Wavelet Diagonalization of Convolution Operators

Fredrik Ekstedt
GÖTEBORG UNIVERSITY
S-412 96 GÖTEBORG
SWEDEN
email:ekstedt@math.chalmers.se

NO 1997-23/ISSN 0347-2809

[0]This research was partially supported by grants ONR N00014-90-J-1343, DARPA N00014-94-11163, and DARPA 36197RTDPS

Abstract

It is well known that wavelets cannot be eigenfunctions of differential operators. We show that for homogeneous convolution operators, one can obtain a diagonal representation using two different biorthogonal wavelet bases, properly adapted to the operator at hand. We generalize this to include many inhomogeneous convolution operators, using "wavelet-like" basis functions, i.e. functions that share all the important properties of classical wavelets but not necessarily are dilates and translates of a single mother wavelet. We also show how to associate a multiresolution structure to these bases, which means that the wavelet transforms involved can be implemented with fast algorithms. Finally, we use these techniques to construct a fast wavelet transform for complex-valued signals that separates positive and negative frequencies, which is important in the analysis of radar signals.

Acknowledgements

Part of this thesis is a joint work with Martin Lindberg. I would like to thank him for many interesting discussions while we were working on this problem. I would also like to thank Björn Jawerth for his support during our stay at the University of South Carolina, and for being a great source of inspiration. Further, I'm grateful to my supervisor Jöran Bergh for all his useful comments and suggestions while reading the manuscript. Finally, I would like to thank the people at Ericsson Microwave Systems, and in particular Claes Hagström, for a very interesting collaboration and for providing data to my numerical experiments.

Contents

Chapter 1

Introduction

The purpose of this introduction is to give an overview of what the thesis is about. While the rest of the thesis is somewhat technical, and requires some familiarity with wavelet analysis, this section is written with the aim of being possible to understand with just some basic knowledge of functional analysis and Fourier analysis. Basically, there are three main important concepts involved in this work:

1. Biorthogonality
2. Wavelets
3. Diagonalization

Below, we will try to explain each of them in an informal way, and at the same time describe the main ideas of our results. We also give an outline of how the thesis is organized. Finally, we discuss some results related to ours.

1.1 Biorthogonality

Biorthogonality is best understood in the language of matrices. Let U be a nonsingular n-by-n matrix and let $\widetilde{U}$ be the transpose of its inverse ($\widetilde{U} = U^{-\mathrm{T}}$). We can write this as

$$U\widetilde{U}^{\mathrm{T}} = I \tag{1.1}$$

Let $u_1, \ldots, u_n$ be the columns of U and $\widetilde{u}_1, \ldots, \widetilde{u}_n$ the columns of $\widetilde{U}$:

$$U = [u_1 \ldots u_n] \quad \text{and} \quad \widetilde{U} = [\widetilde{u}_1 \ldots \widetilde{u}_n],$$

We can think of u_i and $\widetilde{u}_i$ as basis vectors. The basis $\{\widetilde{u}_i\}$ will be called the *dual basis* to the basis $\{u_i\}$. The two bases $\{u_i\}$ and $\{\widetilde{u}_i\}$ together will be referred to as a *biorthogonal system*. We may now interpret U as a change of basis matrix, since any vector x can be written as

$$x = U(\widetilde{U}^{\mathrm{T}} x) = U \begin{pmatrix} \langle x, \widetilde{u}_1 \rangle \\ \vdots \\ \langle x, \widetilde{u}_n \rangle \end{pmatrix} = \sum_{i=1}^{n} \langle x, \widetilde{u}_i \rangle u_i, \tag{1.2}$$

i.e. a linear combination of the basis vectors u_i. The coefficients $\langle x, \widetilde{u}_i \rangle$ are obtained by taking inner products of x with the dual basis vectors $\widetilde{u}_i$. The basis vectors u_i and the dual basis vectors satisfy a *biorthogonality relation*

$$\langle u_i, \widetilde{u}_{i'} \rangle = \delta_{i,i'}, \tag{1.3}$$

which is just another way of writing (1.1). The roles of U and $\widetilde{U}$ may of course be interchanged which is the reason behind the "duality" terminology. In the special case $U = \widetilde{U}$ we have an orthonormal basis (ON-basis). Equations (1.2) and (1.3) can immediately be extended to Hilbert spaces to define *biorthogonal dual Riesz bases*. Two families of vectors $\{u_i\}_{i \in I}$ and $\{\widetilde{u}_i\}_{i \in I}$ in a Hilbert space $\mathcal{H}$ are said to constitute biorthogonal dual Riesz bases if

$$\langle u_i, \widetilde{u}_{i'} \rangle = \delta_{i,i'} \qquad \text{(Biorthogonality)} \tag{1.4}$$

and

$$A \|x\|^2 \leq \sum_{i \in I} |\langle x, \widetilde{u}_i \rangle|^2 \leq B \|x\|^2 . \qquad \text{(Stability)} \tag{1.5}$$

Whenever these conditions hold, each $x \in \mathcal{H}$ can be written as

$$x = \sum_{i \in I} \langle x, \widetilde{u}_i \rangle u_i. \tag{1.6}$$

We should also remark that from (1.4) and (1.5), it also follows that (see [4])

$$B^{-1} \|x\|^2 \leq \sum_{i \in I} |\langle x, u_i \rangle|^2 \leq A^{-1} \|x\|^2 ,$$

and each $x \in \mathcal{H}$ can be written as

$$x = \sum_{i \in I} \langle x, u_i \rangle \widetilde{u}_i.$$

Again, the roles of $\{u_i\}$ and $\{\widetilde{u}_i\}$ can be interchanged (in applications, there can be a significant difference between using a basis or it's dual however). The difference from the finite- dimensional case is (1.5). In a finite-dimensional basis this inequality is always satisfied. But in practice we need

B/A to be of moderate size, so stability questions are of equal importance in the finite-dimensional case. The quotient B/A is called the *condition number* of the matrix $\widetilde{U}^{\mathrm{T}}$ (and U) and is a measure of how much relative errors are amplified when we multiply x with $\widetilde{U}^{\mathrm{T}}$. The computation of the coefficients $\langle x, \widetilde{u}_i \rangle$ is referred to as *analysis* of x in the system $\{u_i\}, \{\widetilde{u}_i\}$, and the summation in (1.6) is called *synthesis*. In the finite-dimensional case, the analysis is realized by the action of $\widetilde{U}^{\mathrm{T}}$ on x, and the synthesis is done by applying U on the coefficient vector $\widetilde{U}^{\mathrm{T}}x$. In a general Hilbert space $\mathcal{H}$, we define the operators

$$\begin{aligned} \widetilde{U}^* &: \mathcal{H} \to \ell^2(I); \qquad & \widetilde{U}^* x &= (\langle x, \widetilde{u}_i \rangle)_{i \in I} \qquad & &\text{“Forward transform”} \\ U &: \ell^2(I) \to \mathcal{H}; \qquad & Uc &= \sum_{i \in I} c_i u_i \qquad & &\text{“Inverse transform”} \end{aligned}$$

performing the analysis and synthesis respectively. With these operators, (1.6) can be written in a more compact way which resembles the matrix notation:

$$x = U(\widetilde{U}^* x), \quad \text{or even simpler} \quad U\widetilde{U}^* = I.$$

Here of course, I is the identity operator.

1.2 Wavelet Analysis

We will be interested in a special kind of biorthogonal dual Riesz bases, namely *biorthogonal wavelet bases* in $L^2(\mathbb{R})$. We will just mention briefly here the basic ideas of wavelets, since chapter 2 contains a short introduction to the subject. For a more thorough description, see one of the references [4] or [13].

1.2.1 Wavelets

The starting point of a wavelet basis is a function ψ with integral zero, and some localization in both time and frequency. This function is called the *mother wavelet*. In figure 1.1 we have plotted some frequently occuring mother wavelets. From this single function we create a whole family of wavelets by making copies of it at different scales and positions:

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k) \tag{1.7}$$

The wavelet $\psi_{j,k}$ is obtained by "stretching" the mother wavelet with a factor 2^{-j} in time and a factor $2^{j/2}$ in amplitude, and finally shifting it $2^{-j}k$ in time. The whole family $\{\psi_{j,k}\}_{j,k}$ thus consists of stretched versions of the mother wavelet located at various positions (see figure 1.2). If we think of the mother wavelet as "living on scale one", $\psi_{j,k}$ will be living on scale 2^{-j}.

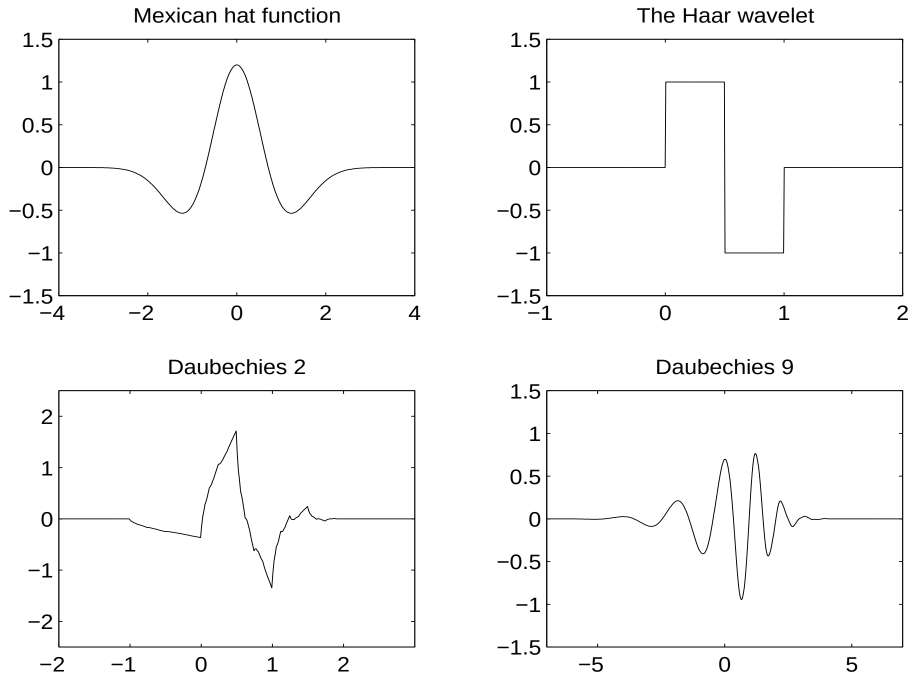


Figure 1.1: Four different mother wavelets.

Note that for fixed scale (fixed j), the wavelets are centred around points $2^{-j}k$ (if we assume that the mother wavelet is centred around 0). So for fine scales (large j), the wavelets will be spread over the real line with a small step 2^{-j}. For large scales (small j), the wavelets will be distributed very sparsely. This enables us to analyze f with different resolutions at different scales; at the fine scales (high frequencies) we analyze f with high resolution, and at the coarse scales we analyze f with low resolution. The representation

$$f = \sum_{j,k} d_{j,k}\psi_{j,k} \tag{1.8}$$

is then very efficient for functions with non-stationary properties. The coefficients $d_{j,k}$ measures the local oscillations at scale 2^{-j} around the point $2^{-j}k$. Where the function is smooth, i.e. does not vary to much, the small-scale coefficients will be almost zero and can therefore be neglected without losing to much information. The small-scale coefficients are only used when they are needed, i.e. where the the function has abrupt changes, transients or highly oscillatory components. This is the idea behind *wavelet compression*.

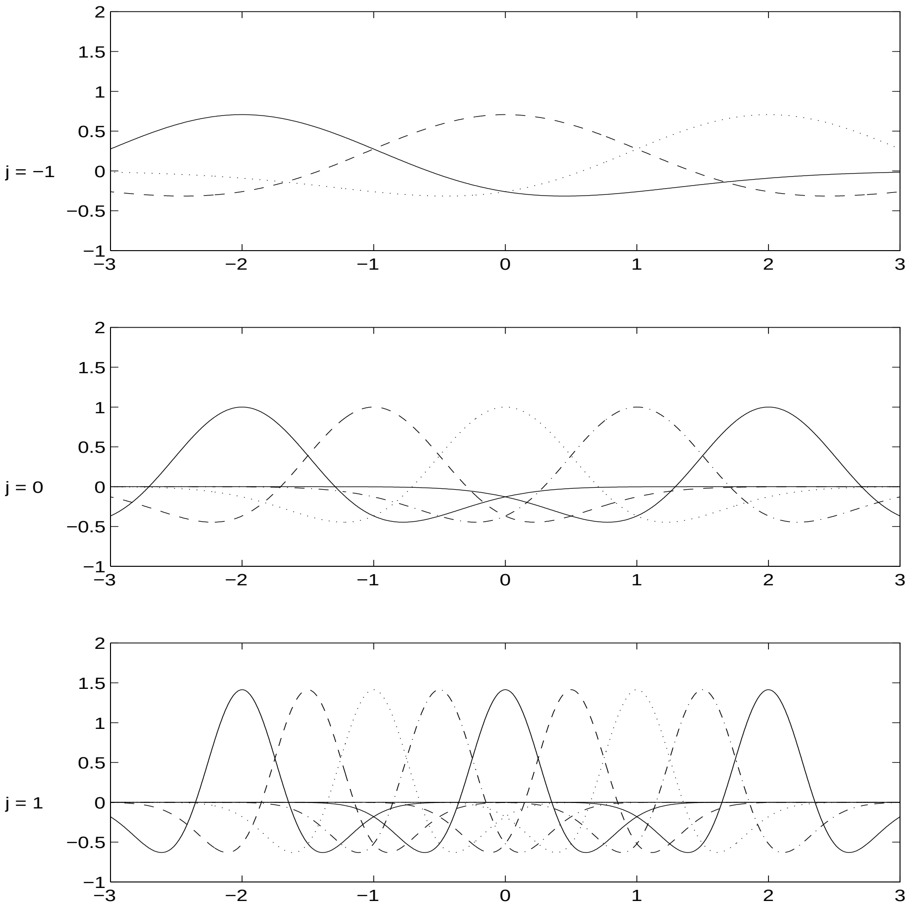


Figure 1.2: The Mexican hat function at three different scales.

1.2.2 Multiresolution Analysis

A natural question at this point is how we compute the coefficients $d_{j,k}$, i.e. what is the dual basis? It turns out that one can construct wavelet bases so that the dual basis is also obtained from a single function, the *dual mother wavelet*, by stretching and moving it around:

$$\widetilde{\psi}_{j,k}(x) = 2^{j/2}\widetilde{\psi}(2^j x - k)$$

These constructions are based on the extremely important concept of a *multiresolution analysis* (MRA), which we will give an informal description of here. We refer to Daubechies book [4] for examples. The central object of an MRA is the *scaling function* φ, which is a localized function with integral one. It is used to approximate functions at a certain scale; we define the *approximation spaces* V_j as

$$V_j = \overline{span\{\varphi_{j,k}\}_k},$$

where of course

$$\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k).$$

The best example to illustrate MRA's with is the *Haar system* which was invented by Haar in 1904. This is the first known wavelet basis constructed. Here the scaling function is the unit box function,

$$\varphi(x) = \begin{cases} 1 & \text{if } 0 \leq x < 1, \\ 0 & \text{otherwise,} \end{cases}$$

see figure 1.3. The V_j-spaces thus consist of functions that are piecewise constant on the dyadic intervals $[2^{-j}k, 2^{-j}(k+1)]$. For a general MRA, a function is approximated at scale 2^{-j} by a projection onto V_j:

$$f \approx f_j := P_j f = \sum_k c_{j,k}\varphi_{j,k}$$

This projection is determined by *dual approximation spaces* $\widetilde{V}_j$ spanned by *dual scaling functions*

$$\widetilde{\varphi}_{j,k}(x) = 2^{j/2}\widetilde{\varphi}(2^j x - k).$$

The scaling function and its duals are biorthogonal to each other in the sense that

$$\langle \varphi_{j,k}, \widetilde{\varphi}_{j,k'} \rangle = \delta_{k,k'}.$$

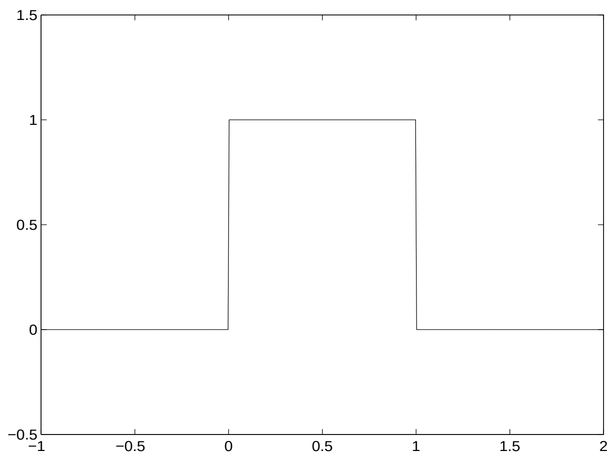


Figure 1.3: The Haar scaling function.

The coefficients $c_{j,k}$ are given by inner products of f with the $\widetilde{\varphi}_{j,k}$'s, i.e.

$$P_j f = \sum_k \langle f, \widetilde{\varphi}_{j,k} \rangle \varphi_{j,k}$$

For the Haar system, which is orthogonal, $V_j = \widetilde{V}_j$ (and $\varphi_{j,k} = \widetilde{\varphi}_{j,k}$). The projection of a function on V_j thus means approximating the function at each interval $[2^{-j}k, 2^{-j}(k+1)]$ with its mean value on that interval. This is shown in figure 1.4 for three different scales.

In applications, where the function f is given by its sample values at $2^{-j}k$ (some j), it is common practice to replace the coefficients $\langle f, \widetilde{\varphi}_{j,k} \rangle$ with the sample values $f(2^{-j}k)$. In that case, the scaling function takes the role of an interpolating function. In fact, some biorthogonal wavelet systems uses B-splines as scaling functions.

Now we impose the following conditions on the approximation spaces V_j in order to derive the wavelets,

1. $V_j \subset V_{j+1}$

2. $\bigcup_j V_j$ is dense in $L^2(\mathbb{R})$

3. $\bigcap_j V_j = \{0\}$

The first condition says that nothing is lost when going from the approximation at one scale to the approximation at the next finer scale. Rather, some details (fluctuations) are added. As we will see, it is those details, or fluctuations, that will sum up to the wavelet decomposition (1.8). The

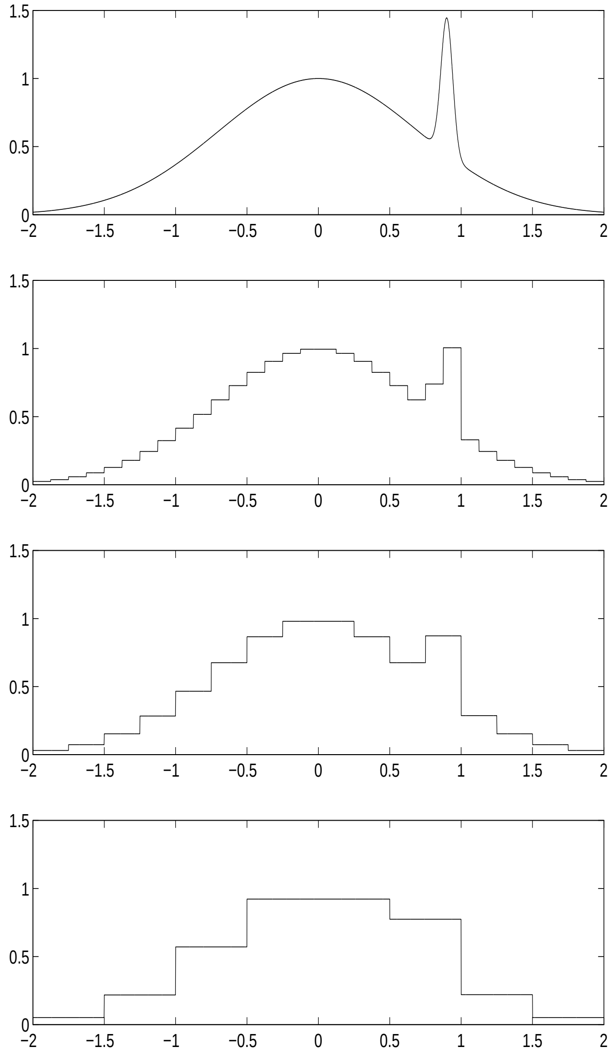


Figure 1.4: Projection in the Haar system onto scales 1/8, 1/4 and 1/2 respectively.

second condition says that f can be approximated arbitrarily well in the V_j-spaces if we take j large enough. Suppose that we have such a fine-scale approximation of f,

$$f_J = P_J f = \sum_k c_{J,k}\varphi_{J,k}$$

where $c_{J,k} = \langle f, \widetilde{\varphi}_{J,k}\rangle$. The wavelets $\{\psi_{j,k}\}$ are constructed so that they for fixed j span the difference between V_{j+1} and V_j. More precisely, let the *detail spaces* be defined by

$$W_j = \overline{span\{\psi_{j,k}\}_k}.$$

We then have (V_j and W_j are not orthogonal in general)

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

We also have *dual detail spaces* spanned by the dual wavelets

$$\widetilde{W}_j = \overline{span\{\widetilde{\psi}_{j,k}\}_k}.$$

The following biorthogonality relations holds for the approximation and detail spaces and their duals, (see figure 1.5)

$$W_j \perp \widetilde{V}_j \quad \text{and} \quad V_j \perp \widetilde{W}_j,$$

which in terms of scaling functions and wavelets becomes

$$\langle \psi_{j,k}, \widetilde{\varphi}_{j,k'}\rangle = 0 \quad \text{and} \quad \langle \varphi_{j,k}, \widetilde{\psi}_{j,k'}\rangle = 0. \tag{1.9}$$

Since $V_J = V_{J-1} \oplus W_{J-1}$, our approximation f_J above can be decomposed into two parts,

$$f_J = f_{J-1} + \delta_{J-1},$$

where $f_{J-1} \in V_{J-1}$ and $\delta_{J-1} \in W_{J-1}$. They can be written as

$$f_{J-1} = \sum_k c_{J-1,k}\varphi_{J-1,k} \quad \text{and} \quad \delta_{J-1} = \sum_k d_{J-1,k}\psi_{J-1,k},$$

with $c_{J-1,k} = \langle f, \widetilde{\varphi}_{J-1,k}\rangle$. From (1.9) it follows that $d_{J-1,k} = \langle f, \widetilde{\psi}_{J-1,k}\rangle$.

Now we turn over to the greatest feature of the MRA; the coefficients $c_{J-1,k}$ and $d_{J-1,k}$ can be computed very fast from the $c_{J,k}$'s by the so called *pyramid algorithm.* Since $\varphi \in V_0 \subset V_1$, the scaling function must satisfy a *scaling relation*

$$\varphi(x) = 2\sum_k h_k \varphi(2x - k). \tag{1.10}$$

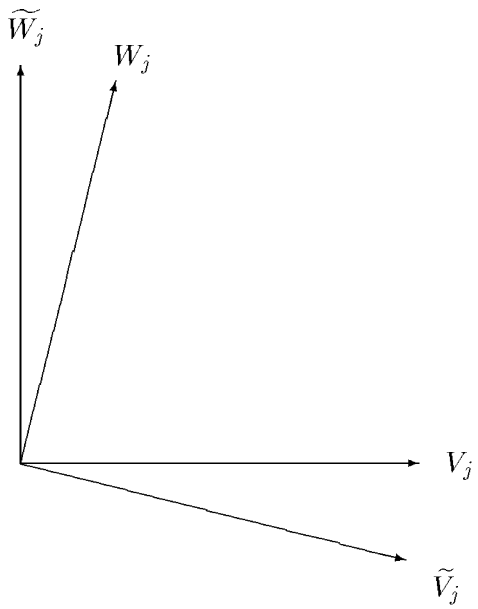


Figure 1.5: The approximation and detail spaces and their duals at level j, where we think of the whole plane as V_{j+1}.

