Lecture 21

Wavelets and multiresolution analysis

Introduction

Relevant section of text by Boggess and Narcowich: 4.1 and 4.2

We begin with a motivation for what is called the multiresolution analysis of functions, with a particular eye toward applications in signal and image processing.

Consider a function \( f \in L^2(\mathbb{R}) \) which will, once again, represent a signal of interest. (Later, we may wish to restrict the support to a finite interval, say, \([0,1]\).) Now suppose that we input this function into a machine that “scans” \( f(t) \) with a sensor to produce approximations to it with a given “refinement” or “resolution.” One possibility – an idealized one – is that the sensor produces disjoint, piecewise constant approximations to \( f(t) \) over contiguous intervals on \( \mathbb{R} \). The coarser the sensor, the longer the “pieces” that comprise the approximation to \( f(t) \). Or, the other way around, the finer the sensor, the shorter the pieces that comprise the approximation to \( f(t) \).

This is certainly an idealization – no sensor can do this. There will either be some kind of overlap between intervals or some loss of information near the endpoints of each interval. For the moment, we neglect this difficulty and consider piecewise constant approximations.

Of course, there are many ways to approximate a function \( g(t) \) over an interval \([a,b]\). In what follows, we shall assume that the sensor yields the best constant approximation in the \( L^2 \) sense, i.e., the mean value of \( g \) over \([a,b]\):

\[
\overline{g}_{[a,b]} = \frac{1}{b-a} \int_a^b g(t) \, dt.
\]  

From a generalized Fourier series perspective, this is equivalent to projecting \( g \) onto the following orthonormal basis element of \( L^2[a,b] \):

\[
\hat{e}_1(t) = \frac{1}{\sqrt{b-a}},
\]  

i.e.,

\[
g(t) \approx \overline{g}_{[a,b]} = \langle g, \hat{e}_1 \rangle \hat{e}_1.
\]  

We shall consider a particular family of such refinements, namely, piecewise-constant approximations \( f_k \) to functions \( f \in L^2(\mathbb{R}) \) over intervals of length \( 2^{-k} \). For example:
1. \( f_0(t) \): approximation to \( f(t) \) by piecewise constant functions over intervals of integer length, i.e., \([l, l + 1], l \in \mathbb{Z}\)

2. \( f_1(t) \): approximation to \( f(t) \) by piecewise constant functions over intervals of half-integer length, i.e., \([m/2, (m + 1)/2], m \in \mathbb{Z}\).

These two approximations to a function \( f(t) \) are presented in the figure below.

Piecewise-constant approximations \( f_0(t) \) and \( f_1(t) \) to \( f(t) \).

The “detail” \( f_d(t) \) between \( f_0(t) \) and \( f_1(t) \).

Clearly, \( f_1(t) \) is a more refined approximation to \( f(t) \) than \( f_0(t) \) is. (And, in general, if \( m < n \), then \( f_n(t) \) is a more refined approximation to \( f(t) \) than \( f_m(t) \) is.) One may ask the following questions, which are not unrelated:

1. How are \( f_0 \) and \( f_1 \) related?

2. Can we characterize the additional information contained in \( f_1 \) over that in \( f_0 \)?
One attempt to answer the above questions is to study the difference between the two approximations, i.e.,

$$f_d(t) = f_1(t) - f_0(t).$$

(4)

$f_d(t)$ is what we have to add to the lower-resolution approximation $f_0(t)$ in order to obtain the higher-resolution approximation $f_1(t)$, i.e.,

$$f_1(t) = f_0(t) + f_d(t).$$

(5)

As we’ll discuss below, $f_d(t)$ is known as the “detail” – the information that is contained in $f_1(t)$ which is not contained in $f_0(t)$.

A plot of $f_d(t)$, shown above, is indeed interesting. The graph of $f_d(t)$ is piecewise constant over half-intervals, as expected. But what may not have been expected is the fact that the two pieces of the graph of $f_d(t)$ over each interval $[k, k + 1]$, $k \in \mathbb{Z}$, are symmetrically places above and below the $x$-axis. In fact, each “piece” of the graph of $f_d(t)$ over the interval $[k, k + 1]$ can be expressed as an appropriate multiple of $\psi(t - k)$, where the detail or wavelet function $\psi(t)$ is defined as

$$\psi(t) = \begin{cases} 
1, & 0 \leq t < 1/2, \\
-1, & 1/2 \leq t < 1, \\
0, & t \notin [0,1].
\end{cases}$$

(6)

The graph of $\psi(t)$ is sketched below.

The “detail” or “wavelet” basis function $\psi(t)$.

A little extra work will show that this should, in fact, be the case: The average value of $f(t)$ – which is $f_0(t)$ – over each interval $I_k = [k, k + 1]$ should lie between the average values of $f(t)$ over the two half-intervals of $I_k$, the two values of $f_1(t)$ over $I_k$. By construction, this implies that the mean of $f_d(t)$ over each interval $I_k$ should be zero.
Such questions are important in the subband coding of signals and their progressive transmission. It is often desired to have a set of coefficients that represent various resolutions of a signal, in such a way that one needs only to add higher resolution coefficients to the already existing lower resolution ones. The generalized Fourier series explored earlier in this course is an example of such a system. However, Fourier series are generally defined over entire intervals of support of a function. Wavelets generally have more localized support.

A more mathematical analysis

We now analyze mathematically the relationship between the approximations \( f_0 \) and \( f_1 \) – and, in general, \( f_k \) and \( f_{k+1} \).

Let us introduce the following space of functions:

\[
V_0 = \{ f \in L^2(\mathbb{R}) : f(t) \text{ is constant over the interval } [k,k+1), \forall k \in \mathbb{Z} \}. \tag{7}
\]

It can be shown (an exercise in advanced analysis) that \( V_0 \) is a closed, linear subspace of \( L^2(\mathbb{R}) \). Furthermore, the following countably-infinite set of functions spans \( V_0 \):

\[
\phi_{0k}(t) = I_{[k,k+1)}(t), \tag{8}
\]

where, for convenience, we have introduced the indicator function of a set \( S \subset \mathbb{R} \) as

\[
I_S(t) = \begin{cases} 
1, & t \in S, \\
0, & t \notin S. 
\end{cases} \tag{9}
\]

Note: The first subscript of \( \phi_{0k}(t) \), i.e., “0”, refers to the resolution level, i.e., \( V_0 \). The second subscript, “\( k \)”, actually enumerates the basis functions. Some of these functions are plotted below. It should not be difficult to see that the functions \( \phi_{0k}(t) \) form an orthonormal set on \( V_0 \), i.e.,

\[
\langle \phi_k, \phi_l \rangle = \delta_{kl}, \tag{10}
\]

since they do not overlap with each other unless \( k = l \). We may write that

\[
V_0 = \text{span}\{ \phi_{0k}(t), k \in \mathbb{Z} \} \cap L^2(\mathbb{R}). \tag{11}
\]

(technically, there should be a bar over the “span” to indicate “closure” of the set, but we’ll omit this detail.) Note the intersection with the space \( L^2(\mathbb{R}) \): We must ensure that functions in \( V_0 \) are also in \( L^2(\mathbb{R}) \), especially if we are going to measure distances between them.
Some orthonormal basis functions $\phi_{0k}(t)$ of the space $V_0$.

