Mathematical signal analysis: wavelets, Wigner distribution and a seismic application
Oonincx, P.J.

Link to publication

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 3

The Wavelet Transform

In this chapter we study a transform that analyses signals in time/space and scale, the wavelet transform. In the first section we introduce the continuous wavelet transform (CWT) on $L^2(\mathbb{R}^n)$, particularly on $L^2(\mathbb{R})$. Definitions and properties of this transform are given. Furthermore, we discuss a group theoretical approach for the CWT. In the second section we consider the discrete wavelet transform (DWT). After an introduction by means of a sampled CWT, we consider its relation with the concepts of Multiresolution Analysis (MRA) and filter bank theory. A fast algorithm to compute the DWT and its use on discrete-time functions in $l^2(\mathbb{Z})$ are discussed.

This chapter is mainly based on existing literature on the Wavelet transform [23, 45, 51, 55, 59, 62] and its relation to filter banks by means of an MRA [16, 43, 60, 94]. Therefore, we will refer several times to the existing literature throughout this chapter. However, we also add some new ideas.

3.1 The Continuous Wavelet Transform

The CWT of $f \in L^2(\mathbb{R})$ is a linear operator defined by

$$W_\psi[f](a,b) = \frac{1}{\sqrt{a}} \int_{\mathbb{R}} f(x) \psi \left( \frac{x-b}{a} \right) dx,$$

(3.1)

for some $\psi \in L^2(\mathbb{R})$ and $a \in \mathbb{R}^+$ and $b \in \mathbb{R}$. By introducing the dilation operator $D_a$ on $L^2(\mathbb{R})$ by

$$D_a[f](x) = \frac{1}{\sqrt{|a|}} f \left( \frac{x}{a} \right),$$

(3.2)
for some $a \in \mathbb{R}$ we can write (3.1) also as
\[ \mathcal{W}_\psi[f](a, b) = (f, T_b D_a \psi), \] (3.3)
with $T_b$ the shift operator as introduced in (2.12). Using this notation we prove that $\mathcal{W}_\psi[f]$ is continuous on $\mathcal{H}$ for all $f \in L^2(\mathbb{R})$, with $\mathcal{H} = \mathbb{R}^+ \times \mathbb{R}$.

**Lemma 3.1.1** $\mathcal{W}_\psi[f]$ is continuous on $\mathcal{H}$ for all $f \in L^2(\mathbb{R})$.

**Proof**

Using the notation $\psi_{a,b} = T_b D_a \psi$, with $\psi \in L^2(\mathbb{R})$, we prove this lemma in two parts. First we give a proof for continuous functions $\psi$ with compact support and next for arbitrary $\psi \in L^2(\mathbb{R})$. Assume $\psi \in L^2_{\text{comp}}(\mathbb{R})$, then $\psi$ is uniformly continuous. So
\[ \|\psi_{a,b} - \psi_{a',b'}\|_2 \to 0, \quad (a, b) \to (a', b'). \]

Now, choose $\epsilon > 0$ and take $g \in L^2(\mathbb{R})$ such that $\|g - \psi\| < \epsilon/3$, with $\psi$ continuous and compactly supported. Then
\[ \|g_{a,b} - g_{a',b'}\|_2 = \|g_{a,b} - \psi_{a,b} + \psi_{a,b} - \psi_{a',b'} + \psi_{a',b'} - g_{a',b'}\|_2 \]
\[ \leq \|g_{a,b} - \psi_{a,b}\|_2 + \|\psi_{a,b} - \psi_{a',b'}\|_2 + \|g_{a',b'} - \psi_{a',b'}\|_2 \]
\[ \leq 2\|g - \psi\|_2 + \|\psi_{a,b} - \psi_{a',b'}\|_2. \]

To establish the proof, we take $(a, b) - (a', b')$ small such that $\|\psi_{a,b} - \psi_{a',b'}\|_2 < \epsilon/3$. \qed

Furthermore, from (3.3) we obtain by Schwarz's inequality
\[ |\mathcal{W}_\psi[f](a, b)| \leq \|f\|_2 \cdot \|\psi\|_2 \quad \forall a \in \mathbb{R}^+ \forall b \in \mathbb{R}, \]
which yields $\mathcal{W}_\psi f \in L^\infty(\mathcal{H})$, for all $f \in L^2(\mathbb{R})$. Later on we shall show that under certain conditions on $\psi$, $\mathcal{W}_\psi : L^2(\mathbb{R}) \to L^2(\mathcal{H}, a^{-2} dbda)$ is an isometry up to a constant that depends on $\psi$. Here $L^2(\mathcal{H}, a^{-2} dbda)$ denotes the space of all Euclidean square integrable functions on $L^2(\mathcal{H})$ with respect to the measure $a^{-2} dbda$. Besides, by Parseval's theorem we can write (3.3) in terms of the Fourier transforms of $f$ and $\psi$
\[ \mathcal{W}_\psi[f](a, b) = (\mathcal{F}f, \mathcal{F}T_b D_a \psi) \]
\[ = \sqrt{a} \int_\mathbb{R} \hat{f}(\omega) \overline{\hat{\psi}(a\omega)} e^{ib\omega} d\omega. \] (3.4)

A third way to write (3.1) is given by means of a convolution product. Take $\tilde{\psi}(x) = \psi(-x)$, then
\[ \mathcal{W}_\psi[f](a, b) = (f * D_a \tilde{\psi})(b). \] (3.5)
Figure 3.1: Two admissible wavelets: a) the Haar wavelet, b) the Mexican hat.

The CWT can be generalized into a transform on $L^2(\mathbb{R}^n)$ by replacing the scalar $a$ by a non-singular matrix $A \in \mathbb{R}^{n \times n}$ and the scalar $b$ by a vector $b \in \mathbb{R}^n$, i.e.,

$$W_\psi[s](A,b) = \frac{1}{\sqrt{\det(A)}} \int_{\mathbb{R}^n} s(x)\hat{\psi}(A^{-1}(x-b)) \, dx.$$  

In the literature mostly $A = aI$ is taken, see e.g. [23, 59]. For this choice, properties for the multi-dimensional CWT can be proved in a rather straightforward way using properties of the one-dimensional CWT. Murenzi followed a different approach in [66]. He introduced a multi-dimensional CWT based on a dilation operator that involves a non-singular matrix $A \in \mathbb{R}^{n \times n}$, translations in $\mathbb{R}^n$ and a rotation operator. Also properties of this CWT resemble the properties we will deduce for the one dimensional CWT in the sequel of this chapter.

**Definition 3.1.2** A function $\psi \in L^2(\mathbb{R})$ which satisfies the admissibility condition

$$0 < C_\psi = 2\pi \int_{\mathbb{R}^+} \frac{|\hat{\psi}(a\omega)|^2}{a} \, da < \infty,$$  

(3.6)

for almost all $\omega \in \mathbb{R}$ is called an (admissible) wavelet.

