Wavelets: A family of orthobases for $L_2(\mathbb{R})$

Wavelets are a class of orthobases that represent a signal using the differences between nested approximations at different scales.

It is best to first look at a couple of examples, and then generalize.

Example 1: Shannon Wavelets

$$B_\pi = \{\text{signals bandlimited to } \pi\}$$

We know that if

$$\phi_0(t) = \frac{\sin(\pi t)}{\pi t}$$

then $\{\phi_0(t-k)\}_{k\in\mathbb{Z}}$ is an $\perp$-basis for B_π.

Recall that given an arbitrary $x(t)$

$$\langle x(t), \phi_0(t-k)\rangle = \text{lowpass filter } x(t), \text{ then sample at } t=n$$

The <u>closest</u> bandlimited signal to $x(t)$ can be written as

$$\hat{x}(t) = \sum_{k=-\infty}^{\infty} \langle x(t), \phi_0(t-k) \rangle \phi_0(t-k)$$

We will write this as

$$\hat{x}(t) = P_{B_\pi}[x(t)]$$

<u>projection</u> onto the space B_π

Now suppose we <u>compress</u> (in time) the sincs by a factor of 2. Set

$$\phi_1(t) = \sqrt{2} \cdot \phi_0(2t)$$

chosen to make $\|\phi_1\|_{L_2} = 1$

$$= \sqrt{2} \, \frac{\sin(2\pi t)}{2\pi t}$$

$$= \sqrt{\frac{1}{2}} \cdot \frac{\sin(2\pi t)}{2\pi t}$$

We have seen before that the set

$$\{\phi_1(t - k/2)\}_{k \in \mathbb{Z}} = \{\sqrt{2}\, \phi_0(2t-k)\}_{k=-\infty}^{\infty}$$

(sincs w/ bandwith 2π at half-integer shifts)

is an orthonormal basis for $B_{2\pi}$.

Similarly, if we define

$$\phi_2(t) = \sqrt{2}\,\phi_1(2t) = 2\phi_0(4t)$$

then

$$\{\phi_2(t-k/4)\}_{k\in\mathbb{Z}} = \{2\phi_0(4t-k)\}_{k\in\mathbb{Z}}$$

is an $\perp$-basis for $B_{4\pi}$.

For any $j\geq 0$, if we define

$$\phi_j(t) = 2^{j/2}\phi_0(2^j t)$$

then

$$\{\phi_j(t-2^{-j}k)\}_{k\in\mathbb{Z}} = \{2^{j/2}\phi_0(2^j t-k)\}_{k\in\mathbb{Z}}$$

is an $\perp$-basis for $B_{2^j\pi}$.

It should be clear that

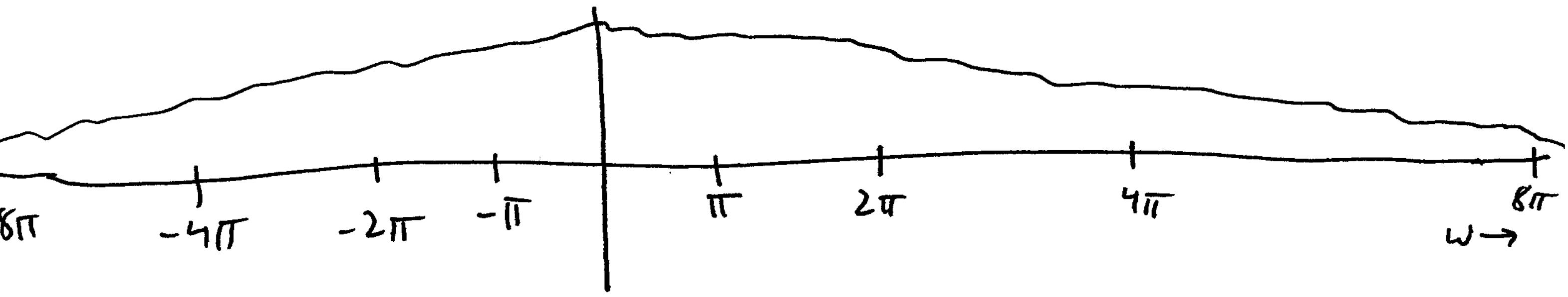
$$P_{B_{2^j\pi}}[x(t)] = \sum_{k=-\infty}^{\infty} \langle x(t), \phi_j(t-2^{-j}k)\rangle\, \phi_j(t-2^{-j}k)$$

gets closer to $x(t)$ as $j\to\infty$

$$\lim_{j\to\infty} P_{B_{2^j\pi}}[x(t)] = x(t)$$

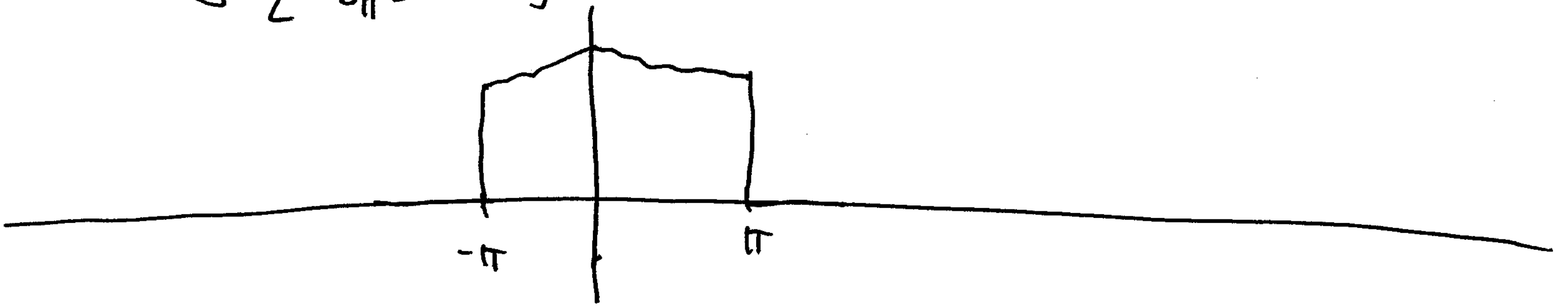
In the Fourier domain, if

$X(j\Omega) =$



then

$\mathcal{F}\{P_{B_\pi}[x(t)]\} =$



$\mathcal{F}\{P_{B_{2\pi}}[x(t)]\} =$

etc.

⋮

It is clear that $B_\pi \subset B_{2\pi}$.

Now consider the difference between the spaces B_π and $B_{2\pi}$

$$W_0 = B_{2\pi} \ominus B_\pi$$

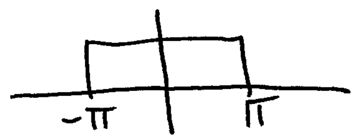
$$= \{ \text{signals in } B_{2\pi} \text{ that are } \underline{\text{orthogonal}} \text{ to } B_\pi \}$$

$$= \{ \text{signals whose Fourier transform is supported on } \pi \le |\omega| \le 2\pi \}$$

We can also find an $\perp$-basis for W_0.

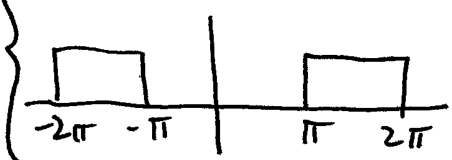
Just as it is true that integer shifts of

$$\phi_0(t) = \mathcal{F}^{-1}\left\{ \quad \right\}$$



are an $\perp$-basis for B_π, the integer shifts of

$$\psi_0(t) = \mathcal{F}^{-1}\left\{ \quad \right\}$$



are an $\perp$-basis for W_0.

(You will solidify this on the next HW.)

Define

$$\psi_0(t) = 2\cos\left(\frac{3\pi t}{2}\right)\frac{\sin(\pi t/2)}{\pi t}$$

then

$\{\psi_0(t-k)\}_{k\in\mathbb{Z}}$ is an $\perp$-basis for W_0.

Given an arbitrary $x(t) \in L_2(\mathbb{R})$, we now have two ways to write the approximation of $x(t)$ in $B_{2\pi}$:

$$P_{B_{2\pi}}[x(t)] = \sum_{k=-\infty}^{\infty} \langle x(t), \phi_1(t-k/2)\rangle \phi_1(t-k/2)$$

and

$$P_{B_{2\pi}}[x(t)] = P_{B_\pi}[x(t)] + P_{W_0}[x(t)]$$

(this equality holds since $B_\pi \perp W_0$, i.e. everything in B_π is $\perp$ to everything in W_0)

$$= \sum_{k=-\infty}^{\infty} \langle x(t), \phi_0(t-k)\rangle \phi_0(t-k) + \sum_{k=-\infty}^{\infty} \langle x, \psi_0(t-k)\rangle \psi_0(t-k)$$

i.e.

$$\text{Approx. in } B_{2\pi} = \text{Approx in } B_\pi + \underbrace{\text{Approx in } W_0}_{\substack{\text{difference between} \\ \text{approx in } B_{2\pi} \,\&\, B_\pi}}$$

Likewise, we can define

$$W_j = B_{2^{j+1}\pi} \ominus B_{2^j\pi}$$

$$= \{\text{signals bandlimited to } |\omega| \in [2^j\pi, 2^{j+1}\pi]\}$$

If we set

$$\psi_j(t) = 2^{j/2}\psi_0(2^j t)$$

then

$$\{\psi_j(t - 2^{-j}k)\}_{k\in\mathbb{Z}} = \{2^{j/2}\psi_0(2^j t - k)\}_{k\in\mathbb{Z}}$$

is an $\perp$-basis for W_j.

Now we can write for any $x(t) \in L_2(\mathbb{R})$

$$x(t) = \underbrace{\underbrace{\underbrace{P_{B_\pi}[x(t)] + P_{W_0}[x(t)]}_{P_{B_{2\pi}}[x(t)]} + P_{W_1}[x(t)]}_{P_{B_{4\pi}}[x(t)]} + P_{W_2}[x(t)]}_{P_{B_{8\pi}}[x(t)]} + \cdots$$

or

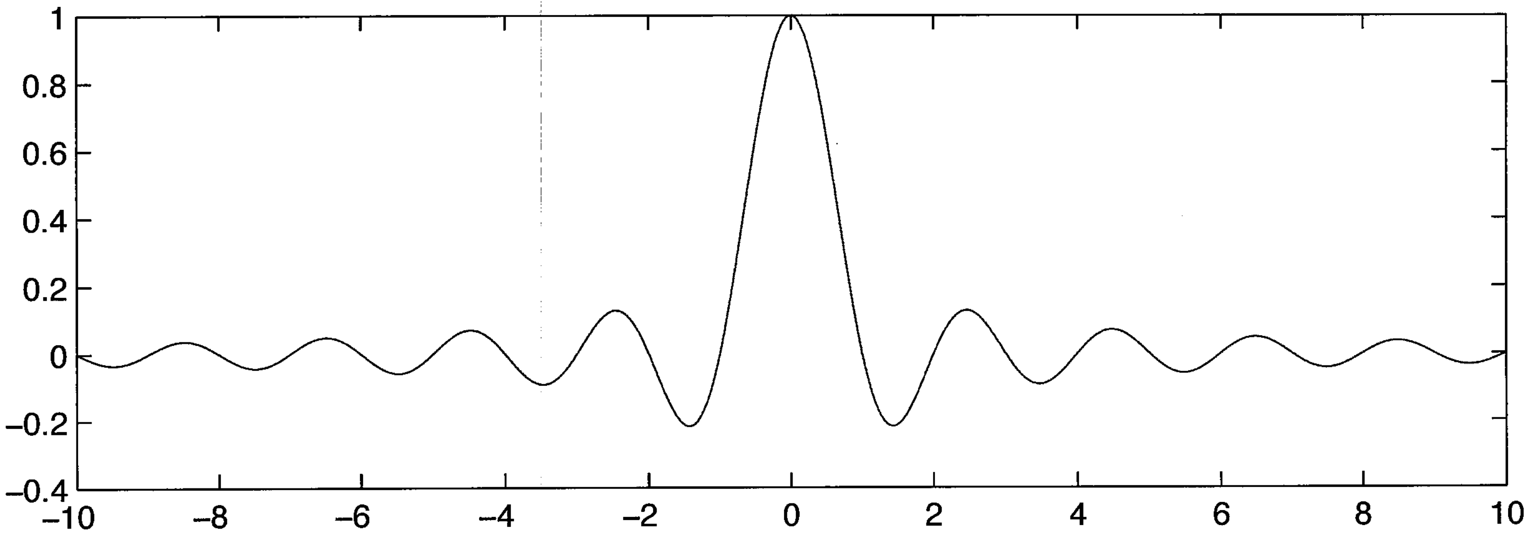
$$
\begin{aligned}
x(t) = &\sum_{k=-\infty}^{\infty} \langle x(t), \phi_0(t-k) \rangle \phi_0(t-k) \\
&+ \sum_{k=-\infty}^{\infty} \langle x(t), \psi_0(t-k) \rangle \psi_0(t-k) \\
&+ \sum_{k=-\infty}^{\infty} \langle x(t), \psi_1(t-k/2) \rangle \psi_1(t-k/2) \\
&+ \sum_{k=-\infty}^{\infty} \langle x(t), \psi_2(t-k/4) \rangle \psi_2(t-k/4) \\
&+ \cdots \\
= &\sum_{k=-\infty}^{\infty} \langle x(t), \phi_{0,k}(t) \rangle \phi_{0,k}(t) + \sum_{j\geq 0} \sum_{k=-\infty}^{\infty} \langle x(t), \psi_{j,k}(t) \rangle \psi_{j,k}(t)
\end{aligned}
$$

