\hat{G}_n$ were examined in some detail. This examination permitted us to present results concerning the strong and weak invertibility of the operator $\mathcal{G}_n \equiv \mathcal{G}|_{V_n}$. An examination of the behavior of $\hat{G}_n$ lead to a convergence result in the case where the convolution kernel satisfies

$$|\hat{g}(\omega)| > 0$$

and

$$|\hat{g}(\omega_1)| \leq |\hat{g}(\omega_2)|$$

for all ω_1 and ω_2 satisfying $|\omega_1| \geq |\omega_2| \geq \Omega > 0$. This result is somewhat restrictive and further investigation of the function $\hat{G}_n$ may lead to results that are valid for a larger class of kernels g.

A typical characteristic of many inverse problems is the trade off between accuracy of approximation and sensitivity to noise in the data. This property leads us to a question regarding the choice of resolution n. That is, given the approximation

$$u_n^\delta = \sum_k \langle y^\delta(t), \tilde{\xi}_n(t - k/2^n)\rangle \phi^{n,k},$$

where $y^\delta = y + \delta$, we need to be able to choose n so that the error of approximation

$$e_n^1 = \|u - u_n^\delta\|$$

and the error due to noise

$$e_n^2 = \|\mathcal{G}_n^{-1}\delta\|$$

are both as small as is possible. Further examination of the dependence of $\hat{G}_n$ on n could lead to a method to choose the resolution which parallels Morozov's Discrepancy Principle (see [29, page 228]).

Of course there is interest in the extension of the results of Chapter 5 to include other common integral operators. However, such work will depend on the kernels of these operators and complications, not found in the current work, may arise.

The work done in the first few sections of Chapter 6 is closely related to the some of the work done in Chapter 5 and hence, similar questions arise. Once again, a method for the choice of resolution level has not been discussed. In this case, a more in depth examination of the operator $\mathcal{P}$ is needed. We point out that the approach considered in Chapter 6 may have some advantages with regard to the choice of resolution level. Recall that the scaling functions $\phi^{n,k}$ can be thought of as low pass filters. Roughly speaking, the width of these filters is proportional to 2^n and hence the pass band of $\phi^{n,k}$ doubles as the resolution is increased from n to $n+1$. Depending on the g and the noise level $\|\delta\|$, is may turn out that the optimal width for the inverse problem at hand will lie in between $C2^n$ and $C2^{n+1}$. C a constant depending on ϕ.

In such a case, the work found in [10] may be of use. In this paper, a splitting technique, which allows for a finer partition of the frequency axis, than the partition induced by the functions $\phi^{n,k}$, is introduced. An incorporation of the ideas presented in [10] with our own work in Chapter 6 may lead to a method for the choice of resolution which can be fine-tuned to the noise and the kernel of the problem under consideration.

With regard to the latter part of Chapter 6, we have presented results which concern a comparison of the methods of C-generalized and ordinary regularization. Moreover, results about the corresponding generalized solutions are also given. These results are of a very general nature and any future research should attempt to exploit the properties of wavelet bases as well as the properties of any particular kernel g under consideration.

Another possibility for future research into multiresolution based regularization is to make use of the idea of a time frequency localization operator (see [14]). Recall that the operator $\mathcal{C}_\lambda$ is given by

$$\mathcal{C}_\lambda = P_0 + \sum_{j\geq 0} \lambda_j Q_j.$$

If we choose to allow for some type of spatial discrimination, then one possible generalization of the operator above is

$$\mathcal{C}_\lambda f = \sum_k \lambda_k \langle f, \phi^{0,k}\rangle \phi^{0,k} + \sum_{j\geq 0}\sum_k \lambda_k^j \langle f, \psi^{j,k}\rangle \psi^{j,k}.$$

which bears some similarity to the time frequency localization operators introduced by Daubechies.

