Chapter 1

Wavelet Transforms and Applications

This chapter reviews some of the background related to wavelet concepts, which will be used extensively in the following chapters. This can be found in many books and papers at many different levels of exposition. Some of the standard books are [Chui, 1992, Daubechies, 1992, Mallat, 1998, Meyer, 1993 and Vetterli M and Kovacevic, 1995]. Introductory papers include [Graps, 1995, Strang, 1994 and Vidakovic, 1994], and more technical ones are [Cohen and Kovacevic, 1996, Mallat, 1996 and Strang and Strela, 1995]. Different applications of wavelets in a wide variety of fields are also discussed in this chapter.

1.1 What makes wavelets useful in signal processing?

In signal processing, the representation of signals plays a fundamental role. David Marr elaborated in [Marr, 1982] on this topic. Most of the signals in practice are time domain signals in their raw format. This representation is not always the best representation of the signal for most of the signal processing related applications. In many cases, the useful information is hidden in the frequency content (spectral components) of the signal. Few of the frequency transforms are Fourier transform, Hilbert transform, Short time Fourier transform (STFT), Wigner distributions and Wavelet Transform.

1.1.1 Fourier Transform: Fourier transform (FT) of a given time domain signal \( x(t) \) is given below

\[
X_{FT}(\omega) = \int_{-\infty}^{\infty} x(t) \cdot e^{-j\omega t} \cdot dt
\]

(1.1)

Where \( \omega \) and \( t \) stands for angular frequency and time respectively and exponential term \( e^{j\omega t} \) is a basis function. This exponential term can be written as
Hence in FT basis functions are Sine and Cosine functions, which are having infinite support (length). FT gives the frequency information of the signal, which means that it tells us about the different frequency component present in the signal, but it fails to give information regarding where in time those spectral components appear because FT is computed by multiplying the signal $x(t)$ with an exponential term with infinite support, at some frequency $f$, and then integrated over all times (Eq.1.1). In FT there is no frequency resolution problem in the frequency domain, i.e. we know exactly what frequencies exist; similarly there is no time resolution problem in time domain, since we know the value of the signal at every instant of time. Conversely, the time resolution in the frequency domain, and the frequency resolution in time domain are zero, since there is no information about them.

This makes Fourier representation is suitable for analyzing stationary signals where spectral components do not change with time but, inadequate when it comes to analyzing transient or non-stationary signals i.e. signals with time varying spectra. In signal and image processing, concentrating on transients (like, e.g., image discontinuities) is a strategy for selecting the most essential information from often an overwhelming amount of data. In order to facilitate the analysis of transient signals, i.e., to localize both the frequency and the time information in a signal, numerous transforms and bases have been proposed (see e.g., [Mallat, 1998 and Vetterli M and Košević, 1995]). Among those, in signal processing the Wavelet and the Short Term Fourier Transform (STFT) or windowed Fourier transform or Gabor transform are quite standard. Let us briefly discuss about these two transforms.

### 1.1.2 Short Term Fourier Transform:

This transform is similar to Fourier transform. In STFT, the signal is divided into small portions, where these portions of the signal are assumed to be stationary. For this purpose a smooth window function $w(t)$ (typically Gaussian) is chosen. The signal is multiplied by a window function and the Fourier integral is applied to the windowed signal. For a signal $x(t)$, the STFT is [Mallat, 1998].

$$X_{STFT}(\tau, \omega) = \int_{-\infty}^{\infty} [x(t) \cdot w(t-\tau)] \cdot e^{-j\omega t} dt$$

(1.3)
CHAPTER 1: Wavelet Transforms and Applications

Where, \( e^{j\omega t} \) is Fourier transform basis function, \( \tau \) and \( \omega \) are time and frequency parameters respectively. Note that the basis functions of a STFT expansion are \( w(t) \) modulated by a sinusoidal wave and shifted in time; the modulation frequency changes while the window remains fixed. Here window is of finite length, which may not be suitable to all frequencies. If the chosen window is narrow it results in good time resolution but poor frequency resolution. Conversely, a wide window causes good frequency resolution but poor time resolution. Furthermore wide windows violate the condition of stationary.

To analyze transient signal of various supports and amplitudes in time, it is necessary to use time-frequency elements with different sizes for different time locations. For example, in the case of high frequency structures, which vary rapidly in time, we need higher time resolution to accurately trace the trajectory of the changes; on the other hand, for lower frequency, we will need a relatively higher absolute frequency to give a better measurement on the value of frequency. The next section shows that wavelet transform provide a natural representation which satisfies these requirements as illustrated in Fig (1.1).