where

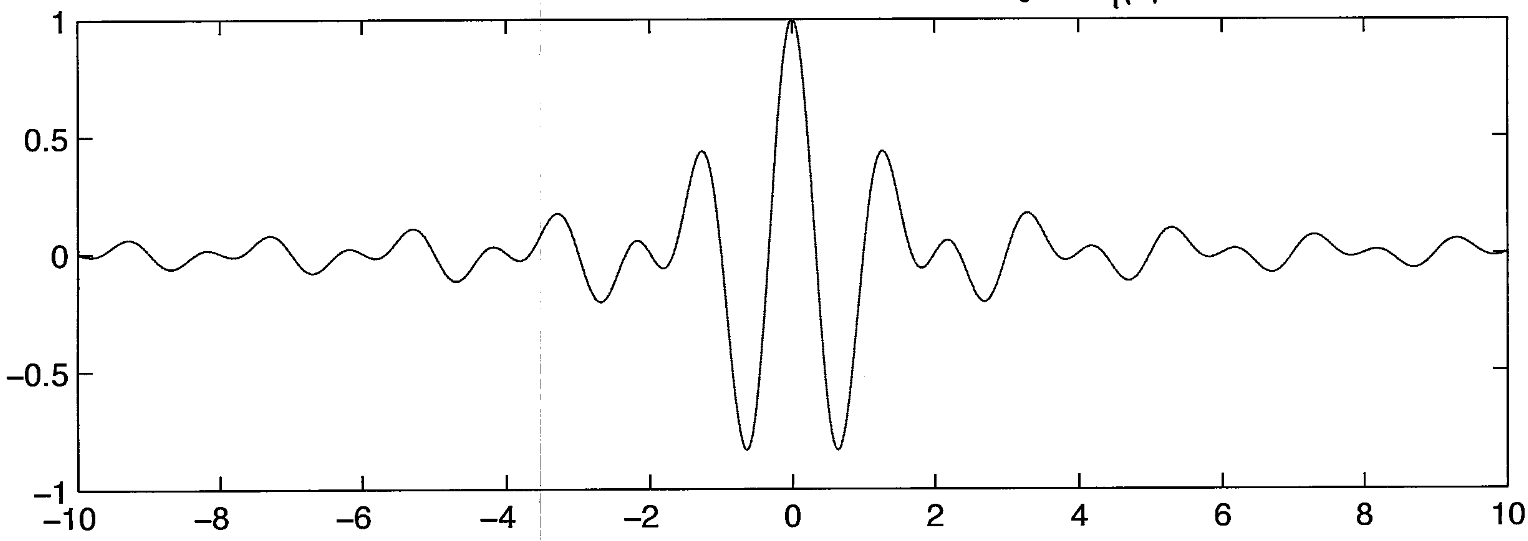
$$
\begin{aligned}
\phi_{0,k}(t) &= \phi_0(t-k) \\
\psi_{j,k}(t) &= 2^{j/2}\psi_0(2^j t - k)
\end{aligned}
$$

So $\{\phi_{0,k}(t)\}_{k\in\mathbb{Z}} \cup \{\psi_{j,k}(t)\}_{j\geq 0, k\in\mathbb{Z}}$
is an orthonormal basis for $L_2(\mathbb{R})$

$$\phi_0(t) = \frac{\sin(\pi t)}{\pi t}$$



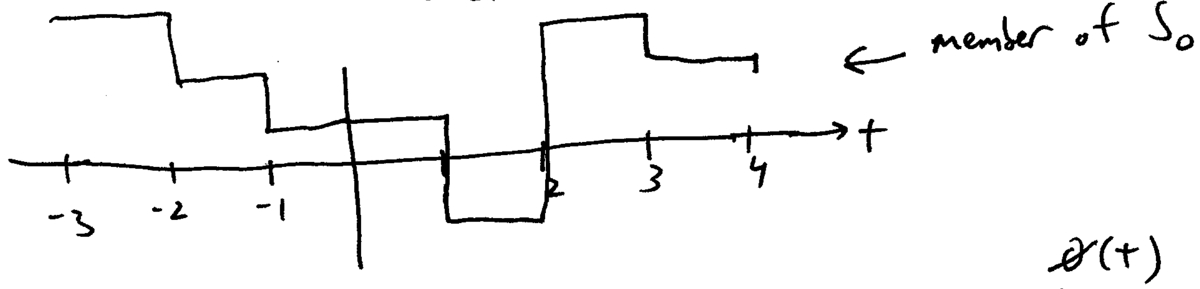
$$\psi_0(t) = 2\cos\left(\frac{3\pi t}{2}\right)\frac{\sin(\pi t/2)}{\pi t}$$

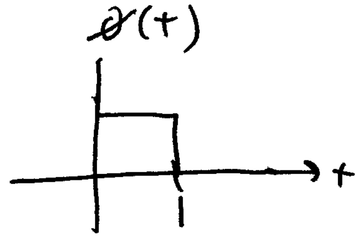


Example: Haar Wavelets

Above, we defined an $\perp$-basis by looking at the difference between approximations at subsequent dyadic scales.
We can do something very similar with piecewise constant functions in the time domain.

Let $S_0 = \{$ signals that are piecewise constant between the integers $\}$





Set $$\phi_0(t) = \begin{cases} 1 & 0 \leq t \leq 1 \\ 0 & \text{else} \end{cases}$$

It is clear that

$\{\phi_0(t-k)\}_{k\in\mathbb{Z}}$ is an $\perp$-basis for S_0

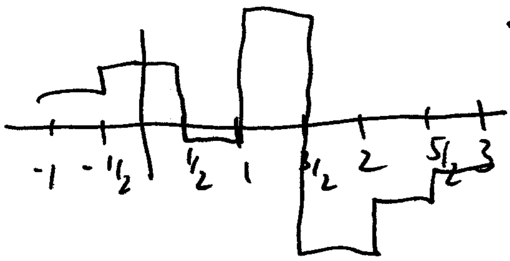
64

Also, given an arbitrary $x(t) \in L_2(\mathbb{R})$, we can find the best piecewise constant approximation to $x(t)$ by projecting onto S_0:

$$P_{S_0}[x(t)] = \sum_{k=-\infty}^{\infty} \langle x(t), \phi_0(t-k) \rangle \, \phi_0(t-k)$$

Now let

$$S_j = \{ \text{signals that are piecewise constant on segments of length } 2^{-j} \}$$



← member of S_1

If we set

$$\phi_j(t) = 2^{j/2} \phi_0(2^j t)$$

then

$$\{ \phi_j(t - 2^{-j}k) \}_{k \in \mathbb{Z}} = \{ 2^{j/2} \phi_0(2^j t - k) \}_{k \in \mathbb{Z}}$$

is an $\perp$-basis for S_j.

Also

$$\lim_{j\to\infty} P_{S_j}[x(t)] = x(t)$$

From Burrus et al:

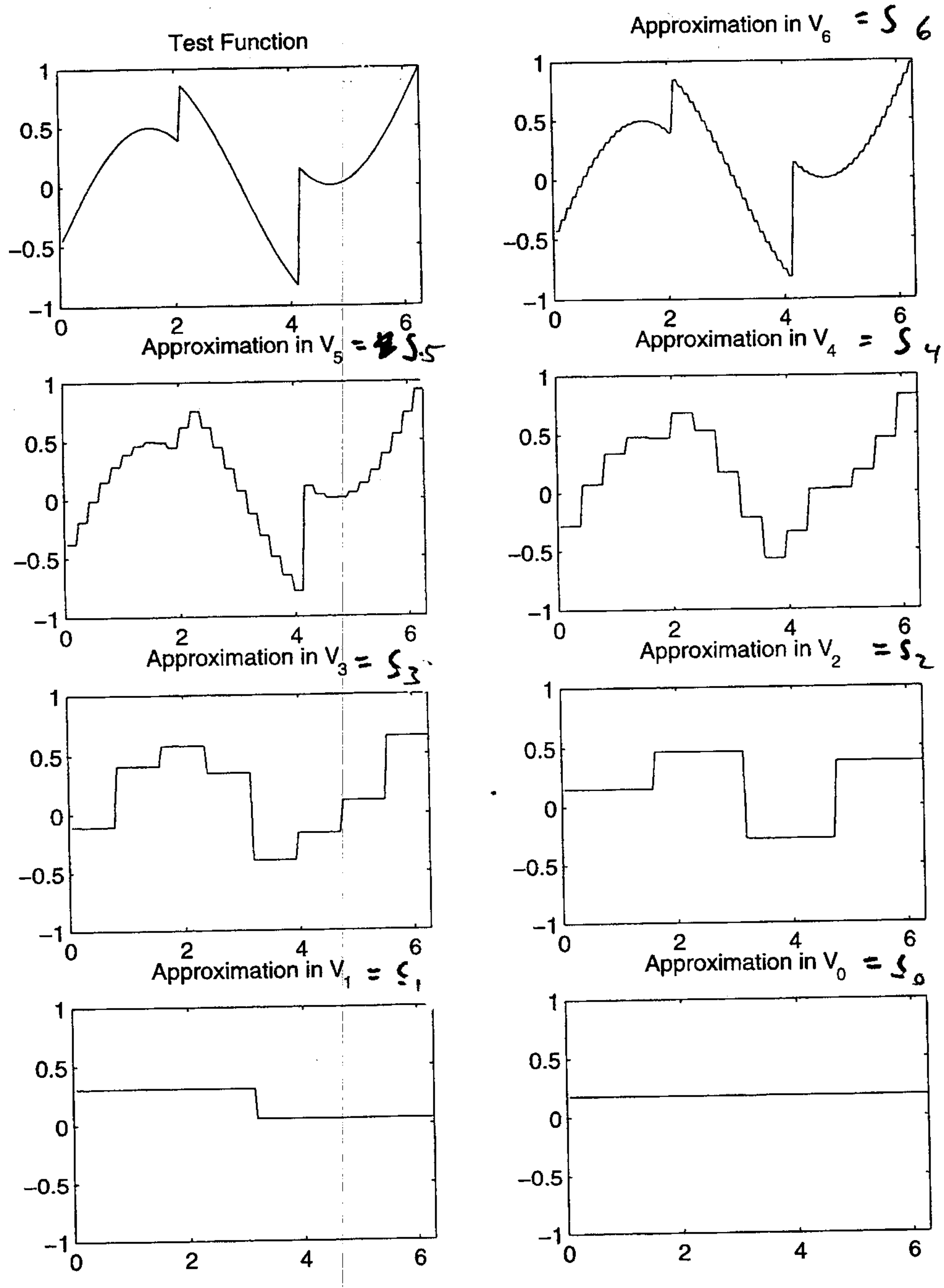


Figure 2.15. Haar Function Approximation in $\mathcal{V}_j$

66

Now set

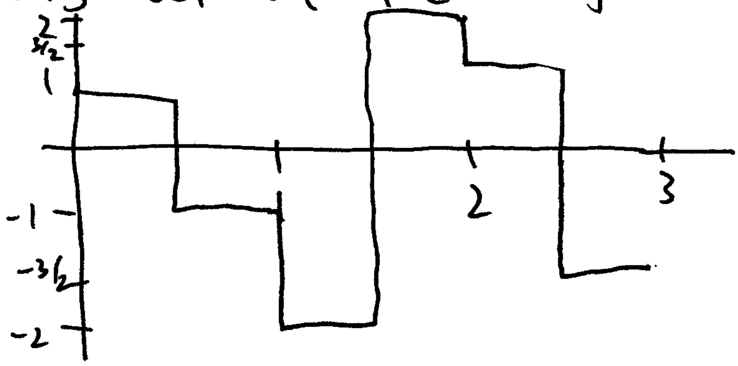
$$W_0 = S_1 \ominus S_0$$

$$= \{\text{signals in } S_1 \text{ that are } \perp \text{ to } S_0\}$$

and more generally

$$W_j = S_{j+1} \ominus S_j$$

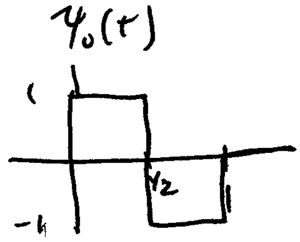
W_0 contains signals that are piecewise constant between the half integers <u>and</u> have zero mean on intervals between the integers



member of W_0

If we define

$$\psi_0(t) = \begin{cases} 1 & 0 \le t \le 1/2 \\ -1 & 1/2 \le t \le 1 \\ 0 & \text{else} \end{cases}$$



then $\{\psi_0(t-k)\}_{k \in \mathbb{Z}}$ is an $\perp$-basis for W_0

If we define

$$\psi_j(t) = 2^{j/2}\psi_0(2^j t)$$

then

$$\{\psi_j(t - 2^{-j}k)\}_{k\in\mathbb{Z}} = \{2^{j/2}\psi_0(2^j t - k)\}_{k\in\mathbb{Z}}$$

is an $\perp$-basis for W_j.

We can now write any $x(t) \in L_2(\mathbb{R})$ as

$$x(t) = \underbrace{\underbrace{\underbrace{P_{S_0}[x(t)] + P_{W_0}[x(t)]}_{P_{S_1}[x(t)]} + P_{W_1}[x(t)]}_{P_{S_2}[x(t)]} + P_{W_2}[x(t)]}_{P_{S_3}[x(t)]} + \cdots$$

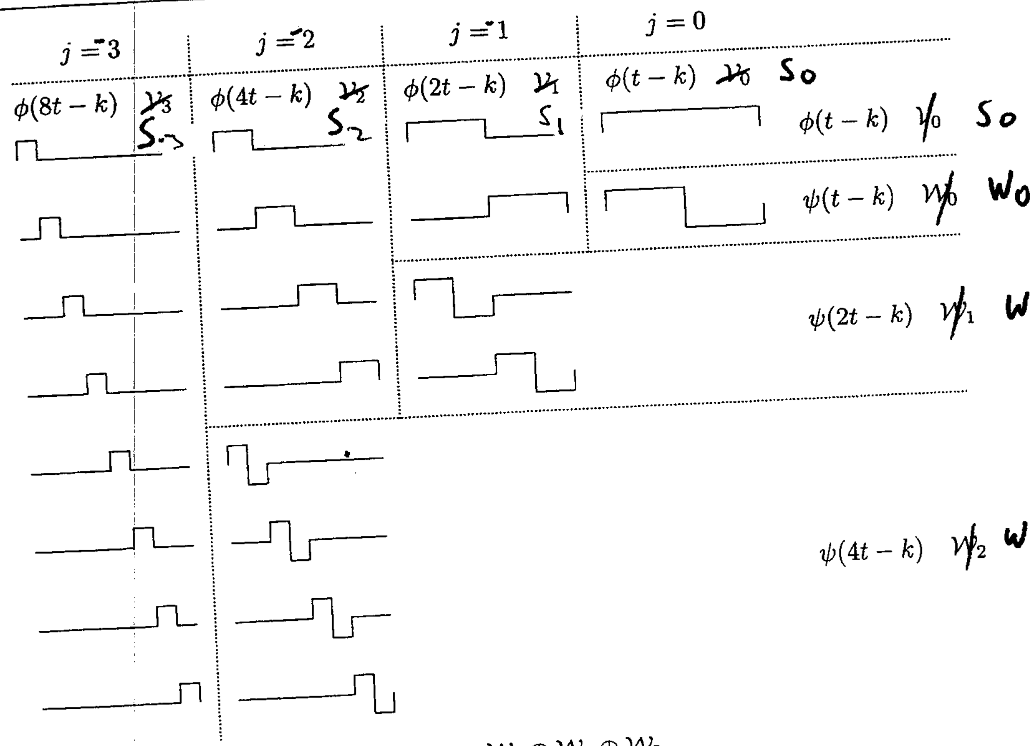
i.e.

$$x(t) = \sum_{k=-\infty}^{\infty} \langle x(t), \phi_{0,k}(t)\rangle \phi_{0,k}(t) + \sum_{j\geq 0}\sum_{k=-\infty}^{\infty} \langle x, \psi_{j,k}\rangle \psi_{j,k}(t)$$

where as before

$$\phi_{0,k}(t) = \phi_0(t-k)$$
$$\psi_{j,k}(t) = \psi_j(t - 2^{-j}k) = 2^{j/2}\psi_0(2^j t - k)$$

Burrus et. al.



$$\mathcal{V}_3 = \mathcal{V}_0 \oplus \mathcal{W}_0 \oplus \mathcal{W}_1 \oplus \mathcal{W}_2$$

$$S_{-3} = S_0 \oplus W_0 \oplus W_{-1} \oplus W_{-2}$$

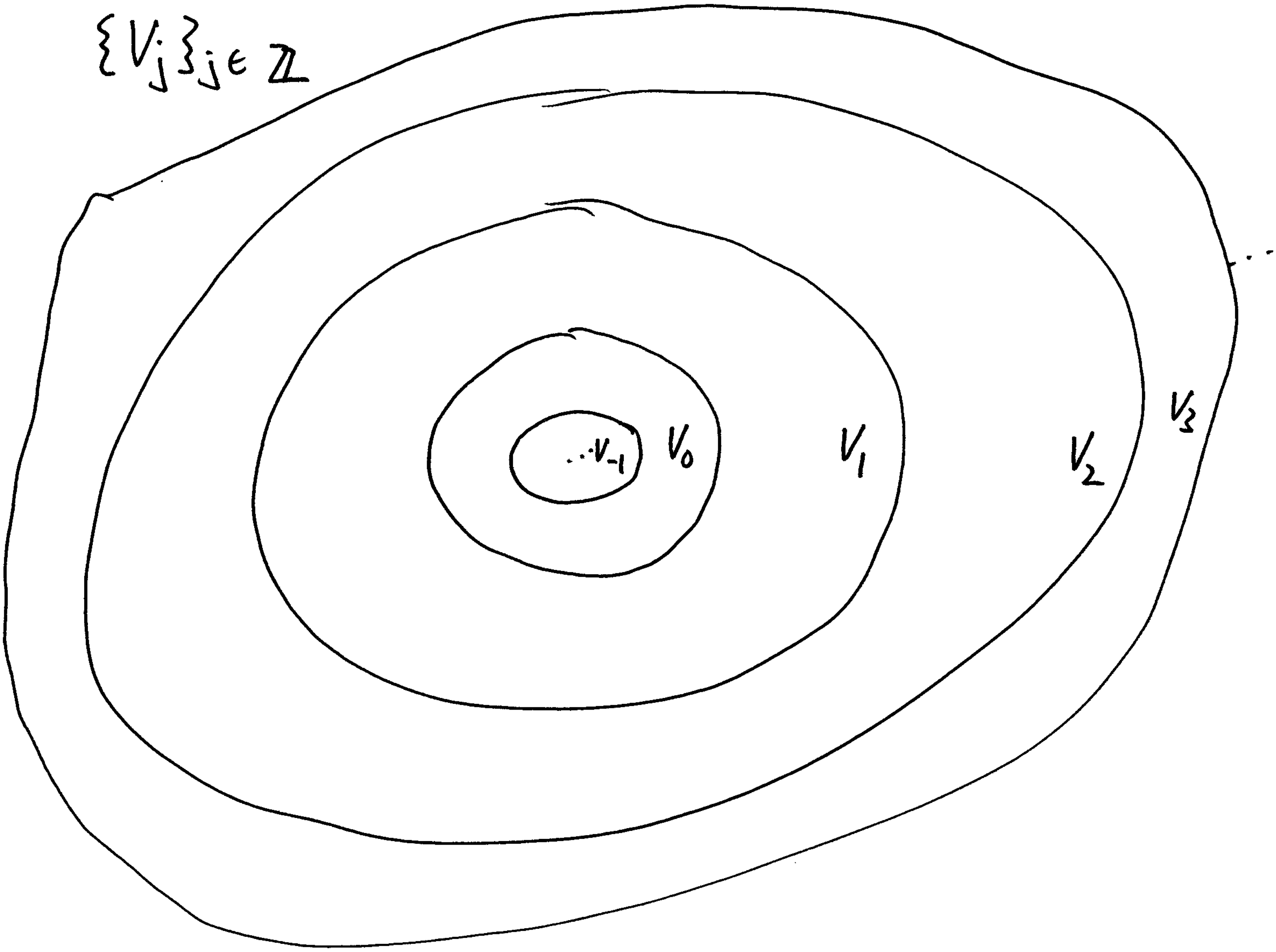
Figure 2.16. Haar Scaling Functions and Wavelets Decomposition of $\mathcal{V}_3$

69

Multiresolution Approximations (MRAs)

Abstracting the key properties of the Shannon and Haar wavelet systems above will give us a general framework for orthonormal wavelet transforms.