By integrating (1.10) and using that the scaling function has a non-vanishing integral we get that

$$\sum_k h_k = 1.$$

For the wavelet we have in the same way

$$\psi(x) = 2 \sum_k g_k \varphi(2x - k) \tag{1.11}$$

where

$$\sum_k g_k = 0.$$

For the Haar system, we have $h_0 = h_1 = 1/2$, all other h_k zero, and $g_0 = 1/2, g_1 = -1/2$, all other g_k zero (see figure 1.6). The same relations hold

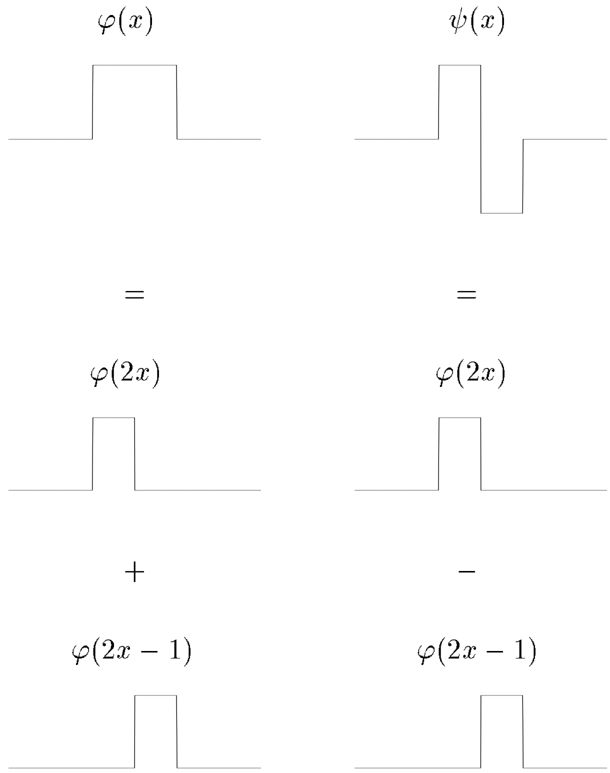


Figure 1.6: Scaling relations for the Haar system.

for the dual scaling function and wavelet,

$$\widetilde{\varphi}(x) = 2 \sum_k \widetilde{h}_k \widetilde{\varphi}(2x - k) \quad \text{and} \quad \widetilde{\psi}(x) = 2 \sum_k \widetilde{g}_k \widetilde{\varphi}(2x - k). \tag{1.12}$$

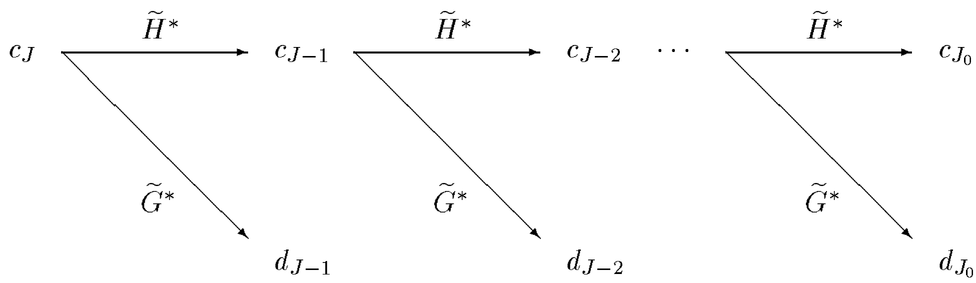


Figure 1.7: The pyramid algorithm.

where

$$\sum_k \widetilde{h}_k = 1 \quad \text{and} \quad \sum_k \widetilde{g}_k = 0. \tag{1.13}$$

We write out (1.12) at scale 2^{-J},

$$\widetilde{\varphi}_{J-1,k} = \sqrt{2}\sum_l \widetilde{h}_{l-2k}\widetilde{\varphi}_{J,l} \quad \text{and} \quad \widetilde{\psi}_{J-1,k} = \sqrt{2}\sum_l \widetilde{g}_{l-2k}\widetilde{\varphi}_{J,l}. \tag{1.14}$$

Let us get back to the decomposition $f_J = f_{J-1} + \delta_{J-1}$, or

$$\sum_l c_{J,l}\varphi_{J,l} = \sum_l c_{J-1,l}\varphi_{J-1,l} + \sum_l d_{J-1,l}\psi_{J-1,l}.$$

Taking inner products with $\widetilde{\varphi}_{J-1,k}$ and using (1.14) we get

$$c_{J-1,k} = \sqrt{2}\sum_l \widetilde{h}_{l-2k}c_{J,l}.$$

Doing similarly for the wavelets we get

$$d_{J-1,k} = \sqrt{2}\sum_l \widetilde{g}_{l-2k}c_{J,l}.$$

These formulas can be viewed in the following way: First we convolve the coefficient sequence $c_J = (c_{J,k})_k$ with the *low-pass filter* $\sqrt{2}\,\widetilde{h}_k$ and the *high-pass filter* $\sqrt{2}\,\widetilde{g}_k$ (or actually with $\sqrt{2}\,\widetilde{h}_{-k}$ and $\sqrt{2}\,\widetilde{g}_{-k}$ to be precise) to get the sequences

$$\sqrt{2}\sum_l \widetilde{h}_{l-k}c_{J,l} \quad \text{and} \quad \sqrt{2}\sum_l \widetilde{g}_{l-k}c_{J,l}.$$

The use of the terminology low-pass and high-pass filter is motivated by (1.13), which tells us that convolution with $\widetilde{h}_k$ is averaging and convolution with $\widetilde{g}_k$ is some kind of differencing. Next we *subsample* these sequences, i.e. we throw away the odd coefficients (put $2k$ instead of k) and end up with the coefficient sequences $c_{J-1,k}$ and $d_{J-1,k}$. For the Haar system, this means taking pairwise mean values and differences of the $c_{J,k}$:s, see figure 1.8. This

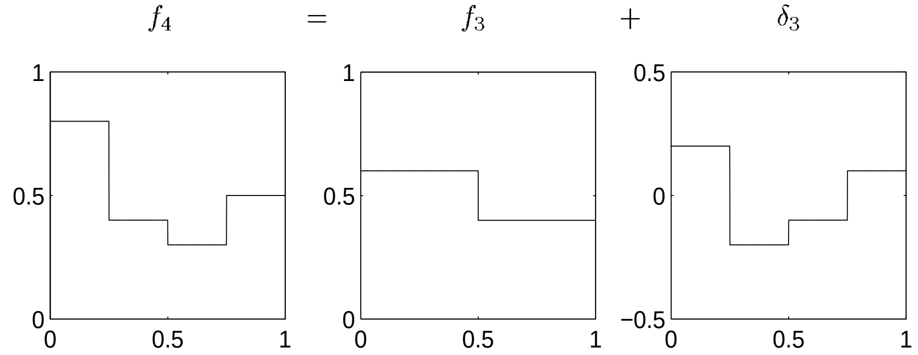


Figure 1.8: One decomposition step in the Haar system.

procedure is now recursively repeated on c_{J-1}, and this is known as the pyramid algorithm or the Forward Wavelet Transform (FWT). In figure 1.7 this is shown schematically, where the $\widetilde{H}^*$ and $\widetilde{G}^*$ means filtering with $\widetilde{h}_{-k}$ and $\widetilde{g}_{-k}$, *and* subsampling. After a while we are left with the decomposition

$$(1.15) \quad f_J = \delta_{J-1} + \delta_{J-2} \ldots + \delta_{j_0} + f_{j_0} = \sum_{j=j_0}^{J-1} \sum_k d_{j,k} \psi_{j,k} + \sum_k c_{j_0,k} \varphi_{j_0,k}$$

Theoretically, we need to run the pyramid algorithm *ad infinitum* in order to obtain the decomposition (1.8). We would also have to start at an "infinitely fine" scale. In practice however, we always stop at a coarsest scale j_0 and end up with the decomposition (1.15).

There is also a fast recursive algorithm to reconstruct the $c_{J,k}$:s from the wavelet coefficients $d_{j,k}$ and coarse-scale coefficients $c_{j_0,k}$. The reconstruction formula is

$$c_{j,k} = \sqrt{2} \left(\sum_l h_{k-2l} c_{j-1,l} + \sum_l g_{k-2l} d_{j-1,l} \right).$$

This can be seen as convolutions of $c_{j-1,l}$ and $d_{j-1,l}$ with the filters $\sqrt{2}\, h_l$ and $\sqrt{2}\, g_l$, after first inserting zeroes between each coefficient ($k - 2l$ instead

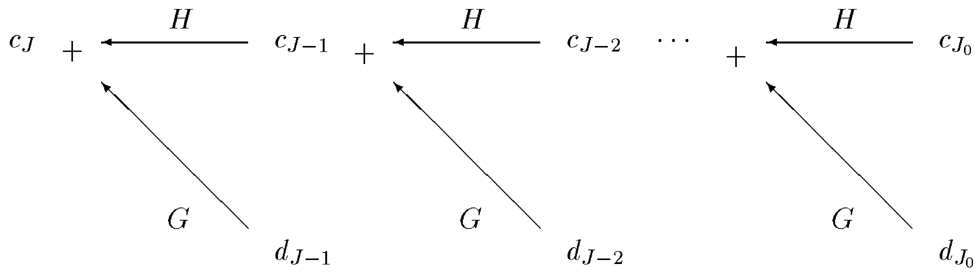


Figure 1.9: The inverse pyramid algorithm.

of $k-l$), and finally summing up. This is shown in figure 1.9, where H and G means both inserting zeroes (*up-sampling*) and filtering.

1.3 Diagonalization

1.3.1 Diagonalization in a Single Basis

A matrix A is said to be diagonalizable if it can be written as

$$A = U\Lambda\widetilde{U}^{\mathrm{T}} \tag{1.16}$$

where $U\widetilde{U}^{\mathrm{T}} = I$ and Λ is a diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$ (the eigenvalues of Λ). If we again view $U = [u_1 \ldots u_n]$ and $\widetilde{U}^{\mathrm{T}} = [\widetilde{u}_1 \ldots \widetilde{u}_n]$ as dual biorthogonal bases, we can express (1.16) in an other way by saying that A becomes a diagonal matrix in the basis $u_1, \ldots, u_n$. This can also be displayed in the "vector-level notation":

$$Ax = \sum_{i=1}^{n} \lambda_i \langle x, \widetilde{u}_i \rangle u_i \tag{1.17}$$

Indeed we see that Ax can be computed by analyzing x in the basis u_i, multiplying each coefficient $\langle x, \widetilde{u}_i \rangle$ with the corresponding eigenvalue λ_i and hence getting the coefficients of Ax in the basis $u_1, \ldots, u_n$. Finally we synthesize Ax by the summation in (1.17). Of course we can have $U = \widetilde{U}$, i.e. diagonalization in an ON-basis. The famous spectral theorem states that this can be achieved if the matrix A is symmetric. Orthogonality is an attractive feature since it guarantees numeric stability in the analysis and synthesis steps (the condition number is 1). The notion of diagonalization extends naturally to Hilbert spaces; we say that an operator A is diagonalized by the biorthogonal system $\{u_i\}, \{\widetilde{u}_i\}$ if

$$Ax = \sum_{i \in I} \lambda_i \langle x, \widetilde{u}_i \rangle u_i$$

which is very similar to (1.17), the difference being that we have to make sure that the series converges. We can also get an analogy to (1.16) by using the operators U and $\widetilde{U}^*$:

$$A = U \Lambda \widetilde{U}^*, \tag{1.18}$$

where Λ is the operator that multiplies each coefficient $\langle x, \widetilde{u}_i \rangle$ with the eigenvalue λ_i:

$$\Lambda : \ell^2(I) \to \ell^2(I), \quad \Lambda c = (\lambda_i c_i)_{i \in I}$$

1.3.2 The Singular Value Decomposition

Not all matrices can be diagonalized in the sense of (1.16). However, there is a completely general way of writing a matrix in diagonal form, called the *singular value decomposition* (SVD). Here, given any matrix A (which need not be quadratic but we will restrict ourselves to square matrices here) there exist two orthogonal matrices U and V such that

$$A = U \Sigma V^{\mathrm{T}}. \tag{1.19}$$

The matrix Σ is diagonal, with diagonal elements σ_i called the *singular values* of A. For a proof and some further interesting features of the SVD, we refer to the book by Golub and Van Loan [8]. With $U = [u_1 \ldots u_n]$ and $V = [v_1 \ldots v_n]$ we can write (1.19) as

$$Ax = \sum_{i=1}^{n} \sigma_i \langle x, v_i \rangle u_i \tag{1.20}$$

With the SVD, Ax is computed by first analyzing x in the ON-basis $\{v_i\}$, then multiplying each coefficient with the corresponding singular value σ_i, and finally performing synthesis in the ON-basis $\{u_i\}$. Note that we use different bases for analysis and synthesis here. The SVD generalizes in an obvious way to general Hilbert spaces, the summation in (1.20) is just taken over the index set I, and in (1.19) we use the operator notation instead. We should stress here that not all operators posses a singular value decomposition, in fact, the existence of an SVD is equivalent to the existence of a spectral decomposition of A^*A.

1.3.3 The Wavelet-Vaguelette Decomposition

The singular value decomposition has been used widely for solving linear inverse problems, i.e. given y, solve $Ax = y$. This can be done by the SVD through the inversion formula (from now on we switch freely between matrix and operator notation)

$$x = A^{-1}y = \sum_i \sigma_i^{-1} \langle y, u_i \rangle v_i.$$

A disadvantage with this approach is that the basis functions used are completely determined by the operator A and may not be well adapted to the class of functions we want to work with. Often, the singular functions have a non-local nature (orthogonal polynomials and spherical harmonics for instance) while the functions to be reconstructed have local features such as edges, transients etc. (A more detailed discussion on this matter can be found in [6]).

What we really would like to have is a diagonal representation like (1.20) were we have some freedom to choose the bases $\{u_i\}$ and $\{v_i\}$ so that they are adapted to the class of objects under study. This can be achieved if we are willing to give up orthogonality and instead work with two biorthogonal systems $\{u_i\}, \{\widetilde{u}_i\}$ and $\{v_i\}, \{\widetilde{v}_i\}$. In other words, we try to write

$$A = U\mathcal{K}\widetilde{V}^{\mathrm{T}}. \tag{1.21}$$

Again, $\mathcal{K}$ is a diagonal matrix; we will call its diagonal elements $\kappa_1, \ldots, \kappa_n$ *quasi-singular values* of A after Donoho [6]. Note that they are by no means uniquely determined by A since the representation (1.21) is highly nonunique. As a matter of fact, we can chose the matrix U in any way we want, then the product $\mathcal{K}\widetilde{V}^{\mathrm{T}}$ is uniquely determined. Therefore, we can chose a proper biorthogonal system $\{u_i\}, \{\widetilde{u}_i\}$ that is well adapted to the objects we want to compute. We then construct a new biorthogonal system $\{v_i\}, \{\widetilde{v}_i\}$ through

$$v_i = \kappa_i A^{-1} u_i \quad \text{and} \quad \widetilde{v}_i = \frac{1}{\kappa_i} A^{\mathrm{T}} \widetilde{u}_i \tag{1.22}$$

where we chose the quasi-singular values κ_i such that the new basis functions v_i have comparable size. Now we can express Ax as

$$Ax = \sum_{i=1}^{n} \langle Ax, \widetilde{u}_i \rangle u_i = \sum_{i=1}^{n} \langle x, A^{\mathrm{T}} \widetilde{u}_i \rangle u_i = \sum_{i=1}^{n} \kappa_i \langle x, \widetilde{v}_i \rangle u_i.$$

Informally, one can argue in the following way to motivate the construction (1.22): If u_i are "good" basis vectors for representing the vectors Ax, then $v_i = \kappa_i A^{-1} u_i$ should be "good" for representing the vectors x.

In this thesis we will be concerned with (1.22) when $\{u_i\}, \{\widetilde{u}_i\}$ are biorthogonal dual wavelet bases. This is what Donoho [6] defines as the *wavelet-vaguelette decomposition.* We now turn over to the notation that will be used throughout the rest of the thesis. The operator we want to diagonalize will be denoted by K since it will often be a convolution operator. The wavelets will as before be denoted by $\psi_{j,k}$ and $\widetilde{\psi}_{j,k}$, the new basis functions (v_i and $\widetilde{v}_i$ before) will now be denoted by $\psi^{\mathrm{K}}_{j,k}$ and $\widetilde{\psi}^{\mathrm{K}}_{j,k}$. Being overly thorough, we write out (1.22) with this new notation.

$$\psi^{\mathrm{K}}_{j,k} = \kappa_{j,k} K^{-1} \psi_{j,k} \quad \text{and} \quad \widetilde{\psi}^{\mathrm{K}}_{j,k} = \frac{1}{\overline{\kappa_{j,k}}} K^* \widetilde{\psi}_{j,k} \tag{1.23}$$

Donoho uses the name *vaguelettes* after Meyer [12] for the functions $\psi^{\mathrm{K}}_{j,k}$ and $\widetilde{\psi}^{\mathrm{K}}_{j,k}$ since they for many operators K are "almost" wavelets or "wavelet-like". We will always refer to them as wavelets throughout the paper, though. In fact, for homogeneous convolution operators they are wavelets in the classical sense (after proper normalization), that is

$$\psi^{\mathrm{K}}_{j,k}(x) = 2^{j/2} \psi^{\mathrm{K}}(2^j x - k) \quad \text{and} \quad \widetilde{\psi}^{\mathrm{K}}_{j,k}(x) = 2^{j/2} \widetilde{\psi}^{\mathrm{K}}(2^j x - k).$$

So far, there is nothing new here. The contribution of this thesis is that for some operators K, the multiresolution structure can be carried over from the original wavelets to the new ones, which enables us to develop fast algorithms for the analysis step. In other words, we will define filters and scaling functions associated with the new wavelet basis. This will be the basic theme of the thesis.

1.3.4 Outline

In chapter 2-3, which is a paper written together with Martin Lindberg [7], we consider the case of homogeneous convolution operators, i.e. convolutions with $\widehat{k}(\omega) = \omega^n$ (derivatives) or $\widehat{k}(\omega) = |\omega|^\alpha$ (Riesz potentials). In this case we show that if the original wavelet bases originates from a multireslution analysis, we can construct scaling functions and filters generating the new wavelets. This is also carried out in higher dimensions for wavelets defined on arbitrary lattice structures.

In chapter 4 this is extended to convolution operators that are *asymptotically homogeneous*, i.e. their Fourier multipliers satisfies

$$\widehat{k}(\omega) \sim |\omega|^\alpha \qquad \text{as} \quad |\omega| \to \infty$$

and

$$\widehat{k}(\omega) \sim |\omega|^\beta \qquad \text{as} \quad \omega \to 0$$

For these operators, we can construct a "generalized" multiresolution analysis" with different filters and scaling functions on each scale. We will still have the translation invariant structure on each level (scale) though. The asymptotic homogeneity will ensure that the new system "converges" to classical wavelet bases as $j \to \pm\infty$, which will guarantee stability under some mild conditions on the original basis. We will also extend the higher-dimensional result to inhomogeneous and/or anisotropic convolution operators in a similar way.

In chapter 5 we give up the translation invariance and consider integral operators,

$$Kf(x) = \int k(x,y)f(y)dy.$$

and general differential operators. In this case we will generate a new MRA with filters and scaling functions depending on both scale and position.

Finally, in chapter 6 we give an application of the above methods, by constructing a wavelet transform for complex-valued radar signals that separates positive and negative frequencies.

1.4 Related Work

This thesis can be seen as a natural extension of Donoho's work about the wavelet-vaguelette decomposition [6]. He does not develop any algorithms for computing the vaguelettes coefficients $\langle f, \widetilde{\psi}^{\mathrm{K}}_{j,k}\rangle$ though. This is now possible since we have derived a multiresolution analysis, and therefore fast algorithms, for the new wavelet basis. For the derivative operator, this was done by Daubechies in [5]. Her construction relies on the very explicit factorizations of the filter functions and is not clear how to extend her construction to more general operators or to non-separable higher dimensional wavelets. Another method for computing the coefficients in the new system was developed by Kolaczyk [11]. He uses Meyer wavelets since they have a very explicit representation in the Fourier domain, where convolution operators are easy to handle. Computing the vaguelette coefficients then comes down to computing certain projections of f in the frequency domain.

A different approach for the numerical computation of operators with wavelets was developed by Beylkin, Coifman and Rokhlin [1]. They actually use two different methods, the *standard representation* and the *non-standard representation*. The standard representation just means expressing the operator K in a wavelet basis, i.e. working with the (infinite) matrix $\langle K\psi_{jk}, \widetilde{\psi}_{j'k'}\rangle_{jk,j'k'}$. They show that for a large class of operators, including differential operators and Calderon-Zygmund operators, this matrix is almost diagonal in the sense that the elements decay very fast away from the diagonal. The non-standard representation corresponds to expanding the kernel $k(x,y)$ of an integral operator in a two-dimensional wavelet basis. A way of performing matrix-vector multiplication (operators application) and matrix multiplication (operator composition) in this representation is also derived.

In a paper by Jawerth and Sweldens [10], techniques similar to ours are used to obtain diagonal stiffness matrices in Wavelet-Galerkin schemes. Given a symmetric, positive differential operator $L = C^*C$, the Wavelet-Galerkin method for solving $Lu = f$ is to find an approximative u in a finite-dimensional space spanned by wavelets $\Psi_{j,k}$ such that

$$\langle Cu, C\widetilde{\Psi}^*_{j,k}\rangle = \langle f, \widetilde{\Psi}^*_{j,k}\rangle,$$

for wavelets $\widetilde{\Psi}^*_{j,k}$ in another finite-dimensional wavelet space. Writing $u = \sum_{j',k'} x_{j',k'}\Psi_{j',k'}$ leads to the linear system

$$Ax = b,$$

where

$$A_{(j,k),(j',k')} = \langle C\Psi_{j',k'}, C\widetilde{\Psi}^*_{j,k}\rangle \quad \text{and} \quad b_{j,k} = \langle f, \widetilde{\Psi}^*_{j,k}\rangle.$$

The wavelets $\Psi_{j,k}$ plays the same role as our "original wavelets"; they are used to represent the function $L^{-1}f$ we want to reconstruct. The wavelets $\widetilde{\Psi}^*_{j,k}$ corresponds to the vaguelettes, or "new wavelets" $\widetilde{\psi}^{\mathrm{K}}_{j,k}$ which are used to analyze f. As in the wavelet-vaguelette decomposition, they are chosen so that the matrix A is diagonal. There is a main difference between the two constructions however. We start with the original wavelets and construct the new wavelets through the operator L:

$$\psi^{\mathrm{K}}_{j,k} = L\psi_{j,k} \quad \text{and} \quad \widetilde{\psi}^{\mathrm{K}}_{j,k} = L^{-*}\widetilde{\psi}_{j,k}.$$

In [10], the starting point is also a biorthogonal system $\{\psi_{j,k}\}, \{\widetilde{\psi}_{j,k}\}$. Those wavelets are not used at all in the computations, their role is to define new wavelets

$$\Psi_{j,k} = C^{-1}\psi_{j,k} \quad \text{and} \quad \widetilde{\Psi}^*_{j,k} = C^{-1}\widetilde{\psi}_{j,k}.$$

It is immediately clear that this construction makes the matrix A diagonal. To see the connection with the wavelet-vaguelette decomposition, define wavelets dual to those above through

$$\widetilde{\Psi}_{j,k} = C^*\widetilde{\psi}_{j,k} \quad \text{and} \quad \Psi^*_{j,k} = C^*\psi_{j,k}.$$

We then have

$$\Psi^*_{j,k} = L\Psi_{j,k} \quad \text{and} \quad \widetilde{\Psi}^*_{j,k} = L^{-*}\widetilde{\Psi}_{j,k}.$$

Here we assume enough regularity of ψ and $\widetilde{\psi}$ in order to apply the differential operator C. But we actually never use the wavelets $\Psi^*_{j,k}$ and $\widetilde{\Psi}_{j,k}$, so ψ and $\widetilde{\psi}$ need not have any regularity at all. In some of the examples in [10], Haar-like systems are used. From there, new scaling functions and filters are defined so that the new wavelets $\Psi_{j,k}$ and $\widetilde{\Psi}^*_{j,k}$ come with a multiresolution structure and fast algorithms.

Chapter 2

Diagonalization of homogeneous convolution operators

For homogeneous convolution operators, we can obtain a diagonal representation using two wavelet bases, properly adapted to the operator. We will derive scaling functions and filters for the new wavelet basis, and show that the new basis is stable under suitable assumptions on the original wavelets.

2.1 Wavelets

In this section we review basic wavelet theory mainly to fix the notation and we refer to [4] for proofs and more details.

2.1.1 Multiresolution Analysis.

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

1. $V_j \subset V_{j+1}$,
2. $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}$,
3. $f(x) \in V_0 \Leftrightarrow f(x+1) \in V_0$,
4. $\bigcup_j V_j$ is dense in $L^2(\mathbb{R})$ and $\bigcap_j V_j = \{0\}$,
5. There exists a *scaling function* $\varphi \in V_0$ such that the collection $\{\varphi(x-l) : l \in \mathbb{Z}\}$ is a Riesz basis of V_0.

It is immediate that the collection of functions $\{\varphi_{j,l} : l \in \mathbb{Z}\}$, with $\varphi_{j,l}(x) = 2^{j/2}\varphi(2^j x - l)$, is a Riesz basis of V_j. From the definition of the MRA it follows that the scaling function must satisfy the *dilation equation*

$$\varphi(x) = 2\sum_l h_l \varphi(2x - l) \quad \text{or} \quad \widehat{\varphi}(\omega) = H(\omega/2)\widehat{\varphi}(\omega/2), \tag{2.1}$$

where H is a 2π-periodic function defined by $H(\omega) = \sum_l h_l e^{-il\omega}$. If the scaling function belongs to $L^1(\mathbb{R})$ it is, under very general conditions, uniquely defined by the dilation equation and the normalization

$$\int \varphi(x)dx = 1 \quad \Leftrightarrow \quad \widehat{\varphi}(0) = 1.$$

We will always assume that this is the case and from (2.1) we then have $H(0) = 1$.

2.1.2 Approximation

The spaces V_j will be used to approximate functions. This will be done by defining appropriate projections onto these spaces. Since the union of all the V_j is dense in $L^2(\mathbb{R})$, we are guaranteed that any function can be approximated arbitrarily close by such projections.

If we want to write any polynomial of degree smaller than N as a linear combination of the scaling function and its translates then the scaling function must satisfy the Strang-Fix conditions,

$$\begin{aligned} &\widehat{\varphi}(0) = 1, \quad \text{and} \\ &\widehat{\varphi}^{(p)}(2\pi k) = 0 \quad \text{for} \quad k \neq 0,\ 0 \leq p < N. \end{aligned} \tag{2.2}$$

From (2.1) it follows that $H(\omega)$ must have a root of multiplicity N at $\omega = \pi$.

2.1.3 Wavelets

By W_j we will denote a space complementing V_j in V_{j+1}, i.e.