**Another note:** We have just developed something that you have not seen before: a set of basis functions that are *not* supported on an entire interval of interest, e.g., sin and cos functions, but rather on subintervals. There is a price to pay for this, since we have to deal with many more functions in our space. But the benefit of this construction is that the basis functions are *localized*: We can now operate on selected portions of a signal if desired. This was not possible with Fourier basis functions.

Back to our space $V_0$: From the above construction, if $u(t) \in V_0$, then

$$u(t) = \sum_{k \in \mathbb{Z}} c_{0k} \phi_{0k},$$  

(12)

where

$$c_{0k} = \langle f, \phi_{0k} \rangle,$$

(13)

since the $\phi_{0k}$ are orthonormal.

Now, it should be clear that the approximation $f_0(t)$ to $f(t)$ defined earlier belongs to $V_0$. But recall the way in which we constructed $f_0(t)$, with reference to Eq. (3): The constants $c_{0k}$ are the average values of $f(t)$ over the intervals $[k, k+1)$ on which are supported the functions $\phi_{0k}$. Therefore,

$$c_{0k} = \langle f, \phi_{0k} \rangle.$$

(14)

In other words,

$$f_0(t) = \sum_{k \in \mathbb{Z}} \langle f, \phi_{0k} \rangle \phi_{0k}.$$  

(15)

We have arrived at the following important result:
There is another very important point. It is quite obvious that the functions $\phi_{0k}(t)$ are translations of each other. We’ll use $\phi_{00}(t)$ as a reference and call it $\phi(t)$. Then

$$\phi_{0k}(t) = \phi(t - k).$$

We may then define the space $V_0$ as follows:

$$V_0 = \text{span}\{\phi(t - k), k \in \mathbb{Z}\} \cap L^2(\mathbb{R}),$$

In this case, integer translations of the function $\phi(t)$ provide a basis for $V_0$. The reference function $\phi(t)$ is obviously important, and will be referred to as the *Haar scaling function* for this particular multiresolution analysis.

Now introduce the space of higher-resolution functions,

$$V_1 = \{f \in L^2(\mathbb{R}) : f(t) \text{ is constant over the interval } \left[\frac{k}{2}, \frac{k+1}{2}\right), \forall k \in \mathbb{Z}\}.$$

Proceeding in the same way as above, we could write that the following countably-infinite set of functions provides an orthonormal basis on $V_1$:

$$\phi_{1k}(t) = \sqrt{2} \ I_{[k/2, (k+1)/2)}(t).$$

This implies that

$$V_1 = \text{span}\{\phi_{1k}(t), k \in \mathbb{Z}\} \cap L^2(\mathbb{R}).$$

Some basis functions $\phi_{1k}(t)$ are sketched below. Note that the multiplicative factor of $\sqrt{2}$ must be introduced because the functions are supported on intervals of width $1/2$.

Notice that the $\phi_{1k}$ are translated copies of $\phi_{10}$. More precisely,

$$\phi_{1k}(t) = \phi_{10}(t) \left( t - \frac{k}{2} \right), \quad k \in \mathbb{Z}.$$  

We could substitute this result into the expression for $V_1$ in Eq. (20) above. But what is more important is that they also are *dilated and translated* copies of the scaling function $\phi(t)$:

$$\phi_{1k}(t) = \sqrt{2} \ \phi(2t - k).$$
Some orthonormal basis functions $\phi_{1k}(t)$ of the space $V_1$.

This might be a little clearer if we rewrite the function on the right as follows:

$$
\phi_{1k}(t) = \sqrt{2} \phi \left( 2 \left( t - \frac{k}{2} \right) \right).
$$

(23)

The function on the right is the function $\phi(t)$, first shrunk by a factor of $1/2$ and then translated by $k/2$ to the right.

Just as we did for $V_0$, we may write

$$
V_1 = \text{span}\{ \sqrt{2} \phi(2t - k), \ k \in \mathbb{Z} \} \cap L^2(\mathbb{R}).
$$

(24)

In other words, if $x(t) \in V_1$, then

$$
x(t) = \sum_{k \in \mathbb{Z}} c_{1k} \phi_{1k}(t),
$$

(25)

where

$$
c_{1k} = \langle x, \phi_{1k} \rangle.
$$

(26)

But from Eq. (22), we may also write that

$$
x(t) = \sum_{k \in \mathbb{Z}} c_{1k} \sqrt{2} \phi(2t - k).
$$

(27)

The function $f_1(t)$, which was the piecewise constant function approximation to $f(t)$ over intervals of half-integer length, clearly belongs to $V_1$. Moreover, it is the best approximation to $f$ in $V_1$, i.e.,

$$
f_1(t) = \sum_{k \in \mathbb{Z}} \langle f, \phi_{1k} \rangle \phi_{1k}.
$$

(28)
We can also go “backwards,” and consider the space of functions that are one step coarser in resolution: the space $V_{-1}$ composed of functions that are constant over intervals of length 2:

$$V_{-1} = \{ f \in L^2(\mathbb{R}) : f(t) \text{ is constant over the interval } [2k, 2k + 2), \forall k \in \mathbb{Z} \}. \quad (29)$$

It should not be too difficult to see that

$$V_{-1} = \text{span}\{ \frac{1}{\sqrt{2}} \phi \left( \frac{t}{2} - k \right), \ k \in \mathbb{Z} \} \cap L^2(\mathbb{R}). \quad (30)$$

In other words, if $x(t) \in V_{-1}$, then

$$x(t) = \sum_{k \in \mathbb{Z}} c_{-1,k} \frac{1}{\sqrt{2}} \phi \left( \frac{t}{2} - k \right). \quad (31)$$

In general, we may now define the general resolution levels $V_J$, $J = \cdots, -1, 0, 1, \cdots$, as follows:

$$V_J = \{ f \in L^2(\mathbb{R}) : f(t) \text{ is constant over the interval } \left[ \frac{k}{2^J}, \frac{k+1}{2^J} \right), \forall k \in \mathbb{Z} \}. \quad (32)$$

This space is spanned by appropriate dilations and translations of the reference scaling function $\phi(t)$:

$$V_J = \text{span}\{ 2^{J/2} \phi(2^J t - k), \ k \in \mathbb{Z} \} \cap L^2(\mathbb{R}). \quad (33)$$
Lecture 22

Wavelets and multiresolution analysis (cont’d)

The resolutions $V_j$ are nested

Recall the following:

1. The space $V_0$ is composed of all $L^2(\mathbb{R})$ functions that are constant over intervals of unit length, $[k, (k + 1)), k \in \mathbb{Z}$.

2. The space $V_1$ is composed of all $L^2(\mathbb{R})$ functions that are constant over intervals of half-integer length, $[j/2, (j + 1)/2), j \in \mathbb{Z}$.