Note that all $\psi \in L^2(\mathbb{R})$ are admissible wavelets if $\psi \neq 0$, $\hat{\psi}$ differentiable in 0 and $\hat{\psi}(0) = 0$. Furthermore, the set of admissible wavelets is dense in $L^2(\mathbb{R})$, which is not
very hard to prove, see e.g. [59].

We introduce two functions that satisfy the admissibility condition, namely the Haar wavelet and the Mexican hat.

**Example 3.1.3** The Haar wavelet is defined by

\[
\psi(x) = \begin{cases} 
1, & x \in [0, 1/2), \\
-1, & x \in [1/2, 1), \\
0, & \text{otherwise.}
\end{cases}
\]  

(3.7)

The Haar wavelet is depicted in Figure 3.1.a. Later we will see that the Haar wavelet is admissible, since it is compactly supported and \( \int_R \psi(x) \, dx = 0 \). These two conditions are sufficient to guarantee that \( \psi \) is admissible. However, here we show that the Haar wavelet is an admissible wavelet by computing

\[
\hat{\psi}(\omega) = \frac{1}{\sqrt{2\pi}} \left( \int_0^{1/2} e^{i\omega x} \, dx - \int_{1/2}^1 e^{-i\omega x} \, dx \right) = \frac{1}{\sqrt{2\pi}} \left( \frac{1 + e^{-i\omega} - 2e^{-i\omega/2}}{i\omega} \right),
\]

and so

\[
\left| \frac{\hat{\psi}(\omega)}{a} \right|^2 = \frac{\left| 1 + e^{-i\omega} - 2e^{-i\omega/2} \right|^2}{a^3\omega^2} = \frac{\left| e^{-i\omega/2} \right|^2 \cdot \left| e^{i\omega/4} - e^{-i\omega/4} \right|^4}{a^3\omega^2} = 16 \frac{\sin^4(\omega/4)}{a^3\omega^2}.
\]

Integrating by parts yields

\[
C_\psi = \lim_{N \to \infty} \frac{\sin^{4}(x)}{x^3} \, dx
\]

\[
= \lim_{N \to \infty} \left[ \frac{- \sin^4(x)}{2x^2} \right]_0^{N\omega/4} + \lim_{N \to \infty} \frac{1}{4} \int_0^{N\omega/4} \frac{2 \sin(2x) - \sin(4x)}{x^2} \, dx
\]

\[
= \lim_{N \to \infty} \frac{\sin(4x) - 2 \sin(2x)}{4x} \bigg|_0^{N\omega/4} + \lim_{N \to \infty} \int_0^{N\omega/4} \frac{\cos(2x) - \cos(4x)}{x} \, dx
\]

\[
= \lim_{N \to \infty} \left( \int_0^{N\omega/4} \frac{\cos(2x)}{x} \, dx - \int_0^{N\omega/2} \frac{\cos(4x) - 1}{x} \, dx \right)
\]

\[
= \lim_{N \to \infty} C(\ln(N\omega/4) - \ln(N\omega/4) - C(\ln(N\omega/2) + \ln(N\omega/2)) = \ln 2,
\]
Figure 3.2: The CWT using the Haar wavelet: a) a contour plot of $W_\psi[f](a, b)$ with $\psi$ the Haar wavelet, b) the original signal $f$.

where $Ci$ denotes the cosine integral, see [95].

We used the Haar wavelet to compute the CWT of the function $f$ as given in (2.18). In Figure 3.2.a the contour plot of this CWT is depicted. In this plot maxima of $W_\psi[f]$ can be observed around scale $a = 80$ for $b \in [1, 2]$ and around scale $a = 20$ for $b \in [2, 3]$. This difference in scaling behaviour is due to the difference in frequency at the corresponding intervals $x \in [1, 2]$ and $x \in [2, 3]$. Note that the frequency of $f$ increases by a factor of 4 going from one interval to the other and that the scale corresponds to reciprocal values in frequency. Finally, we observe that the energy is also spread outside the interval $b \in [1, 3]$. Moreover, energy is spread more widely for increasing scales, which is due to the convolution product (3.5).

**Example 3.1.4** The Mexican hat $\psi$ is defined by

$$\psi(x) = -\frac{d^2}{dx^2}e^{-x^2/2} = (1 - x^2)e^{-x^2/2}.$$ (3.8)

The Mexican hat is depicted in Figure 3.1.b. Since $\mathcal{F}[f'](\omega) = i\omega \hat{f}(\omega)$ and

$$\int_{\mathbb{R}} e^{-x^2/2}e^{-ix\omega} \, dx = e^{-\omega^2/2},$$
Figure 3.3: The CWT using the Mexican hat: a) a contour plot of $\mathcal{W}_\psi[f](a,b)$ with $\psi$ the Mexican hat, b) the original signal $f$.

we get $\hat{\psi}(\omega) = \frac{1}{\sqrt{2\pi}} \omega^2 e^{-\omega^2/2}$. So

$$C_\psi = 2\pi \int_{\mathbb{R}^+} \frac{|\hat{\psi}(a\omega)|^2}{a} \, da = \int_{\mathbb{R}^+} a^3 \omega^2 e^{-a^2 \omega^2} \, da = 1/2 \int_{\mathbb{R}^+} ye^{-y} \, dy = 1/2.$$ 

Also with the Mexican hat we have computed the CWT of $f$, as defined in (2.18). A contour plot of this CWT is shown in Figure 3.3a. As in Figure 3.2a maxima of $\mathcal{W}_\psi[f]$ can be observed. However, here maxima are located around scale $a = 20$ for $b \in [1, 2]$ and around scale $a = 5$ for $b \in [2, 3]$. The difference in scaling behaviour compared to the CWT using the Haar wavelet is due to a difference in frequency behaviour of both wavelets. It can be seen in Figure 3.1 that frequencies of the Haar wavelet that contain most energy are located around frequencies that are about 4 times higher than the corresponding frequencies of the Mexican hat.