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

and consequently

$$\bigoplus_j W_j = L^2(\mathbb{R}).$$

A function ψ is a *wavelet* if the collection of functions $\{\psi(x - l) : l \in \mathbb{Z}\}$ is a Riesz basis of W_0. The collection of functions $\{\psi_{j,l} : j, l \in \mathbb{Z}\}$ is then a Riesz basis of $L^2(\mathbb{R})$. We define $\mathcal{P}_j$ as the projection onto V_j parallel to V_j^c and $\mathcal{Q}_j$ as the projection onto W_j parallel to W_j^c. A function f can now be written as

$$f(x) = \sum_j \mathcal{Q}_j f(x) = \sum_{j,l} \gamma_{j,l}\psi_{j,l}(x),$$

or if we start from a coarsest scale J as

$$f(x) = \mathcal{P}_J f(x) + \sum_{j=J}^{\infty} \mathcal{Q}_j f(x) = \sum_l \lambda_{J,l}\varphi_{J,l}(x) + \sum_{j=J}^{\infty}\sum_l \gamma_{j,l}\psi_{j,l}(x).$$

Below we will describe how to find the coefficients $\lambda_{j,l}$ and $\gamma_{j,l}$. Since the wavelet $\psi \in W_0 \subset V_1$

$$\psi(x) = 2\sum_l g_l \varphi(2x-l) \quad \text{or} \quad \widehat{\psi}(\omega) = G(\omega/2)\widehat{\varphi}(\omega/2), \tag{2.3}$$

where G is a 2π-periodic function defined by $G(\omega) = \sum_l g_l e^{-il\omega}$.

2.1.4 Biorthogonal Wavelets

In a biorthogonal MRA we have a scaling function φ, a wavelet ψ, a dual scaling function $\widetilde{\varphi}$, and a dual wavelet $\widetilde{\psi}$ such that any function f can be written

$$f(x) = \sum_l \langle f, \widetilde{\varphi}_{J,l}\rangle \varphi_{J,l}(x) + \sum_{j=J}^{\infty}\sum_k \langle f, \widetilde{\psi}_{j,l}\rangle \psi_{j,l}(x). \tag{2.4}$$

For this to be true the scaling functions and wavelets must satisfy the biorthogonality conditions

$$\begin{aligned} \langle \widetilde{\varphi}, \psi(\cdot - l)\rangle = \langle \widetilde{\psi}, \varphi(\cdot - l)\rangle = 0 \quad \text{and} \\ \langle \widetilde{\varphi}, \varphi(\cdot - l)\rangle = \langle \widetilde{\psi}, \psi(\cdot - l)\rangle = \delta_l. \end{aligned}$$

Expressed in the filter functions H, G, $\widetilde{H}$, and $\widetilde{G}$ necessary conditions are given by

$$\begin{array}{lcll} \widetilde{H}(\omega)\overline{H(\omega)} & + & \widetilde{H}(\omega+\pi)\overline{H(\omega+\pi)} & = & 1 \\ \widetilde{G}(\omega)\overline{G(\omega)} & + & \widetilde{G}(\omega+\pi)\overline{G(\omega+\pi)} & = & 1 \\ \widetilde{G}(\omega)\overline{H(\omega)} & + & \widetilde{G}(\omega+\pi)\overline{H(\omega+\pi)} & = & 0 \\ \widetilde{H}(\omega)\overline{G(\omega)} & + & \widetilde{H}(\omega+\pi)\overline{G(\omega+\pi)} & = & 0 \end{array} \tag{2.5}$$

Now, if we define the *modulation matrix* M by

$$M(\omega) = \begin{bmatrix} H(\omega) & H(\omega+\pi) \\ G(\omega) & G(\omega+\pi) \end{bmatrix}, \tag{2.6}$$

and similarly for $\widetilde{M}$, then

$$\widetilde{M}(\omega)\overline{M(\omega)}^{\mathrm{T}} = I.$$

Cramer's rule now states that

$$\widetilde{H}(\omega) = \frac{\overline{G(\omega+\pi)}}{\overline{\Delta(\omega)}} \qquad \widetilde{G}(\omega) = -\frac{\overline{H(\omega+\pi)}}{\overline{\Delta(\omega)}}, \tag{2.7}$$

where $\Delta(\omega) = \det M(\omega)$.

When constructing wavelets one often starts by defining the low-pass filters H and $\widetilde{H}$. Then one defines G and $\widetilde{G}$ through equation (2.7) where $\Delta(\omega)$ is chosen equal to $e^{-i\omega}$.

2.1.5 Vanishing Moments

The moments of the wavelet are defined by

$$\mathcal{N}_p = \int x^p \psi(x)dx \quad \text{with} \quad p \in \mathbb{N},$$

and similarly for the dual wavelet. We recall that if the scaling function reproduces any polynomial of degree smaller than N then $H(\omega)$ has a root of multiplicity N at $\omega = \pi$. From (2.7) we see that this is equivalent to $\widetilde{G}(\omega)$ having a root of multiplicity N at $\omega = 0$. Since $\mathcal{F}(\widetilde{\varphi})(0) = 1$ this is also equivalent to $\mathcal{F}(\widetilde{\psi})(\omega)$ having a root of multiplicity N at $\omega = 0$, i.e. the dual wavelet has N vanishing moments. By a similar argument the wavelet ψ will have $\widetilde{N}$ vanishing moments if the dual scaling function reproduces polynomials of degree smaller than $\widetilde{N}$.

2.2 The New Wavelet Basis

2.2.1 Diagonalization

In a biorthogonal wavelet basis we have a wavelet ψ and a dual wavelet $\widetilde{\psi}$ such that any $f \in L^2(\mathbb{R})$ can be written

$$f = \sum_{j,l} \langle f, \widetilde{\psi}_{j,l} \rangle \psi_{j,l}.$$

For a given linear operator K we would like to expand Kf in this basis. We will consider convolution operators,

$$Kf = k * f \quad \text{or} \quad \widehat{Kf}(\xi) = \widehat{k}(\xi)\widehat{f}(\xi),$$

that preserves the characteristics of a wavelet. If we denote the adjoint of K by K^* we can write Kf as

$$Kf = \sum_{j,l} \langle Kf, \widetilde{\psi}_{j,l} \rangle \psi_{j,l} = \sum_{j,l} \langle f, K^* \widetilde{\psi}_{j,l} \rangle \psi_{j,l}.$$

We will now describe how the coefficients $\langle f, K^*\widetilde{\psi}_{j,l}\rangle$ can be calculated in a fast and numerically stable way by analyzing f in a new biorthogonal wavelet basis. We define this new basis by the relations

$$\widetilde{\psi}^{\mathrm{K}} = K^*\widetilde{\psi} \quad \text{and} \quad \psi^{\mathrm{K}} = K^{-1}\psi, \tag{2.8}$$

and in the Fourier domain we have

$$\widehat{\widetilde{\psi}^{\mathrm{K}}}(\omega) = \overline{\widehat{k}(\omega)}\widehat{\widetilde{\psi}}(\omega) \quad \text{and} \quad \widehat{\psi^{\mathrm{K}}}(\omega) = \frac{1}{\widehat{k}(\omega)}\widehat{\psi}(\omega). \tag{2.9}$$

For the moment we assume that the wavelets, ψ and $\widetilde{\psi}$, and the operator K are such that these new functions are well defined and below we will show that under certain assumptions these functions in fact form a biorthogonal wavelet basis. Our goal is to find a condition on the operator K such that the following relation holds

$$K^*\widetilde{\psi}_{j,l} = \overline{\kappa}_j\widetilde{\psi}^{\mathrm{K}}_{j,l},$$

since then the wavelet coefficients of Kf, $\langle f, K^*\widetilde{\psi}_{j,l}\rangle = \kappa_j\langle f, \widetilde{\psi}^{\mathrm{K}}_{j,l}\rangle$. The constant κ_j is independent of l since K^* is translation invariant. So for this to hold true K^* must be invariant under dyadic dilations, up to the constant κ_j. Let us therefore define the dyadic dilation operator D_j as

$$D_j f(x) = 2^{j/2} f(2^j x) \quad \text{or} \quad \widehat{D_j f}(\omega) = 2^{-j/2}\widehat{f}(2^{-j}\omega).$$

The dilation invariance of K^* means that

$$K^* D_j f = \overline{\kappa}_j D_j K^* f$$

or in the Fourier domain

$$\overline{\widehat{k}(\omega)}2^{-j/2}\widehat{f}(2^{-j}\omega) = \overline{\kappa}_j 2^{-j/2}\overline{\widehat{k}(2^{-j}\omega)}\widehat{f}(2^{-j}\omega).$$

From this we arrive at the following condition on K

$$\kappa_j = \frac{\widehat{k}(\omega)}{\widehat{k}(2^{-j}\omega)} \quad \text{is constant.} \tag{2.10}$$

We observe that $\kappa_j = \kappa_1^j$ so if we let $\kappa = \kappa_1$, we have $\kappa_j = \kappa^j$. It is now also clear that the new functions ψ^{K} and $\widetilde{\psi}^{\mathrm{K}}$ are biorthogonal. To conclude, by analyzing f in the new wavelet basis and multiplying the wavelet coefficients by κ^j, we get the wavelet coefficients of Kf in the original wavelet basis:

$$Kf = \sum_{j,l} \kappa^j \langle f, \widetilde{\psi}^{\mathrm{K}}_{j,l}\rangle \psi_{j,l}, \tag{2.11}$$

and this is what we refer to as a diagonalization of the operator K. Examples of operators satisfying (2.10) are

1. Differentiation and integration
   $\widehat{k}(\omega) = (i\omega)^\alpha, \alpha \in \mathbb{Z}$. Here $\kappa = 2^\alpha$.

2. The Riesz potential
   $\widehat{k}(\omega) = |\omega|^\alpha, \alpha \in \mathbb{R}$. Here $\kappa = 2^\alpha$.

3. The Hilbert transform
   $\widehat{k}(\omega) = -i \operatorname{sgn} \omega$. Here $\kappa = 1$.

These three types of operators are essentially exhaustive. This follows if we consider continuous solutions of (2.10) for positive and negative ω separately since then we must have $\widehat{k}(\omega) = C\omega^\alpha$, for some constant C.

2.2.2 Admissibility Conditions

Let us now return to the question of whether the new wavelets as given by (2.9) are well defined. Given the operator K we would like to find suitable conditions on the original wavelets. First we assume that $\widehat{k}(\omega) = |\omega|^\alpha$ and that $\alpha > 0$, since any other choices of $\widehat{k}$ and α are treated similarly. We know that the new wavelets must have at least one vanishing moment each and from (2.9) we then see that ψ must have at least $\lfloor \alpha + 1 \rfloor$ vanishing moments. In this case there is no additional requirement on the number of vanishing moments on $\widetilde{\psi}$. On the other hand if $\widetilde{\psi} \in W^s(\mathbb{R})$ we realize that $\widetilde{\psi}^{\mathrm{K}} \in W^{s-\alpha}(\mathbb{R})$ so we must have $s \geq \alpha$ to have $\widetilde{\psi}^{\mathrm{K}} \in L^2(\mathbb{R})$. If $\alpha < 0$ the roles of ψ^{K} and $\widetilde{\psi}^{\mathrm{K}}$ are simply interchanged.

2.2.3 Decay and Compact Support

Finally, let us discuss the rate of decay of the new wavelets. In most applications we are interested in compactly supported wavelets and this corresponds to transfer functions that are finite impulse response filters. Non-compactly supported wavelets are also useful in practice if they have rapid decay.

If the original wavelets have compact support we know that their Fourier transforms are smooth. Looking at the definition of the new wavelets in the Fourier domain (2.9) we conclude that a necessary condition for the new wavelets to be compactly supported is that $\widehat{k}(\omega)$ is smooth for $\omega = 0$. This will only be the case when the operator is differentiation or integration. Indeed when this is the case it is obvious that the new wavelets are also compactly supported.

When $\alpha > 0$ is not an even integer we have, after a moment's consideration,

$$\widehat{\widetilde{\psi}^{\mathrm{K}}} \in C^{N-1+\lceil\alpha\rceil}(\mathbb{R}), \tag{2.12}$$

where N is the number of vanishing moments of the dual wavelet. Now, if the Fourier transform of $\widetilde{\psi}^{\mathrm{K}}$ and its derivatives were also in $L^1(\mathbb{R})$ the

Riemann-Lebesgue lemma would give us the following estimate on the rate of decay of $\widetilde{\psi}^{\mathrm{K}}$

(2.13) $$x^n \widetilde{\psi}^{\mathrm{K}}(x) \to 0 \quad \text{as} \quad |x| \to \infty, \qquad \text{for} \quad 0 \le n \le N - 1 + \lceil \alpha \rceil.$$

Let us therefore find out when this is actually the case. We assume that the dual wavelet $\widetilde{\psi} \in C^{\widetilde{M}}(\mathbb{R})$ and that it is compactly supported. It follows that $\widetilde{\psi}^{(m)} \in L^1$ for $0 \le m \le \widetilde{M}$ and

$$\omega^m \widehat{\widetilde{\psi}}(\omega) \to 0 \quad \text{as} \quad |\omega| \to \infty, \qquad \text{for} \quad 0 \le m \le \widetilde{M}.$$

From this it follows that there is a constant C such that

$$|\widehat{\widetilde{\psi}}(\omega)| \le C(1 + |\omega|)^{-\widetilde{M}}.$$

Actually, this holds for all derivatives of the Fourier transform of $\widetilde{\psi}$ since $x^p \widetilde{\psi}(x)$ and its $\widetilde{M}$ first derivatives are in $L^1(\mathbb{R})$ for all $p \in \mathbb{N}$, i.e.

$$|\frac{\partial^p}{\partial \omega^p} \widehat{\widetilde{\psi}}(\omega)| \le C(1 + |\omega|)^{-\widetilde{M}} \qquad \text{for} \quad p \in \mathbb{N}.$$

From the definition of $\widetilde{\psi}^{\mathrm{K}}$ and (2.12) we get

$$|\frac{\partial^p}{\partial \omega^p} \widehat{\widetilde{\psi}^{\mathrm{K}}}(\omega)| \le C(1 + |\omega|)^{\alpha - \widetilde{M}} \qquad \text{for} \quad 0 \le p \le N - 1 + \lceil \alpha \rceil.$$

That is, $\partial^p \widehat{\widetilde{\psi}^{\mathrm{K}}} \in L^1(\mathbb{R})$ if $\alpha - \widetilde{M} < -1$. So the decay estimate (2.13) holds when $\widetilde{M} > 1 + \alpha$. Rewriting the decay estimate we have thus arrived at

(2.14) $$|\widetilde{\psi}^{\mathrm{K}}(x)| \le C(1 + |x|)^{1 - N - \lceil \alpha \rceil} \qquad \text{if} \quad \widetilde{M} > 1 + \alpha \quad \text{where} \quad \widetilde{\psi} \in C^{\widetilde{M}}.$$

For the wavelets ψ and ψ^{K} a similar argument as above gives $\widehat{\psi^{\mathrm{K}}} \in C^{\widetilde{N} - 1 - \lfloor \alpha \rfloor}(\mathbb{R})$, where $\widetilde{N}$ is the number of vanishing moments of the wavelet. The rate of decay of the wavelet ψ^{K} is then given by

(2.15) $$|\psi^{\mathrm{K}}(x)| \le C(1 + |x|)^{1 - \widetilde{N} + \lfloor \alpha \rfloor} \qquad \text{if} \quad M > 1 - \alpha \quad \text{where} \quad \psi \in C^M.$$

We see that the number of vanishing moments of the original wavelets determines the rate of decay of the new wavelets. Again, if $\alpha < 0$ the roles of ψ^{K} and $\widetilde{\psi}^{\mathrm{K}}$ are interchanged.

2.2.4 Stability

Finally, let us investigate the stability of the new wavelet basis. Following Donoho ([6], p 108-109), the following conditions are sufficient to guarantee stability

1. $\psi(x) \leq C(1+|x|)^{-1-\epsilon}$, for some constant C and $\epsilon > 0$.
2. $\int \psi(x) = 0$.
3. $\psi \in C^{\beta}$ for some $\beta > 0$.

and similarily for the dual mother wavelet. Thus, all we need is one vanishing moment, and some minimal decay and regularity. We know from the previous section that the first condition is satisfied if the mother wavelet has Hölder exponent greater than $1-\alpha$ and at least $\lfloor\alpha\rfloor+3$ vanishing moments, and if the dual mother wavelet has Hölder exponent greater than $1+\alpha$ and at least $3-\lceil\alpha\rceil$ vanishing moments. The second condition is satiesfied under much weaker conditions. Using standard Fourier techniques, we see that the last requirement is fulfilled if $\widetilde{M} > 1+\alpha$ and $M > 1-\alpha$. To summarize, we will get a new stable wavelet basis if

1. $\widetilde{M} > 1+\alpha$ and $M > 1-\alpha$, where $\widetilde{\psi} \in C^{\widetilde{M}}$ and $\psi \in C^{M}$.
2. $\widetilde{N} > 2+\lfloor\alpha\rfloor$ and $N > 2-\lceil\alpha\rceil$, where $\widetilde{N}$ is the numeber of vanishing moments for ψ and vice versa.

2.3 The New Multiresolution Analysis

We will now describe how to associate a multiresolution analysis with the new biorthogonal wavelet basis. We will do this by defining a new pair of scaling functions. It is natural to try with

$$\widehat{\widetilde{\varphi}^{\mathrm{K}}}(\omega) = \overline{\widehat{\ell}(\omega)}\widehat{\widetilde{\varphi}}(\omega) \quad \text{and} \quad \widehat{\varphi^{\mathrm{K}}}(\omega) = \frac{1}{\widehat{\ell}(\omega)}\widehat{\varphi}(\omega), \tag{2.16}$$

where $\widehat{\ell}$ is an unknown function. Biorthogonality of the original scaling functions then implies biorthogonality of the new scaling functions φ^{K} and $\widetilde{\varphi}^{\mathrm{K}}$

$$\langle\widetilde{\varphi}^{\mathrm{K}}, \varphi^{\mathrm{K}}(\cdot - l)\rangle = \langle\widetilde{\varphi}, \varphi(\cdot - l)\rangle = \delta_l.$$

The scaling functions must also be biorthogonal to the wavelets

$$\langle\widetilde{\varphi}^{\mathrm{K}}, \psi^{\mathrm{K}}(\cdot - l)\rangle = \langle\widetilde{\psi}^{\mathrm{K}}, \varphi^{\mathrm{K}}(\cdot - l)\rangle = 0,$$

which, expressed in the filter functions, amounts to

$$\begin{array}{rcl} \widetilde{H}^{\mathrm{K}}(\omega)\overline{G^{\mathrm{K}}(\omega)} + \widetilde{H}^{\mathrm{K}}(\omega+\pi)\overline{G^{\mathrm{K}}(\omega+\pi)} &=& 0 \quad \text{and} \\ \widetilde{G}^{\mathrm{K}}(\omega)\overline{H^{\mathrm{K}}(\omega)} + \widetilde{G}^{\mathrm{K}}(\omega+\pi)\overline{H^{\mathrm{K}}(\omega+\pi)} &=& 0. \end{array}$$

From the dilation equation for $\widetilde{\varphi}$ and (2.16) we get

$$\widehat{\widetilde{\varphi}^{\mathrm{K}}}(\omega) = \frac{\overline{\widehat{\ell}(\omega)}}{\overline{\widehat{\ell}(\omega/2)}}\widetilde{H}(\omega/2)\widehat{\widetilde{\varphi}^{\mathrm{K}}}(\omega/2) = \widetilde{H}^{\mathrm{K}}(\omega/2)\widehat{\widetilde{\varphi}^{\mathrm{K}}}(\omega/2),$$

and by similar arguments we get the following expressions for the new filter functions

$$H^{\mathrm{K}}(\omega) = \frac{\widehat{\ell}(\omega)}{\widehat{\ell}(2\omega)} H(\omega), \qquad G^{\mathrm{K}}(\omega) = \frac{\widehat{\ell}(\omega)}{\widehat{k}(2\omega)} G(\omega),$$
$$\widetilde{H}^{\mathrm{K}}(\omega) = \frac{\overline{\widehat{\ell}(2\omega)}}{\overline{\widehat{\ell}(\omega)}} \widetilde{H}(\omega), \qquad \widetilde{G}^{\mathrm{K}}(\omega) = \frac{\overline{\widehat{k}(2\omega)}}{\overline{\widehat{\ell}(\omega)}} \widetilde{G}(\omega).$$

Biorthogonality between $\widetilde{\varphi}^{\mathrm{K}}$ and $\psi^{\mathrm{K}}(\cdot - l)$ is then equivalent to

$$\frac{\overline{\widehat{\ell}(2\omega)}}{\overline{\widehat{\ell}(\omega)}} \widetilde{H}(\omega) \frac{\overline{\widehat{\ell}(\omega)}}{\overline{\widehat{k}(2\omega)}} \overline{G(\omega)} + \frac{\overline{\widehat{\ell}(2\omega + 2\pi)}}{\overline{\widehat{\ell}(\omega + \pi)}} \widetilde{H}(\omega + \pi) \frac{\overline{\widehat{\ell}(\omega + \pi)}}{\overline{\widehat{k}(2\omega + 2\pi)}} \overline{G(\omega + \pi)} = 0.$$

Since $\widetilde{H}(\omega)\overline{G(\omega)} + \widetilde{H}(\omega + \pi)\overline{G(\omega + \pi)} = 0$ this is equivalent to

$$\frac{\widehat{\ell}(2\omega)}{\widehat{k}(2\omega)} = \frac{\widehat{\ell}(2\omega + 2\pi)}{\widehat{k}(2\omega + 2\pi)}.$$

This means that $\widehat{\ell}(\omega)$ must be chosen so that

$$m(\omega) = \frac{\widehat{k}(\omega)}{\widehat{\ell}(\omega)} \quad \text{is } 2\pi\text{-periodic.} \tag{2.17}$$

If we can find such an $\widehat{\ell}(\omega)$ all of the biorthogonality conditions will be satisfied. It is still not clear how we should define this function though. However, we have not considered the approximation properties, or the Strang-Fix conditions, of the new scaling functions. If we substitute (2.17) into (2.16) we get

$$\overline{m(\omega)}\widehat{\widetilde{\varphi}^{\mathrm{K}}}(\omega) = \overline{\widehat{k}(\omega)}\widehat{\widetilde{\varphi}}(\omega) \quad \text{and} \quad \widehat{k}(\omega)\widehat{\varphi^{\mathrm{K}}}(\omega) = m(\omega)\widehat{\varphi}(\omega). \tag{2.18}$$

Since we know that $\widehat{\widetilde{\varphi}^{\mathrm{K}}}(0) = \widehat{\varphi^{\mathrm{K}}}(0) = 1$ we must have

$$\frac{m(\omega)}{\widehat{k}(\omega)} \to 1 \text{ as } \omega \to 0.$$

From this we see that we must find a 2π-periodic function $m(\omega)$ that matches $\widehat{k}(\omega)$ at $\omega = 0$, i.e. it should have the same number of zeros at $\omega = 0$.

We make the following more or less canonical choice

$$m(\omega) = \widehat{k}(-i(e^{i\omega} - 1)), \tag{2.19}$$

since $-i(e^{i\omega} - 1) = \omega + o(\omega)$ as $\omega \to 0$. This is indeed a natural choice because the determinants of the modulation matrices of the original and new system will only differ by a multiplicative constant

$$\Delta(\omega) = \widehat{k}(i)\kappa^2 \Delta^{\mathrm{K}}(\omega). \tag{2.20}$$

This choice implies that $m(\omega)$ is a sort of discretized version of $\widehat{k}(\omega)$. From (2.18) we then see, via the Strang-Fix conditions, that the approximation properties of the new scaling functions are exactly related, as they should be, to the number of vanishing moments of the new wavelets. If we write the Fourier series of $m(\omega)$ as

$$m(\omega) = \sum_l m_l e^{-il\omega}$$

we see that

$$K^* \widetilde{\varphi}(x) = \sum_l \overline{m_l} \widetilde{\varphi}^{\mathrm{K}}(x + l). \tag{2.21}$$

For K being the derivative operator this becomes

$$\widetilde{\varphi}'(x) = \widetilde{\varphi}^{\mathrm{K}}(x) - \widetilde{\varphi}^{\mathrm{K}}(x-1),$$

i.e. differentiation in the original system corresponds to a finite difference in the new system.

At this point we have constructed a new biorthogonal multiresolution analysis with the wavelets ψ^{K} and $\widetilde{\psi}^{\mathrm{K}}$ and with the scaling functions φ^{K} and $\widetilde{\varphi}^{\mathrm{K}}$. This means that we can decompose any function in this new basis using the fast wavelet transform. We also know the relation between the wavelet coefficients of f and Kf in the new and original basis, respectively. In a numerical computation we always stop the decomposition at a coarsest scale and we are thus also interested in finding a relation between the scaling function coefficients of Kf and f. Expanding Kf in the original basis we get

$$Kf(x) = \sum_l \langle f, K^* \widetilde{\varphi}_{J,l} \rangle \varphi_{J,l}(x) + \sum_{j=J}^{\infty} \sum_l \langle f, K^* \widetilde{\psi}_{j,l} \rangle \psi_{j,l}(x).$$

Using (2.21) it is easy to verify that

$$\langle f, K^* \widetilde{\varphi}_{j,l} \rangle = \kappa^j \sum_n m_n \langle f, \widetilde{\varphi}^{\mathrm{K}}_{j,l-n} \rangle. \tag{2.22}$$

We note that this formula can be seen as a discretized version of the operator K acting on the subspace V_j.