Now note that all functions in $V_0$ may be considered as special cases of functions in $V_1$: it just happens that the functions have the same values over adjacent intervals $[j/2, (j + 1)/2)$ and $[(j + 1)/2, j + 1)$, i.e., $[j, j + 1)$. This implies that

$$V_0 \subset V_1.$$  \hfill(34)

We can apply this reasoning to any contiguous pair of subspaces $V_j$ and $V_{j+1}$, i.e.,

$$V_j \subset V_{j+1}, \quad j \in \mathbb{Z}. \hfill(35)$$

Therefore, in general, we have the nesting relations

$$\cdots \subset V_{j-1} \subset V_j \subset V_{j+1} \subset \cdots. \hfill(36)$$

But more on this later. We now focus on the particular nesting $V_0 \subset V_1$.

The nesting relation $V_0 \subset V_1$

Let us now return to the Haar scaling function $\phi(t)$. Recall that the set of all integer translates $\phi(t - k), k \in \mathbb{Z}$, forms an orthonormal basis in $V_0$. But from the nesting relation $V_0 \subset V_1$, it follows that $\phi(t)$ is also an element in $V_1$.

Of course it is! We can simply split its graph into two equal and nonoverlapping pieces, as shown in the following figure:
The scaling relation \( \phi(t) = \phi(2t) + \phi(2t - 1) \) satisfied by the Haar scaling function.

The left piece supported on \([0, 1/2]\) may be viewed as the graph of \( \phi(t) \) shrunk by a factor of 2, i.e., \( \phi(2t) \). And the right piece may be viewed as the graph of \( \phi(t) \) shrunk by a factor of 2 and then translated to the right by \( 1/2 \), i.e., \( \phi(2t - 1) \). These two pieces do not overlap and so we can write,

\[
\phi(t) = \phi(2t) + \phi(2t - 1), \quad t \in [0, 1].
\] (37)

There is a more mathematical way of showing that this relation should exist, as we shall now show.

Let's consider the scaling function \( \phi(t) \) again. As we said above, it is an element of \( V_0 \). But recall that \( V_0 \) is a subset of \( V_1 \). This means that we should be able to express it as a linear combination of the basis functions \( \phi_{1k} \) that span \( V_1 \), i.e.,

\[
\phi(t) = \sum_{k \in \mathbb{Z}} h_k \phi_{1k}(t).
\] (38)

Once again, there is little choice. Since \( \phi(t) = 0 \) for \( t \notin [0, 1) \), it follows that \( h_k = 0 \) for \( k \neq 0, 1 \). For \( t \in [0, 1/2) \), the only contribution to the sum is from \( h_0 \phi_{10}(t) = h_0 \sqrt{2} \). Since \( \phi(t) = 1 \) on this interval, it follows that \( h_0 = \frac{1}{\sqrt{2}} \). Likewise for \( t \in [1/2, 1) \), the only contribution to the sum is from \( h_1 \phi_{11}(t) = h_1 \sqrt{2} \). Once again, \( \phi(t) = 1 \) so that \( h_1 = \frac{1}{\sqrt{2}} \). The net result:

\[
\phi(t) = h_0 \phi_{10}(t) + h_1 \phi_{11}(t) = \frac{1}{\sqrt{2}} \phi_{10}(t) + \frac{1}{\sqrt{2}} \phi_{11}(t).
\] (39)

This equation may look different from Eq. (37) but it is, in fact, identical to it.

That being said, it is often customary to work with another form of the above equation, in which we acknowledge that the \( \phi_{1k}(t) \) functions are dilated and translated copies of \( \phi(t) \), namely,

\[
\phi(t) = \sum_{k \in \mathbb{Z}} h_k \sqrt{2} \phi(2t - k).
\] (40)
This equation is often called the multiresolution analysis or refinement equation that is satisfied by the scaling function. It is sometimes simply called the scaling equation. The nonzero coefficients \( h_k \) are known as the scaling coefficients. For the Haar system once again,

\[
h_k = \begin{cases} 
\frac{1}{\sqrt{2}}, & \text{if } k \in \{0, 1\}, \\
0 & \text{otherwise.}
\end{cases}
\] (41)

Substituting these values into Eq. (40) yields

\[
\phi(t) = \frac{1}{\sqrt{2}} \cdot \sqrt{2} \phi(2t) + \frac{1}{\sqrt{2}} \cdot \sqrt{2} \phi(2t - 1) = \phi(2t) + \phi(2t - 1), \quad t \in [0, 1].
\] (42)

which is precisely the scaling equation in Eq. (37).

At the risk of repetition, these coefficients characterize the Haar scaling function. It is the nonzero values of the scaling coefficients \( h_k \) that characterizes the multiresolution system. We’ll see later that we won’t even have to know the explicit form of the scaling function \( \phi \) – it will suffice to know the (nonzero) \( h_k \) values. In many practical applications, the scaling function \( \phi \) has compact support, implying that the set of nonzero \( h_k \) coefficients is finite. However, there are examples of scaling coefficients (and associated mother wavelet functions) which are supported on the entire real line, as we’ll see later.

Wavelets

We consider once again the nesting relation,

\[
V_0 \subset V_1,
\] (43)

but now ask the question, “What is the set of \( L^2(\mathbb{R}) \) functions that are elements of \( V_1 \) but not \( V_0 \)? In a Hilbert space setting, this is the orthogonal complement of the space \( V_0 \), denoted as \( V_0^\perp \):

\[
W_0 = V_0^\perp = \{ x \in V_1 \mid \langle x, y \rangle = 0 \text{ for all } y \in V_0 \}.
\] (44)

Using this definition, it can be shown that \( W_0 \) is a closed linear space:

1. If \( x_1, x_2 \in W_0 \), then \( c_1 x_1 + c_2 x_2 \in W_0 \), for \( c_1, c_2 \in \mathbb{R} \),

2. If the sequence \( \{x_n\} \subset W_0 \) is convergent with limit \( \bar{x} \), then \( \bar{x} \in W_0 \).
Note also that the sets $V_0$ and $W_0$ are not totally disjoint: they both contain the element $x = 0$, since they are both linear spaces.

The spaces $V_0$ and $W_0$ are orthogonal to each other: Think of the $x$- and $y$-axes in $\mathbb{R}^2$. Or better yet, in a function space setting, think of the orthogonal spaces $c_1 \sin t$ and $c_2 \sin 2t$ on $[-\pi, \pi]$. By construction, the space $V_1$ is the direct sum of $V_0$ and $W_0$, i.e.,

$$V_1 = V_0 \oplus W_0.$$  \hfill (45)

This means that any element $z \in V_1$ may be written as the unique sum

$$z = x + y,$$  \hfill (46)

where $x \in V_0$ and $y \in W_0$ are, respectively the projections, or best approximations, of $z$ in $V_0$ and $W_0$, respectively.