To understand the admissibility condition we give some necessary and sufficient conditions on $\psi$ such that it is a wavelet. A necessary condition on $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ such that it satisfies (3.6) can be derived by applying Lemma 2.1.2 on $C_\psi$. Since $\psi \in L^1(\mathbb{R})$ this
Theorem 3.1.5

Let \( f \in L^2(\mathbb{R}) \) and let \( \psi \in L^2(\mathbb{R}) \) be an admissible wavelet. Then

\[
\int_{\mathbb{R}} |f(x)|^2 \, dx = 1/C_\psi \int_{\mathbb{R}} |\mathcal{W}_\psi[f](a,b)|^2 \, db \frac{da}{a^2}.
\]  

Proof

From (3.5) we get

\[
|\mathcal{W}_\psi[f](a,b)|^2 = |(f * \mathcal{D}_a \hat{\psi})(b)|^2.
\]

We integrate both the left-hand side and the right-hand side of this relation over \( \mathbb{H} \) and apply Plancherel's formula (2.7) on the right-hand side. Now, we arrive cf. [55] at

\[
\int_{\mathbb{R}} |\mathcal{W}_\psi[f](a,b)|^2 \, db \frac{da}{a^2} = 2\pi \int_{\mathbb{R}} |\hat{f}(\omega)|^2 \left|\hat{\psi}(a\omega)\right|^2 \, d\omega \frac{da}{a}.
\]

So

\[
\int_{\mathbb{R}} |\mathcal{W}_\psi[f](a,b)|^2 \, db \frac{da}{a^2} = \int_{\mathbb{R}} |\hat{f}(\omega)|^2 \left(2\pi \int_{\mathbb{R}^+} \frac{|\hat{\psi}(a\omega)|^2 \, da}{a}\right) \, d\omega
\]

\[= C_\psi \int_{\mathbb{R}} |\hat{f}(\omega)|^2 \, d\omega = C_\psi \|f\|_2^2,
\]
using Fubini's theorem and Plancherel's formula.

Relation (3.10) can be seen as Plancherel's formula for the CWT. As a result of this relation we get

\[ f \in L^2(\mathbb{R}) \implies \mathcal{W}_\psi f \in L^2(\mathcal{H}, a^{-2} \, db \, da), \]

for any \( \psi \) for which \( C_\psi < \infty \). Moreover, by polarization, (3.10) yields immediately Parseval's formula for the CWT

\[
\int_{\mathbb{R}} f(x) \overline{g(x)} \, dx = \frac{1}{C_\psi} \int_{\mathcal{H}} \mathcal{W}_\psi[f](a,b) \overline{\mathcal{W}_\psi[g](a,b)} \, db \, da, \tag{3.11}
\]

for \( f, g \in L^2(\mathbb{R}) \), if \( \psi \) is admissible. Formula (3.11) gives

\[
\int_{\mathbb{R}} f(x) \overline{g(x)} \, dx = \frac{1}{C_\psi} \int_{\mathcal{H}} \mathcal{W}_\psi[f](a,b) \int_{\mathbb{R}} \frac{1}{\sqrt{a} \, g(x)} \psi \left( \frac{x-b}{a} \right) \, dx \, db \, da \, a^2,
\]

\[
= \int_{\mathbb{R}} \left( \frac{1}{C_\psi} \int_{\mathcal{H}} \mathcal{W}_\psi[f](a,b) \psi \left( \frac{x-b}{a} \right) \, db \, da \, a^2 \sqrt{a} \right) \overline{g(x)} \, dx,
\]

for all \( g \in L^2(\mathbb{R}) \). This proves the following formal reconstruction theorem.

**Theorem 3.1.6** Let \( f \in L^2(\mathbb{R}) \) and let \( \psi \in L^2(\mathbb{R}) \) be an admissible wavelet. Then

\[
f(x) = 1/C_\psi \int_{\mathcal{H}} \mathcal{W}_\psi[f](a,b) \psi \left( \frac{x-b}{a} \right) \, db \, da \, a^2 \sqrt{a}, \text{ weakly in } L^2(\mathbb{R}). \tag{3.12}
\]

Relation (3.12) should be interpreted by means of the first identity derived above before Theorem 3.1.6. A stronger result holds for \( f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \) and \( \dot{f} \in L^1(\mathbb{R}) \). With these assumptions on \( f \) it can be proved, see e.g. [55], that (3.12) holds pointwise. Furthermore, under these assumptions, for each \( x \in \mathbb{R} \), both the inner integral (over \( b \)) and the outer integral (over \( a \)) are necessarily absolutely convergent, while the double integral in (3.12) is not necessarily absolutely convergent. Note that \( f(x) \) is well-defined for all \( x \in \mathbb{R} \), since \( f \) is continuous due to Lemma 2.1.2.

We have shown that \( \mathcal{W}_\psi \) maps all \( f \in L^2(\mathbb{R}) \) into some \( \mathcal{W}_\psi f \in L^2(\mathcal{H}, a^{-2} \, db \, da) \). Moreover, we have shown that we can reconstruct \( f \) from \( \mathcal{W}_\psi f \), if \( \psi \) is an admissible wavelet. However the range of \( \mathcal{W}_\psi \), \( \text{Ran}(\mathcal{W}_\psi) \) is only a subset of \( L^2(\mathcal{H}, a^{-2} \, db \, da) \). To characterize \( \text{Ran}(\mathcal{W}_\psi) \) for admissible wavelets \( \psi \) we derive from (3.11) using Fubini's theorem

\[
\mathcal{W}_\psi[f](a,b) = (f, T_b D_a \psi) = \frac{1}{C_\psi \sqrt{a}} \int_{\mathbb{R}} \int_{\mathcal{H}} \mathcal{W}_\psi[f](u,v) \psi \left( \frac{x-v}{u} \right) \overline{\psi \left( \frac{x-b}{a} \right)} \, dv \, du \, da \, u^2 \sqrt{u} \, dx.
\]
The Continuous Wavelet Transform

\[
\begin{align*}
\mathcal{W}_\psi[f][u,v] &= \frac{1}{C_\psi} \int_{\mathcal{H}} \int_{\mathbb{R}} \mathcal{W}_\psi[f](u,v) \psi \left( \frac{x - b}{a} \right) dx dv \frac{du}{u^2 \sqrt{u}} \\
&= \int_{\mathcal{H}} k_\psi(a,b;u,v) \mathcal{W}_\psi[f](u,v) dv \frac{du}{u^2},
\end{align*}
\]

with \( k_\psi(a,b;u,v) = (T_a^b D_u \psi, T_b^a D_u \psi) / C_\psi \) the reproducing kernel. Remark, that

\[
k_\psi(a,b;u,v) = \mathcal{W}_\psi[T_b^a D_u \psi](u,v).
\]

Therefore, \( k_\psi(a,b;u,v) \in L^2(\mathcal{H}, u^{-2} dv du) \), for all \((a,b) \in \mathcal{H}\). A necessary condition on \( h \in L^2(\mathcal{H}, a^{-2} db da) \) such that it is in \( \text{Ran}(\mathcal{W}_\psi) \) is given by

\[
h(a,b) = \int_{\mathcal{H}} k_\psi(a,b;u,v) h(u,v) dv \frac{du}{u^2}. \tag{3.13}
\]

Note that \( h \) is continuous on \( \mathcal{H} \) by Lemma 3.1.1. Therefore (3.13) is well defined.