An MRA is a sequence of <u>nested</u> <u>subspaces</u> $\{V_j\}_{j\in\mathbb{Z}}$



that obey certain properties:

① $V_j \subset V_{j+1}$

② $\lim_{j\to\infty} V_j = L_2(\mathbb{R})$

(i.e. $\lim_{j\to\infty} P_{V_j}[x(t)] = x(t) \quad \forall\, x(t) \in L_2(\mathbb{R})$)

③ $\lim_{j\to -\infty} V_j = \{0\}$ (only the zero function)

④ $\forall\, j,k \in \mathbb{Z} \quad x(t) \in V_j \iff x(t - 2^{-j}k) \in V_j$

(if V_j contains a signal, it also contains shifts of that signal of integer multiples of 2^{-j})

⑤ $x(t) \in V_j \iff x(2t) \in V_{j+1}$

⑥ There exists a $\phi_0(t) \in V_0$ such that

$\{\phi_0(t-k)\}_{k\in\mathbb{Z}}$ is an $\perp$-basis for V_0

Combined with the other conditions, 6 tells us that

$$\{2^{j/2}\phi_0(2^j t-k)\}_{k\in\mathbb{Z}} \text{ is an } \perp\text{-basis for } V_j$$

As such, we can compute the best approximation to $x(t)$ in V_j using

$$P_{V_j}[x(t)] = \sum_{k=-\infty}^{\infty} \langle x(t), \phi_{j,k}(t)\rangle \phi_{j,k}(t)$$

where

$$\phi_{j,k}(t) = 2^{j/2}\phi_0(2^j t-k)$$

Moreover, if we define

$$\begin{aligned} W_j &= V_{j+1} \ominus V_j \\ &= \text{everything in } V_{j+1} \text{ that is } \perp \text{ to } V_j \end{aligned}$$

then it can be shown that there exists a

$$\psi_0(t) \in W_0$$

such that

$$\{\psi_0(t-k)\}_{k\in\mathbb{Z}} \text{ is an } \perp\text{-basis for } W_0$$

Combined with the properties above, this means that

$$\{2^{j/2}\psi_0(2^jt-k)\}_{k\in\mathbb{Z}} \text{ is an } \perp\text{-basis for } W_j$$

As before, we will use the notation

$$\psi_{j,k}(t) = 2^{j/2}\psi_0(2^jt-k).$$

Nomenclature

The V_j are called scaling spaces, and the $\phi_{j,k}(t)$ are called scaling functions.

The W_j are called wavelet spaces (or detail spaces) and the $\psi_{j,k}(t)$ are called wavelets.

Decomposing $L_2(\mathbb{R})$

We also have that

$$L_2(\mathbb{R}) = \underbrace{\underbrace{V_0 \oplus W_0}_{V_1} \oplus W_1}_{V_2 \;\cdots} \oplus W_2 \oplus W_3 \oplus \cdots$$

or in general we can start at any scale J.

$$L_2(\mathbb{R}) = V_J \oplus W_J \oplus W_{J+1} \oplus W_{J+2} \oplus \cdots$$

We can decompose any $x(t) \in L_2(\mathbb{R})$ as

$$x(t) = P_{V_0}[x(t)] + P_{W_0}[x(t)] + P_{W_1}[x(t)] + P_{W_2}[x(t)] + \cdots$$

$$= \sum_{k=-\infty}^{\infty} \langle x(t), \phi_{0,k}(t)\rangle \phi_{0,k}(t) + \sum_{j\geq 0} \sum_{k=-\infty}^{\infty} \langle x(t), \psi_{j,k}(t)\rangle \psi_{j,k}(t)$$

$$= \sum_{k} c_{0,k}\, \phi_{0,k}(t) + \sum_{j\geq 0} \sum_{k=-\infty}^{\infty} d_{j,k}\, \psi_{j,k}(t)$$

We call

$c_{j,k} = \langle x, \phi_{j,k} \rangle$ the scaling coefficients at scale j

and $d_{j,k} = \langle x, \psi_{j,k} \rangle$ the wavelet coefficients at scale j

The mapping

$$x(t) \longmapsto \{c_{0,k}\}_{k\in\mathbb{Z}},\ \{d_{j,k}\}_{j\geq 0,\, k\in\mathbb{Z}}$$

is called the wavelet transform.

Again, we can initiate this at any scale J, not just $J=0$.

Moving from scale to scale

Let's look closely at the relationship between scales $j=0$ and $j=-1$ (what we learn can be immediately generalized to j and $j-1$ for any j).

First, since $V_{-1} \subset V_0$, we know we can write

$$\phi_{-1}(t) = \frac{1}{\sqrt{2}}\phi_0(t/2)$$

as a linear combination of $\{\phi_0(t-\ell)\}_\ell$.

That is, a <u>dilated</u> version of $\phi_0(t)$ can be written as a combination of <u>shifts</u> of $\phi_0(t)$

$$\begin{aligned}\phi_{-1}(t) = \frac{1}{\sqrt{2}}\phi_0(t/2) &= \sum_\ell \left\langle \frac{1}{\sqrt{2}}\phi_0(t/2), \phi_0(t-\ell)\right\rangle \phi_0(t-\ell) \\ &= \sum_\ell h[\ell]\,\phi_0(t-\ell)\end{aligned}$$

The sequence $h[\ell]$ contains the <u>expansion</u> <u>coefficients</u> for $\phi_{-1}(t)$ in V_0.

It is true (and you can check this) that the same $h[\ell]$ can be used at every scale.

In general

$$\phi_{j-1,0}(t) = \frac{1}{\sqrt{2}}\phi_{j,0}(t) = \sum_{\ell} h[\ell]\,\phi_{j,\ell}(t)$$

We can do the same thing for any other shift of $\phi_{-1}(t)$

$$\begin{aligned}
\phi_{-1,k}(t) &= \frac{1}{\sqrt{2}}\phi_0(t/2-k) \\
&= \sum_{\ell} \left\langle \tfrac{1}{\sqrt{2}}\phi_0(t/2-k), \phi_0(t-\ell)\right\rangle \phi_0(t-\ell) \\
&= \sum_{\ell} \left\langle \tfrac{1}{\sqrt{2}}\phi_0(t/2), \phi_0(t-\ell+2k)\right\rangle \phi_0(t-\ell) \\
&= \sum_{\ell} h[\ell-2k]\,\phi_0(t-\ell)
\end{aligned}$$

and in general

$$\phi_{j-1,k}(t) = \sum_{\ell} h[\ell-2k]\,\phi_{j,\ell}(t)$$

These relationships give us a very nice way to move from an approximation in V_0

$$P_{V_0}[x(t)] = \sum_k c_{0,k}\,\phi_{0,k}(t)$$

to an approximation in V_{-1}

$$P_{V_{-1}}[x(t)] = \sum_k c_{-1,k}\,\phi_{-1,k}(t)$$

Notice:

$$\begin{aligned} c_{-1,k} &= \langle x(t), \phi_{-1,k}(t) \rangle \\ &= \langle x(t), \sum_\ell h[\ell-2k]\,\phi_{0,\ell}(t) \rangle \\ &= \sum_\ell h[\ell-2k]\,\langle x(t), \phi_{0,\ell}(t) \rangle \\ &= \sum_\ell h[\ell-2k]\,c_{0,\ell} \end{aligned}$$

To make this look a little more familiar, we write

$$c_{0,\ell} = c_0[\ell]$$

Then

$$c_{-1}[k] = \sum_{\ell} h[\ell-2k]\, c_0[\ell]$$

$$= \Big(c_0[k] * h[-k] \Big) \text{ down sampled by } 2$$

$$c_0[k] \longrightarrow \boxed{H_0} \longrightarrow \boxed{\downarrow 2} \longrightarrow c_{-1}[k]$$

where H_0 has impulse response

$$h_0[k] = h[-k] = \langle \phi_{-1}(t), \phi_0(t+k) \rangle$$

$$= \langle \tfrac{1}{\sqrt{2}} \phi_0(t/2), \phi(t+k) \rangle$$

Similarly, since $\psi_{-1}(t) \in V_0$

$$\psi_{-1}(t) = \sum_{\ell} \underbrace{\langle \psi_{-1}(t), \phi_0(t-\ell) \rangle}_{g[\ell]} \phi(t-\ell)$$

and

$$\psi_{-1,k}(t) = \sum_{\ell} \langle \tfrac{1}{\sqrt{2}} \psi_0(t/2 - k), \phi_0(t-\ell) \rangle \phi_0(t-\ell)$$

$$= \sum_{\ell} g[\ell-2k]\, \phi_0(t-\ell)$$

Thus

$$P_{W_{-1}}[x(t)] = \sum_k d_{-1,k}\, \psi_{-1,k}(t)$$

where

$$\begin{aligned}
d_{-1,k} &= \langle x(t), \psi_{-1,k}(t) \rangle \\
&= \langle x(t), \sum_\ell g[\ell-2k]\, \phi_0(t-\ell) \rangle \\
&= \sum_\ell g[\ell-2k] \langle x(t), \phi_0(t-\ell) \rangle \\
&= \sum_\ell g[\ell-2k]\, c_0[\ell]
\end{aligned}$$

So

$$c_0[k] \longrightarrow \boxed{H_1} \longrightarrow \boxed{\downarrow 2} \longrightarrow d_1[k]$$

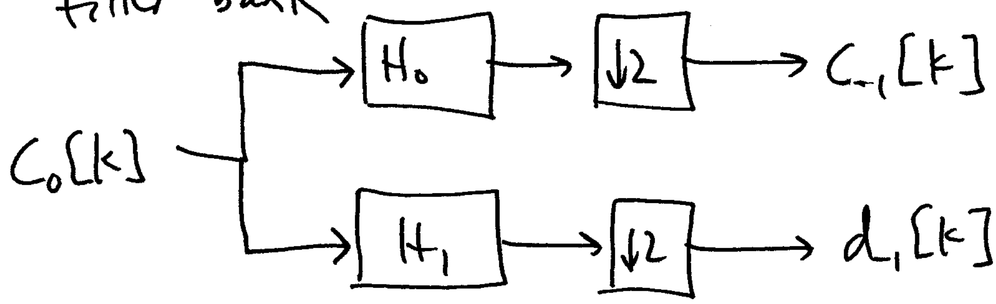
where H_1 has impulse response

$$h_1[k] = g[-k] = \left\langle \tfrac{1}{\sqrt{2}} \psi_0(t/2), \phi_0(t+k) \right\rangle$$

Thus we can interpret

$$P_{V_0}[x(t)] = P_{V_{-1}}[x(t)] + P_{W_{-1}}[x(t)]$$

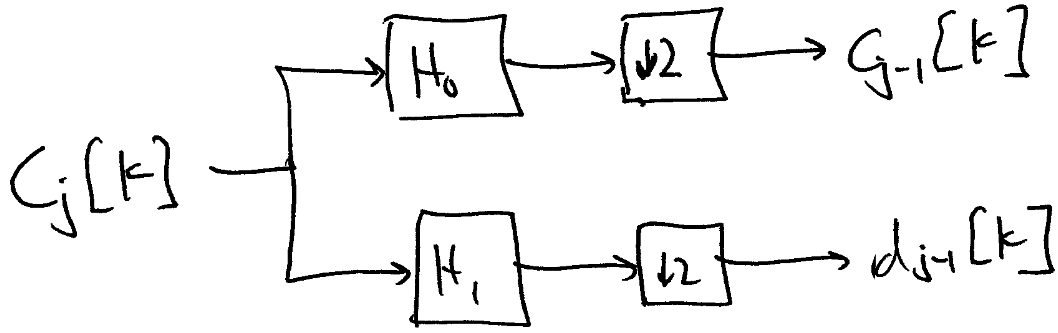
as a filter bank



and more generally

$$P_{V_j}[x(t)] = P_{V_{j-1}}[x(t)] + P_{W_{j-1}}[x(t)]$$

as



Wavelet basis functions and filter banks

Given $\phi_0(t)$, $\psi_0(t)$ we know how to construct the corresponding filters:

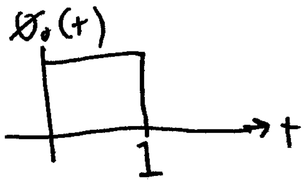
$$h[n] = \langle \phi_{-1}(t), \phi_0(t-n) \rangle$$
$$= \langle \tfrac{1}{\sqrt{2}} \phi_0(t/2), \phi_0(t-n) \rangle$$

$$g[n] = \langle \psi_1(t), \phi_0(t-n) \rangle$$
$$= \langle \tfrac{1}{\sqrt{2}} \psi(t/2), \phi_0(t-n) \rangle$$

and set

$$h_0[n] = h[-n] \qquad h_1[n] = g[-n]$$

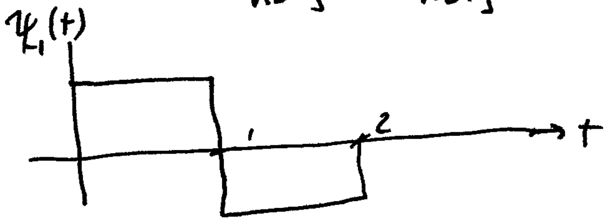
Example: Haar



,

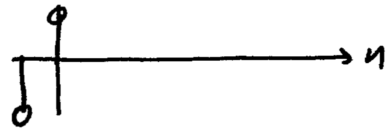
$\phi_{-1}(t) = \frac{1}{\sqrt{2}} \phi_0(t/2)$

$$\phi_{-1}(t) = \underbrace{\frac{1}{\sqrt{2}}}_{h[0]} \phi_0(t) + \underbrace{\frac{1}{\sqrt{2}}}_{h[1]} \phi_0(t-1)$$



$$\psi_{-1}(t) = \underbrace{\frac{1}{\sqrt{2}}}_{g[0]} \phi_0(t) \underbrace{- \frac{1}{\sqrt{2}}}_{g[1]} \phi_1(t)$$

$h_0[n] =$ 

$h_1[n] =$ 

This is just a shifted version of what we used for the Haar filterbank earlier in the course (h_1 is also negated, but this makes no effective difference since it's symmetric).