If the sequence $\{\lambda_k : k \in \mathbb{Z}\}$ and $\{\lambda_k^j : j \geq 0, k \in \mathbb{Z}\}$ satisfy

$$\sum_k \lambda_k^{-2} < \infty$$

and

$$\sum_{j\geq 0} \sum_k \left(\lambda_k^j\right)^{-2} < \infty,$$

then it can be shown that $\mathcal{C}_\lambda^{-1}$ is a Hilbert-Schmidt operator and is therefore compact. The minimizer u_α^λ, of the functional

$$F(u) = \|\mathcal{G}u - y\|^2 + \alpha \|\mathcal{C}_\lambda u\|^2$$

must satisfy the Euler equation

$$\left(\mathcal{G}^*\mathcal{G} + \alpha\mathcal{C}_\lambda^2\right) u_\alpha^\lambda = \mathcal{G}^* y$$

or equivalently

$$\mathcal{C}_\lambda \left((\mathcal{G}\mathcal{C}_\lambda^{-1})^*(\mathcal{G}\mathcal{C}_\lambda^{-1}) + \alpha I\right) \mathcal{C}_\lambda u_\alpha^\lambda = \mathcal{G}^* y. \tag{7.1}$$

Since $\mathcal{G}$ is bounded, we have that $\mathcal{G}\mathcal{C}_\lambda^{-1}$ is compact and this leads to the possibility of using a singular value decomposition to solve equation 7.1

Finally, we point out that in [14], time frequency localization operators are defined through the use of the continuous wavelet transform. It is possible that we could define the operator $\mathcal{C}_\lambda$ so that its inverse is a time frequency localization operator of the type discussed in [14]. This approach may have the advantage that difficulties, such as the lack of translation invariance of wavelet bases, are avoided.

Appendix A

Daubechies scaling functions

In this appendix, we will show that if ϕ is a Daubechies scaling function, then

$$\left|\hat{\phi}(\omega)\right| \geq C > 0$$

for all $\omega \in [-\pi, \pi]$. This proof makes use of the fact that if $H(\omega)$ is the DFT of the two scale sequence $\{h_k : k = 0, \ldots 2N-1\}$, corresponding to ϕ, then $H(\omega)$ is bounded below on the interval $\omega \in [-\pi/2, \pi/2]$. We will also examine the behavior the function $\left|\hat{\phi}\right|$ for ω near 0 and show that

$$1 - \left|\hat{\phi}(\omega)\right|^2 = D_N \omega^{2N} + O(\omega^{2N+2}),$$

for some positive constant D_N, where $N \in \mathbb{N}$ is the number of vanishing moments of the corresponding wavelet.

In [14, page 171], it is shown that the modulus of $H(\omega)$ can be written as

$$|H(\omega)|^2 = 2\cos^{2N}(\omega/2) P_N(\sin^2(\omega/2)), \tag{A.1}$$

where P_N is a polynomial of degree $N-1$, $N \geq 2$, given by

$$P_N(x) = \sum_{k=0}^{N-1} \binom{N-1+k}{k} x^k. \tag{A.2}$$

There are more general choices for the polynomial A.2. However, these choices lead to longer discrete filters[1] and hence, the choice A.2 is generally preferred. The following lemma allows us to conclude that if $|H(\omega)|$ is of the form A.1, then $|H(\omega)| \geq D > 0$ when $\omega \in [-\pi/2, \pi/2]$.

Lemma A.1 *If $|H(\omega)|^2$ is as in A.1, then*

$$\frac{d}{d\omega}|H(\omega)|^2 \leq 0$$

for $\omega \in [0, \pi]$.

Proof. If we let $x = \sin^2(\omega/2)$, then since

$$\frac{dx}{d\omega} = \frac{1}{2}\sin(\omega) \geq 0,$$

for all $\omega \in [0, \pi]$, it is enough to show that the polynomial

$$Q(x) = (1-x)^N P_N(x)$$

is such that

$$Q'(x) \leq 0$$

for $x \in [0, 1]$.