Moreover, it can be shown [51], that (3.13) is also a sufficient condition on \( h \) to be in \( \text{Ran}(\mathcal{W}_\psi) \). Resuming,

\[
\text{Ran}(\mathcal{W}_\psi) = \{ h \in L^2(\mathcal{H}, a^{-2} db da) \mid h(a,b) = \int_{\mathcal{H}} k_\psi(a,b;u,v) h(u,v) dv \frac{du}{u^2} \}.
\]

Obviously, \( \text{Ran}(\mathcal{W}_\psi) \) is closed due to Theorem 3.1.5. Combining this result with Theorem 3.1.5 yields that the transform \( f \mapsto \mathcal{W}_\psi f / \sqrt{C_\psi} \) is a Hilbert space isometry from \( L^2(\mathcal{H}) \) onto \( \text{Ran}(\mathcal{W}_\psi) \) as given above.

We observe that the results on the reproducing kernel and the range of \( \mathcal{W}_\psi \) as presented here are similar to the results we presented in Section 2.2, where we considered the WFT.

In Chapter 2 we have seen that the WFT and the Wigner distribution are related to the Heisenberg group by means of the Schrödinger representation. Also the CWT is related to a group, namely the Lie group \( G \), which is identified with \( \mathcal{H} \) with the multiplication law

\[
(a_1,b_1)(a_2,b_2) = (a_1 a_2, a_1 b_2 + b_1).
\tag{3.14}
\]

We observe that the affine-linear group \( G_\alpha \) identified with \( \mathbb{R}^2 \) with multiplication law (3.14) is isomorphic to \( \mathbb{Z}_2 \ltimes G \). The left and right Haar measures \( (\mu_L \text{ and } \mu_R) \) of \( G \) are Borel measures for which \( \mu_L(gE) = \mu_L(E) \) and \( \mu_R(Eg) = \mu_R(E) \) for all \( g \in G \) and all Borel sets \( E \subseteq G \). A straightforward calculation yields \( d\mu_L(a,b) = a^{-2} db da \) and \( d\mu_R(a,b) = a^{-1} db da \).
We introduce a representation of $G$ in $U(L^2(\mathbb{R}))$ by
\[
\pi(a, b) = T_b D_a, \quad \forall (a, b) \in G.
\]

Obviously $\pi$ is a group homomorphism. Furthermore, it is continuous in the strong operator topology of $U(L^2(\mathbb{R}))$ and $\pi(a, b)$ is unitary for all $(a, b) \in G$. In this setting a function $h \in L^2(\mathbb{R})$ is called an admissible vector if
\[
(a, b) \mapsto (f, \pi(a, b)h)_2 \in L^2(H, \mu_L) \quad \forall f \in L^2(\mathbb{R}).
\]

Since $(f, \pi(a, b)h) = \mathcal{W}_h[f](a, b)$ Theorem 3.1.5 states that $h$ is an admissible vector if and only if $C_h < \infty$. We already observed that the set of admissible wavelets is dense in $L^2(\mathbb{R})$ and therefore the set of all $h \in L^2(\mathbb{R})$ that satisfy (3.15) is dense in $L^2(\mathbb{R})$ as well. The following theorem shows that $\pi$ is irreducible.

**Theorem 3.1.7** Let $\pi$ be the unitary representation of $G$ in $U(L^2(\mathbb{R}))$ by
\[
\pi(a, b) = T_b D_a, \quad \forall (a, b) \in G.
\]

Then $\pi$ is irreducible.

**Proof**
We assume $\pi$ is reducible. Then there exists a closed linear subspace $V \subset L^2(\mathbb{R})$, with $V \neq \{0\}$ and $V \neq L^2(\mathbb{R})$, such that
\[
\pi(a, b)V \subset V \quad \text{for all } (a, b) \in G.
\]

Then there exists non-trivial vectors $g \in V$ and $f \in V^\perp$ such that $(f, \pi(a, b)g)_2 = 0$ for all $(a, b) \in G$. Now
\[
\|f\|_2^2 C_g = (f, \pi(a, b)g)_2 = 0,
\]
yields $f = 0$ or $g = 0$, which is in contradiction with $f \neq 0$ and $g \neq 0$. \qed

To conclude, the continuous wavelet transform with a wavelet $\psi$ is a unitary irreducible representation of the Lie group $G$ with admissible vector $\psi$. More detailed studies on the wavelet transform and group theory can be found in [38, 39, 66].

### 3.2 The Discrete Wavelet Transform

In the previous section we considered the CWT. We showed that this integral transform is able to analyse signals both in time/space and scale. Moreover, it turned out that such signals can be recovered from their CWT. In this section we consider the problem of calculating efficiently the wavelet transform of a function and reconstructing it efficiently from its transform.
A first approach is to compute the wavelet transform only for a discrete subset \( L \subset \mathcal{H} \), e.g.

\[
L = \{(a_0^m, n b_0 a_0^m) \mid n, m \in \mathbb{Z}\},
\]

for some \( a_0 > 1 \) and \( b_0 > 0 \) and to replace the double integral in (3.12) by a double sum over \( L \). Using such an approach we have to show that the integral can indeed be replaced by a double sum without loss of information. Furthermore, reconstruction of a function \( f \in L^2(\mathbb{R}) \) by means of this double sum should depend continuously on \( f \). This kind of stability is guaranteed if the tuple \((\psi, a_0, b_0)\) generates a wavelet frame, i.e.,

\[
m_F \|f\|_2^2 \leq \sum_{(a,b) \in L} |\mathcal{W}_\psi[f](a,b)|^2 \leq M_F \|f\|_2^2, \quad \forall f \in L^2(\mathbb{R})
\]

for some constants \( m_F, M_F > 0 \) called the frame bounds. We observe that if \((\psi, a_0, b_0)\) generates a wavelet frame with \( m_F = M_F = 1 \) and if \( \|\psi\|_2 = 1 \) then

\[
\{T_{n b_0 a_0^m} \mathcal{D}_{a_0^m} \psi \mid n, m \in \mathbb{Z}\}
\]

is an orthonormal basis in \( L^2(\mathbb{R}) \) and conversely. In that case we can transform and reconstruct any \( f \in L^2(\mathbb{R}) \) with respect to the lattice \( L \) using a transformation called the discrete wavelet transform (DWT). This DWT provides an efficient algorithm, that does not need to compute \( \mathcal{W}_\psi[f](a,b) \) for all \((a,b) \in L\) by means of inner products. In Section 3.2.2 we will discuss this algorithm, called the pyramid algorithm.

In [22] Daubechies has given necessary and sufficient conditions on \((\psi, a_0, b_0)\) such that (3.17) is a wavelet frame. We summarize some of these conditions in the following theorem.