Example: Shannon (see the homework)

This connection between basis functions and filter banks also goes the other way. Given a CMF, there is a corresponding scaling function $\phi_0(t)$ and wavelet $\psi_0(t)$.

In general, given $h_0[n]$ (and the corresponding $h_1[n]$), $\phi_0(t)$ (and $\psi_0(t)$) cannot be written down explicitly — rather it is defined by the recursion

$$\frac{1}{\sqrt{2}}\phi_0(t/2) = \sum_{\ell} \underbrace{h[\ell]}_{=h_0[-\ell]}\phi_0(t-\ell)$$

In the Fourier domain, this gives us

$$\begin{aligned}\sqrt{2}\,\Phi_0(j2\Omega) &= \sum_{\ell} h[\ell]\,e^{-j\Omega\ell}\,\Phi_0(j\Omega) \\ &= \left(\Phi_0(j\Omega)\right)\cdot\left(\sum_{\ell} h[\ell]\,e^{-j\Omega\ell}\right) \\ &= \Phi_0(j\Omega)\cdot H(e^{j\Omega})\end{aligned}$$

CTFT of $\phi_0(t)$ — 2π-periodic DTFT of $h[n]$

i.e.

$$\Phi_0(j2\Omega) = \frac{1}{\sqrt{2}}H(e^{j\Omega})\,\Phi_0(j\Omega)$$

or

$$
\begin{aligned}
\Phi_0(j\Omega) &= \frac{1}{\sqrt{2}} H(e^{j\Omega/2})\, \Phi_0(j\Omega/2) \\
&= \frac{1}{\sqrt{2}} H(e^{j\Omega/2})\, \frac{1}{\sqrt{2}} H(e^{j\Omega/4})\, \Phi_0(j\Omega/4) \\
&= \frac{1}{\sqrt{2}} H(e^{j\Omega/2})\, \frac{1}{\sqrt{2}} H(e^{j\Omega/4}) \cdot \frac{1}{\sqrt{2}} H(e^{j\Omega/8})\, \Phi_0(j\Omega/8) \\
&\;\;\vdots \\
&= \left(\prod_{p=1}^{P} \frac{1}{\sqrt{2}} H(e^{j2^{-p}\Omega}) \right) \cdot \Phi_0(j2^{-P}\Omega)
\end{aligned}
$$

for all P. As $P \to \infty$, this expression becomes

$$
= \left[\prod_{p=1}^{\infty} \frac{1}{\sqrt{2}} H(e^{j2^{-p}\Omega}) \right] \cdot \Phi_0(j\cdot 0)
$$

So given $h_0[n]$ (and hence $h[n]$ and hence $H(e^{j\Omega})$) we can define $\phi_0(t)$ in the Fourier domain using this <u>infinite</u> <u>product</u>.

Given $\phi_0(t)$, we can always construct

$$\psi_0(t) = \sum_\ell g[\ell]\, \phi_1(t-\ell)$$

$$= \sqrt{2} \sum_\ell g[\ell]\, \phi_0(2t-\ell)$$

where $\quad g[\ell] = h_1[-\ell] = (-1)^{1-\ell} h_0[1-\ell]$.

D4 $\to$ $p=2$

D6 $\to$ $p=3$

⋮

etc.

	n	$h_p[n]$
p=2	0	.482962913145
	1	.836516303738
	2	.224143868042
	3	−.129409522551
p=3	0	.332670552950
	1	.806891509311
	2	.459877502118
	3	−.135011020010
	4	−.085441273882
	5	.035226291882
p=4	0	.230377813309
	1	.714846570553
	2	.630880767930
	3	−.027983769417
	4	−.187034811719
	5	.030841381836
	6	.032883011667
	7	−.010597401785
p=5	0	.160102397974
	1	.603829269797
	2	.724308528438
	3	.138428145901
	4	−.242294887066
	5	−.032244869585
	6	.077571493840
	7	−.006241490213
	8	−.012580751999
	9	.003335725285
p=6	0	.111540743350
	1	.494623890398
	2	.751133908021
	3	.315250351709
	4	−.226264693965
	5	−.129766867567
	6	.097501605587
	7	.027522865530
	8	−.031582039317
	9	.000553842201
	10	.004777257511
	11	−.001077301085
p=7	0	.077852054085
	1	.396539319482
	2	.729132090846
	3	.469782287405
	4	−.143906003929
	5	−.224036184994
	6	.071309219267
	7	.080612609151
	8	−.038029936935
	9	−.016574541631
	10	.012550998556
	11	.000429577973
	12	−.001801640704

	n	$h_p[n]$
p=8	0	.054415842243
	1	.312871590914
	2	.675630736297
	3	.585354683654
	4	−.015829105256
	5	−.284015542962
	6	.000472484574
	7	.128747426620
	8	−.017369301002
	9	−.04408825393
	10	.013981027917
	11	.008746094047
	12	−.004870352993
	13	−.000391740373
	14	.000675449406
	15	−.000117476784
p=9	0	.038077947364
	1	.243834674613
	2	.604823123690
	3	.657288078051
	4	.133197385825
	5	−.293273783279
	6	−.096840783223
	7	.148540749338
	8	.030725681479
	9	−.067632829061
	10	.000250947115
	11	.022361662124
	12	−.004723204758
	13	−.004281503682
	14	.001847646883
	15	.000230385764
	16	−.000251963189
	17	.000039347320
p=10	0	.026670057901
	1	.188176800078
	2	.527201188932
	3	.688459039454
	4	.281172343661
	5	−.249846424327
	6	−.195946274377
	7	.127369340336
	8	.093057364604
	9	−.071394147166
	10	−.029457536822
	11	.033212674059
	12	.003606553567
	13	−.010733175483
	14	.001395351747
	15	.001992405295
	16	−.000685856695
	17	−.000116466855
	18	.000093588670
	19	−.000013264203

From Mallat, Chap 7

These are the Daubechies CMF filter coeffts from Sec I of the notes.

Each corresponds to a different MRA/wavelet system, called

"D-2p"

D2 = Haar

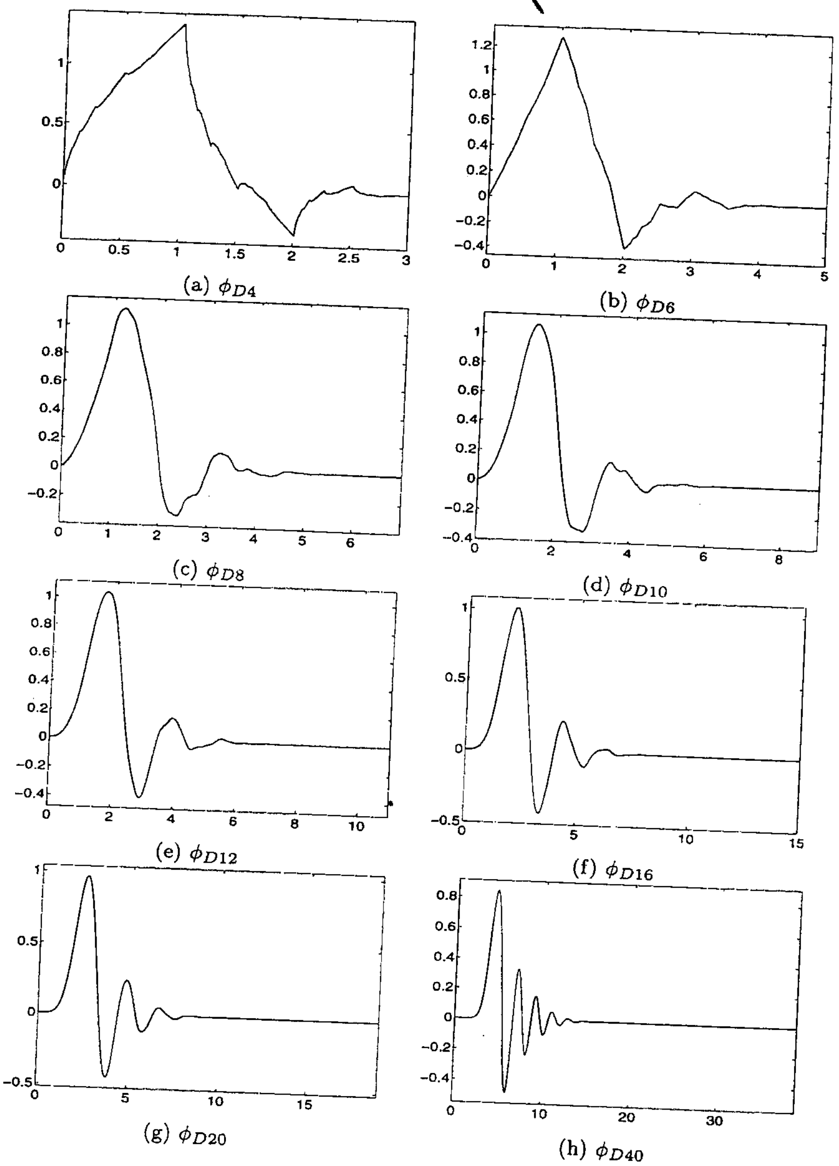
iscrete Moments

:ems in general as
1) and (5.10) that
0) $= m_1(0) = 0$ of
:. Ortho*normality*
equires $m(0) = 1$
legrees of freedom
;th-6 we have two
s we have exactly
$N/2$ zero wavelet
ationships among
xplained in (6.52)

s of the wavelets
ions. Figures 6.1
, $10, 12, 16, 20, 40$.
Chapter 6. The
ıd in [Dau92] or
For $N = 2$, the
ble; for $N = 6$, it
for longer $h(n)$.