We have

$$\begin{aligned} Q'(x) &= -N(1-x)^{N-1}P_N(x) + (1-x)^N P_N'(x) \\ &= (1-x)^{N-1}\left((1-x)P_N'(x) - NP_N(x)\right) \end{aligned}$$

[1]That is, longer two scale sequences.

and, from A.2, we obtain

$$\begin{aligned} Q'(x) &= (1-x)^{N-1}\left[\sum_{k=1}^{N-1} k \binom{N-1+k}{k} x^{k-1} \right. \\ &- \left. \sum_{k=0}^{N-1} k \binom{N-1+k}{k} x^{k} - N\sum_{k=0}^{N-1} \binom{N-1+k}{k} x^{k}\right] \\ &= (1-x)^{N-1}\left[-\binom{2(N-1)}{N-1} x^{N-1} \right. \\ &+ \left. \sum_{k=0}^{N-2}\left\{(k+1)\binom{N+k}{k+1} - (N+k)\binom{N-1+k}{k}\right\} x^{k}\right]. \end{aligned}$$

Since

$$\begin{aligned} (k+1)\binom{N+k}{k+1} &- (N+k)\binom{N-1+k}{k} = (k+1)\frac{(N+k)!}{(k+1)!(N-1)!} \\ &- (N+k)\frac{(N+k-1)!}{k!(N-1)!} = 0, \end{aligned}$$

we find that

$$Q'(x) = -\binom{2(N-1)}{N-1} x^{N-1}(1-x)^{N-1} \leq 0,$$

for $x \in [0,1]$ as required. ∎

In light of Lemma A.1, we see that $|H(\omega)|$ is a non-increasing function in the interval $[0,\pi]$. Since the h_k, $k = 0, 1, \ldots 2N-1$, are real, the function $|H(\omega)|$ is even and

$$|H(\omega)| \geq |H(\pi/2)| = 1.^2 \tag{A.3}$$

We can now prove the following:

[2]The equality follows from the identities $|H(\pi/2)| = |H(3\pi/2)|$ and $|H(\omega)|^2 + |H(\omega+\pi)|^2 = 2$.

Lemma A.2 *If ϕ is a Daubechies scaling function then*

$$\left|\hat{\phi}(\omega)\right| \geq C > 0, \tag{A.4}$$

for all $\omega \in [-\pi, \pi]$.

Proof. A proof follows easily from a contradiction. Suppose that $\hat{\phi}$ has at least one zero $\omega^* \in [-\pi, \pi]$. We use the two scale equation to obtain

$$\hat{\phi}(\omega^*) = \frac{1}{\sqrt{2}} H(\omega^*/2)\hat{\phi}(\omega^*/2)$$

and, since $H(\omega)$ has no zeros in $[-\pi/2, \pi/2]$, we conclude that $\hat{\phi}(\omega^*/2) = 0$. We continue this process and find that

$$\hat{\phi}(\omega^*/2^p) = 0$$

for any $p \in \mathbb{N}$. Since $\hat{\phi}$ is continuous,

$$\hat{\phi}(0) = 0. \tag{A.5}$$

However, ϕ is unimodular with $\hat{\phi}(0) = 1$ and A.5 is a contradiction. We conclude that $\hat{\phi}$ must satisfy the inequality A.4. ∎

We now turn our attention to the behavior of the function $\hat{\phi}$ for ω near zero.

Lemma A.3 *If ϕ is a Daubechies scaling function, then*

$$1 - \left|\hat{\phi}(\omega)\right|^2 = \frac{C_N}{N4^N(4^N-1)}\omega^{2N} + O(\omega^{2N+2}), \tag{A.6}$$

where

$$C_N = \begin{pmatrix} 2(N-1) \\ N-1 \end{pmatrix}.$$

Proof. Recall that

$$Q'(x) = -C_N x^{N-1}(1-x)^{N-1} = -C_N x^{N-1} + O(x^N)$$

and, since $Q(0) = 1$, it follows that

$$Q(x) = 1 - \frac{C_N}{N} x^N + O(x^{N+1}). \tag{A.7}$$

Let $\tilde{H}(\omega) = 1/2\,|H(\omega)|^2$. Since $\tilde{H}(\omega) = Q(\sin^2(\omega/2))$, we have

$$\tilde{H}(\omega) = 1 - \frac{C_N}{N4^N}\omega^{2N} + O(\omega^{2N+2}). \tag{A.8}$$