As stated earlier, we would like to get an idea of what kind of $L^2(\mathbb{R})$ functions “live” in $W_0 \subset V_1$. Therefore, our goal is to find a set of orthonormal basis functions for $W_0$. This turns out to be rather simple for the piecewise-constant case considered above. Recall that $\phi(t) = I_{[0,1]}(t)$ forms a basis for $V_0$. It is not too difficult to see that the function, $\psi(t)$, defined earlier in Eq. (6) and repeated below,

$$\psi(t) = \begin{cases} 
1, & 0 \leq t < 1/2, \\
-1, & 1/2 \leq t < 1,
\end{cases}$$  \hfill (47)

is orthogonal to $\phi(t)$.

\[\begin{array}{c}
\text{The “detail” or “wavelet” basis function } \psi(t) .
\end{array}\]

It is piecewise linear on intervals of half-integer length, so it is an element of $V_1$. And it is orthogonal to $\phi$:

$$\langle \phi, \psi \rangle = \int_0^1 \phi(t) \psi(t) \, dt$$
\[
= \int_0^{1/2} 1 \, dt + \int_{1/2}^{1} (-1) \, dt \\
= 0. \quad (48)
\]

Note also that \( \langle \psi, \psi \rangle = 1: \)

\[
\langle \psi, \psi \rangle = \int_0^1 \psi(t)\psi(t) \, dt \\
= \int_0^{1/2} 1 \, dt + \int_{1/2}^{1} 1 \, dt \\
= 1. \quad (49)
\]

The function \( \psi(t) \) is known at the mother wavelet of the Haar wavelet system, or simply the Haar mother wavelet. The two functions \( \phi(t) \) and \( \psi(t) \) are pictured below for comparison.

![Haar scaling function \( \phi(t) \) and mother wavelet \( \psi(t) \).](image)

Recall that different integer translates of the scaling function \( \phi(t-k) \) are orthogonal to each other, because they are nonzero over different intervals, i.e.,

\[
\langle \phi(t-k_1), \phi(t-k_2) \rangle = \delta_{k_1,k_2}. \quad (50)
\]

For this reason, we saw that the set of all integer translates \( \phi(t-k) \) spans the space \( V_0 \).

A similar situation holds for the mother wavelet function, \( \psi(t) \), which is nonzero only on [0, 1]. Because different integer translates of are nonzero over different intervals, they are orthogonal to each other, i.e.,

\[
\langle \psi(t-k_1), \psi(t-k_2) \rangle = \delta_{k_1,k_2}. \quad (51)
\]
For this reason, the set of all integer translates $\psi(t - k)$, $k \in \mathbb{Z}$ is an orthonormal basis for the space $W_0$. Note also that

$$\langle \phi(t - k_1), \psi(t - k_2) \rangle = 0, \text{ for all } k_1, k_2 \in \mathbb{Z},$$  \hspace{1cm} (52)

reaffirming that $V_0 \perp W_0$.

We can do this a little more mathematically, in the same way as we explored the relation between the nested spaces $V_0$ and $V_1$. Once again, given that the integer translates $\phi(t - k)$ of the scaling function $\phi(t)$ span $V_0$, we would like to find a function $\psi \perp \phi$ whose integer translates $\psi(t - k)$ span $W_0$. Therefore it is sufficient to demand that $\langle \psi, \phi \rangle = 0$.

Since $\psi \in W_0 \subset V_1$, it may be written in terms of the basis functions $\phi_{1k}$ of $V_1$, as follows,

$$\psi(t) = \sum_{k \in \mathbb{Z}} g_k \phi_{1k}(t).$$  \hspace{1cm} (53)

You may compare this equation with Eq. (38) – note that the expansion coefficients are now denoted as $g_k$.

Let us now impose the condition that $\langle \phi, \psi \rangle = 0$. Recalling that

$$\phi(t) = \frac{1}{\sqrt{2}} \phi_{10}(t) + \frac{1}{\sqrt{2}} \phi_{11}(t),$$  \hspace{1cm} (54)

we have

$$\langle \phi, \psi \rangle = \left\langle \frac{1}{\sqrt{2}} \phi_{10}(t) + \frac{1}{\sqrt{2}} \phi_{11}(t), \sum_{k \in \mathbb{Z}} g_k \phi_{1k}(t) \right\rangle.$$  \hspace{1cm} (55)

Recalling that the functions $\phi_{1k}(t)$ form an orthonormal set, i.e.

$$\langle \phi_{1k_1}, \phi_{1k_2} \rangle = \delta_{k_1 k_2},$$  \hspace{1cm} (56)

the only nonzero elements on the RHS of Eq. (55) are given by

$$\langle \phi, \psi \rangle = \frac{1}{\sqrt{2}} g_0 + \frac{1}{\sqrt{2}} g_1 = 0.$$  \hspace{1cm} (57)

This implies that $g_1 = -g_0$, i.e.,

$$\psi(t) = g_0 [\phi_{10}(t) - \phi_{11}(t)].$$  \hspace{1cm} (58)

Since

$$\|\psi\|^2 = \langle \psi, \psi \rangle = (g_0)^2 \langle \phi_{10} - \phi_{11}, \phi_{10} - \phi_{11} \rangle,$$

250
This implies that \( g_0 = \frac{1}{\sqrt{2}} \). Thus our expansion coefficients \( g_k \) in Eq. (53) are given by

\[
g_0 = \frac{1}{\sqrt{2}}, \quad g_1 = -\frac{1}{\sqrt{2}}, \quad g_k = 0, \text{ otherwise.}
\]

(60)

We can also rewrite Eq. (53) as follows,

\[
\psi(t) = \sum_{k \in \mathbb{Z}} g_k \sqrt{2} \phi(2t - k),
\]

(61)

once again acknowledging that the functions \( \phi_{1k} \) are translated and dilated copies of the scaling function \( \phi(t) \). This equation for the Haar wavelet function \( \psi(t) \) may be compared with its scaling function counterpart, Eq. (40).

Finally, we mention that from the orthogonality of the \( \phi(t) \) and \( \psi(t) \) functions, i.e., \( \langle \phi, \psi \rangle = 0 \) and the orthogonality of the \( \phi_{1k} \) functions, it follows, from Eqs. (38) and (53), that

\[
\langle \phi, \psi \rangle = \sum_{k \in \mathbb{Z}} h_k g_k = 0.
\]

(62)

We actually used this relation to determine the coefficients \( g_k \) for the Haar system from the \( h_k \). Go back to Eq. (57): it corresponds to

\[
h_0 g_0 + h_1 g_1 = 0.
\]

(63)

This relation will apply for any other wavelet system. We’ll return to it in later discussions.

A summary of major recent results:

1. The Haar scaling function \( \phi(t) = I_{[0,1]}(t) \) satisfies the scaling relation

\[
\phi(t) = \phi(2t) + \phi(2t - 1),
\]

(64)

which is a special case of the following general equation satisfied by the scaling function associated with a multiresolution analysis:

\[
\phi(t) = \sum_{k \in \mathbb{Z}} h_k \sqrt{2} \phi(2t - k).
\]

(65)
The nonzero coefficients $h_k$ characterize a particular multiresolution analysis. In the Haar case, $h_0 = h_1 = 1/\sqrt{2}$.

2. For each $j \in \mathbb{Z}$, the infinite set of functions

$$\phi_{jk}(t) = 2^{j/2}\phi(2^j t - k), \quad k \in \mathbb{Z},$$

form an orthonormal basis set for the space $V_j$ of piecewise constant $L^2(\mathbb{R})$ functions over intervals of length $2^{-j}$, i.e., $[k/2^j, (k + 1)/2^j]$.