oments that are
:e the degrees of

Examples of scaling functions

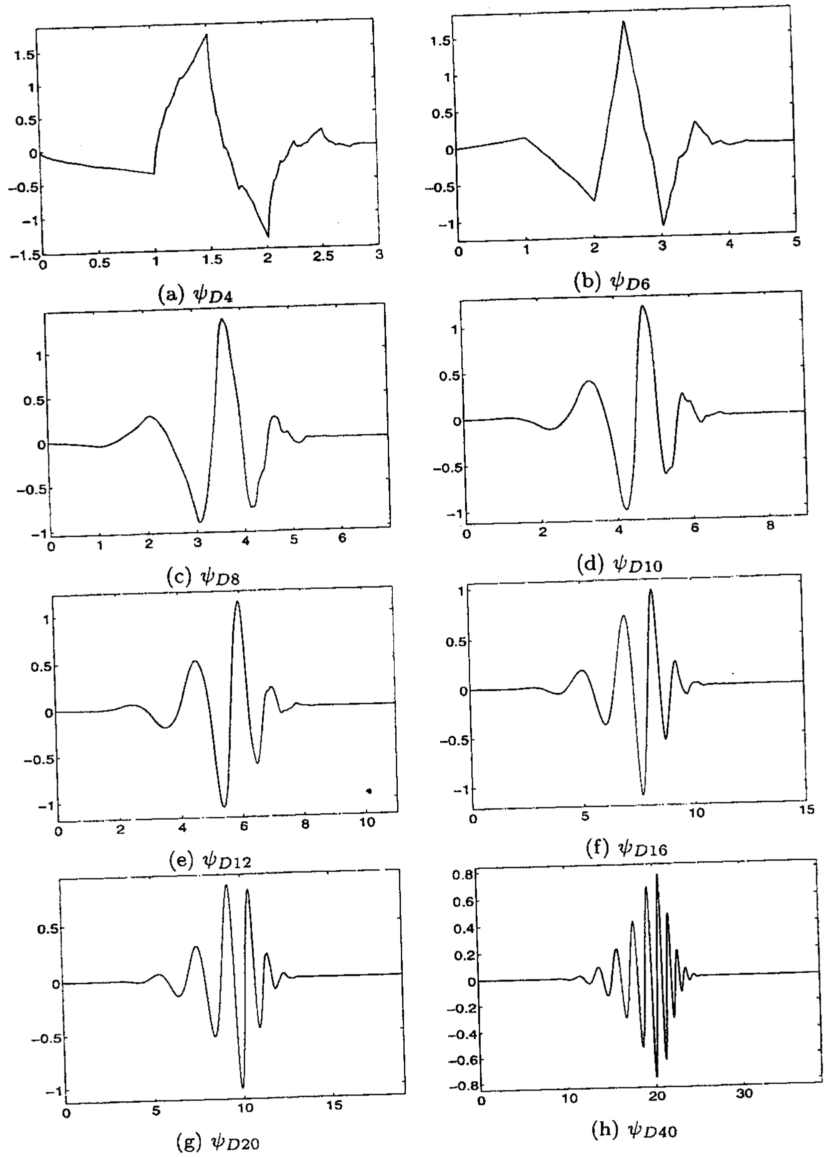


(a) ϕ_{D4} (b) ϕ_{D6} (c) ϕ_{D8} (d) ϕ_{D10} (e) ϕ_{D12} (f) ϕ_{D16} (g) ϕ_{D20} (h) ϕ_{D40}

Figure 6.1. Daubechies Scaling Functions, $N = 4, 6, 8, \ldots, 40$

freedom to maximize the differentiability of $\varphi(t)$ rather than maximize the zero moments. This is not easily parameterized, and it gives only slightly greater smoothness than the Daubechies system [Dau92].

Examples of wavelets (from Burrus et al)



(a) ψ_{D4} (b) ψ_{D6} (c) ψ_{D8} (d) ψ_{D10} (e) ψ_{D12} (f) ψ_{D16} (g) ψ_{D20} (h) ψ_{D40}

Figure 6.2. Daubechies Wavelets, $N = 4, 6, 8, \ldots, 40$

Examples of Daubechies scaling functions resulting from choosing different factors in the spectral factorization of $|H(\omega)|^2$ in (6.18) can be found in [Dau92].

2.7 Examples of Wavelet Expansions

In this section, we will try to show the way a wavelet expansion decomposes a signal and what the components look like at different scales. These expansions use what is called a length-8 Daubechies basic wavelet (developed in Chapter 6), but that is not the main point here. The local nature of the wavelet decomposition is the topic of this section.

These examples are rather standard ones, some taken from David Donoho's papers and web page. The first is a decomposition of a piecewise linear function to show how edges and constants are handled. A characteristic of Daubechies systems is that low order polynomials are completely contained in the scaling function spaces $\mathcal{V}_j$ and need no wavelets. This means that when a section of a signal is a section of a polynomial (such as a straight line), there are no wavelet expansion coefficients $d_j(k)$, but when the calculation of the expansion coefficients overlaps an edge, there is a wavelet component. This is illustrated well in Figure 2.6 where the high resolution scales gives a very accurate location of the edges and this spreads out over k at the lower scales. This gives a hint of how the DWT could be used for edge detection and how the large number of small or zero expansion coefficients could be used for compression.

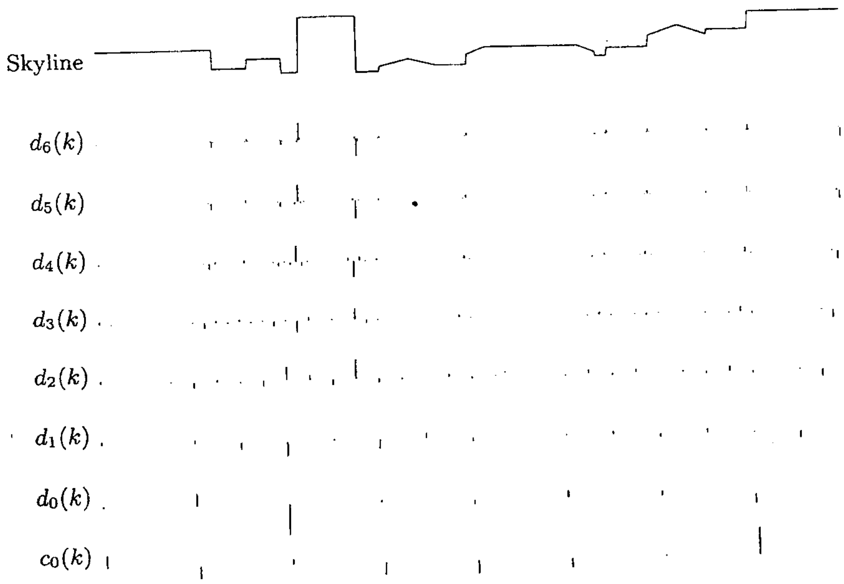


Figure 2.5. Discrete Wavelet Transform of the Houston Skyline, using $\psi_{D8'}$ with a Gain of $\sqrt{2}$ for Each Higher Scale

Figure 2.6 shows the approximations of the skyline signal in the various scaling function spaces $\mathcal{V}_j$. This illustrates just how the approximations progress, giving more and more resolution at higher scales. The fact that the higher scales give more detail is similar to Fourier methods, but the localization is new. Figure 2.7 illustrates the individual wavelet decomposition by showing

88

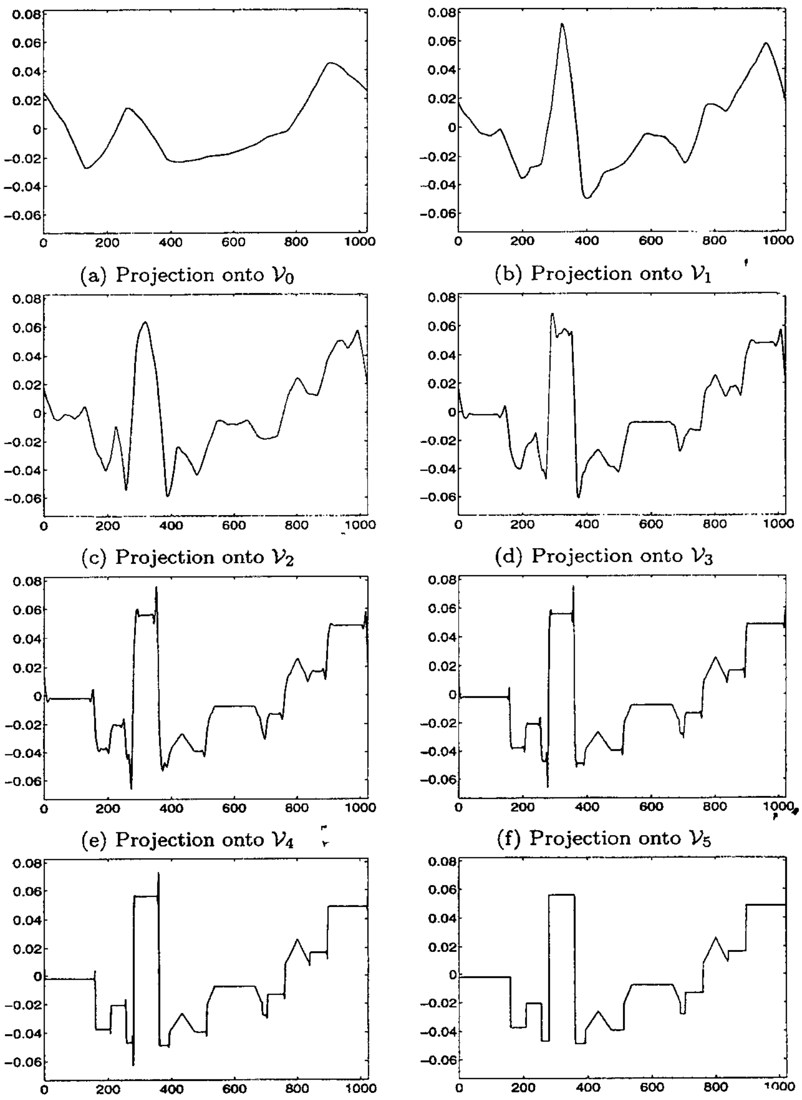
a signal and what called a length-8 . point here. The

s papers and web ges and constants als are completely hat when a section wavelet expansion ps an edge, there a resolution scales lower scales. This e number of small

with a Gain of $\sqrt{2}$

ing function spaces more resolution at urier methods, but osition by showing

the components of the signal that exist in the wavelet spaces $\mathcal{W}_j$ at different scales j. This shows the same expansion as Figure 2.6, but with the wavelet components given separately rather than being cumulatively added to the scaling function. Notice how the large objects show up at the lower resolution. Groups of buildings and individual buildings are resolved according to their width. The edges, however, are located at the higher resolutions and are located very accurately.



(a) Projection onto $\mathcal{V}_0$ (b) Projection onto $\mathcal{V}_1$

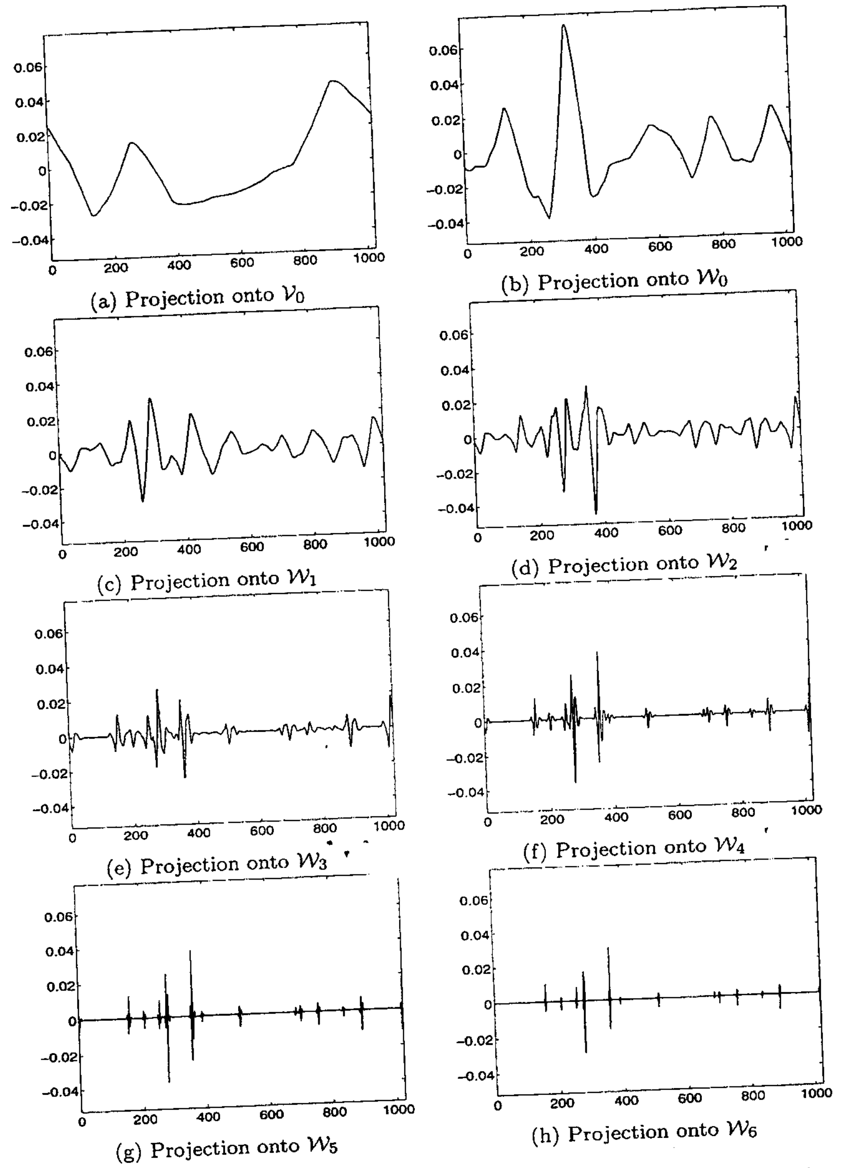
(c) Projection onto $\mathcal{V}_2$ (d) Projection onto $\mathcal{V}_3$