**Theorem 3.2.1** Assume that

1. \( \text{ess} \inf_{|\omega| \in [1,a_0]} \sum_{m \in \mathbb{Z}} |\hat{\psi}(a_0^m \omega)|^2 > 0, \)

2. \( \text{ess} \sup_{|\omega| \in [1,a_0]} \sum_{m \in \mathbb{Z}} |\hat{\psi}(a_0^m \omega)|^2 < \infty, \)

3. \( \sup_{x \in \mathbb{R}} (1 + |x|^2)^{(1+d)/2} h(x) < \infty, \) for some \( \delta > 0 \) with

\[
h(x) = \sup_{|\omega| \in [1,a_0]} \sum_{m \in \mathbb{Z}} |\hat{\psi}(a_0^m \omega)||\hat{\psi}(a_0^m \omega + x)|.
\]

Then there exists an \( N > 0 \) such that

a. \((\psi, a_0, b_0)\) generates a wavelet frame for \( \psi \in L^2(\mathbb{R}) \), \( a_0 > 1 \) and \( 0 < b_0 < N \),

b. \( \forall e > 0 \exists b_0 \in [N, N+e] : (\psi, a_0, b_0) \) does not generate a wavelet frame.
Moreover, if \((\psi, \alpha_0, b_0)\) generates a wavelet frame, then

\[
m_F \leq \frac{C_\psi}{2b_0 \ln \alpha_0} \leq M_F.
\]  
(3.18)

Proof
Cf. [22].

From (3.18) it follows immediately that \(\psi\) is an admissible wavelet if \((\psi, \alpha_0, b_0)\) generates a wavelet frame or an orthonormal basis in \(L^2(\mathbb{R})\). A more elegant and fast way to come to a DWT by means of orthonormal wavelet bases in \(L^2(\mathbb{R})\) is given by the concept of a multiresolution analysis (MRA), which is considered in the sequel of this section. In Chapter 4 we shall return to the notion of frames related to MRA.

### 3.2.1 Multiresolution Analysis in \(L^2(\mathbb{R})\)

The concept of an MRA is due to Mallat [60] and Meyer [62] and was originally used as a signal processing tool by means of perfect reconstruction filter banks [43, 95], which is discussed in Section 3.2.2. We start with the definition of an MRA, following [16, 23, 60].

**Definition 3.2.2** An MRA in \(L^2(\mathbb{R})\) is an increasing sequence of closed subspaces \(V_j, j \in \mathbb{Z}\), in \(L^2(\mathbb{R})\),

\[
\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \cdots,
\]
such that

1. \(\bigcup_{j \in \mathbb{Z}} V_j\) is dense in \(L^2(\mathbb{R})\),
2. \(\bigcap_{j \in \mathbb{Z}} V_j = \{0\}\),
3. \(f \in V_j \iff D^{-1} f \in V_{j-1}, \ \forall j \in \mathbb{Z}\),
4. \(\exists \phi \in L^2(\mathbb{R}) : \{T^k \phi \mid k \in \mathbb{Z}\}\) is an orthonormal basis for \(V_0\),

with \(D := D_2\) and \(T := T_1\) and \(\phi\) real-valued.

For a more general concept of an MRA we can replace Condition 4 by

\[
\exists \phi \in L^2(\mathbb{R}) : \{T^k \phi \mid k \in \mathbb{Z}\}\) is a Riesz basis for \(V_0\).
\]

Characterizations of a Riesz basis are given in [108]. Also the concept of an MRA for \(L^2(\mathbb{R}^n)\) has been described in the literature thoroughly, see [23, 62]. In Chapter 4 we consider this general concept of an MRA in a functional analytical setting. Here we stick at Definition 3.2.2.
By Definition 3.2.2 an orthonormal basis for $V_j$, with $j \in \mathbb{Z}$ fixed, is given by

$$\{ \mathcal{D}^j T^k \phi \mid k \in \mathbb{Z} \},$$

once such function $\phi$ has been found. Such $\phi$ is called scaling function. Since $\mathcal{D}$ is a unitary operator on $L^2(\mathbb{R})$ and $V_0$ is invariant under the action of $T$, the collection

$$\{ \mathcal{D}^{-1} T^k \phi \mid k \in \mathbb{Z} \}$$

is an orthonormal basis for $V_{-1}$. As we also have $\phi \in V_{-1}$, we get

$$\phi = \sum_{k \in \mathbb{Z}} p(k) \mathcal{D}^{-1} T^k \phi,$$  \hspace{1cm} (3.19)

for some real-valued $p \in l^2(\mathbb{Z})$. In Section 3.2.2 we also want $p$ to generate a bounded convolution operator on $l^2(\mathbb{Z})$. Therefore we require $p \in l^1(\mathbb{Z})$. Relation (3.19) is referred to as scale relation and $p$ as scale sequence.

We consider again the inclusion $V_0 \subset V_{-1}$. Obviously we can define a subspace $W_0$ such that $W_0 \cong V_{-1}/V_0$. For a unique definition of $W_0$, we take $W_0 = V_{-1} \ominus V_0^\perp$. Using the invariance of the subspaces $V_j$ under the action of the unitary operator $\mathcal{D}^{-1}$ we arrive in a natural way at the definition of the closed subspaces $W_j \subset L^2(\mathbb{R})$ by putting $W_j = V_{j-1} \ominus V_j^\perp$. Recursively repeating the orthonormal decomposition of some $V_{-j}$ into $V_{-j+1}$ and $W_{-j+1}$ yields

$$V_{-J} = V_{-J+1} \oplus W_{-J+1} = V_{-J+2} \oplus W_{-J+2} \oplus W_{-J+1}$$

$$= \ldots = V_J \oplus \left( \bigoplus_{j=J} W_j \right).$$

Taking $J \to \infty$ and applying Conditions 1 and 2 from Definition 3.2.2 leads to

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

Now assume that we can find a real-valued function $\psi \in V_{-1}$, such that $\{ T^k \psi \mid k \in \mathbb{Z} \}$ is an orthonormal basis for $W_0$. Then $\{ \mathcal{D}^j T^k \psi \mid k \in \mathbb{Z} \}$ is an orthonormal basis for $W_j$, $j \in \mathbb{Z}$. Since the subspaces $W_j$ are chosen to be mutually orthogonal, we then have an orthonormal basis in $L^2(\mathbb{R})$ given by $\{ \mathcal{D}^j T^k \psi \mid j, k \in \mathbb{Z} \}$. By (3.18) the function $\psi$ is a wavelet and

$$\{ \mathcal{D}^j T^k \psi \mid k \in \mathbb{Z} \}$$

is a wavelet basis for $W_j$, for fixed $j \in \mathbb{Z}$. So, using these wavelet bases we are able to decompose any $f \in L^2(\mathbb{R})$ into functions at several scales.
Since the wavelet function \( \psi \) should be in \( V_{-1} \), there exists also a scale relation for \( \psi \)

\[
\psi = \sum_{k \in \mathbb{Z}} q(k) D^{-1} T^k \phi,
\]

(3.20)

for some real-valued \( q \in l^2(\mathbb{Z}) \), the scaling sequence for \( \psi \). As in (3.19) we will also require \( q \in l^1(\mathbb{Z}) \). By this relation, the problem of finding \( \psi \) can be replaced by the problem of finding \( q \) if \( \phi \), and therefore also \( p \), is known. A well-known choice [23] for \( q \) is given by

\[
q(k) = (-1)^k p(1 - k).
\]

(3.21)

Note that we have introduced the concept of an MRA for \( L^2(\mathbb{R}) \) to constitute orthonormal wavelet bases in \( L^2(\mathbb{R}) \) by means of dilations and translations. These bases are equal to the wavelet frames (3.17) defined on the lattice

\[
L_d = \{(2^m, n2^m) \mid m, n \in \mathbb{Z}\}
\]

(3.22)

with frame bounds \( m_F = M_F = 1 \) and \( \|\psi\|_2 = 1 \).