This equation is also obeyed by scaling functions of other multiresolution systems – only the nature of the spaces $V_j$ will change.

3. Eq. (65) is a consequence of the fact that the spaces $V_j$ are nested, i.e.,

$$V_j \subset V_{j+1}, \quad k \in \mathbb{Z}.$$ (67)

4. Focussing on the particular nesting relation $V_0 \subset V_1$, we considered the space

$$W_0 = \{ x \in V_1 \mid \langle x, y \rangle = 0 \text{ for all } y \in V_1 \}.$$ (68)

$W_0$ is the orthogonal complement of $V_0$ in $V_1$. As such, we may express $V_1$ as the direct sum

$$V_1 = V_0 \oplus W_0.$$ (69)

5. We showed that the Haar wavelet function $\psi(t) = I_{[0,1/2)}(t) - I_{[1/2,1)}(t)$ (in other words, $\psi(t) = 1$ for $t \in [0,1/2)$ and $\psi(t) = -1$ for $t \in [1/2,1)$) is a function in $V_1$ that is orthogonal to $\phi(t)$, i.e., $\langle \phi, \psi \rangle = 0$, so $\psi \in W_0$.

The Haar wavelet may also be expressed in terms of the scaling function as follows,

$$\psi(t) = \phi(2t) - \phi(2t - 1),$$ (70)

which is a special case of the more general equation satisfied by mother wavelets of a multiresolution analysis,

$$\psi(t) = \sum_{k \in \mathbb{Z}} g_k \sqrt{2}\phi(t - k).$$ (71)

6. Furthermore, the integer translates of $\psi(t)$, i.e., $\psi(t) = \psi(t - k)$ form an orthonormal basis of $W_0$.  

252
A consequence of the direct product sum \( V_1 = V_0 \oplus W_0 \) is that we need the two sets of functions

\[
\phi_{0k}(t) = \phi(t - k), \quad k \in \mathbb{Z},
\]

\[
\psi_{0k}(t) = \psi(t - k), \quad k \in \mathbb{Z},
\]

(72)
to span the space \( V_1 \). Therefore, any function \( u \in V_1 \), i.e., any \( L^2(\mathbb{R}) \) function that is piecewise constant over the intervals \([k/2, (k + 1)/2)\) of length 1/2, may be expressed as follows,

\[
u(t) = \sum_{k \in \mathbb{Z}} a_k \phi_{0k}(t) + \sum_{k \in \mathbb{Z}} b_k \psi_{0k}(t),
\]

(73)
where

\[
a_k = \langle u, \phi_{0k} \rangle, \quad b_k = \langle u, \psi_{0k} \rangle, \quad k \in \mathbb{Z}.
\]

(74)
The first summation in (73) corresponds to the \( V_0 \) component of \( u \) and the second summation corresponds to the \( W_0 \) component of \( u \).

Let’s now go back to our example at the beginning of this section on wavelets – Lecture 21 – and consider the general \( L^2 \) function \( f(t) \) that was “scanned” at various resolutions. Recall that \( f_0(t) \) was the approximation to \( f(t) \) that was piecewise constant over unit intervals. In other words, \( f_0(t) \) was the best approximation to \( f(t) \) in the space \( V_0 \)! Likewise \( f_1(t) \) was the best approximation to \( f(t) \) in \( V_1 \). Finally, the function \( f_d(t) \) was what you had to add to the lower resolution approximation \( f_0(t) \) to obtain \( f_1(t) \).

In other words, if we let \( u(t) = f_1(t) \), then Eq. (73) represents the equation

\[
f_1(t) = f_0(t) + f_d(t).
\]

(75)
Now recalling the direct sum relation \( V_1 = V_0 \oplus W_0 \), we can state that

1. \( f_0 \) is the projection, or best approximation, of \( f_1 \) on \( V_0 \), and

2. \( f_d \) is the projection, or best approximation, of \( f_1 \) on \( W_0 \).

But we can go one step further. When you project “something”, say “\( X \)”, onto “something else, say “\( Y \)”, you get a result “\( Z \)”. If you project “\( Z \)” onto “\( Y \)”, you get “\( Z \)” again. In other words, after you project something once, you’ll get the same result if you keep projecting. (You may recall hearing the word “idempotent” to describe this property.) We have said that \( f_1 \) is the projection of \( f \) on \( V_1 \). This means that
1. $f_0$ is the projection, or best approximation of the original function $f$ on $V_0$, and

2. $f_d$ is the projection, or best approximation of the original function $f$ on $W_0$.

Therefore, the function $u$ in Eq. (73) corresponds to $f_1$ so that the coefficients in that equation are given by

$$a_k = \langle f, \phi_{0k} \rangle, \quad b_k = \langle f, \psi_{0k} \rangle, \quad k \in \mathbb{Z}.$$  \hspace{1cm} (76)

Once again, this means that the “detail” function $f_d$, i.e., what we have to add to the lower resolution approximation $f_0$ in order to obtain the higher resolution approximation $f_1$, “lives” in the space $W_0$ which is spanned by the Haar wavelet functions $\psi(t-k)$. This is, of course, what we observed in the figures in Lecture 21.

And there is one more important point before we move on! Let’s go back to Eq. (73), which is an expansion of $u \in V_1$ in terms of the orthonormal basis set $\{\phi_{0k}, \psi_{0k}\}_{k \in \mathbb{Z}}$. But we have also stated that the functions $\{\phi_{1k}\}_{k \in \mathbb{Z}}$ comprise an orthonormal basis set in $V_1$, i.e.,

$$u(t) = \sum_{k \in \mathbb{Z}} c_k \phi_{1k}(t).$$  \hspace{1cm} (77)

Recalling that $u = f_1$, the best approximation of $f$ in the space $V_1$, the coefficients $c_k$ are given by

$$c_k = \langle f, \phi_{1k} \rangle.$$  \hspace{1cm} (78)

Clearly, there must be a relationship between the coefficients $\{c_k\}$ and the coefficients $\{a_k\}$ and $\{b_k\}$ of Eq. (73). We shall explore this relationship in the next lecture.
Lecture 23

Wavelets and multiresolution analysis (cont’d)

Higher-order nestings \( V_j \subset V_{j+1} \)

We have just analyzed the nesting relation \( V_0 \subset V_1 \), which led to the direct sum decomposition \( V_1 = V_0 \oplus W_0 \). But we don’t have to stop there – we may now employ the same type of analysis for the next relation \( V_1 \subset V_2 \). We define the space

\[
W_1 = \{ x \in V_2 \mid \langle x, y \rangle = 0 \text{ for all } y \in V_1 \}.
\]

We may then write the decomposition

\[
V_2 = V_1 \oplus W_1.
\]

Recall that \( V_1 \) is spanned by the orthonormal set \( \{ \phi_{1k}(t) = \sqrt{2}\phi(2t - k) \}, k \in \mathbb{Z} \). And recall the orthogonality relation from the previous lecture,

\[
\langle \phi(t - k_1), \psi(t - k_2) \rangle = 0, \text{ for all } k_1, k_2 \in \mathbb{Z}.
\]

Replacing \( t \) with \( 2t \) in the above, and multiplying each function by \( \sqrt{2} \), it follows that

\[
\langle \sqrt{2}\phi(2t - k_1), \sqrt{2}\psi(2t - k_2) \rangle = \langle \phi_{1k_1}, \psi_{1k_2} \rangle = 0, \text{ for all } k_1, k_2 \in \mathbb{Z},
\]

which implies that the functions \( \psi_{1k}(t) = \sqrt{2}\psi(2t - k), k \in \mathbb{Z} \), span the space \( W_1 \).