1.1.3 Wavelet Transform: The wavelet transform (WT) is developed as an alternative approach to the STFT to overcome the resolution problem. Wavelet analysis is done in a similar way to the STFT analysis, in the sense that the signal is multiplied with a function (wavelet), similar to window function in the STFT, and the transform is computed separately for different portions of the time domain signal. However main difference between STFT and the WT is the width of the window is changed while computing transform based on signal spectral components. Hence WT results in varying resolution, i.e. higher frequencies are better resolved in time and lower frequencies are better resolved in frequency as explained below, which is probably the most significant characteristic of the wavelet transform. In addition, basis function is not completely fixed as in the case of other transforms such as Fourier and Laplace Transform etc.

1.1.4 Comparison of Time- Frequency resolutions: Figure (1.1) compares the frequency resolutions of Fourier transform, the windowed Fourier transform (STFT) and the wavelet transform. In a time series with a high resolution in the time domain each point contains information about all frequencies. Due to the convolution properties the opposite is true for the FT of the time series. In this case every point in the frequency
domain contains information from all points in the time domain. The windowed (Short Term) Fourier transform divides the time-frequency plane in rectangular boxes [Nielsen and Wickerhauser, 1996]. The resolution in time is increased at the expense of the frequency resolution. The wavelet transform overcomes this problem by scaling the basis functions relative to their support. The WT needs more time for the detection of low frequencies than for the detection of high frequencies. Using these properties of the WT, it is possible to describe an experimental signal on different frequency levels which leads to Mallat’s Multiresolution Analysis (MRA).

1.2 Historical perspective of wavelet transforms

In the history of mathematics, wavelet analysis shows many different origins [Meyer, 1993]. Much of the work was performed in the 1930s, and, at the time, the separate efforts did not appear to be parts of a coherent theory.

1.2.1 PRE-1930: Before 1930, the main branch of mathematics leading to wavelets began with Joseph Fourier with his theories of frequency analysis. Fourier’s assertion played an essential role in the evolution of the ideas mathematicians had about the functions. He
opened up the door to a new functional universe. After 1807, by exploring the meaning of functions, Fourier series convergence, and orthogonal systems, mathematician gradually were led from their previous notation of *frequency analysis* to the notion of *scale analysis*.

The first recorded mention of what we now call as "wavelet" seems to be in 1909, in a thesis by Alfred Haar. Haar wavelet is compact support wavelet (vanishes outside of a finite interval). Unfortunately, Haar wavelets are not continuously differentiable which some what limits their applications.

1.2.2 THE 1930: In the 1930s, several groups working independently researched the representation of functions using *scale-varying basis functions*. The researchers discovered a function that can vary in scale and can conserve energy when computing energy of a function. Their work provided David Mar with an effective algorithm for numerical image processing using wavelets in the early 1980s.

1.2.3 THE 1980: In 1980, Grossman and Morlet, a physicist and an engineer, broadly defined wavelets in the context of quantum physics. These two researchers provided a way of thinking for wavelets based on physics intuition.

1.2.4 POST 1980: In 1985, Stephane Mallat gave wavelets an additional jump start through his work in digital signal processing. He discovered some relationships between quadrature mirror filters, pyramid algorithms, and orthogonal wavelet bases. Inspired in part by these results, Meyer Y constructed the first non-trivial wavelets. Unlike Haar wavelets, Meyer wavelets are continuously differentiable; however they don’t have compact support. A couple of years later, Ingrid Daubechies used Mallat’s work to construct a set of wavelet orthonormal basis functions that are perhaps the most elegant, and have become the cornerstone of wavelet application.

1.3 Wavelet Representation

Unlike the Fourier transform, whose basis functions are sinusoids, wavelet transforms are based on small waves i.e. wavelets, which are small, localized waves of a particular shape having an average value of zero as shown in Fig (1.2). Note the support of wavelet basis is finite in time whereas the Fourier basis oscillates forever. This allows
wavelets to provide both spatial/time and frequency information (hence time frequency analysis); whereas the non-local Fourier transform gives only frequency information.

**Figure 1.2**: An example of the basis functions of Fourier transforms and Wavelet transforms.

The central idea in wavelet transform is to analyze a signal according to scale. Wavelet analysis means breaking up a signal into scaled and translated versions of wavelet (mother wavelet). One chooses a particular wavelet, stretches it (to meet a given scale) and shifts it, while computing wavelet transform. The Fig (1.3) below shows a signal \( x(t) \) along with the Morlet wavelet at three scales and shifts.