An MRA for \( L^2(\mathbb{R}^m) \) can be defined in a similar way, see e.g. [23, 62]. In Chapter 4 we also consider a framework of such an MRA as well as a setting for an MRA which generates wavelet bases defined on other lattices than (3.22).

### 3.2.2 MRA and Filter Banks

An MRA can be related to filter banks by looking at a one-level decomposition and reconstruction \( V_{j-1} = V_j \oplus W_j \). At the same time this relation yields a scheme to calculate the wavelet transform of a function in \( L^2(\mathbb{R}) \) and to reconstruct it from its transform on \( L_d \) in a fast way. For this decomposition and reconstruction we introduce the orthoprojectors \( P_j \) and \( Q_j \) on \( V_j \) and \( W_j \) respectively. By definition we have

\[
P_{j-1} = P_j + Q_j,
\]

(3.23)

for all \( j \in \mathbb{Z} \). Note that Conditions 1 and 2 from Definition 3.2.2 yield

\[
\lim_{j \to -\infty} P_j f = f,
\]

(3.24)

and

\[
\lim_{j \to \infty} P_j f = 0,
\]

(3.25)

for all \( f \in L^2(\mathbb{R}) \).
The decomposition algorithm:

We assume \( \mathcal{P}_{j-1} f \in V_{j-1} \) is known for a certain \( j \in \mathbb{Z} \). Consequently there exists a sequence \( c_{j-1} \in l^2(\mathbb{Z}) \) such that

\[
\mathcal{P}_{j-1} f = \sum_{k \in \mathbb{Z}} c_{j-1}(k) D^{j-1} \mathcal{T}^k \phi.
\]

Moreover, the sequence \( c_{j-1} \) is given by \( c_{j-1}(k) = (D^{j-1} \mathcal{T}^k \phi, f)_2 \). Following decomposition (3.23) we have

\[
\mathcal{P}_{j-1} f = \sum_{k \in \mathbb{Z}} c_j(k) D^j \mathcal{T}^k \phi + \sum_{k \in \mathbb{Z}} d_j(k) D^j \mathcal{T}^k \psi, \tag{3.26}
\]

with the sequence \( d_j \) given by

\[
Q_j f = \sum_{k \in \mathbb{Z}} d_j(k) D^j \mathcal{T}^k \psi.
\]

Note that the sequences \( c_j \) and \( d_j \) can be calculated respectively by \( c_j(k) = (D^j \mathcal{T}^k \phi, f)_2 \) and \( d_j(k) = (D^j \mathcal{T}^k \psi, f)_2 \). However with the known sequence \( c_{j-1} \) we can also derive using (3.19) and (3.26)

\[
c_j(k) = (\mathcal{P}_{j-1} f, D^j \mathcal{T}^k \phi)_2 = \sum_{n \in \mathbb{Z}} p(n) (\mathcal{P}_{j-1} f, D^{j-1} \mathcal{T}^{2k+n} \phi)_2 = \sum_{m, n \in \mathbb{Z}} c_{j-1}(m) p(n) (D^{j-1} \mathcal{T}^{m} \phi, D^{j-1} \mathcal{T}^{2k+n} \phi)_2 = \sum_{n \in \mathbb{Z}} c_{j-1}(2k+n) p(n) = ((\downarrow 2)[c_{j-1} * \tilde{p}]) (k), \tag{3.27}
\]

with \((c_{j-1} * \tilde{p}) (k) = \sum_{n \in \mathbb{Z}} c_{j-1}(n-k) \tilde{p}(k)\), \( \tilde{p}(n) = p(-n) \) and with \((\downarrow 2)\) the downsampling operator given by

\[
((\downarrow 2)[u]) (k) = u(2k),
\]

for all \( u \in l^2(\mathbb{Z}) \). In the same manner we get

\[
d_j = (\downarrow 2)(c_{j-1} * \tilde{q}). \tag{3.28}
\]

So \( c_j \) and \( d_j \) are obtained from \( c_{j-1} \) by taking \( c_{j-1} \) as input for the linear time-invariant filters \( \tilde{p} \) and \( \tilde{q} \) respectively. After this filtering operation the sequences are downsampled by a factor 2. This can be visualized by means of an analysis part of a two channel filter bank as depicted in Figure 3.4.
Recursively we get expressions for $c_{j+n}$ and $d_{j+n}$ for $n \geq 1$, namely

$$c_{j+n} = ((\downarrow 2)C_p)^n c_j \quad \text{and} \quad d_{j+n} = ((\downarrow 2)C_q)_{n-1} c_j,$$

(3.29)

where $C_u$ denotes convolution with $u \in l^1(\mathbb{Z})$, i.e.

$$C_u[c](n) = \sum_{k \in \mathbb{Z}} u(n - k)c(k),$$

for all $c \in l^2(\mathbb{Z})$. In terms of filter banks we can say that $c_{j+n}$ and $d_{j+n}$ are obtained by iterative use of the analysis part of two channel filter bank of Figure 3.4.

**The reconstruction algorithm:**

Once we have computed a decomposition of $P_{j-1}f$ into $P_jf$ and $Q_jf$ by means of the coefficients $c_j$ and $d_j$, we can also recover $c_{j-1}$ out of $c_j$ and $d_j$ in an efficient way. In order to come to such a reconstruction formula we will represent $P_jf$ and $Q_jf$ in terms of $D^{j-1}T^k\phi$, $k \in \mathbb{Z}$, the basis functions of $V_{j-1}$. For this we introduce for $f \in L^2(\mathbb{R})$ the $l^2$-sequences

$$\alpha_{n,m}(k) = (P_n f, D^m T^k \phi) \quad \text{and} \quad \beta_{n,m}(k) = (Q_n f, D^m T^k \phi)$$

(3.30)