Summary

Before looking at some examples we summarize our results. Given a function f such that

$$f(x) = \sum_k \lambda_{J,l} \varphi^{\mathrm{K}}_{J,l}(x) + \sum_{j=J}^{\infty} \sum_k \gamma_{j,l} \psi^{\mathrm{K}}_{j,l}(x),$$

we can find the expansion of Kf

$$Kf(x) = \sum_k \lambda^{\mathrm{K}}_{J,l} \varphi_{J,l}(x) + \sum_{j=J}^{\infty} \sum_k \gamma^{\mathrm{K}}_{j,l} \psi_{j,l}(x),$$

by the relations

$$\gamma^{\mathrm{K}}_{j,l} = \kappa^j \gamma_{j,l},$$
$$\lambda^{\mathrm{K}}_{J,l} = \kappa^J \sum_n m_n \lambda_{J,l-n}.$$

With our choice of $m(\omega)$ the new filter functions become

$$H^{\mathrm{K}}(\omega) = \frac{\widehat{k}(e^{i\omega}+1)}{\kappa} H(\omega), \qquad G^{\mathrm{K}}(\omega) = \frac{1}{\kappa \widehat{k}(-i(e^{i\omega}-1))} G(\omega),$$
$$\widetilde{H}^{\mathrm{K}}(\omega) = \frac{\overline{\kappa}}{\overline{\widehat{k}(e^{i\omega}+1)}} \widetilde{H}(\omega), \qquad \widetilde{G}^{\mathrm{K}}(\omega) = \overline{\kappa \widehat{k}(-i(e^{i\omega}-1))} \widetilde{G}(\omega).$$

Remark

A more elegant way to derive the new scaling functions and filters was pointed out to us by Patrik Andersson. We start with a biorthogonal system and first find a new pair of scaling functions as follows. Begin with the identity

$$\frac{e^{i\omega}-1}{i\omega} = \prod_{j=1}^{\infty} \frac{e^{i2^{-j}\omega}+1}{2}.$$

Our operators satisfy $\widehat{k}(\omega_1 \omega_2) = \widehat{k}(\omega_1)\widehat{k}(\omega_2)$ and if we apply $\widehat{k}$ to the left and right hand side of the identity we get

$$\frac{\widehat{k}(-i(e^{i\omega}-1))}{\widehat{k}(\omega)} = \prod_{j=1}^{\infty} \frac{\widehat{k}(e^{i2^{-j}\omega}+1)}{\kappa},$$

since $\widehat{k}(2) = \kappa \widehat{k}(1)$ and where we have assumed that $\widehat{k}(1) = 1$. By a repeated application of (2.1) we can write the scaling function as the infinite product

$$\widehat{\varphi}(\omega) = \prod_{j=1}^{\infty} H(2^{-j}\omega),$$

and it follows that

$$\frac{\widehat{k}(-i(e^{i\omega}-1))}{\widehat{k}(\omega)}\widehat{\varphi}(\omega) = \prod_{j=1}^{\infty} \frac{\widehat{k}(e^{i2^{-j}\omega}+1)}{\kappa} H(2^{-j}\omega).$$

That is, if we define a new scaling function by

$$\widehat{\varphi^{\mathrm{K}}}(\omega) = \frac{\widehat{k}(-i(e^{i\omega}-1))}{\widehat{k}(\omega)}\widehat{\varphi}(\omega),$$

it will be associated with the filter

$$H^{\mathrm{K}}(\omega) = \frac{\widehat{k}(e^{i\omega}+1)}{\kappa} H(\omega),$$

and this is exactly the filter we got with our choice of $m(\omega)$ above. Similarly, we get the same filter for the dual scaling function. Now we can define the new wavelets from equation (2.7) and instead of the standard choice of the determinant of the modulation matrix we define Δ^{K} through equation (2.20). Then it is easy to verify that new wavelets are the same as before.

2.4 Examples

2.4.1 Differentiation

For the derivative operator we have

$$K = \frac{d}{dx}, \quad K^* = -\frac{d}{dx}, \quad K^{-1} = \int_{-\infty}^{x} dy,$$

and the new wavelets are thus given by

$$\widetilde{\psi}^{\mathrm{K}}(x) = -\widetilde{\psi}'(x), \qquad \psi^{\mathrm{K}}(x) = \int_{-\infty}^{x} \psi(y)dy.$$

Since $\widehat{k}(\omega) = i\omega$ we have $m(\omega) = e^{i\omega} - 1$ and

$$\widetilde{\varphi}'(x) = \widetilde{\varphi}^{\mathrm{K}}(x) - \widetilde{\varphi}^{\mathrm{K}}(x-1).$$

We also note that the new wavelets and scaling functions are compactly supported if the original wavelets and scaling functions are.

2.4.2 The Ramp Filter

We consider the Riesz potential operator with $\alpha = 1$ as an example, i.e. the ramp filter. Now

$$\widehat{Kf}(\omega) = |\omega|\widehat{f}(\omega) \quad \text{and} \quad \widehat{K^{-1}f}(\omega) = \frac{1}{|\omega|}\widehat{f}(\omega).$$

Since $\widehat{k}(\omega) = |\omega|$ we have

$$m(\omega) = |e^{-i\omega} - 1| \quad \text{and} \quad m_l = \frac{4}{\pi(1-4l^2)}.$$

In this case $\widehat{k}(\omega)$ is not smooth for $\omega = 0$ so the new wavelets and scaling functions will not have compact support. If we start with a biorthogonal basis where the wavelets have several vanishing moments the new wavelets will decay fast though.

2.4.3 The Hilbert Transform

In the case of the Hilbert transform we have

$$\widehat{Kf}(\omega) = -i\,\mathrm{sgn}\,\omega\ \widehat{f}(\omega) \quad \text{or} \quad Kf(x) = \frac{1}{\pi}\ \text{p.v.} \int \frac{f(y)}{x-y} dy.$$

We note that $\widehat{k}(\omega) = -i\ \mathrm{sgn}\,\omega = -i\frac{w}{|\omega|}$ so

$$m(\omega) = -\frac{e^{-i\omega}-1}{|e^{-i\omega}-1|} \quad \text{and} \quad m_l = \frac{1}{\pi(l+1/2)}.$$

We make the interesting observation that we have a convolution with $m_l = 1/(\pi(l+1/2))$ acting on the V_j spaces, i.e. a discretized version of the Hilbert transform. Just as for the ramp filter the new wavelets and scaling functions will not have compact support but if the original wavelets have several vanishing moments the new wavelets will decay fast. See figure 2.1 and 2.2 for an example where the original scaling functions and wavelets where chosen from the 6/10 factorization of the maxflat Daubechies halfband filters.

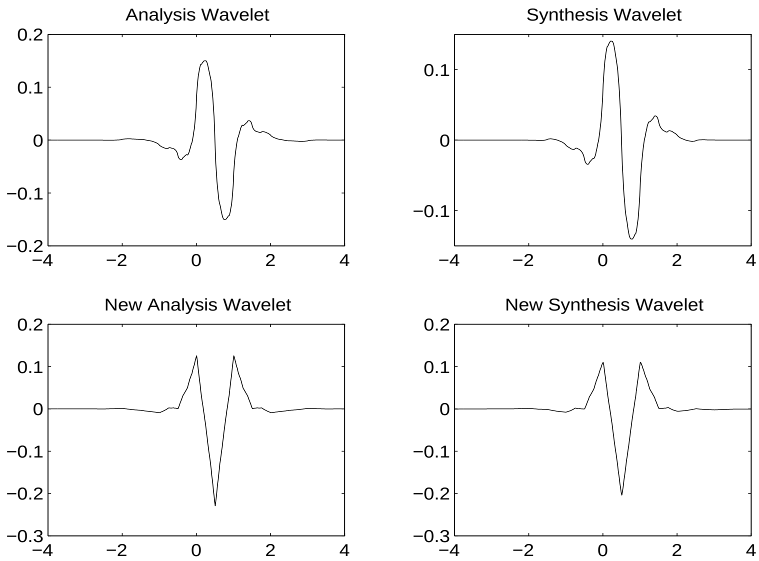


Figure 2.1: The wavelets for the Hilbert transform, (see sect. 2.4.3).

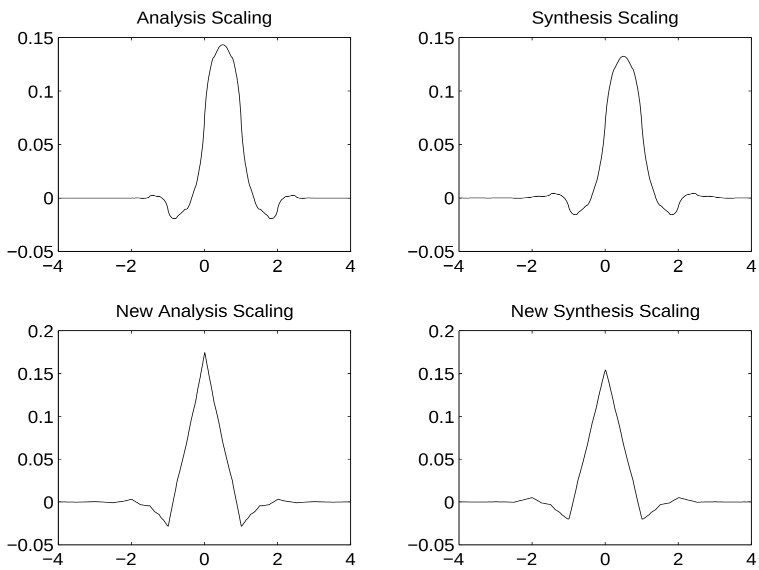


Figure 2.2: The scaling functions for the Hilbert transform.

Chapter 3

Diagonalization in Several Dimensions

In this chapter we will generalize our method to higher dimensional spaces and convolution operators. If both the basis and operator are separable it is easy and straightforward to extend our results of the previous sections. What is interesting though is that we can generalize the diagonalization technique to the case of non-separable wavelet bases.

3.1 Separable Bases and Operators

For a separable multidimensional wavelet basis in $\mathbb{R}^n$ it is easy to see that the previous ideas can be generalized in a straightforward way if the convolution kernel k is also separable, i.e. if

$$\widehat{k}(\omega) = \widehat{k}_1(\omega_1)\cdots\widehat{k}_n(\omega_n), \qquad \text{where} \quad \omega \in \mathbb{R}^n,$$

and all the k_i:s satisfies the diagonalization condition (2.10). In a two-dimensional separable wavelet basis we form the scaling function Φ and the wavelets Ψ_ν by tensor products of a one-dimensional scaling function φ and wavelet ψ

$$\Phi = \varphi \otimes \varphi, \quad \Psi_1 = \varphi \otimes \psi, \quad \Psi_2 = \psi \otimes \varphi, \quad \text{and} \quad \Psi_3 = \psi \otimes \psi.$$

Similarly, we define a dual scaling function $\widetilde{\Phi}$ and dual wavelets $\widetilde{\Psi}_\nu$ if we want a biorthogonal two-dimensional wavelet basis. Starting from such a basis we define the new wavelets analogously with the one-dimensional case

$$\widetilde{\Psi}_\nu^{\mathrm{K}} = K^* \widetilde{\Psi}_\nu \quad \text{and} \quad \Psi_\nu^{\mathrm{K}} = K^{-1}\Psi_\nu.$$

We will form the new scaling functions by defining new one-dimensional scaling functions for each coordinate direction, $i = 1, 2$,

$$\widehat{\widetilde{\varphi}_i^{\mathrm{K}}}(\omega_i) = \overline{\widehat{\ell}_i(\omega_i)}\widehat{\widetilde{\varphi}}(\omega_i) \quad \text{and} \quad \widehat{\varphi_i^{\mathrm{K}}}(\omega_i) = \frac{1}{\overline{\overline{\widehat{\ell}_i(\omega_i)}}}\widehat{\varphi}(\omega_i),$$

where each $\widehat{\ell}_i(\omega)$ is derived from $\widehat{k}_i$ trough (2.17) and (2.19) as before. The new scaling functions Φ^{K} and $\widetilde{\Phi}^{\mathrm{K}}$ are formed by taking tensor products of the new one-dimensional scaling functions

$$\widetilde{\Phi}^{\mathrm{K}} = \widetilde{\varphi}_1^{\mathrm{K}} \otimes \widetilde{\varphi}_2^{\mathrm{K}} \quad \text{and} \quad \Phi^{\mathrm{K}} = \varphi_1^{\mathrm{K}} \otimes \varphi_2^{\mathrm{K}}.$$

The new wavelets become

$$\begin{aligned}
\widetilde{\Psi}_1^{\mathrm{K}} &= \widetilde{\varphi}_1^{\mathrm{K}} \otimes \widetilde{\psi}_2^{\mathrm{K}} \qquad && \text{and} \qquad && \Psi_1^{\mathrm{K}} = \varphi_1^{\mathrm{K}} \otimes \psi_2^{\mathrm{K}},\\
\widetilde{\Psi}_2^{\mathrm{K}} &= \widetilde{\psi}_1^{\mathrm{K}} \otimes \widetilde{\varphi}_2^{\mathrm{K}} && \text{and} && \Psi_2^{\mathrm{K}} = \psi_1^{\mathrm{K}} \otimes \varphi_2^{\mathrm{K}},\\
\widetilde{\Psi}_3^{\mathrm{K}} &= \widetilde{\psi}_1^{\mathrm{K}} \otimes \widetilde{\psi}_2^{\mathrm{K}} && \text{and} && \Psi_3^{\mathrm{K}} = \psi_1^{\mathrm{K}} \otimes \psi_2^{\mathrm{K}}.
\end{aligned}$$

where of course

$$\widetilde{\psi}_i^{\mathrm{K}} = k_i^* * \widetilde{\psi} \quad \text{and} \quad \psi_i^{\mathrm{K}} = k_i^{-1} * \psi.$$

3.2 Non-separable Bases

It is possible to construct non-separable wavelet bases in several dimensions although fairly few such bases have actually been constructed. Non-separable bases are of interest in for example image processing since they are more isotropic than a separable basis, which is strongly oriented along the coordinate axes. For a separable basis in $\mathbb{R}^n$ the underlying structure is the integer lattice $\mathbb{Z}^n$ and the dilation is the same along all coordinate axes. This is not the case for a non-separable basis where we have some other lattice and/or another dilation. Two examples for which non-separable wavelet bases have been constructed are the hexagonal and the Quincunx lattice. Cohen and Daubechies [2] have constructed symmetric biorthogonal wavelets with compact support and arbitrarily high regularity on the Quincunx lattice. The two-dimensional biorthogonal wavelets on the hexagonal lattice by Cohen and Schlenker [3] have symmetry under 30^0 rotations, compact support and some regularity. For a class of lattices with certain tiling properties Strichartz [14] constructs n-dimensional orthogonal wavelets with arbitrarily high regularity but not with compact support. In a forthcoming paper Jawerth and Mao [9] present a general method for the construction of wavelets on lattices.

3.2.1 Lattices

For a standard separable basis in $\mathbb{R}^n$ the underlying structure is the integer lattice $\mathbb{Z}^n$ and the wavelets are generated from $2^n - 1$ mother wavelets ψ_ν, $\nu = 1, 2, \ldots, 2^n - 1$,

$$\psi_{\nu,j,\gamma}(x) = 2^{jn/2}\psi_\nu(D^j x - \gamma), \qquad j \in \mathbb{Z},\ \gamma \in \mathbb{Z}^n,$$

where $D = 2I$ is the dilation matrix.

To construct non-separable wavelet bases we start with a lattice $\mathbf{\Gamma}$ and dilation matrix D. A *lattice* in $\mathbb{R}^n$ is defined as $\mathbf{\Gamma} = \Gamma\mathbb{Z}^n$, where Γ is a nonsingular n-by-n matrix. With

$$\Gamma = \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix}$$

we get the hexagonal lattice in $\mathbb{R}^2$, see figure 3.1. The integer lattice has $\Gamma = I$ and so has the Quincunx lattice. As we will see it is the dilation matrix that distinguishes these lattices from each other. Given a lattice $\mathbf{\Gamma}$ the *dilation matrix* D has to satisfy the requirement $D\mathbf{\Gamma} \subset \mathbf{\Gamma}$. Moreover, all eigenvalues of D must have modulus greater then one so that we are expanding in all directions. We will refer to the lattice $D\mathbf{\Gamma} = D\Gamma\mathbb{Z}^n$ as the *subsampling lattice* of $\mathbf{\Gamma}$. The subsampling lattice of the Quincunx lattice, see figure 3.2, is defined by

$$D = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \quad \text{or} \quad D = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

From the conditions on D it follows that

$$D = \Gamma N \Gamma^{-1},$$

where N is a non-singular integer matrix. If we let $m = |\det D| = |\det N|$ we will see that we get $m-1$ different mother wavelets. In most constructions of non-separable bases the dilation matrix is often diagonal or of the form

$$D = mP,$$

where P is an orthogonal matrix.

3.2.2 Multiresolution Analysis

We define a multiresolution analysis of $L^2(\mathbb{R}^n)$, associated with the lattice $\mathbf{\Gamma}$ and the dilation matrix D, as a sequence of closed subspaces V_j of $L^2(\mathbb{R}^n)$, $j \in \mathbb{Z}$, such that

1. $V_j \subset V_{j+1}$,
2. $f(x) \in V_j \Leftrightarrow f(Dx) \in V_{j+1}$,
3. $f(x) \in V_0 \Leftrightarrow f(x+\gamma) \in V_0,\ \forall \gamma \in \mathbf{\Gamma}$,
4. $\bigcup_j V_j$ is dense in $L^2(\mathbb{R}^n)$ and $\bigcap_j V_j = \{0\}$,
5. There exists a scaling function $\varphi \in V_0$ such that the collection $\{\varphi(x-\gamma) : \gamma \in \mathbf{\Gamma}\}$ is a Riesz basis of V_0.

It is immediate that the collection of functions $\{\varphi_{j,\gamma} : \gamma \in \mathbf{\Gamma}\}$, with $\varphi_{j,\gamma}(x) = m^{j/2}\varphi(D^j x - \gamma)$, is a Riesz basis of V_j. As usual, the scaling function satisfies a dilation equation

$$\varphi(x) = m \sum_{\gamma \in \mathbf{\Gamma}} h_\gamma \varphi(Dx - \gamma), \tag{3.1}$$

and the transfer function of h_γ is defined by $H(\omega) = \sum_{\gamma \in \mathbf{\Gamma}} h_\gamma e^{-i\gamma\cdot\omega}$. To proceed we have to be able to do Fourier analysis on the lattice group $\mathbf{\Gamma}$. We then need to define the dual group of $\mathbf{\Gamma}$. The *dual lattice* of $\mathbf{\Gamma}$ is defined as

$$\mathbf{\Gamma}^* = \{\gamma^* \in \mathbb{R}^n : \gamma \cdot \gamma^* \in \mathbb{Z}, \ \forall \gamma \in \mathbf{\Gamma}\}, \tag{3.2}$$

and the dual group of $\mathbf{\Gamma}$ is then the quotient group $\mathbb{R}^n / 2\pi\mathbf{\Gamma}^*$. From this definition it easy to verify that $\mathbf{\Gamma}^* = \Gamma^{-\mathrm{T}}\mathbb{Z}^n$. Since $D^{\mathrm{T}}\gamma^* \cdot \gamma = \gamma^* \cdot D\gamma$ for every $\gamma \in \mathbf{\Gamma}$ and $\gamma^* \in \mathbf{\Gamma}^*$ we have $D^{\mathrm{T}}\mathbf{\Gamma}^* \subset \mathbf{\Gamma}^*$. From the definition of H we notice that for any $\gamma^* \in \mathbf{\Gamma}^*$

$$H(\omega + 2\pi\gamma^*) = \sum_{\gamma \in \mathbf{\Gamma}} h_\gamma e^{-i\gamma\cdot\omega} e^{-i2\pi\gamma\cdot\gamma^*} = \sum_{\gamma \in \mathbf{\Gamma}} h_\gamma e^{-i\gamma\cdot\omega} = H(\omega),$$

since $\gamma \cdot \gamma^* \in \mathbb{Z}$ for any $\gamma \in \mathbf{\Gamma}$. In other words, $H(\omega)$ is $2\pi\mathbf{\Gamma}^*$-periodic.

For the standard integer lattice $\mathbb{Z}^n$ the dual lattice is also $\mathbb{Z}^n$. The dual lattice of the hexagonal lattice is the hexagonal lattice rotated 30^0 and it is shown in figure 3.1, where also the *Voronoi cell* of $2\pi\mathbf{\Gamma}^*$ around 0, i.e. the set of points closer to 0 than any other $\gamma^* \in 2\pi\mathbf{\Gamma}^*$, is shown. This is the smallest possible set on which $H(\omega)$ is completely determined since it is a representative of the quotient group $\mathbb{R}^n / 2\pi\mathbf{\Gamma}^*$.

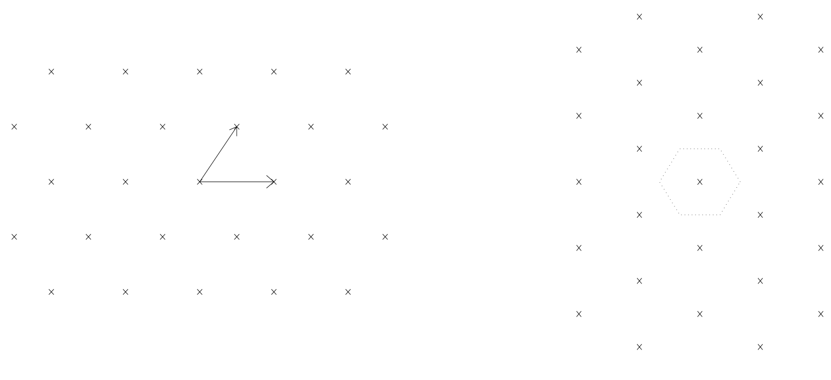


Figure 3.1: The hexagonal lattice and its dual with a Voronoi cell.

Having defined a multiresolution analysis let us now introduce wavelets. In the separable case we need $2^n - 1$ different W_j spaces, and the same

Figure 3.2: The Quincunx lattice and sublattice.

number of mother wavelets, to complement V_j in V_{j+1}. In the non-separable case we need $m-1$ different mother wavelets which follows from the fact that the order of the quotient group $\mathbf{\Gamma}/D\mathbf{\Gamma}$ is m. Each mother wavelet ψ_ν will satisfy a scaling equation

$$\psi_\nu(x) = m \sum_{\gamma\in\mathbf{\Gamma}} g_{\nu,\gamma}\varphi(Dx-\gamma), \quad \nu = 1,\dots,m-1,$$

where the transfer function of each $g_{\nu,\gamma}$ is defined by $G_\nu(\omega) = \sum_{\gamma\in\mathbf{\Gamma}} g_{\nu,\gamma}e^{-i\gamma\cdot\omega}$. The wavelet basis is obtained by taking D-dilates and $\mathbf{\Gamma}$-translates of these mother wavelets $\psi_{\nu,j,\gamma}(x) = m^{j/2}\psi_\nu(D^jx-\gamma)$ for $j\in\mathbb{Z}$ and $\gamma\in\mathbf{\Gamma}$. Let us find the conditions on the filters in order to obtain an orthogonal MRA. First φ and its $\mathbf{\Gamma}$-translates have to be orthogonal

$$\begin{aligned}\delta_{0,\gamma} = \langle\varphi,\varphi(\cdot-\gamma)\rangle &= \int |\widehat{\varphi}(\omega)|^2 e^{-i\gamma\cdot\omega}d\omega \\ &= \sum_{\gamma^*\in\mathbf{\Gamma}^*}\int_{V(2\pi\gamma^*)} |\widehat{\varphi}(\omega)|^2 e^{-i\gamma\cdot\omega}d\omega \\ &= \sum_{\gamma^*\in\mathbf{\Gamma}^*}\int_{V(0)} |\widehat{\varphi}(\omega+2\pi\gamma^*)|^2 e^{-i\gamma\cdot\omega}d\omega \\ &= \int_{V(0)}\sum_{\gamma^*\in\mathbf{\Gamma}^*} |\widehat{\varphi}(\omega+2\pi\gamma^*)|^2 e^{-i\gamma\cdot\omega}d\omega\end{aligned}$$

where $V(2\pi\gamma^*)$ are the Voronoi cells of the lattice $2\pi\mathbf{\Gamma}^*$. This gives us a necessary orthogonality condition on the scaling function

$$\sum_{\gamma^*\in\mathbf{\Gamma}^*} |\widehat{\varphi}(\omega+2\pi\gamma^*)|^2 = \frac{1}{|V(0)|}. \tag{3.3}$$

We will now find out the corresponding condition on the filter function H. In the Fourier domain the dilation equation (3.1) becomes

$$\widehat{\varphi}(\omega) = H(D^{-\mathrm{T}}\omega)\widehat{\varphi}(D^{-\mathrm{T}}\omega). \tag{3.4}$$

Let $\mathbf{\Gamma}_0^* = \{r_1^*, \ldots, r_m^*\}$ be a representative of the quotient group $\mathbf{\Gamma}^*/D^{\mathrm{T}}\mathbf{\Gamma}^*$ so that every $\gamma^* \in \mathbf{\Gamma}^*$ can be written uniquely as $\gamma^* = \gamma_0^* + \gamma_1^*$, where $\gamma_0^* \in \mathbf{\Gamma}_0^*$ and $\gamma_1^* \in D^{\mathrm{T}}\mathbf{\Gamma}^*$. We then have

$$\begin{aligned}
\sum_{\gamma^* \in \mathbf{\Gamma}^*} |\widehat{\varphi}(\omega + 2\pi\gamma^*)|^2 &= \sum_{\gamma_0^* \in \mathbf{\Gamma}_0^*} \sum_{\gamma^* \in D^{\mathrm{T}}\mathbf{\Gamma}^* + \gamma_0^*} |\widehat{\varphi}(\omega + 2\pi\gamma^*)|^2 \\
&= \sum_{\gamma_0^* \in \mathbf{\Gamma}_0^*} \sum_{\gamma^* \in \mathbf{\Gamma}^*} |\widehat{\varphi}(\omega + 2\pi D^{\mathrm{T}}\gamma^* + 2\pi\gamma_0^*)|^2 \\
&= \sum_{\gamma_0^* \in \mathbf{\Gamma}_0^*} |H(D^{-\mathrm{T}}\omega + 2\pi D^{-\mathrm{T}}\gamma_0^*)|^2 \\
&\qquad \times \sum_{\gamma^* \in \mathbf{\Gamma}^*} |\widehat{\varphi}(D^{-\mathrm{T}}\omega + 2\pi\gamma^* + 2\pi D^{-\mathrm{T}}\gamma_0^*)|^2 \\
&= \frac{1}{|V(0)|} \sum_{\gamma_0^* \in \mathbf{\Gamma}_0^*} |H(D^{-\mathrm{T}}\omega + 2\pi D^{-\mathrm{T}}\gamma_0^*)|^2
\end{aligned}$$

by (3.3), (3.4) and the $2\pi\mathbf{\Gamma}^*$-periodicity of $H(\omega)$. Just as in dimension one this leads to the orthogonality condition

$$\sum_{\gamma_0^* \in \mathbf{\Gamma}_0^*} |H(\omega + 2\pi D^{-\mathrm{T}}\gamma_0^*)|^2 = 1 \tag{3.5}$$