And, of course, we don’t have to stop there: In general, for the nesting relation \( V_j \subset V_{j+1} \), we define

\[
W_j = \{ x \in V_{j+1} \mid \langle x, y \rangle = 0 \text{ for all } y \in V_j \}.
\]

This produces the decomposition

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

Recall that the space \( V_j \) is spanned by the orthonormal set \( \{ \phi_{jk}(t) = 2^{j/2}\phi(2^j t - k) \}, k \in \mathbb{Z} \). We now consider the orthogonality relation in Eq. (81), replace \( t \) by \( 2^j t \) and multiplying each function by \( 2^j \) to give

\[
\langle 2^{j/2}\phi(2t - k_1), 2^{j/2}\psi(2t - k_2) \rangle = \langle \phi_{jk_1}, \psi_{jk_2} \rangle = 0, \text{ for all } k_1, k_2 \in \mathbb{Z}.
\]

This implies that the functions \( \psi_{jk}(t) = 2^{j/2}\psi(2^j t - k), k \in \mathbb{Z} \), span the space \( W_j \).

We now summarize the fundamental results of the previous section:
1. $V_0 \subset V_1$, leading to $V_1 = V_0 \oplus W_0$.

2. $V_1 \subset V_2$, leading to $V_2 = V_1 \oplus W_1$.

3. $V_j \subset V_{j+1}$, leading to $V_{j+1} = V_j \oplus W_j$, for any $j \in \mathbb{Z}$.

Now take the relation $V_2 = V_1 \oplus W_1$ and substitute the first relation for $V_1$ to give

$$V_2 = V_1 \oplus W_1 = V_0 \oplus W_0 \oplus W_1.$$  \hfill (86)

In other words, we have decomposed the space $V_2$ into three orthogonal subspaces, $V_0, W_0$ and $W_1$. As such, $V_2$ spanned by a triple set of orthonormal functions:

$$V_2 = \text{span}\{\phi_{0k}, \psi_{0k}, \psi_{1k}, \ k \in \mathbb{Z}\} \cap L^2(\mathbb{R}).$$  \hfill (87)

This implies that a function $u \in V_2$ admits the expansion,

$$u(t) = \sum_{k \in \mathbb{Z}} \langle u, \phi_{0k} \rangle \phi_{0k}(t) + \sum_{k \in \mathbb{Z}} \langle u, \psi_{0k} \rangle \phi_{0k}(t) + \sum_{k \in \mathbb{Z}} \langle u, \psi_{1k} \rangle \phi_{0k}(t).$$  \hfill (88)

The first summation produces the low resolution component $u_0 \in V_0$ of $u$. (Recall $f_0$ in the first figure of Lecture 21.) Adding the terms in the next summation produces the function $u_1 \in V_1$, a higher resolution approximation to $u$. (This was $f_1$ in that example.) Adding the next terms reconstructs $u \in V_2$, etc..

We may now generalize this procedure: For a fixed positive integer $J > 0$, we have

$$V_{J+1} = V_J \oplus W_J$$

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

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

$$\vdots$$

$$V_{J+1} = V_0 \oplus W_0 \oplus W_1 \oplus \cdots \oplus W_{J-1} \oplus W_J.$$  \hfill (89)

The result: The space $V_{J+1}$ has been decomposed into a direct sum of $J + 2$ orthogonal spaces, $V_0$ and $W_j$, $0 \leq j \leq J$. Recalling that the space $W_j$ is spanned by the orthonormal basis functions $\psi_{jk}, k \in \mathbb{Z}$, we may expand an element $u \in V_{J+1}$ as follows,

$$u(t) = \sum_{k \in \mathbb{Z}} a_{0k} \phi_{0k}(t) + \sum_{j=0}^{J} \sum_{k \in \mathbb{Z}} b_{jk} \psi_{jk}(t).$$  \hfill (90)
You may think of each space $W_j$ as a particular “frequency band”: It does not consist of one function over the entire real line, such as $e^{ikt}$ for a fixed $k$, but rather of an infinite set of wavelet functions $\psi_{jk}(t)$ each of which is nonzero only over an interval of length $2^{-j}$.

As $J$ gets larger and larger in Eq. (89), the functions “living” in the space $V_{J+1}$ are getting “finer and finer”, i.e., piecewise constant functions over intervals of length $2^{J+1}$. These would correspond to higher resolution scans $f_{J+1}$ of our function $f(t)$. We would hope that in the limit $J \to \infty$, the function $f_{J+1}$ would converge to $f(t)$, at least in the $L^2$-sense.

This is indeed the case, which we shall not justify here since it is the topic of an advanced course in analysis. We shall simply write that

$$\lim_{J \to \infty} V_J = L^2(\mathbb{R}).$$

(91)

This means that we can write

$$L^2(\mathbb{R}) = V_0 \oplus W_0 \oplus W_1 \oplus \cdots \oplus W_j \oplus \cdots$$

$$= V_0 \oplus \bigoplus_{j=0}^{\infty} W_j. \quad (92)$$

Keeping in mind that $V_0$ is spanned by $\phi_{0k}$, $k \in \mathbb{Z}$, and the $W_j$ are spanned by the functions $\psi_{jk}$, $k \in \mathbb{Z}$, it follows that a function $f \in L^2(\mathbb{R})$ admits the following expansion, in the $L^2$-sense:

$$f(t) = \sum_{k \in \mathbb{Z}} a_{0k} \phi_{0k}(t) + \sum_{j=0}^{\infty} \sum_{k \in \mathbb{Z}} b_{jk} \psi_{jk}(t), \quad (93)$$

where

$$a_{0k} = \langle f, \phi_{0k} \rangle, \quad b_{jk} = \langle f, \psi_{jk} \rangle. \quad (94)$$

The space $V_0$ represents our “lowest resolution scan” of $f$. Note that we didn’t have to start at $V_0$: We could have started at another level, say $V_K$ and proceeded upwards.

And this brings us to the final point! We can go the other way! From our nesting relation

$$V_{j+1} = V_j \oplus W_j, \quad (95)$$

It follows that we can decompose $V_0$ by setting $j = -1$:

$$V_0 = V_{-1} \oplus W_{-1}. \quad (96)$$

By definition,
1. $V_{-1}$ is the space of functions that are constant over intervals $[2k, 2k+2)$ of length 2. It is spanned by the functions $\phi_{-1,k}(t) = 2^{-1/2}\phi(t/2 - k)$.