**Figure 1.3**: An example of wavelet representation. A signal \( x(t) \) along with the Morlet wavelet at three scales and shifts (\( w(2(t+9)), w(t), \) and \( w((t-9)/2) \)). [Rao and Bopadikar, 1998].
Mallat [1989] first showed that the wavelet transform provides the foundation of a powerful new approach to signal processing and analysis, called multiresolution analysis (MRA). MRA theory unifies techniques from several fields, including sub-band coding from signal processing [Woods and O'Neil, 1986], Quadrature mirror filtering (QMF) from speech recognition [Vaidyanathan and Hoang, 1988], and pyramidal image processing [Burt and Adelson, 1983]. Due to the close links to these techniques, the wavelet transform has found many applications [Basseville et al., 1992, Brailean and Katsaggelos, 1995, Chang et al., 2000 and Shapiro, 1993].

1.4 Continuous Wavelet Transform

The continuous wavelet transform (CWT) of a signal \( x(t) \in L^2 \) (space of all square integrable functions) is defined as

\[
W_x(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi^* \left( \frac{t-b}{a} \right) dt = \int_{-\infty}^{\infty} x(t) \psi^* (a,b) dt
\]

Where, \( \psi^*(t) \) denotes the complex conjugate of \( \psi(t) \). As seen in the above Eq. (1.4), the transformed signal is a function of two variables, \( b \) and \( a \), the translation/localization and scale parameters, respectively. \( \psi(t) \) is the transforming function, and it is called the mother wavelet which is prototype for generating the other window functions (baby wavelets) [Rao and Bopadikar, 1998].

\[
\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi \left( \frac{t-b}{a} \right)
\]

Here, \( 1/\sqrt{a} \) is the normalization factor to ensure that all wavelets have the same energy. The term translation is related to the location of a wavelet, as the wavelet is shifted through the signal, which corresponds to the time information in the transform domain. The variation of \( a \) has a dilation effect (when \( a>1 \)) and a contraction effect (when \( a<1 \)) of the mother wavelet. Therefore, it is possible to analyze the long and short period features of the signal or the low and high frequency aspects of the signal.

This transform is called as continuous wavelet transform (CWT), because the scale and localization parameters assume continuous values.
1.4.1 CWT as a correlation: Given two finite signals \( f(t) \) and \( g(t) \in L^2 \), their inner product is given by

\[
\langle f(t), g(t) \rangle = \int f(t)g^*(t)\,dt
\]  

Equation (1.6) leads to

\[
W_x(a,b) = \langle x(t), \psi_{a,b}(t) \rangle
\]  

The cross correlation \( R_{x,y} \) of the two functions \( x(t) \) and \( y(t) \) is defined as:

\[
R_{x,y}(\tau) = \int x(t)y^*(t-\tau)\,dt = \langle x(t), y(t-\tau) \rangle
\]  

Then

\[
W_x(a,b) = \langle x(t), \psi_{a,0}(t-b) \rangle = R_{x,\psi_{a,b}}(b)
\]  

Thus \( W_x(a,b) \) is the cross correlation of the signal \( x(t) \) with the mother wavelet at scale \( a \) and translation \( b \). If \( x(t) \) is similar to the mother wavelet at this scale and translation, then \( W_x(a,b) \) will be large.

1.4.2 Filtering Interpretation: The CWT offers both the time and frequency selectivity. The translating effect will result the time selectivity of the CWT and frequency selectivity is achieved with collection of linear, time variant filters with impulse responses that are dilations of the mother wavelet reflected about the time-axis.

This can be explained from the convolution, which is given as:

\[
h(t) \ast x(t) = \int h(t-\tau)x(\tau)d\tau
\]  

Then

\[
W_x(a,b) = x(b) \ast \psi_{a,0}^*(-b)
\]  

For any given scale \( a \) (frequency \( \sim 1/a \)), the CWT \( W_x(a,b) \) is the output of the filter with the impulse response \( \psi_{a,0}^*(-b) \) to the input \( x(b) \), and i.e. we have a continuum of filters, parameterized by the scale factor \( a \). These filters defined by the mother wavelet are of constant Q, which is the ratio of center frequency to the bandwidth. The Fig (1.4) shows the filters defined by the CWT for different values of \( a \).
The inverse CWT is defined as

\[ x(t) = \frac{1}{C_\psi} \left( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_\psi(a,b) \psi_{a,b}(t) \, db \, da \right) \, a^2 \]  

(1.12)

When using a transform in order to get better insight into the properties of a signal, it should be ensured that the signal can be perfectly reconstructed from its representation. Otherwise the representation may be completely or partly meaningless. For the wavelet transform the condition that must be met in order to ensure perfect reconstruction is [Rao and Bopadikar, 1998]

\[ C_\psi = \int_{-\infty}^{\infty} \left| \psi(\omega) \right|^2 \left| \omega \right|^{-1/2} \, d\omega < +\infty \]  