(e) Projection onto $\mathcal{V}_4$ (f) Projection onto $\mathcal{V}_5$

(g) Projection onto $\mathcal{V}_6$ (h) Projection onto $\mathcal{V}_7$

Figure 2.6. Projection of the Houston Skyline Signal onto $\mathcal{V}$ Spaces using $\phi_{D8'}$

89



(a) Projection onto $\mathcal{V}_0$ (b) Projection onto $\mathcal{W}_0$

(c) Projection onto $\mathcal{W}_1$ (d) Projection onto $\mathcal{W}_2$

(e) Projection onto $\mathcal{W}_3$ (f) Projection onto $\mathcal{W}_4$

(g) Projection onto $\mathcal{W}_5$ (h) Projection onto $\mathcal{W}_6$

Figure 2.7. Projection of the Houston Skyline Signal onto $\mathcal{W}$ Spaces using $\psi_{D8'}$

The second example uses a chirp or doppler signal to illustrate how a time-varying frequency is described by the scale decomposition. Figure 2.8 gives the coefficients of the DWT directly as a function of j and k. Notice how the location in k tracks the frequencies in the signal in a way the Fourier transform cannot. Figures 2.9 and 2.10 show the scaling function approximations and the wavelet decomposition of this chirp signal. Again, notice in this type of display how the "location" of the frequencies are shown.

90

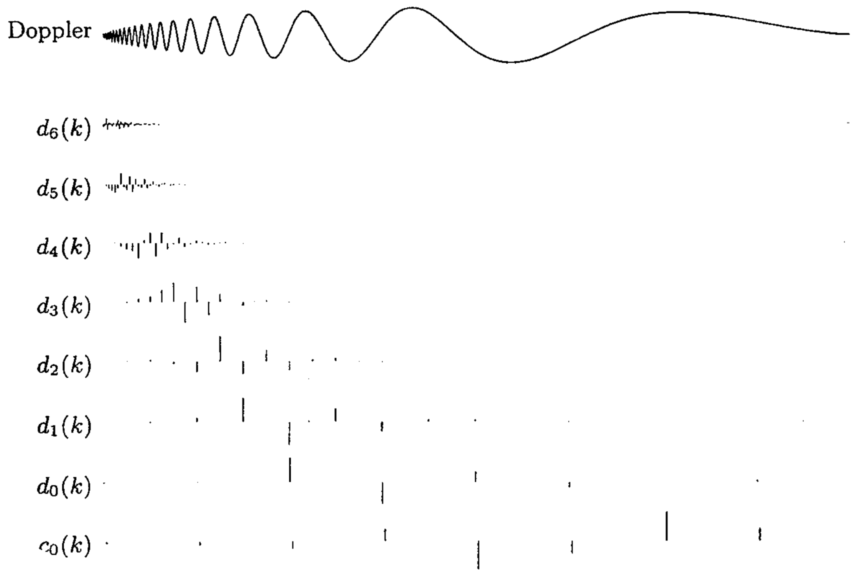


Figure 2.8. Discrete Wavelet Transform of a Doppler, using $\psi_{D8'}$ with a gain of $\sqrt{2}$ for each higher scale.

2.8 An Example of the Haar Wavelet System

In this section, we can illustrate our mathematical discussion with a more complete example. In 1910, Haar [Haa10] showed that certain square wave functions could be translated and scaled to create a basis set that spans L^2. This is illustrated in Figure 2.11. Years later, it was seen that Haar's system is a particular wavelet system.

If we choose our scaling function to have compact support over $0 \leq t \leq 1$, then a solution to (2.13) is a scaling function that is a simple rectangle function

$$\varphi(t) = \begin{cases} 1 & \text{if } 0 < t < 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.42}$$

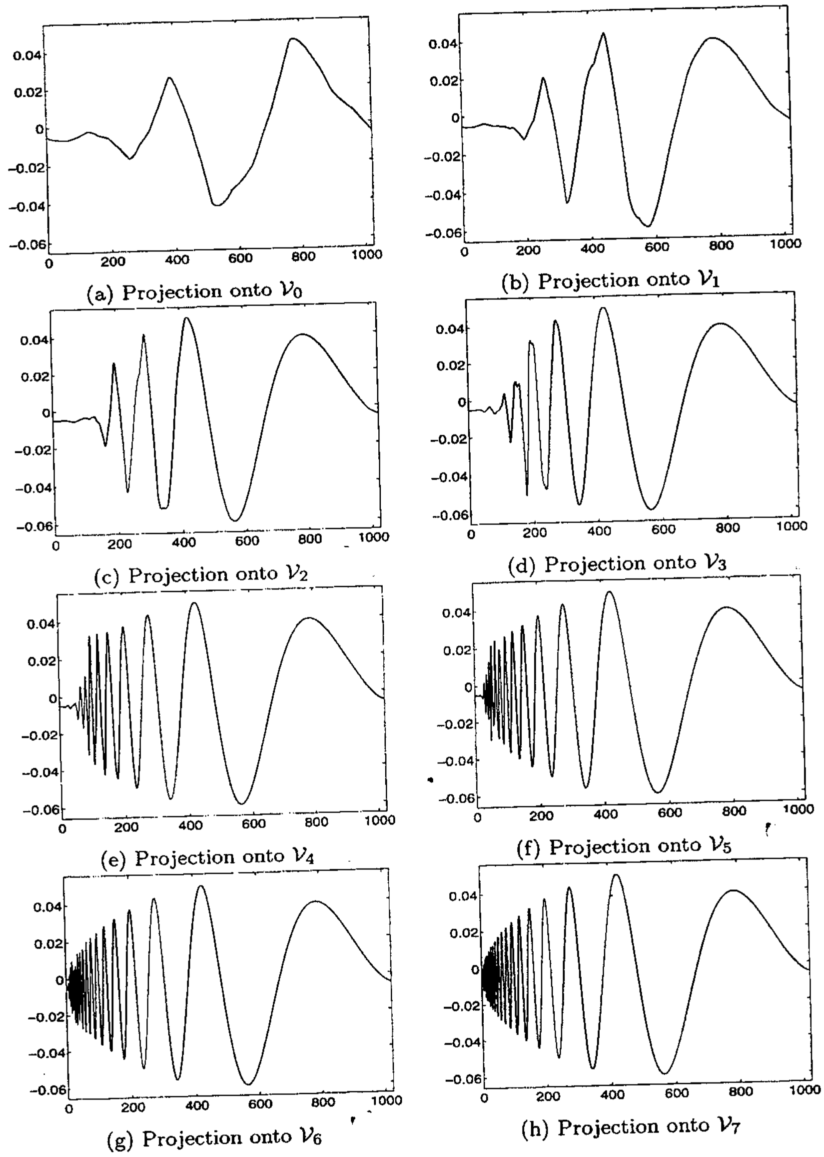
with only two nonzero coefficients $h(0) = h(1) = 1/\sqrt{2}$ and (2.24) and (2.25) require the wavelet to be

$$\psi(t) = \begin{cases} 1 & \text{for } 0 < t < 0.5 \\ -1 & \text{for } 0.5 < t < 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.43}$$

with only two nonzero coefficients $h_1(0) = 1/\sqrt{2}$ and $h_1(1) = -1/\sqrt{2}$.

$\mathcal{V}_0$ is the space spanned by $\varphi(t-k)$ which is the space of piecewise constant functions over integers, a rather limited space, but nontrivial. The next higher resolution space $\mathcal{V}_1$ is spanned by $\varphi(2t-k)$ which allows a somewhat more interesting class of signals which does include $\mathcal{V}_0$. As we consider higher values of scale j, the space $\mathcal{V}_j$ spanned by $\varphi(2^j t - k)$ becomes better able to approximate arbitrary functions or signals by finer and finer piecewise constant functions.

91



(a) Projection onto $\mathcal{V}_0$

(b) Projection onto $\mathcal{V}_1$

(c) Projection onto $\mathcal{V}_2$

(d) Projection onto $\mathcal{V}_3$

(e) Projection onto $\mathcal{V}_4$

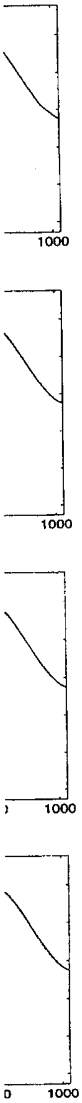
(f) Projection onto $\mathcal{V}_5$

(g) Projection onto $\mathcal{V}_6$

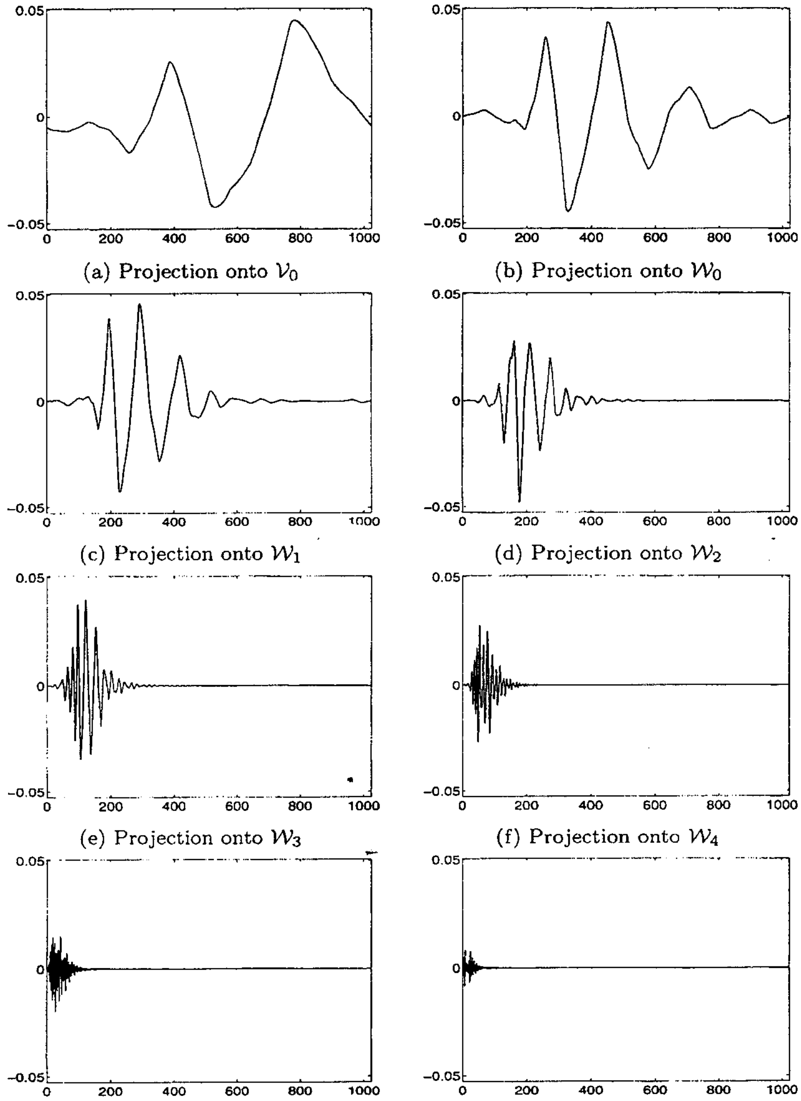
(h) Projection onto $\mathcal{V}_7$

Figure 2.9. Projection of the Doppler Signal onto $\mathcal{V}$ Spaces using $\phi_{D8'}$