Since $P_j f \in V_{j-1}$, we can write

$$P_j f = \sum_{k \in \mathbb{Z}} \alpha_{j,j-1}(k) D^{j-1} T^k \phi.$$
An expression for $\alpha_{j,j-1}$ is found by taking the inner product with $D^{j-1}T^k\phi$ in both the left-hand side and the right-hand side of this equation. This yields in combination with (3.26) and (3.19)

$$\alpha_{j,j-1}(k) = (P_j f, D^{j-1}T^k\phi)_2 = \sum_{n \in \mathbb{Z}} c_j(n) (D^{j-1}T^n\phi, D^{j-1}T^k\phi)_2$$

$$= \sum_{m,n \in \mathbb{Z}} c_j(n)p(m) (D^{j-1}T^{m+2n}\phi, D^{j-1}T^k\phi)_2$$

$$= \sum_{n \in \mathbb{Z}} c_j(n)p(k-2n) = (((\uparrow 2)c_j) \ast p)(k),$$

(3.31)

with $(\uparrow 2)$ the upsampling operator given by

$$(\uparrow 2)[u](k) = \begin{cases} u(k/2), & k \text{ mod } 2 = 0, \\ 0, & \text{otherwise}, \end{cases}$$

for all $u \in L^2(\mathbb{Z})$. We observe that $(\downarrow 2)(\uparrow 2) = I$ and $(\downarrow 2)^* = (\uparrow 2)$. In the same manner we have

$$Q_j f = \sum_{k \in \mathbb{Z}} \beta_{j,j-1}(k) D^{j-1}T^k\phi,$$

with

$$\beta_{j,j-1} = ((\uparrow 2)d_j) \ast q.$$  

(3.32)

So we derived the reconstruction formula

$$c_{j-1} = \alpha_{j,j-1} + \beta_{j,j-1} = ((\uparrow 2)c_j) \ast p + ((\uparrow 2)d_j) \ast q.$$  

Hence, $c_{j-1}$ can be recovered from $\alpha_{j,j-1}$ and $\beta_{j,j-1}$. Here $\alpha_{j,j-1}$ and $\beta_{j,j-1}$ can be seen as the output sequences of the linear time-invariant filters $p$ and $q$ respectively with input sequences $c_j$ and $d_j$ upsampled by a factor 2. This can be visualized by means of a synthesis part of a two channel filter bank as has been depicted also in Figure 3.4.

Recursively we get expressions for $\alpha_{j+n,j}$ and $\beta_{j+n,j}$ for $n \geq 1$, namely

$$\alpha_{j+n,j} = (C_p(\uparrow 2))^n c_{j+n} \quad \text{and} \quad \beta_{j+n,j} = (C_p(\uparrow 2))^{n-1} C_q(\uparrow 2)d_{j+n},$$

(3.33)

In terms of filter banks we can say that $\alpha_{j+n,j}$ and $\beta_{j+n,j}$ are obtained by iterative use of the synthesis part of two channel filter bank of Figure 3.4. This recursive approach to obtain $\alpha_{j+n,j}$ and $\beta_{j+n,j}$ is depicted in Figure 3.5 for $n = 2$.

The algorithm of decomposing and reconstructing functions by means of filter banks related to an MRA is known as the pyramid algorithm.
3.2.3 Implementation of the DWT and its Use for $l^2(\mathbb{Z})$

A problem that appears in decomposing a function $f \in L^2(\mathbb{R})$ at several scales by means of the pyramid algorithm is that the coefficients $c_j$ should be known in order to compute $P_{j+m} f$, $m \in \mathbb{N}$, in a fast way. Computing $c_j$ does not only slow down the algorithm, it can also be impossible if only samples of $f$ are given. The last problem appears if $f$ is given by discrete-time measurements.

An approximation of $c_j$ can be given by

$$c_j(k) = f(2^j k).$$

Note that if only measurements of $f$ are available, we can identify these measurements with $c_j$ assuming that the samples are taken from some $f \in L^2(\mathbb{R})$ at sampling rate $2^j$. The following theorem, which is a generalization of an exercise in [59], shows that this is a good approximation, under certain conditions on $f$ and the scaling function $\varphi$.

**Theorem 3.2.3** Let $f \in L^2(\mathbb{R})$ be Hölder continuous of order $\alpha \in (0, 1]$, i.e.,

$$|f(x) - f(y)| \leq C \cdot |x - y|^\alpha, \quad \forall x, y \in \mathbb{R},$$

(3.34)

for a constant $C > 0$ and let $s_j(k) = f(2^j k)$ for some $j \in \mathbb{Z}$. Let the scaling function $\varphi$ be
The Discrete Wavelet Transform

continuous in \( \mathbb{Z} \) and let \( \phi \) satisfy

\[
\sum_{k \in \mathbb{Z}} \phi(x - k) = 1 \quad \text{a.e.} \quad x \in \mathbb{R},
\]

where the sum converges absolutely almost everywhere, and

\[
\sum_{k \in \mathbb{Z}} |k|^\alpha |\phi(k)| < \infty. 
\]

Then

\[
\forall \varepsilon > 0 \exists j \in \mathbb{Z} \forall n \in \mathbb{Z} \quad |P_j[f](2^j n) - S_j[f](2^j n)| < \varepsilon,
\]

with \( S_j : L^2(\mathbb{R}) \to V_j \) given by

\[
S_j[f](x) = \sum_{k \in \mathbb{Z}} s_j(k) \phi(2^{-j} x - k).
\]

Proof

For all \( j \in \mathbb{Z} \) we can write

\[
|P_j[f](x) - S_j[f](x)| = |P_j[f](x) - f(x) - S_j[f](x) + f(x)|
\leq |P_j[f](x) - f(x)| + |S_j[f](x) - f(x)|.
\]

From the Hölder continuity of \( f \) and Jackson’s inequality, see e.g. [105], it follows that

\[
||P_j f - f||_\infty \leq C \sup_{0 < |h| < 2^j} ||f - T_h f||_\infty
\leq C_0 2^{\alpha j} \to 0 \quad (j \to -\infty),
\]

for some positive constants \( C \) and \( C_0 \). The proof is established by showing that

\[
|S_j[f](2^j n) - f(2^j n)| \to 0 \quad (j \to -\infty).
\]

We derive

\[
|S_j[f](2^j n) - f(2^j n)| = \left| \sum_{k \in \mathbb{Z}} s_j(k) \phi(n - k) - s_j(n) \right|
\leq \sum_{k \in \mathbb{Z}} |f(2^j n) - f(2^j (n - k))| |\phi(k)|
\leq C 2^{j \alpha} \sum_{k \in \mathbb{Z}} |k|^\alpha |\phi(k)| \to 0 \quad (j \to -\infty).
\]
In the previous theorem we used two conditions on \( \phi \), which might seem strong and a bit unfamiliar as well. However, Condition (3.36) is already satisfied if \( \phi \) is compactly supported, which is the case for the well-known Daubechies functions [23] and the spline scaling functions, i.e. B-splines [12]. For \( \phi \) continuous in \( \mathbb{Z} \), sufficient conditions on \( \phi \) such that Condition (3.35) is satisfied are given in the following lemma.