(1.13)

Where, \( \psi(\omega) \) denotes the Fourier transform of the wavelet. This condition is known as the \textit{admissibility condition} for the wavelet \( \psi(t) \). Obviously, in order to satisfy (1.13) the wavelet must satisfy

\[ \psi(0) = \int_{-\infty}^{\infty} \psi(t) \, dt = 0 \]  

(1.14)

Moreover, \( |\psi(\omega)| \) must decrease rapidly for \( |\psi(\omega)| \to 0 \) and for \( |\omega| \to \pi \). That is, \( \psi(t) \) must be a band pass impulse response. Since a band pass impulse response looks like a small wave, the transform is named wavelet transform.

The CWT is highly \textit{redundant}, and is \textit{shift invariant}. It is extensively used for the characterization of signals [Mallat and Hwang, 1992]: the evolution of the CWT magnitude across scales provides information about the local \textit{regularity} of a signal. [Burrus and Gopinath, 1998]
1.5 Discrete wavelets transform

The CWT provides a redundant representation of the signal in the sense that the entire information of $W_x(a, b)$ need not be used to recover the input signal $x(t)$. This redundancy, on the other hand, requires a significant amount of computation time and resources. By sampling $a$ and $b$ on a dyadic grid, a type of non-redundant wavelet representation is developed which is called as Discrete Wavelet Transform. This type of representation is also arises in the context of Multiresolution Resolution Analysis (MRA). DWT is considerably easier to implement when compared to the CWT. The basic concepts of the DWT will be introduced in the following sections along with its properties and algorithms used to implement it.

1.6 Scaling Functions

In order to understand MRA we start by defining the scaling function and then define wavelet in terms of it. Set of scaling functions in terms of integer translates of the basic scaling function is defined as

$$\phi_k(t) = \phi(t-k), \quad k \in \mathbb{Z}, \quad \phi \in L^2$$

(1.15)

The subspace of $L^2(\mathbb{R})$ spanned by these functions is defined as

$$v_0 = \text{span}\{\phi_k(t)\} \quad \text{for all integers} \ k \text{ from } -\infty \text{ to } +\infty$$

(1.16)

This means that

$$f(t) = \sum_k a_k \phi_k(t) \quad \text{for any} \quad f(t) \in v_0$$

(1.17)

A two-dimensional family of functions is generated from the basic scaling function by scaling and translation as

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

(1.18)

Whose span over $k$ is,

$$v_j = \text{span}\{\phi_k(2^j t)\} = \text{span}\{\phi_{j,k}(t)\} \quad \text{for all integers} k \in \mathbb{Z}$$

(1.19)

This means that if $f(t) \in v_j$ then it can be expressed as
\[ f(t) = \sum_k a_k \varphi(2^j t + k) \]  

(1.20)

For \( j > 0 \), the span can be larger since \( \varphi_{j,k}(t) \) is narrow and is translated in smaller steps, therefore can represent finer detail. For \( j < 0 \), \( \varphi_{j,k}(t) \) is wider and translated in large steps. So these wider scaling functions can represent only coarse information, and the space they span is smaller.

1.7 Multiresolution Analysis (MRA)

MRA, as implied by its name, analyzes the signal at different frequencies with different resolutions. MRA is designed to give good time resolution at high frequencies and good frequency resolution at low frequencies. This approach makes sense especially when the signal at hand has high frequency components for short duration and low frequency components for long duration. Fortunately, the signals that are encountered in practical applications are often of this type.

An MRA consists of the nested linear vector spaces with

\[ v_j \subseteq v_{j+1} \quad \text{for all } j \in \mathbb{Z} \]  

(1.21)

Or

\[ \{0\} \subseteq \ldots \subseteq v_{-2} \subseteq v_{-1} \subseteq v_0 \subseteq v_1 \subseteq v_2 \subseteq \ldots \subseteq L^2 \]  

(1.22)

\[ v_{-\infty} = \{0\}, \quad v_{\infty} = L^2 \]  

(1.23)

The space that contains high resolution signals will contain those of lower resolution also.

\[ f(t) \in v_j \quad \iff \quad f(2t) \in v_{j+1} \]  

(1.24)

Above Eq. (1.24) indicates elements in a space are simply scaled versions of the elements in the next space. The relationship of the spanned spaces is illustrated in Fig (1.5)
If \( \phi(t) \) is in \( v_0 \), it is also in \( v_1 \), the space spanned by \( \phi(2t) \). This means \( \phi(t) \) can be expressed in terms of a weighted sum of shifted \( \phi(2t) \) as \cite{Burrus1998}

\[
\phi(t) = \sum_{n} h_0(n) \sqrt{2} \phi(2t - n), \quad n \in \mathbb{Z}
\] (1.25)

Where, the coefficients \( h_0(n) \) are a sequence of real/complex numbers called the scaling filter coefficients (or scaling filter) and \( \sqrt{2} \) maintains the norm of the scaling function with the scale of two. It is called the refinement equation, the MRA equation, recursion equation and dilation equation.