By similar calculations as above we see that orthogonality for the whole multiresolution is equivalent to the m-by-m modulation matrix $M(\omega)$ being unitary, where

$$\begin{aligned}
M(\omega)_{1,l} &= H(\omega + 2\pi D^{-\mathrm{T}}r_l^*) & l &= 1, \ldots, m, \\
M(\omega)_{\nu+1,l} &= G_\nu(\omega + 2\pi D^{-\mathrm{T}}r_l^*) & \nu &= 1, \ldots, m-1.
\end{aligned}$$

In the same way, biorthogonality requires that

$$\widetilde{M}(\omega)\overline{M(\omega)}^{\mathrm{T}} = I.$$

3.3 Diagonalization

Now when we have introduced the appropriate framework and notation for non-separable multidimensional wavelets it is fairly straightforward to generalize the previous diagonalization technique. Let us therefore assume that we are given a biorthogonal wavelet system in $\mathbb{R}^n$ with scaling functions φ

and $\widetilde{\varphi}$, m mother wavelets ψ_ν and $\widetilde{\psi}_\nu$, and an n-dimensional convolution operator K. Just as before we define the new wavelets as

$$\widetilde{\psi}_\nu^{\mathrm{K}} = K^* \widetilde{\psi}_\nu \quad \text{and} \quad \psi_\nu^{\mathrm{K}} = K^{-1} \psi_\nu. \tag{3.6}$$

As before a necessary condition on the operator K to obtain diagonalization is that K commutes with dilations

$$\kappa = \frac{\widehat{k}(D^{\mathrm{T}}\omega)}{\widehat{k}(\omega)} \quad \text{independent of } \omega. \tag{3.7}$$

The wavelet expansion of Kf is then given by

$$Kf = \sum_{\nu=1}^{m} \sum_{j\in\mathbb{Z}} \sum_{l\in\mathbf{\Gamma}} \kappa^j \langle f, \widetilde{\psi}_{\nu,j,l}^{\mathrm{K}} \rangle \psi_{\nu,j,l}. \tag{3.8}$$

Again we try to find new scaling functions by setting

$$\widehat{\widetilde{\varphi}^{\mathrm{K}}}(\omega) = \overline{\widehat{\ell}(\omega)} \widehat{\widetilde{\varphi}}(\omega) \quad \text{and} \quad \widehat{\varphi^{\mathrm{K}}}(\omega) = \frac{1}{\overline{\widehat{\ell}(\omega)}} \widehat{\varphi}(\omega),$$

which gives us the new filter functions

$$H^{\mathrm{K}}(\omega) = \frac{\widehat{\ell}(\omega)}{\widehat{\ell}(D^{\mathrm{T}}\omega)} H(\omega), \qquad G_\nu^{\mathrm{K}}(\omega) = \frac{\widehat{\ell}(\omega)}{\widehat{k}(D^{\mathrm{T}}\omega)} G_\nu(\omega),$$

$$\widetilde{H}^{\mathrm{K}}(\omega) = \frac{\overline{\widehat{\ell}(D^{\mathrm{T}}\omega)}}{\overline{\widehat{\ell}(\omega)}} \widetilde{H}(\omega), \qquad \widetilde{G}_\nu^{\mathrm{K}}(\omega) = \frac{\overline{\widehat{k}(D^{\mathrm{T}}\omega)}}{\overline{\widehat{\ell}(\omega)}} \widetilde{G}_\nu(\omega).$$

By calculations identical with those in the one-dimensional case we see that we get biorthogonal filters if $m(\omega)$ is $2\pi\mathbf{\Gamma}^*$-periodic where, $m(\omega) = \widehat{k}(\omega)/\widehat{\ell}(\omega)$.

As in one dimension $m(\omega)$ must be chosen so that

$$\frac{\widehat{k}(\omega)}{m(\omega)} \to 1 \quad \text{as} \quad \omega \to 0.$$

For simplicity we consider the two-dimensional case only. In analogy with the one-dimensional case we try to write

$$m(\omega) = \widehat{k}(-ia(\omega), -ib(\omega)) \tag{3.9}$$

where $a(\omega)$ and $b(\omega)$ are $2\pi\mathbf{\Gamma}^*$-periodic functions such that $a(\omega) = i\omega_1 + o(|\omega|)$ and $b(\omega) = i\omega_2 + o(|\omega|)$ as $\omega \to 0$. Let γ_1 and γ_2 be the column vectors of Γ and try with

$$\begin{aligned} a(\omega) &= a_1(e^{i\gamma_1\cdot\omega} - 1) + a_2(e^{i\gamma_2\cdot\omega} - 1) \\ &= i(a_1\gamma_1 + a_2\gamma_2)\cdot\omega + o(|\omega|) \qquad \text{as } \omega \to 0. \end{aligned}$$

We now chose a_1 and a_2 such that $(a_1\gamma_1 + a_2\gamma_2)\cdot\omega = \omega_1$, i.e such that

$$\Gamma \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

and similarly for $b(\omega)$. This can be written in the condensed form

$$m(\omega) = \widehat{k}(-i\Gamma^{-\mathrm{T}}(e^{\Gamma^{\mathrm{T}}\omega} - \mathbf{1})), \tag{3.10}$$

where the exponential of the vector $\Gamma^{\mathrm{T}}\omega$ is taken element wise. This choice can be thought of as a one-sided difference approximation of k in the directions γ_1 and γ_2. That is, if k is a directional derivative in one of the directions γ_i, then m will be a one-sided difference approximation in that direction.

3.4 Examples

To conclude the multidimensional case we now consider two examples where the operator is the laplacian and a directional derivative, respectively.

3.4.1 The Laplacian

If $K = \Delta$ we have $\widehat{k}(\omega) = -|\omega|^2$. From the diagonalization condition (3.7) we must have

$$\kappa = \frac{|D^{\mathrm{T}}\omega|^2}{|\omega|^2} \quad \text{independent of } \omega.$$

It follows that the dilation matrix has to be of the form

$$D = mP,$$

for some positive integer m and orthogonal matrix P. We then have $\kappa = m^2$. Now, if the original wavelets have compact support so will the new wavelets. On the other hand we see that the new low-pass filter

$$H^{\mathrm{K}}(\omega) = \frac{1}{\kappa}\frac{|\Gamma^{-\mathrm{T}}(e^{(D\Gamma)^{\mathrm{T}}\omega} - 1)|^2}{|\Gamma^{-\mathrm{T}}(e^{\Gamma^{\mathrm{T}}\omega} - 1)|^2} H(\omega),$$

is infinite impulse response (IIR) since we can not factor the denominator and simplify as in the one-dimensional case. All the other filter functions will also be IIR which means that the new scaling functions do not have compact support. We also note that operators with

$$\widehat{k}(\omega) = |\omega|^{\alpha} \quad \text{where } \alpha \in \mathbb{R},$$

can be handled in the same way as the laplacian.

3.4.2 Directional Derivatives

Now suppose that Kf is the directional derivative of f in the direction v

$$Kf = \frac{\partial f}{\partial v} = \nabla f \cdot v,$$

where v is a unit vector in $\mathbb{R}^n$. Then

$$\widehat{k}(\omega) = i\omega \cdot v,$$

From the diagonalization condition (3.7) we must have

$$\kappa = \frac{D^{\mathrm{T}}\omega \cdot v}{\omega \cdot v} \quad \text{independent of } \omega.$$

It follows that the dilation matrix has to be of the form

$$D = mI,$$

where m is a positive integer. In this case the most reasonable choice would be $m = 2$.

If one of the lattice directions is v, the new filters will be finite. Otherwise, they will be infinite.

Chapter 4

Diagonalization of inhomogeneous convolution operators

In this section we will relax the very restrictive homogeneity condition (2.10). Once we give up homogeneity, we can no longer obtain diagonalization in classical wavelet bases. However, if we are willing to accept different wavelets at different scales, our method can be generalized to convolution operators that are asymptotically homogeneous, for instance, linear differential operators with constant coefficients. We will also generalize the result for non-separable wavelets in several dimensions in a similar way.

4.1 The New Wavelet Basis

4.1.1 The Wavelet-Vaguelette Decomposition

Let us start with recalling the *wavelet-vaguelette decomposition.* The fundamental idea is that we have an operator K and want to expand Kf in a given biorthogonal wavelet basis,

$$Kf = \sum_{j,k} \langle Kf, \widetilde{\psi}_{j,k} \rangle \psi_{j,k} = \sum_{j,k} \langle f, K^* \widetilde{\psi}_{j,k} \rangle \psi_{j,k}.$$

We define the *vaguelettes* as

$$\widetilde{\psi}^{\mathrm{K}}_{j,k} = \frac{1}{\kappa_{j,k}} K^* \widetilde{\psi}_{j,k} \quad \text{and} \quad \psi^{\mathrm{K}}_{j,k} = \kappa_{j,k} K^{-1} \psi_{j,k} \tag{4.1}$$

It's a direct consequence of the definition that the vaguelettes are biorthogonal,

$$\langle \psi^{\mathrm{K}}_{j,k}, \widetilde{\psi}^{\mathrm{K}}_{j',k'} \rangle = \delta_{j,j'} \delta_{k,k'}.$$

We can now write Kf as

$$Kf = \sum_{j,k} \kappa_{j,k} \langle f, \widetilde{\psi}^{\mathrm{K}}_{j,k} \rangle \psi_{j,k}.$$

The constants $\kappa_{j,k}$ are referred to as *quasi-singular values* of K and they should be chosen so that the $\widetilde{\psi}_{j,k}, \widetilde{\psi}^{\mathrm{K}}_{j,k}$ all are of comparable size. This is necessary since we require the vaguelettes to constitute biorthogonal dual Riesz bases of $L^2(\mathbb{R})$.

4.1.2 Asymptotically Homogeneous Convolution Operators

In the previous chapters we saw that if K is a convolutional operator with kernel k, $Kf = k * f$, and if the kernel is homogeneous, i.e.

$$\frac{\widehat{k}(2\omega)}{\widehat{k}(\omega)} = \kappa \quad \text{independent of } \omega,$$

then $\{\widetilde{\psi}_{j,k}\}$ and $\{\widetilde{\psi}^{\mathrm{K}}_{j,k}\}$ are wavelets in the classical sense and $\kappa_j = \kappa^j$. This immediately leads to the idea that if k is "almost" homogeneous in some sense, then the $\{\psi^{\mathrm{K}}_{j,k}\}$ and $\{\widetilde{\psi}^{\mathrm{K}}_{j,k}\}$ will "almost" be wavelets. We will only consider the case of asymptotically homogeneous operators here. With asymptotically homogeneous convolution operators we mean that the Fourier multiplier $\widehat{k}(\omega)$ satisfies

$$\begin{aligned} \widehat{k}(\omega) &\sim |\omega|^{\alpha} \quad (\text{or } \omega^m) \quad \text{as} \quad |\omega| \to \infty \\ \widehat{k}(\omega) &\sim |\omega|^{\beta} \quad (\text{or } \omega^n) \quad \text{as} \quad \omega \to 0 \end{aligned} \tag{4.2}$$

where $\alpha, \beta \in \mathbb{R}$ and $m, n \in \mathbb{Z}$. We also require $\widehat{k}(\omega)$ to have no zeros or singularities other than possibly at $\omega = 0$. Those requirements together with some natural assumptions about regularity and vanishish moments for ψ and $\widetilde{\psi}$ will be enough to ensure that $\psi^{\mathrm{K}}_{j,k}$ and $\widetilde{\psi}^{\mathrm{K}}_{j,k}$ constitute Riesz bases. For simplicity we will assume that the exponents in (4.2) are nonnegative. Negative exponents can be treated similarly by just interchanging the roles of $\{\psi^{\mathrm{K}}_{j,k}\}$ and $\{\widetilde{\psi}^{\mathrm{K}}_{j,k}\}$. A typical example of an asymptotically homogeneous convolution is a linear differential operator with constant coefficients. In this case, m is the order of the highest order derivative term, and n is the order of the lowest order term.

The new wavelets will no longer be wavelets in the classical sense, i.e. dilations and translations of a single mother wavelet. Instead, at each level j, they will be translations of "scale-dependent" mother wavelets $\psi^{\mathrm{K}}_{j,0}$ and $\widetilde{\psi}^{\mathrm{K}}_{j,0}$, due to the translation invariant nature of K. However, we will have the $2^j x - k$-structure asymptotically as $j \to \pm\infty$. To explain this in more detail, let us write out the new dual wavelets (4.1) in the Fourier domain:

$$\widehat{\widetilde{\psi}^{\mathrm{K}}_{j,k}}(\omega) = \frac{\overline{\widehat{k}(\omega)}}{\overline{\kappa_j}}\widehat{\widetilde{\psi}_{j,k}}(\omega) \quad \text{and} \quad \widehat{\psi^{\mathrm{K}}_{j,k}}(\omega) = \frac{\kappa_j}{\widehat{k}(\omega)}\widehat{\psi_{j,k}}(\omega)$$

Note that the quasi-singular values no longer depend on k. For the new dual wavelets we then have (it is enough to consider $k = 0$ only)

$$\widehat{\widetilde{\psi}^{\mathrm{K}}_{j,0}}(\omega) = \frac{\overline{\widehat{k}(\omega)}}{\overline{\kappa_j}}\widehat{\widetilde{\psi}_{j,0}}(\omega) = \frac{\overline{\widehat{k}(\omega)}}{\overline{\kappa_j}}2^{-j/2}\widehat{\widetilde{\psi}}(2^{-j}\omega)$$

Since $\widehat{\widetilde{\psi}}(2^{-j}\omega)$ is concentrated around $\omega = 2^j$, we can approximate $\widehat{k}(\omega)$ with $|\omega|^\alpha$ for large j, which gives us

$$\begin{aligned}\widehat{\widetilde{\psi}^{\mathrm{K}}_{j,0}}(\omega) &\approx \frac{|\omega|^\alpha}{\overline{\kappa_j}}2^{-j/2}\widehat{\widetilde{\psi}}(2^{-j}\omega) \\ &= \frac{1}{\overline{\kappa_j}}\frac{|\omega|^\alpha}{|2^{-j}\omega|^\alpha}2^{-j/2}\left|2^{-j}\omega\right|^\alpha\widehat{\widetilde{\psi}}(2^{-j}\omega) \\ &= \frac{2^{j\alpha}}{\overline{\kappa_j}}2^{-j/2}\widehat{\widetilde{\psi}^{(\alpha)}}(2^{-j}\omega)\end{aligned}$$

where we have used the notation $\widehat{\widetilde{\psi}^{(\alpha)}}(\omega) = |\omega|^\alpha\,\widehat{\widetilde{\psi}}(\omega)$.

In other words, as $j \to \infty$, the dual wavelets will "almost" be dilations and translations of the "mother wavelet" $\widetilde{\psi}^{(\alpha)}$. In the same way, as $j \to -\infty$ they will almost be dilations and translations of $\widetilde{\psi}^{(\beta)}$. For the primal wavelets the same thing will hold except for that the exponents will be $-\alpha$ and $-\beta$ instead. We also get a hint about how we should choose the quasi-singular values. For j large we need to have $\kappa_j \sim 2^{j\alpha}$ and as $j \to -\infty$ we must choose $\kappa_j \sim 2^{j\beta}$.

4.2 The New Multiresolution Structure

Note that on each level, the wavelets are translates of each other. Therefore one can hope for a multiresolution structure with different filters at each level. Following the approach in 2.3 we try to find level-dependent scaling functions

$$\widehat{\widetilde{\varphi}^{\mathrm{K}}_{j,k}}(\omega) = \overline{\widehat{l_j}(\omega)}\widehat{\widetilde{\varphi}_{j,k}}(\omega) \quad \text{and} \quad \widehat{\varphi^{\mathrm{K}}_{j,k}}(\omega) = \frac{1}{\widehat{l_j}(\omega)}\widehat{\varphi_{j,k}}(\omega). \tag{4.3}$$

What about the filters? Simple calculations gives

$$\begin{aligned}\widehat{\varphi^{\mathrm{K}}_{j-1,0}}(\omega) &= \frac{1}{\widehat{l_{j-1}}(\omega)}\widehat{\varphi_{j-1,0}}(\omega)\\ &= \frac{1}{\widehat{l_{j-1}}(\omega)}H(2^{-j}\omega)\widehat{\varphi_{j,0}}(\omega)\\ &= \frac{\widehat{l_j}(\omega)}{\widehat{l_{j-1}}(\omega)}H(2^{-j}\omega)\widehat{\varphi^{\mathrm{K}}_{j,0}}(\omega)\end{aligned}$$

and

$$\begin{aligned}\widehat{\psi^{\mathrm{K}}_{j-1,0}}(\omega) &= \frac{\kappa_{j-1}}{\widehat{k}(\omega)}\widehat{\psi_{j-1,0}}(\omega)\\ &= \frac{\kappa_{j-1}}{\widehat{k}(\omega)}G(2^{-j}\omega)\widehat{\varphi_{j,0}}(\omega)\\ &= \kappa_{j-1}\frac{\widehat{l_j}(\omega)}{\widehat{k}(\omega)}G(2^{-j}\omega)\widehat{\varphi^{\mathrm{K}}_{j,0}}(\omega)\end{aligned}$$

Doing the same thing for the dual wavelets and scaling functions we see that we will get level-dependent filters $h^{\mathrm{K}}_j, g^{\mathrm{K}}_j, \widetilde{h}^{\mathrm{K}}_j$ and $\widetilde{g}^{\mathrm{K}}_j$, whose transfer functions are given by

(4.4)
$$\begin{aligned}H^{\mathrm{K}}_j(\omega) &= \frac{\widehat{l_j}(2^j\omega)}{\widehat{l_{j-1}}(2^j\omega)}H(\omega) \qquad & G^{\mathrm{K}}_j(\omega) &= \kappa_{j-1}\frac{\widehat{l_j}(2^j\omega)}{\widehat{k}(2^j\omega)}G(\omega)\\ \widetilde{H}^{\mathrm{K}}_j(\omega) &= \frac{\overline{\widehat{l_{j-1}}(2^j\omega)}}{\overline{\widehat{l_j}(2^j\omega)}}\widetilde{H}(\omega) \qquad & \widetilde{G}^{\mathrm{K}}_j(\omega) &= \frac{1}{\kappa_{j-1}}\frac{\overline{\widehat{k}(2^j\omega)}}{\overline{\widehat{l_j}(2^j\omega)}}\widetilde{G}(\omega)\end{aligned}$$

For all this to make sense we have to make sure that those are 2π-periodic functions. This is equivalent to the $2^j 2\pi$-periodicity of

$$m_j(\omega) := \frac{\widehat{k}(\omega)}{\widehat{l_j}(\omega)}$$

since this also implies that

$$\frac{\widehat{l_{j-1}}(\omega)}{\widehat{l_j}(\omega)} = \frac{m_j(\omega)}{m_{j-1}(\omega)}$$

is $2^j 2\pi$-periodic. Writing out (4.3) in another way,

(4.5) $$\overline{m_j(\omega)\widehat{\widetilde{\varphi}^{\mathrm{K}}_{j,k}}(\omega)} = \overline{\widehat{k}(\omega)\widehat{\widetilde{\varphi}_{j,k}}(\omega)} \quad \text{and} \quad \widehat{k}(\omega)\widehat{\varphi^{\mathrm{K}}_{j,k}}(\omega) = m_j(\omega)\widehat{\varphi_{j,k}}(\omega)$$

we also see that m_j must match $\widehat{k}(\omega)$ at $\omega = 0$. We therefore make the natural choice

(4.6) $$m_j(\omega) = \widehat{k}(-i2^j(e^{i2^{-j}\omega} - 1))$$

which can be regarded as a discrete approximation of $\widehat{k}(\omega)$ at the scale 2^{-j}. We have to take extra care here, so that $m_j(\omega)$ does not have zeros at other points than $\omega = k2^{-j}2\pi, k \in \mathbb{Z} \setminus \{0\}$, since this will lead to division by zero in (4.4). Contrary to the homogeneous case, $\widehat{k}(\omega)$ can have zeros in the complex plane other than $\omega = 0$ and if one of the "circles" $-i2^j(e^{i2^{-j}\omega} - 1)$ passes through such a zero this will cause the corresponding $m_j(\omega)$ to have a "forbidden" zero. We do not have a general method to solve this problem, but we show below how it can be circumvented in certain cases.

To summarize, we have a new biorthogonal multiresolution analysis with wavelets given by

$$\widehat{\widetilde{\psi}^{\mathrm{K}}_{j,k}}(\omega) = \frac{\overline{\widehat{k}(\omega)}}{\overline{\kappa_j}}\widehat{\widetilde{\psi}_{j,k}}(\omega) \quad \text{and} \quad \widehat{\psi^{\mathrm{K}}_{j,k}}(\omega) = \frac{\kappa_j}{\widehat{k}(\omega)}\widehat{\psi_{j,k}}(\omega),$$

level-dependent scaling functions defined by the relations

$$\overline{\widehat{k}(\omega)}\widehat{\widetilde{\varphi}^{\mathrm{K}}_{j,k}}(\omega) = \overline{\widehat{k}(-i2^j(e^{i2^{-j}\omega} - 1))}\widehat{\widetilde{\varphi}_{j,k}}(\omega)$$

and

$$\widehat{k}(\omega)\widehat{\varphi_{j,k}}(\omega) = \widehat{k}(-i2^j(e^{i2^{-j}\omega} - 1))\widehat{\varphi^{\mathrm{K}}_{j,k}}(\omega)$$

and level-dependent filters given by

$$H^{\mathrm{K}}_j(\omega) = \frac{\widehat{k}(-i2^{j-1}(e^{i2\omega} - 1))}{\widehat{k}(-i2^j(e^{i\omega} - 1))}H(\omega) \qquad G^{\mathrm{K}}_j(\omega) = \frac{\kappa_{j-1}}{\widehat{k}(-i2^j(e^{i\omega} - 1))}G(\omega)$$

$$\widetilde{H}^{\mathrm{K}}_j(\omega) = \frac{\overline{\widehat{k}(-i2^j(e^{i\omega} - 1))}}{\overline{\widehat{k}(-i2^{j-1}(e^{i2\omega} - 1))}}\widetilde{H}(\omega) \qquad \widetilde{G}^{\mathrm{K}}_j(\omega) = \frac{\overline{\widehat{k}(-i2^j(e^{i\omega} - 1))}}{\overline{\kappa_{j-1}}}\widetilde{G}(\omega)$$

With those scaling functions we have an associated "generalized" biorthogonal multiresolution analysis with approximation spaces

$$V^{\mathrm{K}}_j = \overline{span\{\varphi^{\mathrm{K}}_{j,k}\}_k} \quad \text{and} \quad \widetilde{V}^{\mathrm{K}}_j = \overline{span\{\widetilde{\varphi}^{\mathrm{K}}_{j,k}\}_k},$$

and detail spaces

$$W^{\mathrm{K}}_j = \overline{span\{\psi^{\mathrm{K}}_{j,k}\}_k} \quad \text{and} \quad \widetilde{W}^{\mathrm{K}}_j = \overline{span\{\widetilde{\psi}^{\mathrm{K}}_{j,k}\}_k}.$$

These approximation spaces satisfies all the requirements of an "ordinary" biorthogonal multiresolution analysis with the only exception that they are not scaled versions of each other. When it comes to the detail spaces, there is a a technicality that we must take care of. We have to check that the W^{K}_j-spaces complements the V^{K}_j-spaces in the V^{K}_{j+1}-spaces. It is not difficult to see that this is the same as requiring the (level-dependent) modulation matrix to be nonsingular. Let us see if this is the case by looking at its determinant,

$$\begin{aligned}\Delta_j^{\mathrm{K}}(\omega) &= \begin{vmatrix} H_j^{\mathrm{K}}(\omega) & H_j^{\mathrm{K}}(\omega+\pi) \\ G_j^{\mathrm{K}}(\omega) & G_j^{\mathrm{K}}(\omega+\pi) \end{vmatrix} \\ &= H_j^{\mathrm{K}}(\omega)G_j^{\mathrm{K}}(\omega+\pi) - H_j^{\mathrm{K}}(\omega+\pi)G_j^{\mathrm{K}}(\omega) \\ &= \frac{m_j(2^j\omega)}{m_{j-1}(2^j\omega)} m_j(2^j\omega + 2^j\pi) H(\omega) G(\omega+\pi) \\ &\quad - \frac{m_j(2^j\omega+2^j\pi)}{m_{j-1}(2^j\omega+2^j\pi)} m_j(2^j\omega) H(\omega+\pi) G(\omega) \\ &= \frac{m_j(2^j\omega)}{m_{j-1}(2^j\omega)} m_j(2^j\omega + 2^j\pi) \Delta(\omega).\end{aligned}$$

where $\Delta(\omega) = H(\omega)G(\omega+\pi) - H(\omega+\pi)G(\omega)$ is the determinant of the "original" modulation matrix which we know is nonzero. Also, $m_j(2^j\omega)$ has its zeros at $\omega = k2\pi$, $m_{j-1}(2^j\omega)$ at $\omega = k\pi$ and $m_j(2^j\omega + 2^j\pi)$ at odd integer multiples of π. All those zeros are of the same order, so they will always cancel out and $\Delta_j^{\mathrm{K}}(\omega)$ will never be zero. An identical argument shows that the dual detail spaces $\widetilde{W}_j^{\mathrm{K}}$ complement $\widetilde{V}_j^{\mathrm{K}}$ in $\widetilde{V}_{j+1}^{\mathrm{K}}$.