2. $W_{-1} \subset V_0$ is the space of functions orthogonal to $V_{-1}$. It is spanned by the wavelet functions $\psi_{-1,k}(t) = 2^{-1/2}\psi(t/2 - k)$ that are also nonzero over intervals $[2k, 2k+2)$ of length 2.

But we may continue, by decomposing $V_1$, then $V_2$, etc., to obtain

$$V_0 = \cdots \oplus W_{-2} \oplus W_{-1}. \quad (97)$$

We may substitute this result into Eq. (92) to give

$$L^2(\mathbb{R}) = \bigoplus_{k=-\infty}^{\infty} W_j. \quad (98)$$

We have removed the scaling functions $\phi(t)$ from the picture. The space $L^2(\mathbb{R})$ has been expressed as a direct sum of orthogonal subspaces $W_j$ ranging from $j = -\infty$ to $j = \infty$. As such,

The set of functions $\{\psi_{jk}(t)\}, j, k \in \mathbb{Z}$ forms an orthonormal basis of $L^2(\mathbb{R})$.

Consequently, any function $f \in L^2(\mathbb{R})$ admits a unique expansion of the form

$$f(t) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} b_{jk} \psi_{jk}(t), \quad (99)$$

where

$$b_{jk} = \langle f, \psi_{jk} \rangle. \quad (100)$$
Special case: Haar wavelet expansions of functions on a finite interval

In practical applications, we deal with finite signals or images. It is therefore instructive to consider the case of functions supported on a finite interval \([a, b]\). Without loss of generality, we shall let \([a, b] = [0, 1]\). In this case the coarsest possible resolution is that of constant functions on \([0, 1]\), the space \(V_0\). In this case the space of square-integrable functions \(L^2[0, 1]\) can be decomposed as follows

\[
L^2[0, 1] = V_0 \oplus W_0 \oplus W_1 \oplus \cdots. 
\] (101)

Because the interval of support is \([0, 1]\), each subspace will be spanned by a finite number of basis functions:

1. \(V_0\): spanned only by \(\phi_{00}\).
2. \(W_0\): spanned only by \(\psi_{00}\).
3. \(W_1\): spanned by \(\psi_{10}, \psi_{11}\), since these functions are nonzero on intervals of length \(1/2\).
4. \(W_2\): spanned by \(\psi_{2k}, 0 \leq k \leq 3\), since these functions are nonzero on intervals of length \(1/4\).
5. \(W_j\): spanned by \(\psi_{jk}, 0 \leq k \leq 2^j - 1\), since these functions are nonzero on intervals of length \(1/2^j\), etc..

Therefore, the expansion of a function \(f \in L^2[0, 1]\) has the form

\[
f(t) = a_{00}\phi_{00}(t) + \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} b_{jk}\psi_{jk}(t),
\] (102)

where, as before,

\[
a_{00} = \langle f, \phi_{00} \rangle, \quad b_{jk} = \langle f, \psi_{jk} \rangle.
\] (103)

The expansion coefficients are often displayed conveniently in the following tabular form, often referred to as a wavelet coefficient tree,

<table>
<thead>
<tr>
<th>(a_{00})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(b_{00})</td>
</tr>
<tr>
<td>(b_{10})</td>
</tr>
<tr>
<td>(b_{20})</td>
</tr>
<tr>
<td>(b_{30})</td>
</tr>
</tbody>
</table>

\[259\]
Note that the coefficients $b_{ij}$ are arranged in the form of a binary tree that extends downward indefinitely. The width of the box enclosing the coefficients may also be viewed as representing the interval $[0, 1]$. In the (nonoverlapping) Haar wavelet case, the width of the box of each coefficient indicates the support of its corresponding wavelet/scaling function. For example,

1. $a_{00}$ is the coefficient of $\phi_{00}(t)$, which is supported on the entire interval.

2. $b_{33}$ is the coefficient of $\psi_{33}(t)$, which is supported on the subinterval $[3/8, 1/2)$, as represented by the box in the table.

It will also be convenient to express the expansion in Eq. (102) the following form,

$$f_j(t) = f_0(t) + \sum_{j=1}^{\infty} w_j(t),$$

where $f_0(t) = a_{00}\phi_{00}(t)$ is the projection of $f$ on $V_0$ and the detail functions, $w_j \in W_j$, are defined as follows,

$$w_j(t) = \sum_{k=0}^{2^j-1} b_{jk}\psi_{jk}(t).$$

Each detail function $w_j(t)$ is represented by a row of $b_{jk}$ coefficients in the wavelet coefficient tree.

In practical situations, of course, we deal with finite-dimensional samplings of a function $f(t)$, which may be considered as piecewise-constant approximations of it. The projection of a function $f \in L^2[0, 1]$ on the space $V_j$, to be denoted as $f_j$, will be the best piecewise-constant approximation of $f$ over intervals of length $2^{-j}$, i.e., a $2^j$-point sampling of $f$. Recalling that $V_j$ is spanned by the orthonormal basis functions,

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

the projection $f_j$ may be written as follows,

$$f_j(t) = \sum_{k=0}^{2^j-1} a_{jk}\phi_{jk}(t), \quad a_{jk} = \langle f, \phi_{jk} \rangle. $$

The $2^j$ coefficients $(a_{j0}, a_{j1}, \cdots, a_{j, 2^j-1})$ may be viewed as a representation, or digitization, of $f_j$.

But recall that the space $V_j$ may be decomposed as follows,

$$V_j = V_0 \oplus W_0 \oplus W_1 \oplus \cdots \oplus W_{j-1}. $$

260
As such, \( f_j \) may also be expanded as follows,

\[
f_j(t) = f_0(t) + \sum_{i=1}^{j-1} w_i(t)
\]

\[
= a_{00} \phi_{00}(t) + b_{00} \psi_{00}(t) + \sum_{i=1}^{j-1} \sum_{k=0}^{2^i-1} b_{ik} \psi_{ik}(t),
\]  

(109)

where, again,

\[
a_{00} = \langle f, \phi_{00} \rangle, \quad b_{ik} = \langle f, \psi_{ik} \rangle.
\]  

(110)

Eq. (109) is a finite-dimensional truncation of the infinite summation in Eq. (102). The binary tree pictured above is then truncated, with the bottom row comprised of the \( 2^j \) coefficients,

\[
b_{j-1,0} \quad b_{j-1,1} \quad \cdots \quad b_{j-1,2^i-1} \quad b_{j-1,2^i-1-1}.
\]  

(111)

Let us now count up the number of coefficients that comprise the expansion in Eq. (109). We can use the results of the list presented at the beginning of the lecture:

\[
\text{No. of coeffs.} = 1 \, (V_0) + 1 \, (W_0) + 2 \, (W_1) + 4 \, (W_2) + \cdots + 2^{j-1} \, (W_{j-1}) = 2^j.
\]  

(112)

Interestingly enough, this is the same number of coefficients as are used in the \( V_j \) expansion in Eq. (107).

This leads us to an interesting question. Both expansions are valid for the same function. Are the expansion coefficients of these two expansions somehow related? The answer is “Yes” as we shall show in the next lecture.