### 1.8 The Wavelet functions

The important features of signal can be described by a set of wavelet functions \( \psi_{j,l}(t) \), which spans the difference between the spaces spanned by the various scales of the scaling functions.

There are several advantages to requiring that the scaling functions and wavelets be orthogonal, orthogonal basis allow simple calculation of expansion coefficients. The orthogonal component of \( v_j \) in \( v_{j+1} \) is \( w_j \). This means that all members of \( v_j \) are orthogonal to all members of \( w_j \).

\[
\langle \phi_{j,k}(t), \psi_{j,l}(t) \rangle = \int \phi_{j,k}(t) \psi_{j,l}(t) dt = 0 \quad \text{for} \quad j,k,l \in \mathbb{Z}
\] (1.26)

The relationship of the various subspaces can be seen from the following expressions. We start with \( j=0 \), (1.17) becomes

\[
v_0 \subset v_1 \subset v_2 \subset \ldots \subset L^2
\] (1.27)
Wavelet spanned subspace $w_0$ can be defined as

$$v_1 = v_0 \oplus w_0$$  \hspace{1cm} (1.28)

This extends to

$$v_2 = v_0 \oplus w_0 \oplus w_1$$ \hspace{1cm} (1.29)

In general this gives

$$L^2 = v_0 \oplus w_0 \oplus w_1 \oplus \ldots$$ \hspace{1cm} (1.30)

When $v_{0,0}$ the initial space spanned by the scaling function $\phi(t-k)$.

Figure (1.5) pictorially shows relationship of scaling functions and wavelet functions for different scales $j$. The scale of the initial space is arbitrary and could be chosen at a higher resolution (say, $j=10$) or at a lower resolution (say, $j=-10$).

At $j=-\infty$ Eq. (1.30) becomes

$$L^2 = \ldots \oplus w_{-2} \oplus w_{-1} \oplus w_0 \oplus w_1 \oplus w_2 \oplus \ldots$$  \hspace{1cm} (1.31)

Above Eq. (1.31) shows that wavelet functions completely describe the original signal i.e.

$$g(t) = \sum_{j,k} C_{j,k} \psi_{j,k}(t)$$  \hspace{1cm} (1.32)

Another way to describe the relation of $v_0$ to the wavelet space is

$$w_{-\infty} \oplus \ldots \oplus w_{-1} = v_0$$ \hspace{1cm} (1.33)

This again shows that the scale of scaling space is arbitrarily. In practice, it is usually chosen to represent the coarsest detail of interest in a signal.

Since these wavelets reside in the space spanned by the next narrower scaling function, i.e. $w_0 \subset v_1$, they can be represented by a weighted sum of scaling function $\phi(2t)$

$$\psi(t) = \sum_{n} h_1(n) \sqrt{2} \phi(2t-n), \hspace{1cm} n \in Z$$  \hspace{1cm} (1.34)

Where, the coefficients $h_1(n)$ are a sequence of real/complex numbers called the wavelet filter coefficients (or wavelet filter). The wavelet coefficients are required by Orthogonality to be related to the scaling function coefficients $h_0(n)$ by
For finite even length-N \( h_0(n) \) Eq. (1.35) becomes

\[
h_1(n) = (-1)^n h_0(1 - n)
\]

The function generated by Eq. (1.34) gives the prototype or mother wavelet \( \psi(t) \) for a class of expansion functions of the form

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

Where, \( 2^j \) is the scaling of \( t \) (\( j \) is \( \log_2 \) of the scale), \( 2^j k \) is the translation in \( t \), and \( 2^{j/2} \) maintains the \( L^2 \) norm of the wavelet at different scales.