We need to find a relation between the scaling function coefficients of Kf in the original and of f in the new basis, since we always stop the wavelet decomposition at a coarsest scale. Let

$$m_j(\omega) = \sum_l m_{j,l} e^{-il\omega}.$$

Since

$$m_j(\omega)\widehat{\widetilde{\varphi}_{j,k}^{\mathrm{K}}}(\omega) = \widehat{k}(\omega)\widehat{\widetilde{\varphi}_{j,k}}(\omega),$$

we have

$$K^*\widetilde{\varphi}_{j,k} = \sum_l \overline{m_{j,l}}\widetilde{\varphi}_{j,k-l}^{\mathrm{K}},$$

and therefore

$$\langle Kf, \widetilde{\varphi}_{j,k}\rangle = \sum_l m_{j,l}\langle f, \widetilde{\varphi}_{j,k-l}^{\mathrm{K}}\rangle.$$

This leaves us with the following algorithm to compute Kf:

1. Compute the fine-scale coefficients $\langle f, \widetilde{\varphi}_{J,k}^{\mathrm{K}}\rangle$ of f in some V_J^{K}-space (pre-processing).

2. Perform a forward wavelet transform with the level-dependent filters $\widetilde{H}_j$ and $\widetilde{G}_j$ to get the wavelet coefficients $\langle f, \widetilde{\psi}_{j,k}^{\mathrm{K}}\rangle$, $j_0 \le j < J$, in the new basis, and the coarse scale scaling function coefficients $\langle f, \widetilde{\varphi}_{j_0,k}^{\mathrm{K}}\rangle$.

3. Transform into the coefficients of Kf in the original basis, $\langle f, \widetilde{\psi}_{j,k}\rangle = \kappa_j \langle f, \widetilde{\psi}^{\mathrm{K}}_{j,k}\rangle$ and $\langle Kf, \widetilde{\varphi}_{j_0,k}\rangle = \sum_l m_{j_0,l}\langle f, \widetilde{\varphi}^{\mathrm{K}}_{j_0,k-l}\rangle$.

4. Perform an inverse wavelet transform with the filters H and G to get the fine-scale coefficients $\langle Kf, \widetilde{\varphi}_{J,k}\rangle$.

5. Compute sample values of Kf from these coefficients (post-processing).

4.3 Stability, Regularity and Decay of the New Wavelets

In this section we will show that, under suitable assumptions on the original wavelets, the new wavelets will constitute Riesz bases, i.e. they will satisfy the stability condition

$$A\,\|f\|^2 \le \sum_{j,k} |\langle f, \widetilde{\psi}^{\mathrm{K}}_{j,k}\rangle|^2 \le B\,\|f\|^2\,. \tag{4.7}$$

and similarly for the duals.

First, we have to make sure that the new wavelets are valid $L^2(\mathbb{R})$-objects. For the duals, this is equivalent to the dual mother wavelet $\widetilde{\psi}$ belonging to the Sobolev space W^α. In general, if the dual mother wavelet belongs to W^s, the new dual wavelets will belong to $W^{s-\alpha}$. Similarly, the new primal wavelets will "gain" α $L^2(\mathbb{R})$-derivatives. The new primal wavelets will be $L^2(\mathbb{R})$-functions, if the mother wavelet ψ has $\lceil\beta\rceil$ vanishing moments. The decay of the new wavelets can be analyzed in just the same way as in the homogeneous case so we now turn over to the stability issue instead.

Intuitively, it seams clear from (4.2) that if the systems generated from the mother wavelets $\psi^{-\alpha}, \widetilde{\psi}^{\alpha}$ and $\psi^{-\beta}, \widetilde{\psi}^{\beta}$ are stable, so is the new system. Let us be a little bit more precise about this. In the Fourier domain, we have

$$\begin{aligned}
\langle f, \widetilde{\psi}^{\mathrm{K}}_{j,k}\rangle &= \int \widehat{f}(\omega)\widehat{\widetilde{\psi}^{\mathrm{K}}_{j,k}}(\omega)d\omega \\
&= \frac{1}{\kappa_j}\int \widehat{f}(\omega)\widehat{k}(\omega)\widehat{\widetilde{\psi}_{j,k}}(\omega)d\omega \\
&= \frac{1}{\kappa_j}\int \widehat{f}(\omega)\frac{\widehat{k}(\omega)}{|\omega|^\alpha}|\omega|^\alpha\,\widehat{\widetilde{\psi}_{j,k}}(\omega)d\omega.
\end{aligned}$$

For simplicity we assume that the original wavelets have compact support in the frequency domain (this restriction can be removed to the prize of adding some more technicalities). Since

$$\frac{\widehat{k}(\omega)}{|\omega|^\alpha} \to 1 \quad \text{as} \quad |\omega| \to \infty$$

and

$$\kappa_j \to 2^{j\alpha} \quad \text{as} \quad |\omega| \to \infty$$

there is some J such that for j larger than J, $\kappa_j \leq 2^{j\alpha+1}$ and

$$\frac{1}{2} \leq \frac{\widehat{k}(\omega)}{|\omega|^{\alpha}} \leq 2$$

holds on the support of $\widehat{\widetilde{\psi}_{j,k}}(\omega)$. For $j \geq J$ we therefore have

$$|\langle f, \widetilde{\psi}^{\mathrm{K}}_{j,k}\rangle| \leq 4 \left| \int \widehat{f}(\omega) \frac{|\omega|^{\alpha}}{2^{j\alpha}} \widehat{\widetilde{\psi}_{j,k}}(\omega) d\omega \right| = 4\, |\langle f, \widetilde{\psi}^{\alpha}_{j,k}\rangle|$$

If the basis generated by $\widetilde{\psi}^{\alpha}$ is a Riesz basis we then have

$$\sum_{\substack{j>J \\ k\in\mathbb{Z}}} |\langle f, \widetilde{\psi}^{\mathrm{K}}_{j,k}\rangle|^2 \leq B\, \|f\|^2 ,$$

fore some constant B. An identical argument for $j \to -\infty$ gives us

$$\sum_{\substack{j<J' \\ k\in\mathbb{Z}}} |\langle f, \widetilde{\psi}^{\mathrm{K}}_{j,k}\rangle|^2 \leq B\, \|f\|^2 ,$$

for another constant B and some J'. Now we are left with a finite range of scales, so it is enough to prove the corresponding inequality for each scale separately, i.e. we have to show

$$\sum_{k} |\langle f, \widetilde{\psi}^{\mathrm{K}}_{j,k}\rangle|^2 \leq B\, \|f\|^2 ,$$

for $J' \leq j \leq J$. A standard integral inequality gives us

$$\sum_{k} |\langle f, \widetilde{\psi}^{\mathrm{K}}_{j,k}\rangle|^2 \leq \|\widetilde{\psi}^{\mathrm{K}}_{j}\|_1 \sup_{x} \sum_{k} |\widetilde{\psi}^{\mathrm{K}}_{j}(x - 2^{-j}k)|\ \ \|f\|^2 .$$

Under some minimal conditions on the decay of the new wavelets, e.g.

$$|\widetilde{\psi}^{\mathrm{K}}_{j}(x)| \leq \frac{C}{(1+|x|)^{1+\epsilon}},$$

we have thus shown that

$$\sum_{j,k} |\langle f, \widetilde{\psi}^{\mathrm{K}}_{j,k}\rangle|^2 \leq B\, \|f\|^2 .$$

In precisely the same way we can show that

$$\sum_{j,k} \left|\langle f, \psi^{\mathrm{K}}_{j,k}\rangle\right|^2 \leq B\, \|f\|^2 .$$

Together with the biortogonality, this is sufficient to ensure that $\{\psi^{\mathrm{K}}_{j,k}\}$ and $\{\widetilde{\psi}^{\mathrm{K}}_{j,k}\}$ are Riesz bases. As a remark, we note that the stability constants can blow up if the new wavelets have large L^1-norms.

4.4 Some examples

In this section we will consider some examples in order to illustrate some aspects of the new wavelet system. Let us start with $K = I - D$ where I is the identity and D is the derivative operator. The new dual wavelets will be

$$\widetilde{\psi}^{\mathrm{K}}_{j,k} = \frac{1}{\kappa_j}(\widetilde{\psi}_{j,k} + 2^j \widetilde{\psi}'_{j,k})$$

where $\widetilde{\psi}'_{j,k}$ means dilations and translations of $\widetilde{\psi}'$ and *not* derivatives of $\widetilde{\psi}_{j,k}$. We see that $\kappa_j = 1+2^j$ is a suitable choice of quasi-singular values. Another possible choice (see Donoho [6]) is $\kappa_j = max(1, 2^j)$. At the fine scales, the new dual wavelets are essentially dilates and translates of $\widetilde{\psi}'$. For the coarse scales we get the old system. In figure (4.1) we plot these new wavelets for a range of scales. For comparison, they are all normalized to scale one, i.e. we plot $2^{-j/2}\widetilde{\psi}^{\mathrm{K}}_{j,0}(2^{-j}x)$. One can clearly see how these "level-dependent mother wavelets" evolves as j increases, from the original mother wavelet to its derivative. As original (dual) mother wavelet we have used the "Mexican hat function", $\widetilde{\psi}(x) = (1 - 2x^2)e^{-x^2}$.

Let us take a look at the new filters

$$\begin{aligned}
\widetilde{H}^{\mathrm{K}}_j(\omega) &= \frac{1 - 2^j(e^{-i\omega} - 1)}{1 - 2^{j-1}(e^{-i2\omega} - 1)}\widetilde{H}(\omega) \\
\widetilde{G}^{\mathrm{K}}_j(\omega) &= \frac{1 - 2^j(e^{-i\omega} - 1))}{1 + 2^{j-1}}\widetilde{G}(\omega) \\
H^{\mathrm{K}}_j(\omega) &= \frac{1 - 2^{j-1}(e^{i2\omega} - 1)}{1 - 2^j(e^{i\omega} - 1)}H(\omega) \\
G^{\mathrm{K}}_j(\omega) &= \frac{1 + 2^{j-1}}{1 - 2^j(e^{i\omega} - 1))}G(\omega)
\end{aligned}$$

For large j we get

$$\begin{aligned}
\widetilde{H}^{\mathrm{K}}_j(\omega) &\approx \frac{2}{e^{-i\omega} + 1}\widetilde{H}(\omega) \\
\widetilde{G}^{\mathrm{K}}_j(\omega) &\approx 2(1 - e^{-i\omega})\widetilde{G}(\omega) \\
H^{\mathrm{K}}_j(\omega) &\approx \frac{e^{i\omega} + 1}{2}H(\omega) \\
G^{\mathrm{K}}_j(\omega) &\approx \frac{1}{2(1 - e^{-i\omega})}G(\omega)
\end{aligned}$$

i.e. approximately the filter we would get with the derivative operator. As $j \to -\infty$ they will approximately be the old filters.

We mentioned above that problems could arise when $\widehat{k}(\omega)$ has other zeros in the complex plane than the origin. Let us illustrate this by considering

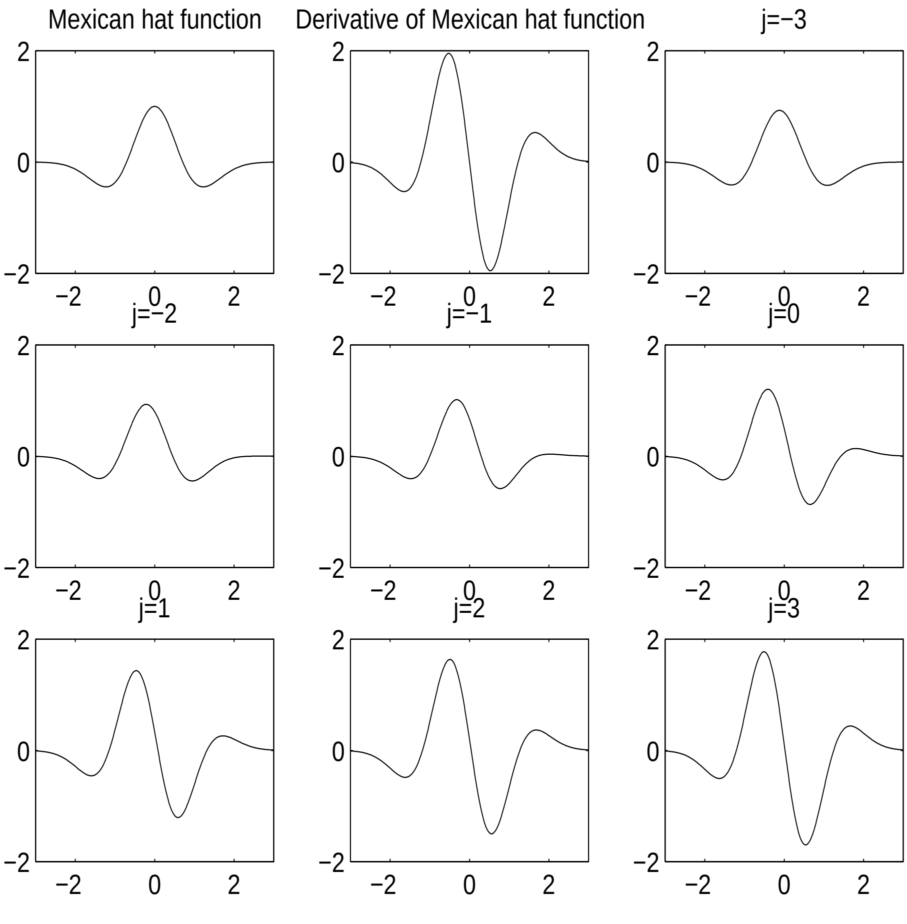

Mexican hat function
Derivative of Mexican hat function
j=−3
j=−2
j=−1
j=0
j=1
j=2
j=3
2
0
−2
−2
0
2

Figure 4.1:

$K = I + D$. Here we have

$$\widehat{k}(\omega) = 1 + i\omega,$$

so we will have a zero at $\omega = i$. Indeed we will get into difficulties; the m_j's becomes

$$m_j(\omega) = 1 + 2^j(e^{i2^{-j}\omega} - 1),$$

and we will have $m_{-1}(\pi/2) = 0$ which will lead to inconsistency in (4.5). This difficulty can be circumvented by choosing m_j as "backward difference approximations" instead;

$$m_j(\omega) = \widehat{k}(i2^j(e^{-i2^{-j}\omega} - 1)).$$

In the complex plane, this amounts to reflecting the circles $-i2^j(e^{i2^{-j}\omega} - 1)$ in the real axis, hence avoiding the zero at $\omega = i$.

Let us also study the example $K = I - D^2$. Here, $\widehat{k}(\omega) = 1 + \omega^2$ has zeros at $\omega = \pm i$ so the trick in the previous example does not apply. However, this time we can get around the problem even more easily, we just choose $\widehat{k}(\omega) = 1 + |\omega|^2$ as the extension to the complex plane. Note that $m_j(\omega) = 1 - 2^j(e^{-i2^{-j}\omega} - 2 + e^{i2^{-j}\omega})$, so we have made a symmetric difference approximation of the second derivative term.

4.5 More on Multidimensional Convolution Cperators.

In this section we will investigate some examples where the diagonalization condition (3.7) does not hold, but we still can diagonalize the operator in some kind of wavelet-like bases. We will in particular consider the case of anisotropic operators when the dilation matrix contains a rotation. The constructions here will be very similar to those above for one-dimensional convolution operators, so some readers might want to skip this section.

4.5.1 Notation

Let us first recall the notation from chapter 3. The wavelets and scaling functions are defined with respect to a sampling lattice $\mathbf{\Gamma} = \Gamma\mathbb{Z}^d$,where d is the dimension, and a dilation matrix D which is required to satisfy $D\mathbf{\Gamma} \subset \mathbf{\Gamma}$ and $m = |detD| > 1$. The scaling functions and their duals are defined as

$$\varphi_{j,\gamma}(x) = m^{d/2}\varphi(D^jx - \gamma) \quad \text{and} \quad \widetilde{\varphi}_{j,\gamma}(x) = m^{d/2}\widetilde{\varphi}(D^jx - \gamma).$$

The wavelets and dual wavelets are defined as D-dilations and $\mathbf{\Gamma}$-translations of $m - 1$ mother wavelets and dual mother wavelets,

$$\psi_{\nu,j,\gamma}(x) = m^{d/2}\psi_\nu(D^jx - \gamma) \quad \text{and} \quad \widetilde{\psi}_{\nu,j,\gamma}(x) = m^{d/2}\widetilde{\psi}_\nu(D^jx - \gamma),$$

for $\nu = 1, \ldots, m-1$. We have scaling equations

$$(4.8)\qquad \begin{aligned} \varphi(x) =&\ m\sum_\gamma h_\gamma \varphi(Dx-\gamma), & \widetilde{\varphi}(x) =&\ m\sum_\gamma \widetilde{h}_\gamma \varphi(Dx-\gamma) \\ \psi_\nu(x) =& m\sum_\gamma g_{\nu,\gamma}\varphi(Dx-\gamma), & \widetilde{\psi}_\nu(x) =& m\sum_\gamma \widetilde{g}_{\nu,\gamma}\varphi(Dx-\gamma) \end{aligned}$$

for some low-pass filters $h_\gamma, \widetilde{h}_\gamma$ and high-pass filters $g_{\nu,\gamma}$ and $\widetilde{g}_{\nu,\gamma}$. In the Fourier domain, the scaling equations (4.8) becomes

$$(4.9)\qquad \begin{aligned} \widehat{\varphi}(\omega) =&\ H(D^{-\mathrm{T}}\omega)\widehat{\varphi}(D^{-\mathrm{T}}\omega), & \widehat{\widetilde{\varphi}}(\omega) =&\ \widetilde{H}(D^{-\mathrm{T}}\omega)\widehat{\widetilde{\varphi}}(D^{-\mathrm{T}}\omega) \\ \widehat{\psi_\nu}(\omega) =& G_\nu(D^{-\mathrm{T}}\omega)\widehat{\varphi}(D^{-\mathrm{T}}\omega), & \widehat{\widetilde{\psi}_\nu}(\omega) =& \widetilde{G}_\nu(D^{-\mathrm{T}}\omega)\widehat{\widetilde{\varphi}}(D^{-\mathrm{T}}\omega) \end{aligned}$$

with the filter functions

$$\begin{aligned} H(\omega) &= \sum_{\gamma\in\boldsymbol{\Gamma}} h_\gamma e^{-i\gamma\cdot\omega} & \widetilde{H}(\omega) &= \sum_{\gamma\in\boldsymbol{\Gamma}} \widetilde{h}_\gamma e^{-i\gamma\cdot\omega} \\ G_\nu(\omega) &= \sum_{\gamma\in\boldsymbol{\Gamma}} g_{\nu,\gamma} e^{-i\gamma\cdot\omega} & \widetilde{G}_\nu(\omega) &= \sum_{\gamma\in\boldsymbol{\Gamma}} \widetilde{g}_{\nu,\gamma} e^{-i\gamma\cdot\omega} \end{aligned}$$

4.5.2 The New Wavelets

Let us now consider a d-dimensional convolution operator $Kf = k * f$. In chapter 3 we saw that we could construct a new wavelet basis with mother wavelets $\psi^{\mathrm{K}} = K^{-1}\psi$ and $\widetilde{\psi}^{\mathrm{K}} = K^*\widetilde{\psi}$ that gives a diagonal representation of K, provided that the operator commutes with D-dilations, up to a constant κ. In the Fourier domain, this amounts to

$$\frac{\widehat{k}(D^{\mathrm{T}}\omega)}{\widehat{k}(\omega)} = \kappa \quad \text{, independent of } \omega.$$

When $D = 2I$, this just requires K to be homogeneous. But there are cases, for instance the Quincunx lattice, where the dilation matrix is a scaling and a rotation, so this requires K to be rotation invariant as well. This rules out operators of the form

$$\widehat{k}(\omega) = |A\omega|^\alpha \text{ (or } \widehat{k}(\omega) = (A\omega)^n) \quad \text{and} \quad \widehat{k}(\omega) = |v\cdot\omega|^\alpha \text{ (or } \widehat{k}(\omega) = (v\cdot\omega)^n)$$

where A is a symmetric, positive definite matrix and v is a unit vector. Of course we can also have sums of these (e.g. linear partial differential operators) in which case the diagonalization condition will be violated even for $D = 2I$. For those operators, the wavelet-vaguelette decomposition will not give us wavelets that are D-dilations and $\mathbf{\Gamma}$-translations of m mother wavelets. We write out the new wavelets once again,

$$\widetilde{\psi}^{\mathrm{K}}_{\nu,j,\gamma} = \frac{1}{\kappa_{\nu,j}} K^* \widetilde{\psi}_{\nu,j,\gamma} \quad \text{and} \quad \psi^{\mathrm{K}}_{\nu,j,\gamma} = \kappa_{\nu,j} K^{-1} \psi_{\nu,j,\gamma} \tag{4.10}$$

Let us look at the Quincunx lattice to understand the difference between this basis and a "real" wavelet basis. In the latter case, we have one mother wavelet ψ that typically is associated with one direction, i.e. it measures oscillations in one specific direction. The dilation matrix is a 45^o rotation and a scaling with $\sqrt{2}$. The family $\psi_{\nu,j,\gamma}$ is thus the mother wavelet rotated to four different directions (horizontal,vertical and the two diagonal directions) and at different scales. For each direction, the scaling factor is 4, i.e. after four D-dilations we have come back to the same direction but we have scaled by $\sqrt{2}$ four times, that is, by a factor 4. The wavelets defined by (4.10) on the other hand, will be different for the different directions, due to the anisotropic nature of K.

4.5.3 The New Multiresolution Analysis

We derive level-dependent filters and scaling functions just as in the previous chapter with the non-homogeneous convolution operators. We try to find level-dependent scaling functions through

$$\widehat{\widetilde{\varphi}^{\mathrm{K}}_{j,\gamma}}(\omega) = \overline{\widehat{l_j}(\omega)} \widehat{\widetilde{\varphi}_{j,\gamma}}(\omega) \quad \text{and} \quad \widehat{\varphi^{\mathrm{K}}_{j,\gamma}}(\omega) = \frac{1}{\widehat{l_j}(\omega)} \widehat{\varphi_{j,\gamma}}(\omega).$$

This will give us new filter whose filter functions are

$$\begin{aligned} H^{\mathrm{K}}_j(\omega) &= \frac{\widehat{l_j}(D^j\omega)}{\widehat{l_{j-1}}(D^j\omega)} H(\omega) & \widetilde{H}^{\mathrm{K}}_j(\omega) &= \frac{\overline{\widehat{l_{j-1}}(D^j\omega)}}{\overline{\widehat{l_j}(D^j\omega)}} \widetilde{H}(\omega) \\ G^{\mathrm{K}}_{\nu,j}(\omega) &= \kappa_{\nu,j-1} \frac{\widehat{l_j}(D^j\omega)}{\widehat{k}(D^j\omega)} G_\nu(\omega) & \widetilde{G}^{\mathrm{K}}_{\nu,j}(\omega) &= \frac{1}{\kappa_{\nu,j-1}} \frac{\overline{\widehat{k}(D^j\omega)}}{\overline{\widehat{l_j}(D^j\omega)}} \widetilde{G}_\nu(\omega) \end{aligned} \tag{4.11}$$

These are $2\pi\mathbf{\Gamma}^*$-periodic if

$$m_j(\omega) = \frac{\widehat{k}(\omega)}{\widehat{l_j}(\omega)}$$

is $2\pi D^j \mathbf{\Gamma}^*$-periodic. Looking at (2.19) and (4.6) suggests thatwe should choose

$$m_j(\omega) = \widehat{k}(-iD^j\Gamma^{-\mathrm{T}}(e^{\Gamma^{\mathrm{T}} D^{-j}\omega} - 1)).$$

We can describe this in words by something like "a one-sided difference approximation at scale m^{-j}, in the lattice directions rotated j times".

Chapter 5

Giving Up Shift Invariance

In this chapter we will sketch some ideas how to extend the methods described in previous chapters to more general operators that need not be translation invariant. We will still have some requirements on the operators under study though. They have to preserve the most important properties of wavelets, such as localization in time and frequency, regularity and vanishing moments. Typical operators of interest are linear differential operators with varying coefficients and integral operators with kernels rapidly decaying away from the diagonal, i.e. the same operators as in [1]. Our presentation will leave a lot of questions to be answered, and more research is needed to realize the ideas described here.