**Lemma 3.2.4** Let \( \phi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \) such that \( \{ T^k \phi : k \in \mathbb{Z} \} \) is an orthonormal set in \( L^2(\mathbb{R}) \) and \( \hat{\phi}(0) = 1/\sqrt{2\pi} \). Then

\[
\sum_{k \in \mathbb{Z}} \phi(x-k) = 1 \text{ a.e. } x \in \mathbb{R},
\]

where the sum converges absolutely almost everywhere.

**Proof**

Take \( g(\omega) = \sum_{l \in \mathbb{Z}} |\hat{\phi}(\omega + 2\pi l)|^2 \). Then \( g \) is \( 2\pi \)-periodic and \( L^1 \). Its Fourier coefficients are given by

\[
c_k = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} g(\omega) e^{-i\omega k} d\omega = \frac{1}{\sqrt{2\pi}} \sum_{l \in \mathbb{Z}} |\hat{\phi}(\omega + 2\pi l)|^2 e^{-i\omega k} d\omega
\]

\[
= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |\hat{\phi}(\omega)|^2 e^{-i\omega k} d\omega = 1/\sqrt{2\pi} (\phi, T^{-k} \phi) = 1/\sqrt{2\pi} \delta_{0,k}.
\]

This yields \( g(\omega) = 1/\sqrt{2\pi} \) for all \( \omega \in \mathbb{R} \), since \( \hat{\phi} \) is continuous. In particular we have \( g(0) = 1/\sqrt{2\pi} \), which leads in combination with \( \hat{\phi}(0) = 1/\sqrt{2\pi} \) to \( \hat{\phi}(2\pi l) = 0 \) for all \( l \in \mathbb{Z} \backslash \{0\} \). Now, put \( h(x) = \sum_{k \in \mathbb{Z}} \phi(x-k) \). Since \( \phi \in L^1(\mathbb{R}) \) this sum converges absolutely. This means that \( h \) is well defined. Furthermore, \( h \) is 1-periodic and \( L^1 \). Its Fourier coefficients can be computed in the same manner as for \( g \). We get \( c_l = \sqrt{2\pi} \hat{\phi}(2\pi l) = \delta_{0,l} \).

Writing down the Fourier series of \( h \) establishes the proof.

The property of \( \phi \) as considered in Lemma 3.2.4 is called the partition of the unity. If \( \phi \) is also continuous in \( k \in \mathbb{Z} \), then the preceding lemma gives sufficient conditions on \( \phi \) such that (3.35) holds. The conditions on \( \phi \) such that this partition is guaranteed are quite natural. In fact, Wojtaszczyk showed in [106] that every scaling function \( \phi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \) of an MRA as defined in Definition 3.2.2 satisfies the conditions in Lemma 3.2.4. Moreover, also scaling functions \( \phi \) for which \( \{ T^k \phi : k \in \mathbb{Z} \} \) is a Riesz basis may satisfy Conditions (3.35) and (3.36). A classical example of such scaling functions \( \phi \) are the cardinal B-spline functions [12], which will be discussed briefly in Section 4.5.
Starting with a discrete-time function, obtained from measurements, we would like to get a decomposition at various discrete-time resolution levels instead of a DWT for $L^2(\mathbb{R})$. A natural way to obtain such a decomposition is to identify a given $s \in l^2(\mathbb{Z})$ with a sequence of coefficients $c_j$ for some $j \in \mathbb{Z}$. So, we construct a function in $L^2(\mathbb{R})$ with $U_j : l^2(\mathbb{Z}) \rightarrow V_j$ by

$$U_j[s](x) = 2^{-j/2} \sum_{k \in \mathbb{Z}} s(k) \phi(2^{-j} x - k),$$

for all $x \in \mathbb{R}$. Note that $P_j U_j = U_j$ and $s(k) = (U_j s, T^j \phi^*)$.

A decomposition of $s$ at level $m$, denoted by $s^{(m)}$, should satisfy

$$U_j s^{(m)} = Q_{j+m} U_j s.$$

Its approximation at level $m$, denoted by $s^{(m)}_{ap}$, should satisfy

$$U_j s^{(m)}_{ap} = P_{j+m} U_j s.$$

These relations hold if and only if $s^{(m)} = \beta_{j+m,j}$, $s^{(m)}_{ap} = \alpha_{j+m,j}$ and $c_j = s$, with $\alpha$ and $\beta$ as in (3.33). Combining this result with (3.29) we arrive at the definition of the DWT for $l^2(\mathbb{Z})$.

**Definition 3.2.5** Let $p$ and $q$ be the scale sequences as given in (3.19) and (3.20). Furthermore, let $G_u = C_u(\uparrow 2)$ for $u \in l^1(\mathbb{Z})$. The $L^2$-DWT of a sequence $s \in l^2(\mathbb{Z})$ at scale $m \in \mathbb{N}$ is given by

$$s^{(m)} = G_p^{m-1} G_q G_q^* (G_p^*)^{m-1} s.$$  

(3.37)

Its approximation at scale $m \in \mathbb{N}$ is given by

$$s^{(m)}_{ap} = G_p^m (G_p^*)^m s.$$  

(3.38)

In the following lemma we come to some useful properties of the $L^2$-DWT, which we already met in the case of the DWT for $L^2(\mathbb{R})$.

**Lemma 3.2.6** For $s \in l^2(\mathbb{Z})$, let $s^{(m)}$ and $s^{(m)}_{ap}$ denote the $L^2$-DWT at level $m$ and its approximation at level $m$ respectively for $m \in \mathbb{N}$. Then

1. $\|s\|_2^2 = \|s^{(M)}\|_2^2 + \sum_{m=1}^{M} \|s^{(m)}\|_2^2$, for all $M \in \mathbb{N}$,

2. $\lim_{m \to \infty} \|s^{(m)}\|_2 = 0.$
Proof
First we observe that $U_j$ is unitary for all $j \in \mathbb{Z}$, which yields

$$
\|s\|_2^2 = \|U_j s\|_2^2 = \|P_j U_j s\|_2^2 \\
= \|(P_{j+M} + \sum_{m=1}^{M} Q_{j+m}) U_j s\|_2^2 = \|P_{j+M} U_j s\|_2^2 + \sum_{m=1}^{M} \|Q_{j+m} U_j s\|_2^2 \\
= \|U_j s_{ap}^{(M)}\|_2^2 + \sum_{m=1}^{M} \|U_j s(m)\|_2^2 = \|s_{ap}^{(M)}\|_2^2 + \sum_{m=1}^{M} \|s(m)\|_2^2.
$$

Using (3.25) we get

$$
\|s(m)\|_2^2 = \|U_j s^{(m)}\|_2^2 = \|Q_{j+m} U_j s\|_2^2 \\
= \|P_{j+m-1} U_j s\|_2^2 - \|P_{j+m} U_j s\|_2^2 \to 0 \ (m \to -\infty),
$$

with $j \in \mathbb{Z}$ fixed. \qed