If we let $\Phi(\omega) = \left|\hat{\phi}(\omega)\right|^2$, then

$$\Phi(\omega) = \tilde{H}(\omega/2)\Phi(\omega/2)$$

and the Leibniz rule for the differentiation of products yields

$$\begin{aligned}\Phi^{(p)}(\omega) = \frac{1}{2^p}(\tilde{H}(\omega/2)\Phi^{(p)}(\omega/2) + p\tilde{H}'(\omega/2)\Phi^{(p-1)}(\omega/2) &+ \cdots \\ &+ \tilde{H}^{(p)}(\omega/2)\Phi(\omega/2).)\end{aligned}$$

which implies

$$\Phi^{(p)}(0) = \frac{1}{2^p}\left(\Phi^{(p)}(0) + p\tilde{H}'(0)\Phi^{(p-1)}(0) + \cdots + \tilde{H}^{(p)}(0)\right).$$

If we solve the above for $\Phi^{(p)}(0)$, then we obtain

$$\Phi^{(p)}(0) = \frac{1}{2^p - 1}\left(p\tilde{H}'(0)\Phi^{(p-1)}(0) + \cdots + \tilde{H}^{(p)}(0)\right). \tag{A.9}$$

In view of A.8, we can use A.9 to conclude that

$$\Phi^{(p)}(0) = 0$$

for all $p = 1, 2, \ldots, 2N-1$ and that

$$\Phi^{(2N)}(0) = \frac{\tilde{H}^{(2N)}(0)}{4^N - 1} = -\frac{(2N)!C_N}{N4^N(4^N - 1)}$$

and therefore,

$$\Phi(\omega) = 1 - \frac{C_N}{N4^N(4^N - 1)}\omega^{2N} + O(\omega^{2N+2}),$$

from which we obtain A.6. ∎

Bibliography

[1] R. S. Anderssen. The linear functional strategy for improperly posed problems. volume 77 of *International Series of Numerical Mathematics*, pages 11–30. Basel, 1986. Birkhauser.

[2] N. Aronszajn. Theory of reproducing kernels. *Transactions of the American Mathematical Society*, 68:337–404, 1950.

[3] M. Bertero. Regularization methods for linear inverse problems. In G. Talenti, editor, *Inverse Problems*, number 1225 in Lecture Notes in Mathematics, pages 52–167, Berlin, 1986. Springer-Verlag.

[4] G. Beylkin. On representation of operators in bases of compactly supported wavelets. *SIAM Journal of Numerical Analysis*, 6(6):1716–174, 1992.

[5] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms I. *Communications on Pure and Applied Mathematics*, XLIV:141–183, 1991.

[6] Ronald N. Bracewell. *The Fourier Transform and Its Applications*. McGraw-Hill, New York, 1984.

[7] D. C. Champeney. *A Handbook of Fourier Theorems*. Cambridge University Press, Cambridge, 1987.

[8] Bor-Shen Chen and Chin-Wei Lin. Multiscale Wiener filter for the restoration of fractal signals: Wavelet filter bank approach. *IEEE Transactions on Signal Processing*, 42(11):2972–2982, 1994.

[9] Charles K. Chui. *An Introduction to Wavelets*. Academic Press, San Diego, 1992.

[10] A. Cohen and Ingrid Daubechies. On the instability of biorthogonal wavelet packets. *SIAM Journal of Mathematical Analysis*, 24(5):1340–1354, 1993.

[11] A. Cohen and Ingrid Daubechies. Orthonormal bases of compactly supported wavelets III. better frequency resolution. *SIAM Journal of Mathematical Analysis*, 24(2):520–527, 1993.

[12] Ingrid Daubechies. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, VXLI:909–996, 1988.

[13] Ingrid Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*, 36(5):961–1005, 1990.

[14] Ingrid Daubechies. *Ten Lectures on Wavelets*. SIAM, Philadelphia, 1992.

[15] Stanley R. Deans. *The Radon Transform and Some of Its Applications*. Wiley, New York, 1983.

[16] Joe DeStefano and Tim Olson. Wavelet localization of the Radon transform in even dimensions. preprint, July 1992.

[17] John W. Dettman. *Applied Complex Variables*. Dover. New York. 1984.

[18] David L. Donoho. Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition. preprint. April 1992.

[19] Patrick Flandrin. On the spectrum of fractional Brownian motion. *IEEE Transactions on Information Theory*, 35(1):197–199, 1989.

[20] Gerald B. Folland. *Fourier Analysis and Its Applications*. Wadsworth. Belmont, 1992.

[21] C. W. Groetsch. *Generalized Inverses of Linear Operators*. Marcel Dekker. New York. 1977.