According to Eq. (1.30) any function \( g(t) \in L^2(R) \) could be written as a series expansion in terms of the scaling function and wavelets.

\[
g(t) = \sum_{k=-\infty}^{\infty} c(k) \phi_k(t) + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} d(j,k) \psi_{j,k}(t)
\]

In this expansion, the first summation in Eq. (1.38) gives a function that is a low resolution or coarse approximation of \( g(t) \) where as second summation adds higher or finer resolution as index \( g(t) \) increases, which adds increasing detail. This is analogous to Fourier series where the higher frequency terms contain the detail of the signal. [Burrus et al., 1998]

### 1.9 Signal Representation in DWT

The discrete wavelet transform (DWT) is the signal expansion into (bi)-orthogonal wavelet basis. In this transform scale is sampled at dyadic steps \( a \in \{2^j: j \in \mathbb{Z}\} \), and the position is sampled proportionally to the scale \( b \in \{k2^j: j, k \in \mathbb{Z}\} \). Since

\[
L^2 = \oplus_{j_0} w_{j_0} \oplus w_{j_0+1} \oplus \ldots
\]

Using equations (1.18) and (1.37), a more general statement of the expansion Eq. (1.38) can be given by

\[
g(t) = \sum_{k} c_{j_0}(k) \phi_{j_0,k}(k) + \sum_{j} \sum_{k} d_j(k) \psi_{j,k}(k)
\]

Or

\[
g(t) = \sum_{k} c_{j_0}(k) \phi_{j_0,k}(k) + \sum_{j} \sum_{k} d_j(k) \psi_{j,k}(k)
\]
Where, \( j_0 \) could be arbitrarily chosen in between \(-\infty\) and \( \infty \), the choice of \( j_0 \) sets the coarsest scale whose space is spanned by \( \varphi_{j_0, \alpha}(t) \). The rest of \( L^2(\mathbb{R}) \) is spanned by the wavelets which provide the high resolution details of the signal.

The set of expansion coefficients \( c_{j_0}(k) \) and \( d_{j}(k) \) are called DWT of \( g(t) \). If the wavelet system is orthogonal, these coefficients can be calculated by inner products.

\[
\begin{align*}
    c_j(k) &= \left\langle g(t), \varphi_{j,k}(t) \right\rangle = \int g(t) \varphi_{j,k}(t) \, dt \quad (1.42) \\
    d_j(k) &= \left\langle g(t), \psi_{j,k}(t) \right\rangle = \int g(t) \psi_{j,k}(t) \, dt \quad (1.43)
\end{align*}
\]

Even for the worst case signal, the wavelet expansion coefficients drop off rapidly as \( j \) and \( k \) increase. This is why the DWT is efficient for signal and image compression.

The DWT is similar to a Fourier series but, in much more flexible and informative. It can be made periodic like a Fourier series to present periodic signals efficiently. However, unlike a Fourier series, it can be used directly on non-periodic transient signals with excellent results.

### 1.10 Parseval’s theorem

If the scaling and wavelet functions form an orthogonal basis Parseval’s theorem relates the energy of the signal \( g(t) \) to the energy in each of the components and wavelet coefficients. This is the one reason why orthogonality is important.

For the general wavelet expansion of (1.38) or (1.41), Parseval’s theorem

\[
\int |g(t)|^2 \, dt = \sum_{j=-\infty}^{\infty} |c_j(L)|^2 + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} |d_{j,k}(k)|^2 \quad (1.44)
\]

### 1.11 Filter banks and the DWT

In practical applications we never deal with scaling functions and wavelets. Only the coefficients \( h_0(n) \) and \( h_1(n) \) in the dilation equations (1.25) and (1.34) and \( c_{j_0}(k) \) and \( d_{j}(k) \) in expansion equations (1.38), (1.42) and (1.43) need to be considered and they can be viewed as digital filters and digital signals respectively [Gopinath and Burrus, 1992 and Vaidyanathan, 1992] While it is possible to develop most of the results of wavelet theory using filter banks only.
1.11.1 Wavelet Analysis by Multirate Filtering: The relationship between the expansion coefficients at a lower scale level in terms of higher scale is given by following equations. [Burrus et al., 1998]

Relationship for scaling coefficients

\[ c_j(k) = \sum_m h_0(m-2k)c_{j+1}(m) \]  

(1.45)

Corresponding relationship for the wavelet coefficients is

\[ d_j(k) = \sum_m h_1(m-2k)d_{j+1}(m) \]  

(1.46)

These equations show that the scaling (Approximation) and wavelet (detail) coefficients at different levels of scale can be obtained by convolving the expansion coefficients at scale \( j \) by the time reversed recursion coefficients \( h_0(-n) \) and \( h_1(-n) \) then down sampling or decimating (taking every other term, the even terms) to give the expansion coefficients at the next level of \( j-1 \). These structures implement Mallat’s algorithm

The implementation of equations (1.45) and (1.46) is shown in Fig (1.6) Where down pointing arrows is indicate down sampling and left two boxes denote FIR (finite impulse response) filtering or a convolution by \( h_0(-n) \) or \( h_1(-n) \).