5.1 General Multiresolution Analyses

Due to the lack of shift invariance, we have to consider a more general notion of multiresolution analysis where the scaling functions can depend both on scale and position. Following Sweldens [15], we define a multiresolution analysis as a sequence of subspaces of $L^2(\mathbb{R})$ such that

1. $V_j \subset V_{j+1}$
2. $\overline{\bigcup_j V_j} = L^2(\mathbb{R})$
3. $\bigcap_j V_j = 0$
4. Each V_j has a basis of *scaling functions* $\{\varphi_{j,k}\}_k$

We still think of V_j as an approximation space at scale 2^{-j} and of $\varphi^{K}_{j,k}$ as living at scale 2^{-j} and position $2^{-j}k$. We also keep using the name scaling functions for the $\varphi_{j,k}$'s even if they cannot be written as linear combinations of scaled versions of themselves. By W_j we denote "detail spaces" complementing V_j in V_{j+1}. A collection of functions $\psi_{j,k}$ are called *wavelets* if they are Riesz bases for the W_j spaces, i.e.

$$W_j = \overline{span\{\psi_{j,k}\}_k} \quad \text{and} \quad A\|f\|^2 \leq \sum_k |\langle f, \psi_{j,k}\rangle|^2 \leq B\|f\|^2.$$

A *dual multiresolution analysis* has approximation spaces $\widetilde{V}_j$ and *dual scaling functions* $\widetilde{\varphi}_{j,k}$ which are dual to the scaling functions $\{\varphi_{j,k}\}_k$, i.e.

$$\langle \varphi_{j,k}, \widetilde{\varphi}_{j,k'}\rangle = \delta_{k,k'}. \tag{5.1}$$

The dual detail spaces $\widetilde{W}_j$ are spanned by the *dual wavelets* which are biorthogonal to the wavelets,

$$\langle \psi_{j,k}, \widetilde{\psi}_{j',k'}\rangle = \delta_{j,j'}\delta_{k,k'}. \tag{5.2}$$

The approximation and detail spaces and their duals are related by

$$W_j \perp \widetilde{V}_j \quad \text{and} \quad V_j \perp \widetilde{W}_j.$$

The corresponding relations for the scaling functions and the wavelets are

$$\langle \psi_{j,k}, \widetilde{\varphi}_{j,k'}\rangle = 0 \quad \text{and} \quad \langle \varphi_{j,k}, \widetilde{\psi}_{j,k'}\rangle = 0. \tag{5.3}$$

We refer to (5.1)-(5.3) together as the *biorthogonality conditions.*

We will now turn over to the *filters.* Since $\varphi_{j,k} \in V_j \subset V_{j+1}$ the scaling functions must satisfy *refinement relations*

$$\varphi_{j,k} = \sum_l h_{j,k,l}\varphi_{j+1,2k+l}, \tag{5.4}$$

for some coefficients $h_{j,k,l}$. For fixed j we define the level-dependent *lowpass filter* as the operator $H_j : \ell^2 \to \ell^2$ defined by

$$(H_j c)_k = \sum_l h_{j,k,l} c_{2k+l}. \tag{5.5}$$

Note that H_j is not a shift invariant filter. Similarly we must also have coefficients $\widetilde{h}_{j,k,l}$, $g_{j,k,l}$ and $\widetilde{g}_{j,k,l}$ such that

$$\begin{aligned} \widetilde{\varphi}_{j,k} &= \sum_l \widetilde{h}_{j,k,l}\widetilde{\varphi}_{j+1,2k+l}, \\ \psi_{j,k} &= \sum_l g_{j,k,l}\varphi_{j+1,2k+l}, \\ \widetilde{\psi}_{j,k} &= \sum_l \widetilde{g}_{j,k,l}\widetilde{\varphi}_{j+1,2k+l}. \end{aligned} \tag{5.6}$$

Define lowpass and highpass filters $\widetilde{H}_j, G_j$ and $\widetilde{G}_j$ as in (5.5). With a slight abuse of notation we can then write (5.4) and (5.6) as

$$\begin{array}{rclcrcl} \varphi_j & = & H_j \varphi_{j+1} & \quad & \widetilde{\varphi}_j & = & \widetilde{H}_j \widetilde{\varphi}_{j+1} \\ \psi_j & = & G_j \varphi_{j+1} & \quad & \widetilde{\psi}_j & = & \widetilde{G}_j \widetilde{\varphi}_{j+1} \end{array} \tag{5.7}$$

In terms of these filters, the biorthogonality conditions becomes

$$\begin{array}{rclcl} \widetilde{H}_j H_j^* & = & \widetilde{G}_j G_j^* & = & I, \\ \widetilde{G}_j H_j^* & = & \widetilde{H}_j G_j^* & = & 0, \\ H_j^* \widetilde{H}_j & + & G_j^* \widetilde{G}_j & = & I. \end{array} \tag{5.8}$$

5.2 The New Biorthogonal System

We define the new wavelets as before,

$$\widetilde{\psi}^{\mathrm{K}}_{j,k} = \frac{1}{\kappa_{j,k}} K^* \widetilde{\psi} \quad \text{and} \quad \psi^{\mathrm{K}}_{j,k} = \kappa_{j,k} K^{-1} \psi_{j,k}.$$

Let us try to find a biorthogonal multiresolution analysis corresponding to these new wavelets. We do so by glancing at the shift-invariant case. There we had that the new scaling functions $\varphi^{\mathrm{K}}_{j,k}$, $\widetilde{\varphi}^{\mathrm{K}}_{j,k}$ were related to the old one through

$$\overline{m_j}(\omega) \widehat{\widetilde{\varphi}^{\mathrm{K}}_{j,k}}(\omega) = \overline{\widehat{k}}(\omega) \widehat{\widetilde{\varphi}_{j,k}}(\omega) \quad \text{and} \quad \widehat{k}(\omega) \widehat{\varphi^{\mathrm{K}}_{j,k}}(\omega) = m_j(\omega) \widehat{\varphi_{j,k}}(\omega),$$

where

$$m_j(\omega) = \sum_l m_{j,l} e^{-il\omega}.$$

In the time (or spatial) domain this becomes

$$\sum_l m_{j,-l} \widetilde{\varphi}^{\mathrm{K}}_{j,k+l} = K^* \widetilde{\varphi}_{j,k} \quad \text{and} \quad \sum_l m_{j,l} \varphi_{j,k+l} = K \varphi^{\mathrm{K}}_{j,k}.$$

We can think of the m_j's as discrete approximations of the operator at scale 2^{-j}. Inspired by this we try to define new scaling functions through

$$\begin{array}{rcl} \sum_l m_{j,l,k} \widetilde{\varphi}^{\mathrm{K}}_{j,k+l} & = & K^* \widetilde{\varphi}_{j,k}, \\ \sum_l m_{j,k,l} \varphi_{j,k+l} & = & K \varphi^{\mathrm{K}}_{j,k}. \end{array} \tag{5.9}$$

Again we think of the matrices m_j as discrete approximations of K. The main question here is of course how to chose the coefficients $m_{j,k,l}$. Ignoring that for the moment we instead move on to derive the new filters. First we note that we can write (5.9) in operator notation

$$M_j^* \widetilde{\varphi}^{\mathrm{K}}_j = K^* \widetilde{\varphi}_j \quad \text{and} \quad M_j \varphi_j = K \varphi^{\mathrm{K}}_j.$$

with the obvious definition of the operators M_j. Assuming that the M_j's are invertible we have

$$\begin{aligned} K\varphi_j^{\mathrm{K}} &= M_j\varphi_j = M_jH_j\varphi_{j+1} = M_jH_jM_{j+1}^{-1}M_{j+1}\varphi_{j+1} \\ &= M_jH_jM_{j+1}^{-1}K\varphi_{j+1}^{\mathrm{K}} = KM_jH_jM_{j+1}^{-1}\varphi_{j+1}^{\mathrm{K}} \end{aligned}$$

This implies that

$$\varphi_j^{\mathrm{K}} = M_jH_jM_{j+1}^{-1}\varphi_{j+1}^{\mathrm{K}}$$

The same kind of calculations gives us that

$$\begin{aligned} \widetilde{\varphi}_j^{\mathrm{K}} &= M_j^{-*}\widetilde{H}_jM_{j+1}^{*}\widetilde{\varphi}_{j+1}^{\mathrm{K}} \\ \psi_j^{\mathrm{K}} &= \mathcal{K}_jG_jM_{j+1}^{-1}\varphi_j^{\mathrm{K}} \\ \widetilde{\psi}_j^{\mathrm{K}} &= \mathcal{K}_j^{-1}\widetilde{G}_jM_{j+1}^{*}\widetilde{\varphi}_j^{\mathrm{K}} \end{aligned}$$

Here $\mathcal{K}_j$ is the diagonal operator

$$(\mathcal{K}_jd)_k = \kappa_{j,k}d_k$$

We see that we get a new biorthogonal system with filters defined by the expressions

$$(5.10) \qquad \begin{array}{rclrcl} H_j^{\mathrm{K}} &=& M_jH_jM_{j+1}^{-1} & \widetilde{H}_j^{\mathrm{K}} &=& M_j^{-*}\widetilde{H}_jM_{j+1}^{*} \\ G_j^{\mathrm{K}} &=& \mathcal{K}_jG_jM_{j+1}^{-1} & \widetilde{G}_j^{\mathrm{K}} &=& \mathcal{K}_j^{-1}\widetilde{G}_jM_{j+1}^{*} \end{array}$$

It is straightforward to show that these new filters satisfies the biorthogonality conditions (5.8).

At this point a number of questions arises about the new wavelets, scaling functions and filters, such as stability, regularity, decay and vanishing moments. We do not intend to answer this in general but rather consider some interesting examples in more detail.

5.3 Differential Operators

We will consider a particular differential operator here, but we believe that the ideas extend to more general differential operators. The operator we will study is $K = DaD$ where a is a positive and differentiable function with bounded derivative, $|a'(x)| \leq C$. Note that this operator is self-adjoint. We take the original wavelet basis as a classical one, i.e. generated from mother wavelets ψ and $\widetilde{\psi}$ by dilations and translations. First we have to chose the quasi-singular values. The new dual wavelets are

$$\begin{aligned} \widetilde{\psi}_{j,k}^{\mathrm{K}}(x) &= a(x)D^2\widetilde{\psi}_{j,k}(x) + a'(x)D\widetilde{\psi}_{j,k}(x) \\ &= 4^ja(x)(\widetilde{\psi}'')_{j,k}(x) + 2^ja'(x)(\widetilde{\psi}')_{j,k}(x) \end{aligned}$$

The natural assumption on a is that it should have a uniformly continuous derivative. Intuitively, as $j \to \infty$, the variation of a' (and a) on the support of $\widetilde{\psi}_{j,k}$ can then uniformly (over k) be made arbitrarily small and we can approximate a and a' with constants. A proper choice of quasi-singular values would then be $\kappa_{j,k} = max(a(2^{-j}k)4^j, \left|a'(2^{-j}k)\right| 2^j)$. This approximation is only valid for fine-scale wavelets and we cannot count on the new system to be stable as $j \to -\infty$. However, in practice we always have a coarsest scale so this is not really a problem.

Assuming the additional condition that $|a'(x)|/a(x)$ is bounded we can in fact show that the new system is stable as $j \to \infty$.

Let us now consider vanishing moments. Suppose that the original dual wavelets have N vanishing moments,

$$\int x^n \widetilde{\psi}_{j,k}(x)dx = 0, \quad \text{for} \quad 0 \le n < N.$$

For the new wavelets we will have

$$\begin{aligned}\int (x - 2^{-j}k)^n \widetilde{\psi}^{\mathrm{K}}_{j,k}(x)dx &= \frac{1}{\kappa_{j,k}} \int (x - 2^{-j}k)^n Da(x) D\widetilde{\psi}_{j,k}(x)dx \\ &= \frac{2^j n}{\kappa_{j,k}} \int (x - 2^{-j}k)^{n-1} a(x) (\widetilde{\psi}')_{j,k}(x)dx\end{aligned}$$

By the mean value theorem,

$$a(x) = a(2^{-j}k) + (x - 2^{-j}k)a'(\xi_{j,k}),$$

and hence

$$\begin{aligned}\int (x - 2^{-j}k)^n \widetilde{\psi}^{\mathrm{K}}_{j,k}(x)dx &= \int (x - 2^{-j}k)^{n-1} (\widetilde{\psi}')_{j,k}(x)dx \\ &\quad + \frac{2^j n}{\kappa_{j,k}} \int (x - 2^{-j}k)^{n-1} a'(\xi_{j,k}) (\widetilde{\psi}')_{j,k}(x)dx \\ &= \frac{2^j n}{\kappa_{j,k}} \int (x - 2^{-j}k)^{n-1} a'(\xi_{j,k}) (\widetilde{\psi}')_{j,k}(x)dx,\end{aligned}$$

for $0 < n < N+2$, since the first term is zero due to the vanishing moments of the original system. For $n = 0$ the integral is exactly zero. We now assume for simplicity that the original wavelets are supported on the intervals $I_{j,k} = [2^{-j}(k - N), 2^{-j}(k + N)]$. By Cauchy-Schwartz inequality we then have

$$\begin{aligned}\left|\int (x - 2^{-j}k)^n \widetilde{\psi}^{\mathrm{K}}_{j,k}(x)dx\right|^2 &\le \left(\frac{2^j n}{\kappa_{j,k}}\right)^2 (max\,|a'(x)|)^2 \int_{I_{j,k}} x^{2n-2}dx\, \|\widetilde{\psi}'\|^2 \\ &= \left(\frac{2^j n}{\kappa_{j,k}}\right)^2 \frac{N^{2n-1}}{2n-1} (max\,|a'(x)|)^2 \|\widetilde{\psi}'\|^2,\end{aligned}$$

and therefore

$$(5.11)\qquad \left|\int (x-2^{-j}k)^n \widetilde{\psi}^{\mathrm{K}}_{j,k}(x)dx\right| \le \frac{2^j n}{\kappa_{j,k}} \frac{N^{n-1/2}}{\sqrt{2n-1}} max\,|a'(x)|\ \|\widetilde{\psi}'\|.$$

We see that even if we cannot expect vanishing moments of the new dual wavelets, at the fine scales they will be "almost orthogonal" to Taylor polynomials around their central points, since $\kappa_{j,k} \sim 4^j$ as $j \to \infty$.

Let us have a look at the M_j's. We choose them as two-sided difference approximations

$$(5.12)\qquad m_{j,k,l} = a(2^{-j}(k+1))\delta_{l,k+1} - 2a(2^{-j}k)\delta_{l,k} + a(2^{-j}(k-1))\delta_{l,k-1}$$

Of course, there are a lot of questions to investigate; what will the new filters look like, will the cascade algorithm converge, and if so, what will the new scaling functions look like? If we want to use this method to solve differential equations, we also need to incorporate boundary conditions. These questions will be a matter of further study.

5.4 Calderón-Zygmund Operators

Let us consider Calderón-Zygmund operators, i.e. integral operators

$$Kf(x) = \int k(x,y)f(y)dy$$

where the kernel k satisfies the estimates

$$(5.13)\qquad \begin{aligned} |k(x,y)| &\le \frac{C}{|x-y|} \\ \left|\partial_x^N k(x,y)\right| + \left|\partial_y^N k(x,y)\right| &\le \frac{C}{|x-y|^{N+1}} \end{aligned}$$

We assume that the original dual wavelets have N vanishing moments. This gives us some decay properties of the new wavelets, for if we expand the kernel in a Taylor series in its first argument,

$$\begin{aligned} k(x,y) &= k(2^{-j}k,y) + (x-2^{-j}k)\partial_1 k(2^{-j}k,y) + \frac{(x-2^{-j}k)^2}{2}\partial_1^2 k(2^{-j}k,y) \\ &+ \cdots + \frac{(x-2^{-j}k)^{N-1}}{(N-1)!}\partial_1^{N-1}k(2^{-j}k,y) + \frac{(x-2^{-j}k)^N}{N!}\partial_1^N k(\xi_{j,k},y), \end{aligned}$$

where $\left|\xi_{j,k} - 2^{-j}k\right| \le \left|x-2^{-j}k\right|$, we have

$$\begin{aligned} \widetilde{\psi}^{\mathrm{K}}_{j,k}(x) &= \int k(y,x)\widetilde{\psi}_{j,k}(y)dy \\ &= \frac{1}{N!}\int y^N \partial_1^N k(\xi_{j,k},x)\widetilde{\psi}_{j,k}(y)dy. \end{aligned}$$

We again assume that the original wavelets are supported on the intervals $I_{j,k} = [2^{-j}(k-N), 2^{-j}(k+N)]$. For x outside this interval we have

$$\left|\partial_1^N k(\xi_{j,k}, x)\right| \leq \frac{C}{|\xi_{j,k} - x|} \leq \frac{C}{|2^{-j}N - x|^{N+1}}$$

and therefore

$$\left|\widetilde{\psi}^{\mathrm{K}}_{j,k}(x)\right| \leq \frac{N^{N+1/2}}{N!\sqrt{2N+1}} \frac{C}{|2^{-j}N - x|^{N+1}}$$

We now turn over to the choice of M_j. We recall the relation

$$\sum_l m_{j,l,k} \widetilde{\varphi}^{\mathrm{K}}_{j,k+l} = K^* \widetilde{\varphi}_{j,k}, \tag{5.14}$$

between the original and new scaling functions. In the shift-invariant case, we can write this in the Fourier domain

$$\overline{m_j}(\omega)\widehat{\widetilde{\varphi}^{\mathrm{K}}_{j,k}}(\omega) = \overline{\widehat{k}}(\omega)\widehat{\widetilde{\varphi}_{j,k}}(\omega) \quad \text{and} \quad \widehat{k}(\omega)\widehat{\varphi^{\mathrm{K}}_{j,k}}(\omega) = m_j(\omega)\widehat{\varphi_{j,k}}(\omega),$$

where

$$m_j(\omega) = \sum_l m_{j,l} e^{-il\omega}.$$

Here our choice of m_j was guided by letting ω approach zero. We therefore take integrals in (5.14)

$$\begin{aligned}
\sum m_{j,l,k} \int \widetilde{\varphi}^{\mathrm{K}}_{j,k+l} &= \int K^* \widetilde{\varphi}_{j,k} \\
&= \int \left(\int k(y,x) \widetilde{\varphi}^{\mathrm{K}}_{j,k+l}(y) dy \right) dx \\
&= \int \widetilde{K}(y) \widetilde{\varphi}^{\mathrm{K}}_{j,k+l}(y) dy \\
&\approx \widetilde{K}(2^{-j}k) 2^{-j/2},
\end{aligned}$$

where

$$\widetilde{K}(y) = \int k(y,x) dx.$$

The above approximation is valid for large j. As $j \to \infty$, we would like to have

$$\int \widetilde{\varphi}^{\mathrm{K}}_{j,k} = 2^{-j/2}.$$

This gives us

$$\sum m_{j,l,k} \approx \widetilde{K}(2^{-j}k) = \int k(2^{-j}k, y)dy \approx \sum_l k(2^{-j}k, 2^{-j}(k+l+1/2)).$$

We therefore make the choice

$$m_{j,k,l} = k(2^{-j}k, 2^{-j}(k+l+1/2)).$$

which indeed seems to be a reasonable approximation of K at scale 2^{-j}. Again, there are lot of details here to investigate, vanishing moments of the new wavelets for instance.

Chapter 6

Wavelet Transforms for Complex Radar Signals

In this chapter we will use the techniques developed in the previous chapters to adapt the wavelet transform to complex radar signals. For complex-valued signals, positive and negative frequencies do no longer contain the same information. Moreover, for the complex radar signals we will study, positive and negative frequencies have a significant physical interpretation, and it is therefore of interest to separate them when analyzing these signals. This problem was given to us by Claes Hagström of the radar section at Ericsson Microwave Systems.

6.1 Complex Radar Signals

Below we will give a simplified description of complex radar signals and how they arise. The purpose is to explain what positive and negative frequencies means physically, and we do not intend to give the complete picture of radar technology here. The signal transmitted from a pulse radar is a "rectangular pulse train" with a certain pulse width and pulse repetition frequency, modulated by a carrier cosine wave with some carrier frequency ω_c. The Fourier transform of this modulated pulse train is sketched in figure 6.1. (Note that the "bumps" around $\pm\omega_c$ do not look exactly like that, but as an illustration, this figure will do). If this signal hits a target (aeroplane, helicopter etc), it is partly reflected and will be received by the radar antenna. If the target is moving in a direction towards the radar, the spectrum will be shifted to higher frequencies due to the Doppler effect. If the target is moving away from the radar, it will be shifted to lower frequencies. If the target has a complex structure with moving components there may be contributions in both ways and the spectrum will be fairly unsymmetric around ω_c. (There is also a copy centred around $-\omega_c$, but it contains no further information and will be neglected). An example of how this might

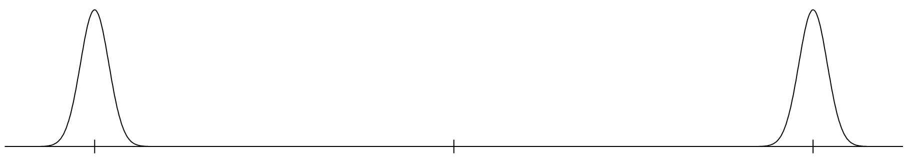


Figure 6.1: Fourier transform of transmitted signal.

look like is shown in figure 6.2. The received signal is shifted in frequency

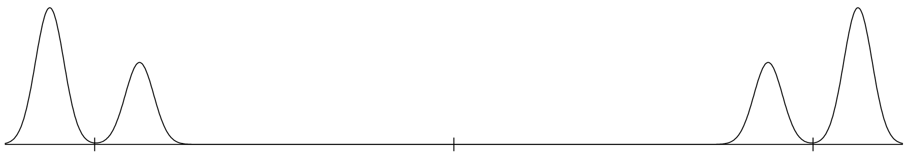


Figure 6.2: Fourier transform of received signal.

by ω_c, i.e. it is multiplied by $e^{-i\omega_c t}$, and the spectrum is therefore centred around zero, see figure 6.3 (the part around $-2\omega_c$ is removed at this stage by a low-pass filter). Notice that the new spectrum is in general unsymmetric and that the corresponding modulated signal is complex. Moreover, negative frequencies corresponds to targets moving away from the radar and vice versa. Therefore it is of importance being able to distinguish between positive and negative frequencies when analyzing this signal. This is of course automatically achieved with traditional Fourier methods. A disadvantage with the Fourier transform is that it does not give any spatial resolution and this is why we are interested in wavelet transforms because they give us resolution both in space and frequency.

6.2 Complex Wavelet Transforms

6.2.1 The complex discrete wavelet transform

Let $\{\psi_{j,k}\}$ be an orthonormal basis of (real) wavelets, where

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

and ψ is the mother wavelet. Let $f(t)$ be the complex-valued signal described above and $d_{j,k} = \langle f, \psi_{j,k} \rangle$ its wavelet coefficients. We can interpret the

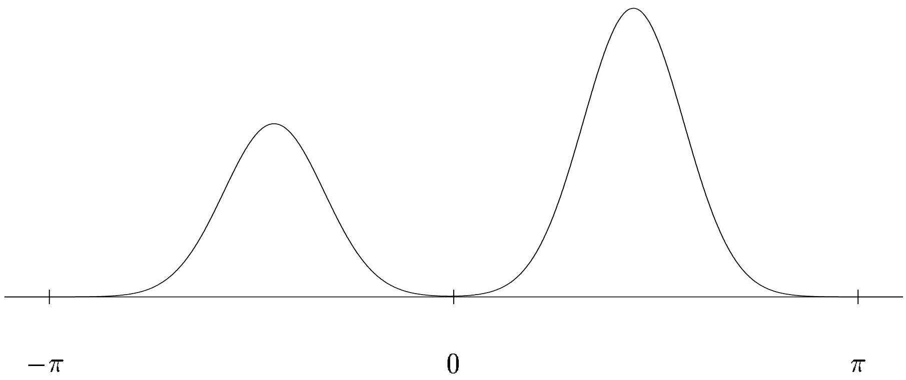


Figure 6.3: Fourier transform of complex radar signal.

$|d_{j,k}|^2$'s as an energy distribution, since by orthogonality

$$\int |f(t)|^2\, dt = \sum_{j,k} |d_{j,k}|^2\,.$$

The wavelet $\psi_{j,k}$ is essentially supported in the frequency intervals $2^{j-1} \le |\xi| \le 2^j$ and in the time interval $[2^{-j}k, 2^{-j}(k+1)]$. Hence $|d_{j,k}|^2$ is a measure of the energy in these frequency bands and in this time interval. But we have no information about how the energy is separated into positive and negative frequencies, since each coefficient $c_{j,k}$ mixes contributions from the frequency intervals $[-2^j, -2^{j-1}]$ and $[2^{j-1}, 2^j]$. What one would like to have are wavelet functions supported in positive (or negative) frequencies only. The natural way to achieve this is to separate the mother wavelet into two parts:

$$\widehat{\psi^+}(\omega) = \theta(\omega)\widehat{\psi}(\omega) \quad \text{and} \quad \widehat{\psi^-}(\omega) = \theta(-\omega)\widehat{\psi}(\omega)$$

where θ is Heaviside's step function. From these "mother wavelets" we define new wavelet basis functions

$$\psi^+_{j,k}(t) = 2^{j/2}\psi^+(2^jt - k) \quad \text{and} \quad \psi^-_{j,k}(t) = 2^{j/2}\psi^-(2^jt - k).$$

It is immediate that $\{\psi^\nu_{j,k}\}$, $\nu = +$ or $-$, is an orthonormal basis in $L^2(\mathbb{R})$. The analysis of f in this basis, i.e. the computation of the coefficients $\langle f, \psi^\nu_{j,k}\rangle$ will be referred to as the *complex wavelet transform*. It is equivalent to first splitting f into positive and negative frequencies respectively and then analyzing each part in the basis $\{\psi_{j,k}\}$. The positive and negative frequency parts of f can be obtain by multiplication with Heaviside in the frequency domain, or in the time domain as

$$f \pm iHf.$$

Here, H stands for the Hilbert transform. Now we want to compute the wavelet transform, which we denote by $\mathcal{W}$. From chapter 2 we know that there exist another wavelet transform $\mathcal{W}^{\mathrm{H}}$ such that $\mathcal{W}H = \mathcal{W}^{\mathrm{H}}$ and we therefore have

$$\mathcal{W}(f \pm iHf) = \mathcal{W}f \pm i\mathcal{W}Hf = \mathcal{W}f \pm i\mathcal{W}^{\mathrm{H}}f.$$

The new wavelet basis have mother wavelets defined by

$$\widetilde{\psi}^{\mathrm{H}} = H^*\widetilde{\psi} \quad \text{and} \quad \psi^{\mathrm{H}} = H^{-1}\psi.$$

The new filters have filter functions

$$\begin{aligned} H^{\mathrm{H}}(\omega) &= \frac{e^{i\omega}+1}{|e^{i\omega}+1|}H(\omega) & \widetilde{H}^{\mathrm{H}}(\omega) &= \frac{|e^{i\omega}+1|}{e^{-i\omega}+1}\widetilde{H}(\omega) \\ G^{\mathrm{H}}(\omega) &= -\frac{|e^{i\omega}-1|}{e^{i\omega}-1}H(\omega) & \widetilde{G}^{\mathrm{H}}(\omega) &= -\frac{e^{-i\omega}+1}{|e^{i\omega}-1|}\widetilde{G}(\omega) \end{aligned} \tag{6.1}$$

We now have the following algorithm for the complex wavelet transform of f:

1. Compute the wavelet coefficients $\langle f, \widetilde{\psi}_{j,k}\rangle$ and the coarse-scale coefficients $\langle f, \widetilde{\varphi}_{j_0,k}\rangle$ with the filters $\widetilde{H}$ and $\widetilde{G}$.