[22] C. W. Groetsch. *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*. Pitman Publishing, London. 1984.

[23] C. W. Groetsch. *Inverse Problems in the Mathematical Sciences*. Vieweg. Wiesbaden. 1993.

[24] Christopher E. Heil and David F. Walnut. Continuous and discrete wavelet transforms. *SIAM Review*, 31(4):628–666. 1989.

[25] M. Holschneider. Inverse Radon transforms through inverse wavelet transforms. *Inverse Problems*, 7:853–861, 1991.

[26] Yitzhak Katznelson. *An Introduction to Harmonic Analysis*. Dover, New York. 1976.

[27] Erwin Kreyszig. *Introductory Functional Analysis With Applications*. Krieger. Malabar, 1989.

[28] Wayne M. Lawton. Necessary and sufficient conditions for construcing orthonormal wavelet bases. *Journal of Mathematical Physics*, 32(1):57–61. 1991.

[29] L. P. Lebedev, I. I. Vorovich, and G. M. L. Gladwell. *Functional Analysis Applications in Mechanics and Inverse Problems*. Kluwer, Dordrecht, 1996.

[30] Jun Liu. A multiresolution method for distributed parameter estimation. *SIAM Journal of Scientific Computing*. kak(2):389–405. 1993.

[31] Jun Liu. A sensitivity analysis for least-squares ill-posed problems using the Haar basis. *SIAM Journal of Numerical Analysis*, 31(5):1486–1496, 1994.

[32] Jun Liu, Beatrice Guerrier, and Christine Benard. A sensitivity decomposition for the regularized solution of inverse heat conduction problems by wavelets. *Inverse Problems*, 11:1177–1187, 1995.

[33] Jian Lu. Dennis Healy, Jr., and John B. Weaver. Signal recovery and wavelet reproducing kernels. *IEEE Transactions on Signal Processing*. 42(7):1845–1849. 1994.

[34] Stephane G. Mallat. Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbb{R})$. *Transactions of the American Mathematical Society*, 315(1):69–87, 1989.

[35] Elias Masry. The wavelet transform of stochastic processes with stationary increments and its application to fractional Brownian motion. *IEEE Transactions on Infomation Theory*, 39(1):260–264, 1993.

[36] Yves Meyer. *Wavelets and Operators*. Cambridge University Press, Cambridge, 1992.

[37] Yves Meyer. *Wavelets: Algorithms and Applications.* SIAM, Philadelphia, 1993.

[38] Arch W. Naylor and George R. Sell. *Linear Operator Theory in Engineering and Science.* Springer-Verlag, New York, 1982.

[39] Athanasios Papoulis. *Probability, Random Variables and Stochastic Processes.* McGraw-Hill, New York, 1991.

[40] Sediki M. Riad. The deconvolution problem: An overview. *Proceedings of the IEEE*, 74(1):83–85, 1986.

[41] Mark J. Shensa. The discrete wavelet transform: Wedding the à trous and Mallat alorithms. *IEEE Transactions on Signal Processing*, 40(10):2464–2482, 1992.

[42] Jean-Luc Starek and Albert Bijaoui. Filtering and deconvolution by the wavelet transform. *Signal Processing*, 35:195–211, 1994.

[43] David Walnut. Applications of Gabor and wavelet expansions to the Radon transform. preprint, June 1991.

[44] Gilbert G. Walter. *Wavelets and Other Orthogonal Systems With Applications.* CRC Press, Boca Raton, 1994.

[45] Gregory W. Wornell and Alan V. Oppenheim. Estimation of fractal signals from noisy measurements using wavelets. *IEEE Tranactions on Signal Processing*, 40(3):611–623, 1992.

[46] Xiang-Gen Xia and M. Z. Nashed. The Backus-Gilbert method for signals in reproducing kernel Hilbert spaces and wavelet subspaces. *Inverse Problems*, 10:785–804, 1994.

[47] Robert M. Young. *An Introduction to Nonharmonic Fourier Series*. Academic Press. New York, 1980.

[48] Jun Zhang and Gilbert G. Walter. A wavelet-based KL-like expansion for wide-sense stationary random processes. *IEEE Transactions on Signal Processing*, 42(7):1737–1745, 1994.