Here FIR filter implemented by \( h_0 \) is a low pass filter and the one implemented by \( h_1 \) is a high pass filter. The average number of data points out of this system is the same as the number in. The number is doubled by having two filters, and then it is halved by decimation back to the original number. This means no information has been lost in this system and hence perfect reconstruction is possible. The aliasing occurring in the upper bank can be cancelled by the lower bank. This is the idea behind the perfect reconstruction in filter bank theory [Vaidyanathan, 1992 and Fliege, 1994].
This process (splitting, filtering, and decimation) is called as \textit{wavelet decomposition}. We can repeat this \textit{wavelet decomposition} on scaling coefficients for iterating the filter bank. Figures (1.7 and 1.8) show three stages Two-Band Analysis tree and Frequency Bands for the Analysis tree.

The first stage of two banks divide the spectrum of $x[n]$ into two equal bands (low pass and high pass), resulting in scaling (approximation) coefficients and wavelet (detailed) coefficients at lower scale. The second stage then divides that low pass band into another low pass and high pass bands and so on. This results in a logarithmic set of bandwidths as illustrated in Fig (1.8) These are called “constant-Q” filters (CQF) in filter bank language because the ratio of the band width to the center frequency of the band is constant. We can observe that down sampling assures that reconstructed signal has the same number of samples as the original one.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{three_stage_two_band_analysis_tree.png}
\caption{Three stage Two-Band Analysis Tree}
\end{figure}
Figure 1. 8: Frequency Bands for the Analysis tree

Figure 1. 9: DWT of a chirp signal, notice how location in $k$ traces the frequencies in the signal in a way the Fourier transform can not. [Matlab, 2004]
Figure (1.9) shows original signal (chirp signal, which has a time-varying frequency) and corresponding approximation and detailed coefficients. Notice how location in \( k \) traces the frequencies in the signal in a way the Fourier transform can not. This suggests that wavelet transform is well suited for time-frequency analysis.

1.11.2 Wavelet Synthesis by Multirate Filtering: A reconstruction of the original fine scale coefficients of the original signal can be made from a combination of the scaling function (approximation) and wavelet coefficients (detail) coefficients at a coarse resolution.

We start with two sets of coefficients \( c_j(k) \) and \( d_j(k) \) at scale index \( j \) and produce the coefficients at scale index \( j+1 \). We have following relation, [Burrus et al., 1998]

\[
c_{j+1}(k) = \sum_m c_j(m)h_0(k - 2m) + \sum_m d_j(m)h_1(k - 2m)
\] (1.47)

This equation can be evaluated by up-sampling the \( j \) scale coefficient sequence \( c_j(k) \) (which means double its length by inserting zeros between each term), then convolving with the scaling filter coefficients \( h_0(n) \). The same is done to the \( j \) level wavelet sequence \( d_j(k) \) and the results are added to give the \( j+1 \) level scaling coefficients. This structure is show in Fig (1.10); this combining process can be extended to any level by combining the appropriate scale wavelet coefficients. The resulting two-scale tree is show in Fig (1.11)

![Figure 1.10: One stage Two-Band Synthesis Bank](image-url)
1.12 Wavelet Families

There are number of standard basis functions which can be used as a mother wavelet functions in wavelet transforms. Figure (1.12) illustrates some of the commonly used wavelet functions. Haar wavelet is the one of the oldest and simplest wavelet.

Therefore, any discussion of wavelets starts with the Haar wavelet. Daubechies wavelets are the most popular wavelets. The Haar, Daubechies, Symlets and Coiflets are compactly supported orthogonal wavelets. The Meyer, Morlet and Mexican Hat wavelets are symmetric in shape.
1.13 Properties of MRA filters, scaling and wavelet functions

1. Orthogonality of Wavelets: The baby wavelets are orthogonal and have unit energy

\[ \int_{-\infty}^{\infty} \psi_{j,k}(t) \psi_{m,n}(t) \, dt = \delta(j - m) \delta(k - n) \quad (1.48) \]

2. Orthogonality of scale: The translates \( \phi(t-k) \) \(-\infty < k < \infty \) are orthogonal and have unit energy \( \langle \phi(t), \phi(t) \rangle \)

\[ \int_{-\infty}^{\infty} \phi(t) \phi(t-n) \, dt = \delta(n) \quad (1.49) \]

3. Completeness: The translates \( \phi(t-k) \) \(-\infty < k < \infty \), span the same space as wavelets

\[ \omega_{j,k}(t), -\infty < j < 0 \quad \text{and} \quad -\infty < k < \infty \quad (1.50) \]