2. Compute the wavelet coefficients $\langle f, \widetilde{\psi}^{\mathrm{H}}_{j,k}\rangle$ and the coarse-scale coefficients $\langle f, \widetilde{\varphi}^{\mathrm{H}}_{j_0,k}\rangle$ with the filters $\widetilde{H}^{\mathrm{H}}$ and $\widetilde{G}^{\mathrm{H}}$.

3. Get the wavelet coefficients and fine-scale coefficients of Hf in the original system through the formulas $\langle Hf, \widetilde{\psi}_{j,k}\rangle = \langle f, \widetilde{\psi}^{\mathrm{H}}_{j,k}\rangle$ and $\langle Hf, \widetilde{\varphi}_{j_0,k}\rangle = \sum_l m_l \langle f, \widetilde{\varphi}^{\mathrm{H}}_{j_0,k-l}\rangle$ where $m_l = \frac{1}{\pi(l+1/2)}$ (see section 2.4).

4. Now it is simple to get the complex wavelet coefficients,
$d^+_{j,k} = \langle f, \widetilde{\psi}^+_{j,k}\rangle = \langle f, \widetilde{\psi}_{j,k}\rangle + i\langle f, \widetilde{\psi}^{\mathrm{H}}_{j,k}\rangle$ and
$d^-_{j,k} = \langle f, \widetilde{\psi}^-_{j,k}\rangle = \langle f, \widetilde{\psi}_{j,k}\rangle - i\langle f, \widetilde{\psi}^{\mathrm{H}}_{j,k}\rangle$

There are some technical difficulties associated with this method. First, the new filters will not be of finite length since the filter functions in (6.1) are not smooth at $\omega = 0$. We then have to truncate the filters. We also have to truncate the convolution of the coarse-scale coefficients with the discrete version m_l of the Hilbert transform. We will discuss this more when we consider some numerical examples below.

6.2.2 The Complex Continuous Wavelet Transform.

In some situations when a detailed analysis of a signal is desired, one needs a more fine-tuned instrument than the usual discrete wavelet transform. It can be provided by the *continuous wavelet transform.* This just consists of analyzing the signal at *all* scales a and positions b, by looking at the correlation of the signal with a family of wavelets, at all scales and positions. This family is obtained from a single mother wavelet ψ, just as in the discrete wavelet transform, by dilations and translations,

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\,\psi\left(\frac{x-b}{a}\right) \quad \text{where} \quad a > 0,\ b \in \mathbb{R}.$$

The continuous wavelet transform of a signal f is thus defined by

$$\mathcal{W}f(a,b) = \int f(t)\psi_{a,b}(t)dt. \tag{6.2}$$

There is also an inverse continuous transform, or reconstruction formula, to compute the signal from its continuous wavelet transform:

$$f(t) = \frac{1}{C_\psi}\int_0^\infty \int_{-\infty}^\infty \mathcal{W}f(a,b)\psi_{a,b}(t)\frac{dadb}{a^2}. \tag{6.3}$$

where

$$C_\psi = 2\pi \int_0^\infty |\widehat{\psi}(\omega)|^2 \frac{d\omega}{\omega},$$

which of course needs to be finite. This reconstruction formula is not of major interest in applications, where we use the continuous wavelet transform merely to analyze signals. In practice, we only compute $\mathcal{W}f(a,b)$ for a finite set of values of a and b, but contrary to the discrete wavelet transform, where $a = 2^{-j}$ and $b = 2^{-j}k$, we can chose this discrete set of values more freely, and adapt it to the signal at hand. This enables us to "zoom in" on transients and edges. We also have a lot more freedom in choosing the mother wavelet, since it no longer has to fit into a multiresolution structure.

We still need fast algorithms to compute the $\mathcal{W}f(a,b)$'s. Regarding (6.2) as a convolution, we get

$$\begin{aligned}\mathcal{W}f(a,b) &= \int f(t)\psi_{a,b}(t)dt \\ &= \frac{1}{\sqrt{a}}\int f(t)\psi\left(\frac{t-b}{a}\right)dt \\ &= \frac{1}{\sqrt{a}}(f * \psi_a)(b) \\ &= \sqrt{a}\,\mathcal{F}^{-1}\Big(\widehat{f}(\omega)\widehat{\psi}(a\omega)\Big)(b),\end{aligned}$$

where we have used the notation

$$\psi_a(t) = \psi\left(\frac{t}{a}\right).$$

To implement the continuous wavelet transform we have to do the following:

1. Store the Fourier transform of the mother wavelet $\widehat{\psi}(\omega)$ with high precision.

2. For a given signal f, do an FFT to get $\widehat{f}(\omega)$.

3. For each scale a, multiply $\widehat{f}(\omega)$ with $\sqrt{a}\,\widehat{\psi}(a\omega)$ and do an inverse FFT

For a discrete signal of length N, this requires $(\#a + 1)N \log_2 N + \#aN$ operations, where $\#a$ is the number of scales we are looking at.

Since we are working in the frequency domain, it is easy to deal with complex-valued signals and to separate positive and negative frequencies. In step 2 above we separate $\widehat{f}(\omega)$ into two parts,

$$\widehat{f^+}(\omega) = \theta(\omega)\widehat{f}(\omega) \quad \text{and} \quad \widehat{f^-}(\omega) = \theta(-\omega)\widehat{f}(\omega),$$

and we then do step 3 on each of them separately. This gives us the *complex continuous wavelet transform.*

6.3 Numerical Examples.

It is about time to study some numerical examples. The signals we will work with were provided by the radar section at Ericsson Microwave Systems. The first signal contains radar echoes from the traffic on a road. The absolute value of this signal is shown in figure 6.4, where also the power spectrum is plotted. From this spectrum we see that is a large component around zero frequency, probably due to reflections from buildings and from the ground. There are also some smaller "bumps" at positive and negative frequencies. This is reflections from cars moving at different speeds and directions. In figure 6.5 we plot a Mallat-style multiresolution display of the complex wavelet decomposition. Let us explain exactly what we see there. For the positive and negative frequency part separately, we compute projections onto the W_j-spaces from the complex wavelet coefficients. We then plot the absolute value of these projections, where the coarsest scale is plotted on top. At the coarse scales, we see the echoes from the surroundings, buildings, parked cars etc. At the finer scales, we can clearly see cars moving away from the radar at $t = 0.2$ and $t = 0.9$. There is also a car moving towards the radar at $t = 0.2$ and several between $t = 0.5$ and $t = 0.8$.

The second signal is the radar echo from a helicopter. The absolute value is shown in figure 6.6, together with the power spectrum. In figure 6.7, we plot the multiresolution display as described above. Here, one can clearly see the echo from the helicopter body at low negative frequencies, so the helicopter is slowly moving away from the radar. At the fine scales, we see reflections from the four rotating airfoils, two in each direction.

We used Coiflet-3 as our original (orthogonal) wavelet basis, since they reduce the effect of entering sample values $f(2^{j_0}k)$ instead of the coarse-scale coefficients $\langle f, \widetilde{\varphi}_{j_0,k} \rangle$. We choosed to truncate the new filters after 31 coefficients.

The MATLAB routines we used to compute the complex wavelet transform and to generate the plots are available at
`ftp://ftp.math.chalmers.se/pub/users/ekstedt/Complex.tar.gz`
They work together with WaveLab v .700.

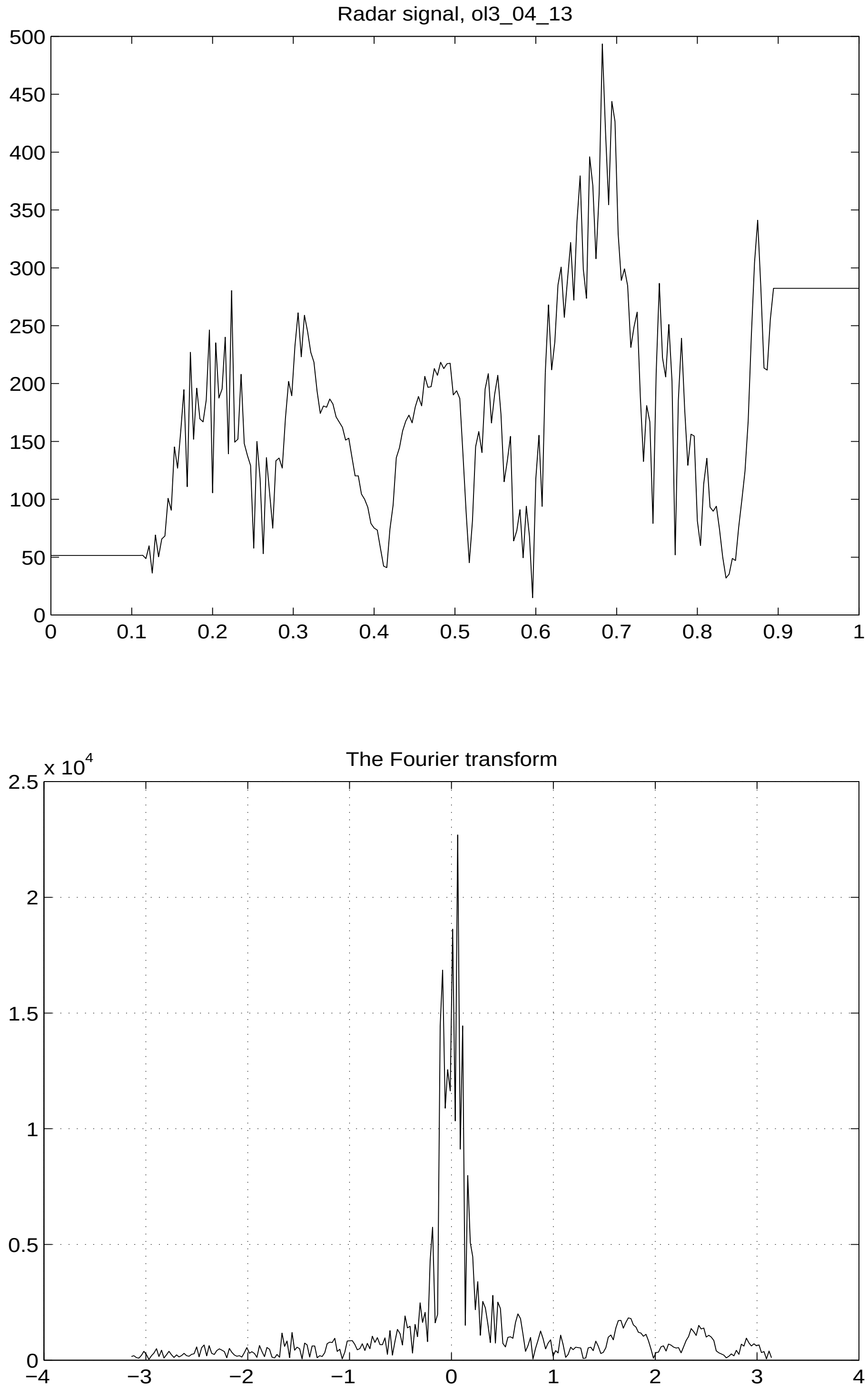
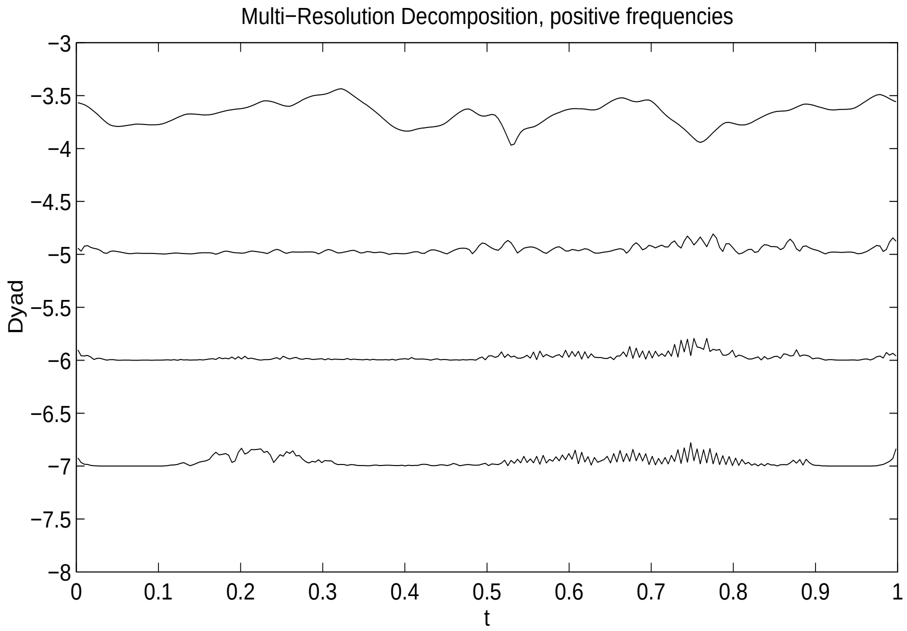


Figure 6.4: The radar signal (on top), and its Fourier transform.



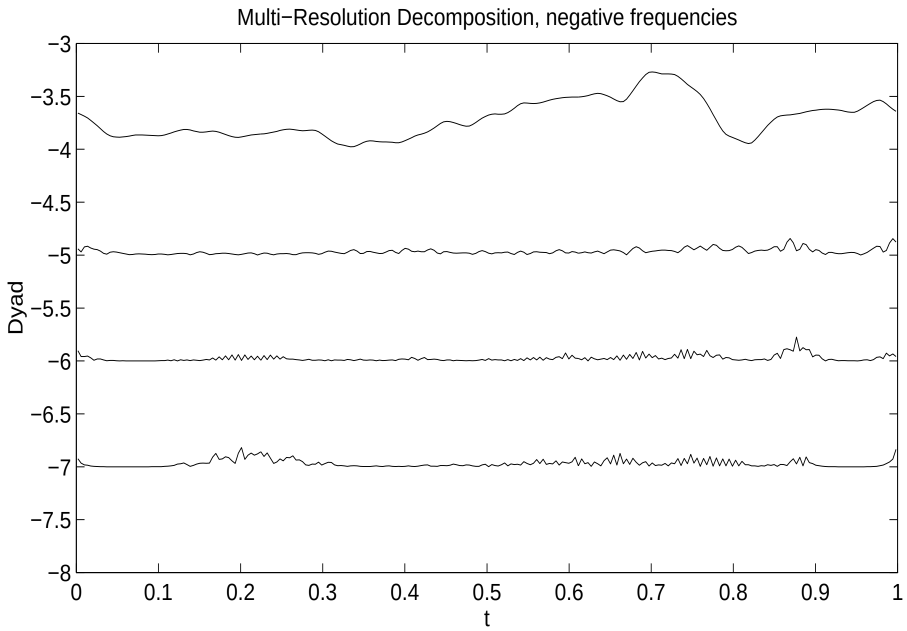


Figure 6.5: A multiresolution plot of its complex wavelet transform.

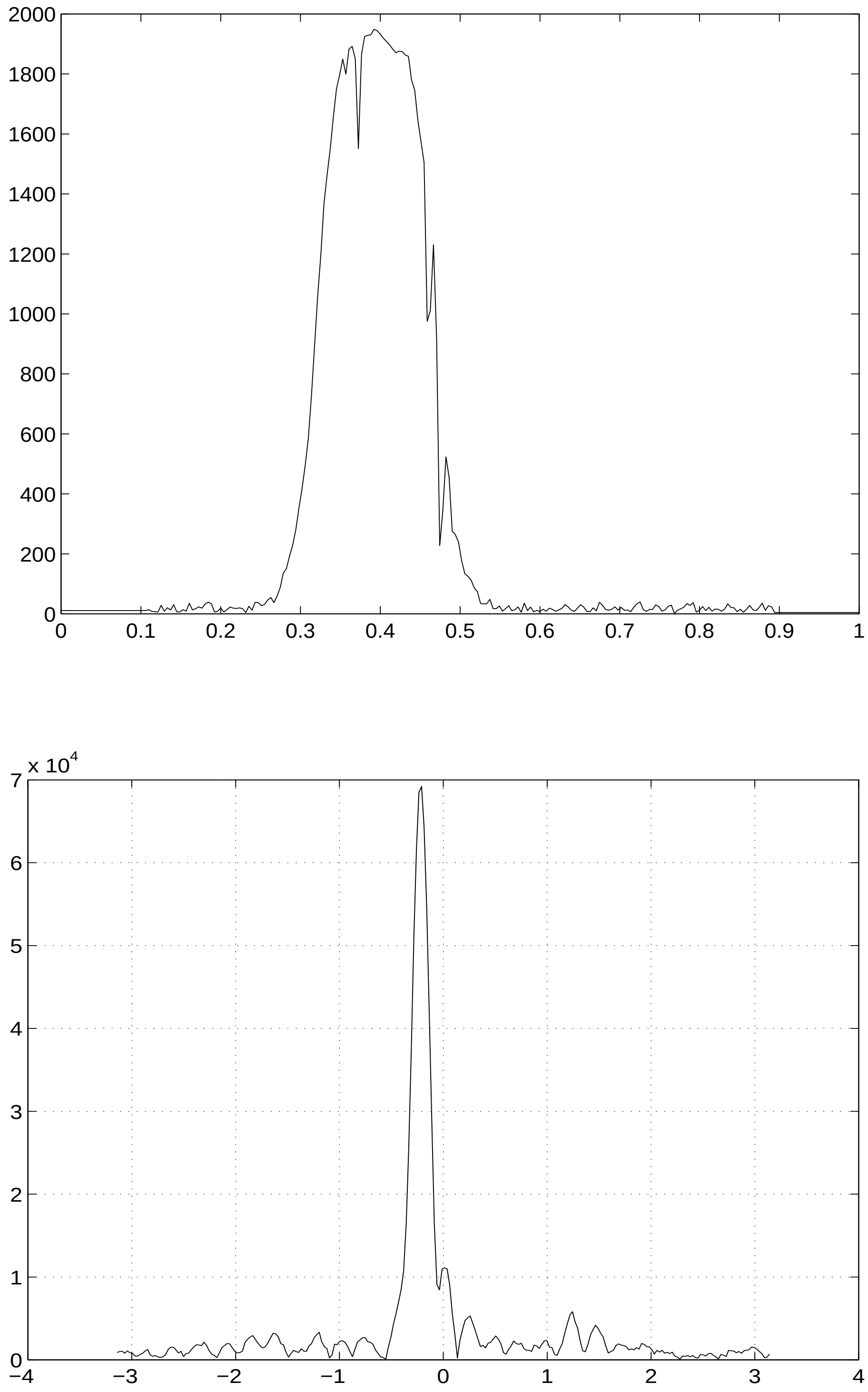


Figure 6.6: The radar signal (on top), and its Fourier transform.

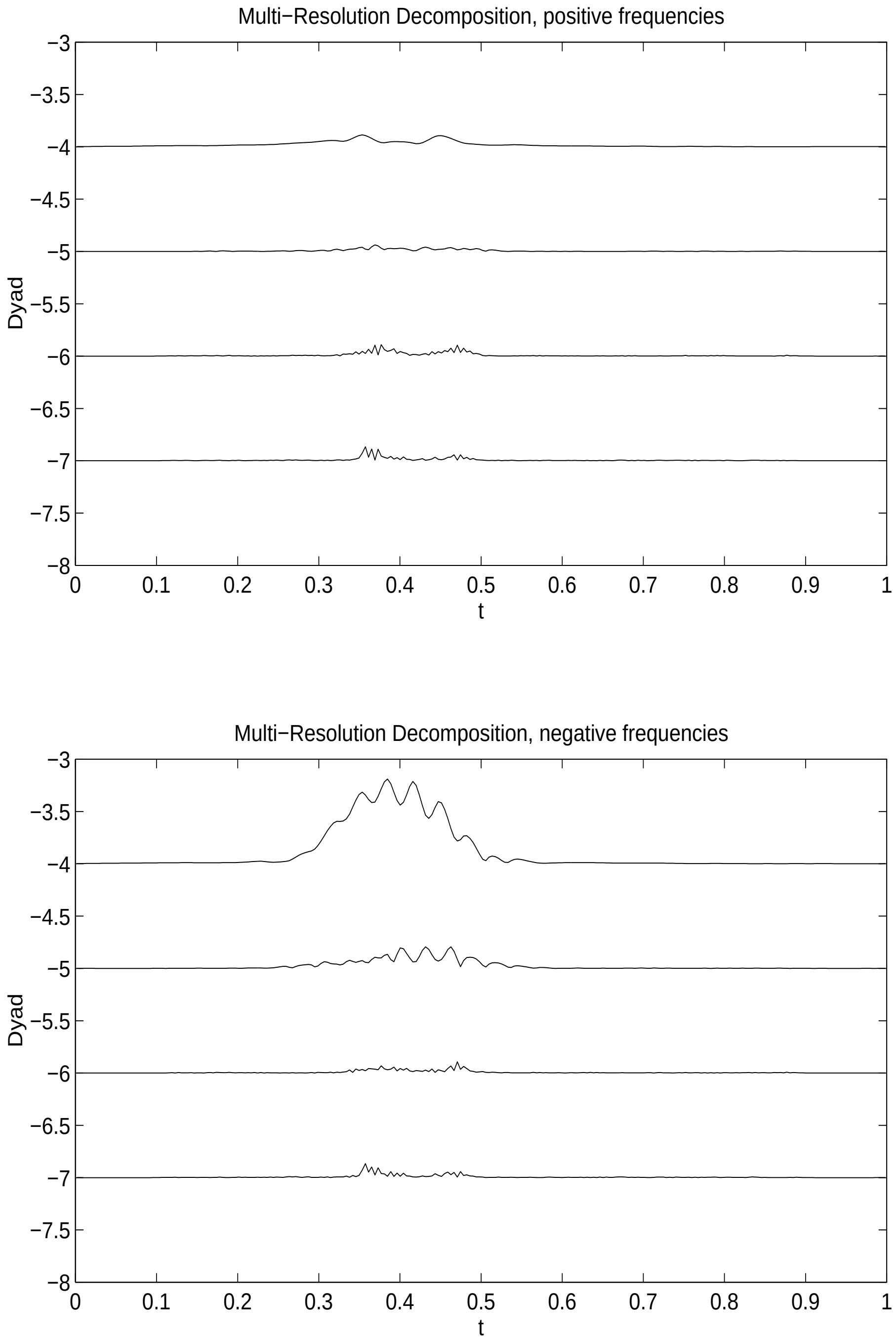


Figure 6.7: A multiresolution plot of its complex wavelet transform.

Bibliography

[1] G. Beylkin, R. Coifman, and V. Rokhlin., *Fast wavelet transforms and numerical algorithms I*, Comm. Pure and Appl. Math. **44** (1991), 141–183.

[2] A. Cohen and I. Daubechies, *Non-separable bidimensional wavelet bases*, Rev. Mat. Iberoamer. **9** (1993), 51–138.

[3] A. Cohen and J-M Schlenker, *Compactly supported bidimensional wavelet bases with hexagonal symmetry*, Constr. Approx. **9** (1993), 209–236.

[4] I. Daubechies, *Ten lectures on wavelets*, SIAM, 1992.

[5] I. Daubechies, *Two recent results on wavelets: wavelet bases for the interval, and biorthogonal wavelets diagonalizing the derivative operator*, Recent Advances in Wavelet Analysis (L. Schumaker and G. Webb, eds.), Wavelet analysis and its applications, vol. 3, Academic Press, 1994, pp. 237–257.

[6] D. Donoho, *Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition*, Applied and Computational Harmonic Analysis **2** (1995), 101–126.

[7] F. Ekstedt and M. Lindberg, *Diagonalization of homogeneous linear operators in biorthogonal wavelet bases*, Preprint,Chalmers Univ. of Techn., Göteborg, Sweden, 1997.

[8] G. Golub and C. van Loan, *Matrix computations*, 2 ed., The John Hopkins University Press, 1989.

[9] B. Jawerth and Y. Mao, *Compactly supported higher dimensional wavelet bases with invariance of certain group actions*, Preprint,Univ. of South Carolina, 1997.

[10] B. Jawerth and W. Sweldens, *Wavelet multiresolution analyses adapted for the fast solution of boundary value ordinary differential equations*, Proceedings of the sixth Copper Mountain Multigrid Conference, April 1994.

[11] E. Kolaczyk, *Wavelet methods for the inversion of certain homogenous linear operators in the presence of noisy data*, Ph.D. thesis, Dept. of statistics, Stanford University, October 1994.

[12] Y. Meyer, *Wavelets and operators*, Cambridge Univ. Press, 1993, English translation from a french text.

[13] G. Strang and T. Nguyen, *Wavelets and filter banks*, Wellesley-Cambridge Press, 1996.

[14] R. Strichartz, *Wavelets and self-affine tilings*, Constr. Approx. **9** (1993), 327–346.

[15] W. Sweldens, *The lifting scheme: A construction of second generation wavelets*, To appear in SIAM Journal on Mathematical Analysis, 1995.