**Example:** In order to illustrate the above principles, we consider the following function on \([0, 1]\),

\[
f(x) = \begin{cases} 
8(x - 0.6)^2 + 1, & 0 \leq x < 0.6, \\
8(x - 0.6)^2 + 3, & 0.6 \leq x < 1.
\end{cases}
\]  

(113)

Starting on the next page are shown a series of plots of projections \( f_j \in V_j \) to \( f \), each with a plot of \( f(x) \) for comparison. The top figure shows the projection \( f_5(x) \in V_5 \) (piecewise-constant approximation over intervals of length 1/32). The next row of plots shows: \( f_4 \in V_4 \) (left) along with the detail function \( w_4 = f_5 - f_4 \in W_4 \) that must be added to \( f_4 \) to produce \( f_5 \). The rows continue with plots of \( f_3 \in V_3, f_2 \in V_2, f_1 \in V_1 \) and \( f_0 \in V_0 \) along with their associated detail functions \( w_i \in W_i \) for \( i = 3, 2, 1, 0 \), respectively.
Projections $f_j \in V_j$, $j = 5, 4, 3, 2, 1, 0$, to $f(x)$ given in text, along with associated detail functions $f_{j+1} - f_j \in W_j$ for $j = 4, 3, 2, 1$.  

262
\[ f_2 \in V_2 \]

\[ w_2 = f_3 - f_2 \in W_2 \]

\[ f_1 \in V_1 \]

\[ w_1 = f_2 - f_1 \in W_1 \]

\[ f_0 \in V_0 \]

\[ w_0 = f_1 - f_0 \in W_0 \]
The function \( f_0 \in V_0 \) is the best constant approximation to \( f \), given by the mean of \( f \) over \([0, 1]\),

\[
f_0(x) = \int_0^1 f(x) \, dx = \frac{191}{75} = 2.546.
\]  

(114)

This function was constructed in order to show the effects of a discontinuity on the projections and detail functions. You can see the effect in the plot of the detail function \( w_4(t) \), the function that must be added to \( f_4(t) \) to produce \( f_5(t) \). There is a significant component of \( f_4(t) \) in the vicinity of the discontinuity at 0.6.

The wavelet coefficient tree (to two decimals), i.e., the coefficients \( a_{00} \) and \( b_{ik} \) associated with the projection \( f_5 \in V_5 \), are presented in the table below.

<table>
<thead>
<tr>
<th>( a_{00} \approx 2.55 )</th>
<th>( b_{00} \approx -0.40 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( b_{10} \approx 0.49 )</td>
<td>( b_{11} \approx -0.49 )</td>
</tr>
<tr>
<td>( b_{20} \approx 0.24 )</td>
<td>( b_{21} \approx 0.11 )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>0.09</th>
<th>0.07</th>
<th>0.05</th>
<th>0.03</th>
<th>-0.13</th>
<th>-0.01</th>
<th>-0.04</th>
<th>-0.06</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
</tr>
</tbody>
</table>

The sharp discontinuity of \( f(t) \) must be reflected in the wavelet coefficient tree: after all, it is the \( b_{jk} \) coefficients that determine the detail function \( w_j \in W_j \) (more on this later). Note that the coefficient \( b_{49} = -0.20 \) stands out in magnitude from the other coefficients \( b_{4k} \) in its row. The support of the wavelet function \( \psi_{49}(t) \) is \( [9/16, 10/16] = [0.5625, 0.625] \). The discontinuity at 0.6 is contained in this interval. We shall comment further on this feature in a future lecture.

We conclude this discussion by pointing out that in the above table, the coefficient \( a_{00} \) represents the constant approximation \( f_0 \in V_0 \) to \( f \). Using the two coefficients \( a_{00} \) and \( b_{00} \) produces the approximation \( f_1 \in V_1 \). And then adding the two coefficients \( b_{10} \) and \( b_{11} \) produces the approximation \( V_2 \). The procedure may be continued in a straightforward manner.

**Fourier series approximations to \( f(x) \) as a comparison**

It is interesting to compare the above results with the approximations to the function \( f(x) \) in Eq. (113) afforded by Fourier series. We consider the even extension of \( f(x) \) defined over \([0, 1]\) so that the
representation to \( f(x) \) is a Fourier cosine series of the form

\[
 f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos(n\pi x). \tag{115}
\]

As was the case for the \( V_0 \) approximation to \( f(x) \), the coefficient \( a_0 \) (which essentially multiplies the constant scaling function \( \phi(x) = 1 \)) is given by the mean value of \( f(x) \) over \([0, 1]\), i.e.,

\[
a_0 = \int_0^1 f(x) \, dx = \frac{191}{79} \approx 2.55. \tag{116}
\]

The higher-order coefficients \( a_n \) may be computed as follows,

\[
a_n = 2 \int_0^1 f(x) \cos(n\pi x) \, dx, \quad n = 1, 2, \ldots. \tag{117}
\]

We define the following partial sum approximations to \( f(x) \),

\[
 S_N(x) = a_0 + \sum_{n=1}^{N} a_n \cos(n\pi x). \tag{118}
\]

(Technically, the partial sums contain \( N + 1 \) terms.) In the next figure, we show the partial sum approximations \( S_N(x) \) to \( f(x) \) for \( N = 32, 16, 8, 4 \) and 2. In terms of numbers of coefficients/basis functions employed, this selection corresponds closely to the approximations to \( f(x) \) in the spaces \( V_5 \), \( V_4 \), \( V_3 \), \( V_2 \) and \( V_1 \) for the Haar basis shown earlier.

As expected, the accuracy of the approximation appears to improve with \( N \). Also as expected, there is a Gibbs phenomenon in the vicinity of the point of singularity, \( x = 0.6 \), of \( f(x) \). One may note other similarities and differences between these approximations and those afforded by the Haar multiresolution basis, but we won’t spend any time on them here. One fundamental difference between the two approximation schemes which is worthy of mention, however, is concerned with the respective expansion coefficients of the two schemes. Because of the localized support of the Haar basis functions, the location of the singularity of the function is reflected in the behaviour of the Haar coefficients, as has already been discussed: At a given resolution level, i.e., \( W_j \), the coefficient \( a_{jk} \) of the wavelet function \( \psi_{jk}(x) \) that is located on an interval containing the singularity \( x = 0.6 \) is significantly larger than the other coefficients in that level. Because all basis functions \( \cos(n\pi x) \) of the Fourier cosine series are supported over the entire interval \([0, 1]\), they cannot not provide any information about the location of the singularity. In the following figure are plotted the magnitudes \(|a_n|\) of the coefficients \( a_0 \) to \( a_{32} \) employed in these computations. As expected, they demonstrate a general decay. But they provide no information regarding the location of the singularity of \( f(x) \) at \( x = 0.6 \). This was one reason that wavelets became quite attractive in signal/image analysis.
Fourier series approximations to $f(x)$

Approximations to $f(x)$ in Eq. (113) yielded by partial sums $S_N(x)$ of the Fourier series expansion in Eq. (115).
Plot of magnitudes $|a_n|$ of coefficients in the Fourier series representation of the function $f(x)$ in Eq. (113).