4. Double Shift Orthogonality of the Filters

\[ \int_{-\infty}^{\infty} \phi(t) \phi(t-n) \, dt = \delta(n) = \sum_k h_0(k) h_0(k-2n) \quad (1.51) \]

\[ \int_{-\infty}^{\infty} \omega(t) \omega(t-n) \, dt = \delta(n) = \sum_k h_1(k) h_1(k-2n) \quad (1.52) \]

These equations (1.51-1.52) are called the double shift orthogonality relations of the filters. They lead to a number of other properties of the filters. They are

a) If we take \( n=0 \)

\[ \sum_k h_0(k)^2 = 1 \quad \text{and} \quad \sum_k h_1(k)^2 = 1 \quad (1.54) \]

b) If we integrate both sides of the equation we will get

\[ \sum_k h_0(k) = \sqrt{2} \quad \text{and} \quad \sum_k h_1(k) = 0 \quad (1.55) \]

c) The even and odd terms of both filters are

\[ \sum_k h_0(2k) = \sum_k h_0(2k+1) = \frac{1}{\sqrt{2}} \quad (1.56) \]

\[ \sum_k h_1(2k) = \pm \sum_k h_1(2k+1) = \frac{1}{\sqrt{2}} \quad (1.57) \]
5. Support of the scaling function: The support of a function $\phi(t)$ is the range of $t$ where the function is non-zero. The recursion equation imposes a restriction on the support of the scaling function. If $N+1$ is the length of low pass filter $h_0(n)$, then the support of $\phi(t)$ is the in the interval $0 \leq t \leq N$.

1.14 Applications of Wavelet Transforms

There is a wide range of applications for wavelet transforms in different fields ranging from signal processing to biometrics, and the list is still growing.

One of the prominent applications is Data compression [Rao and Bopadikar, 1998]. Apart from its original intention of analyzing non-stationary signals; wavelets have been most successful in image processing and compression applications. Due to the compact support of the basis functions used in wavelet analysis, wavelets have good energy concentration properties. In DWT, the most prominent information in the signal appears in high amplitudes and the less prominent information appears in very low amplitudes. Data compression can be achieved by discarding these low amplitudes. The wavelet transforms enables high compression ratios with good quality of reconstruction. Recently, the wavelet transforms have been chosen for the JPEG 2000 compression standard [Marcellin et al., 2000 and Rabbani and Joshi, 2002].

Compression property has been further explored by Iain Jonstone and David Donoho [Jonstone and Donoho, 1995a] and they have devised the wavelet shrinkage denoising (WSD). The idea behind WSD is based on recognizing the noise level, which will show itself at finer scales, and discarding the coefficient that fall below a certain threshold at these scales will remove the noise.

Wavelets also find applications in speech processing, which reduces transmission time in mobile applications. They are used in edge detection, feature extractions, speech recognition, echo cancellation, and others. They are promising for real time audio and video compression applications. Wavelets have numerous applications in digital communications, study of distant universes [Bijaoui et al., 1996], fractal analysis, turbulence analysis [Meyer, 1993] and financial analysis.

Wavelets have often been employed to analyze wind disturbances such as gravity waves [Shimomai et al., 1996] and to remove ground and intermittent clutters, such as due to airplane echoes, in the atmospheric radar data [Jordon et al., 1997, Boisse et al., 1999 and Lehmann & Teschke, 2001] using standard wavelets such as Daubechies 20.
Wavelets place important role in Biomedical Engineering owing to the nature of all biological signals being non-stationary. Further wavelets are useful the analysis of ECG (electro cardiogram) for diagnosing cardiovascular disorders and of electro encephalogram (EEG) for diagnosing neurophysiologic disorders, such as seizure detection, or analysis of evoked potentials for detection of Alzheimer’s disease [Polikar et al., 1997]. Wavelets have also been used for the detection of micro calcifications in mammograms and processing of computer tomography [CT] and magnetic resonance image [MRI]. The popularity of wavelet transforms is growing because of its ability to reduce distortion in the reconstructed signal while retaining all the significant features present in the signal.

As mentioned above, though elegant and powerful wavelet based tools are being applied in number of areas, their application to radar signal processing has been rather limited. Considering the vastness of the area of radar signal processing it appears that wavelet base techniques haven’t been applied to their full potential in this area. The main objective of the this work is to explore wavelet transform based signal processing to atmospheric radar i.e. MST radar with a view to extract Doppler spectra from the noisy data with improved signal to noise ratio to extend height coverage and improve the accuracy of the parameters extracted from the spectra.