92



ng $\phi_{D8'}$



(a) Projection onto $\mathcal{V}_0$ (b) Projection onto $\mathcal{W}_0$

(c) Projection onto $\mathcal{W}_1$ (d) Projection onto $\mathcal{W}_2$

(e) Projection onto $\mathcal{W}_3$ (f) Projection onto $\mathcal{W}_4$

(g) Projection onto $\mathcal{W}_5$ (h) Projection onto $\mathcal{W}_6$

Figure 2.10. Projection of the Doppler Signal onto $\mathcal{W}$ Spaces using $\psi_{D8'}$

93

The Daubechies wavelets, illustrated on the past few pages, have the following properties

① $\{ \phi_0(t-k) \}_{k\in\mathbb{Z}} \cup \{ 2^{j/2}\psi_0(2^j t - k) \}_{j\geq 0,\, k\in\mathbb{Z}}$
is an ⊥-basis for $L_2(\mathbb{R})$.
This basically follows from the "perfect reconstruction" and "orthogonality" conditions for the corresponding CMF.

② Vanishing moments, $\psi_0(t)$ in the D_{2p} system has p vanishing moments. This means that there are no polynomial (of order $<p-1$) components in the wavelet spaces W_j.
It follows from $h_1[n]$ having vanishing moments in discrete time (see homework)

$\int \psi_0(t)\cdot t^q \, dt = 0$ for $q = 0, \ldots, p-1$

③ Support size. The scaling function $\phi_0(t)$ and wavelet function are local in time. Their support size is $2p-1$ (closely related to the lengths of the FIR filters h_0, h_1).

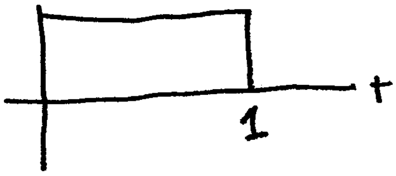
What are wavelets good for?

At the end of the day, the wavelet transform is just another way to "take apart" a signal.

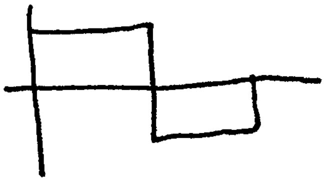
For $x_c(t)$ on $[0,1]$

$$x_c(t) = c_{0,0}\phi_{0,0}(t) + \sum_{j\geq 0}\sum_{k} d_j[k]\,\psi_{j,k}(t)$$

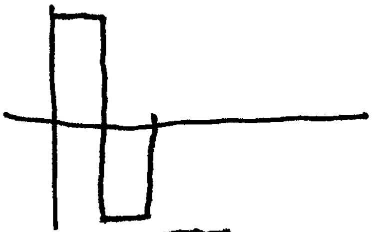
$\phi_{0,0}(t) =$



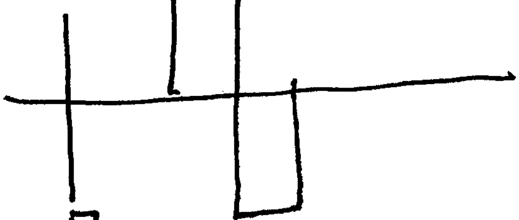
$\psi_{0,0}(t) =$



$\psi_{1,0}(t) =$

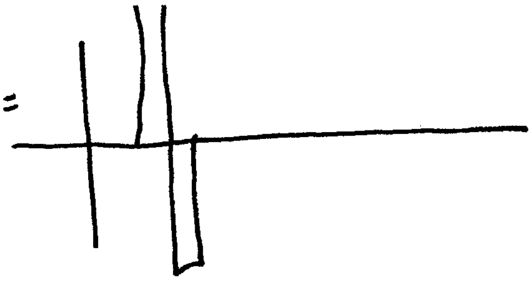


$\psi_{1,1}(t) =$



$\psi_{2,0}(t) =$

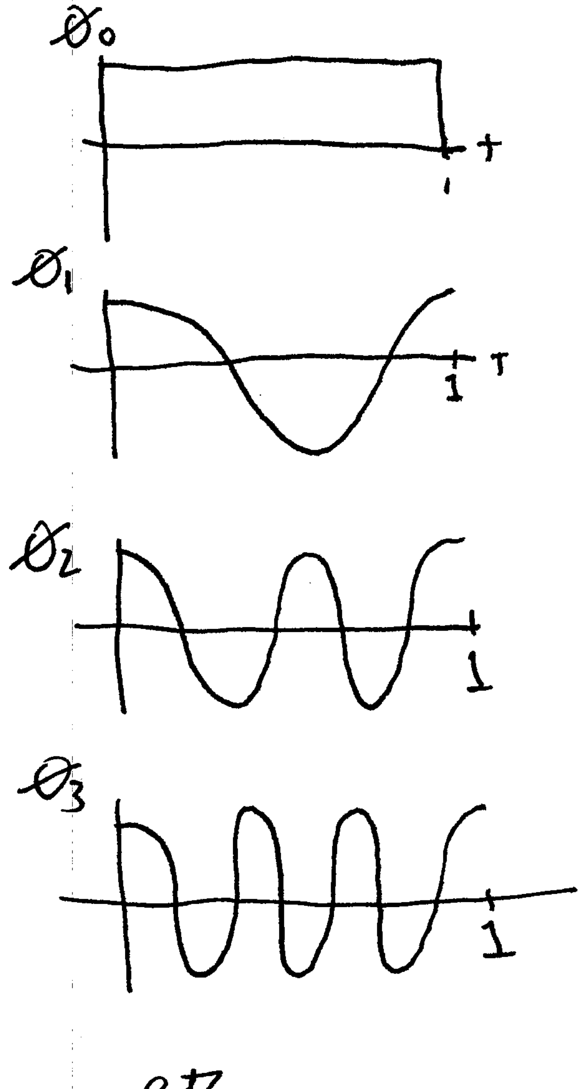
$\psi_{2,1}(t) =$



etc
$\vdots$

Compare to Fourier Series basis functions

$$x_c(t) = \sum_k a_k \underbrace{e^{j2\pi kt}}_{=\phi_k(t)}$$

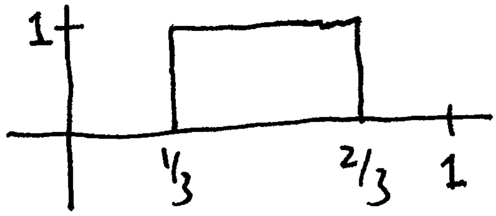


(real parts)

etc.

Why would we want to use wavelets instead of sinusoids (which we already know a lot about)?

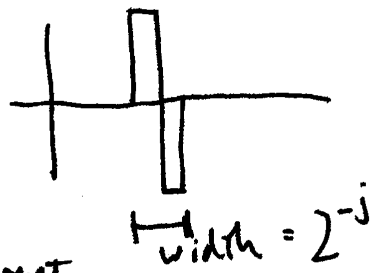
Say $x_c(t) =$



Questions:

① The (Haar) wavelet coefficient at scale j and shift k is

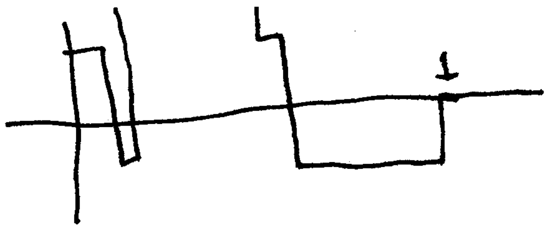
$$d_j[k] = \langle x_c, \psi_{j,k} \rangle$$



There are 2^j wavelets (different shifts) at scale j.

For fixed j, how many of the $d_j[k]$ are non-zero?

② Say $x_c(t) =$



What is an upper bound on the number of non-zero $d_j[k]$? (Again for a fixed j)

Moral: For piecewise constant functions, very few of the terms in

$$x_c(t) = c_{0,0}\,\phi_{0,0}(t) + \sum_{j\geq 0}\sum_{k} d_j[k]\,\underbrace{\psi_{j,k}(t)}_{=\text{Haar wavelet}}$$

are non-zero.

What about the Fourier series expansion?

$$x_c(t) = \sum_{k} a_k e^{j2\pi kt}$$

How many of the a_k are non-zero?

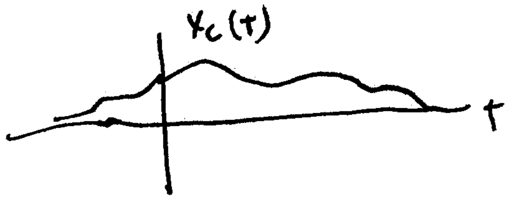
For piecewise constant signals, Haar approximations are **much** more accurate than Fourier approximations.

The energy "compacts" onto a much smaller set in the Haar domain than in the Fourier domain, making piecewise constant functions very "simple".

→ We can ignore most of the wavelet coefficients and not lose anything.

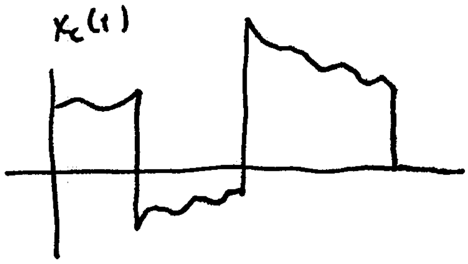
This makes wavelets very useful for things like compression, noise removal, and feature detection.
("nonlinear filtering")

As a general rule, signals which are smooth



have roughly the same amount of compaction in the wavelet and Fourier domains.

But signals which are <u>piecewise</u> <u>smooth</u>



are <u>far</u> more compacted in the wavelet domain.

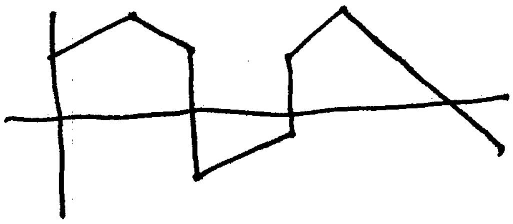
Piecewise smooth signals come up in all kinds of applications:

- digital photography
- seismic exploration (i.e. finding oil)
- medical imaging
- radar imaging and remote sensing

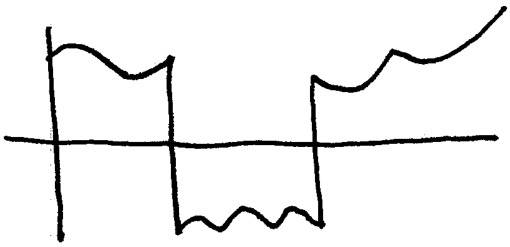
Beyond piecewise constant

Haar wavelets: 

- good for piecewise constant
- not quite as good for piecewise linear



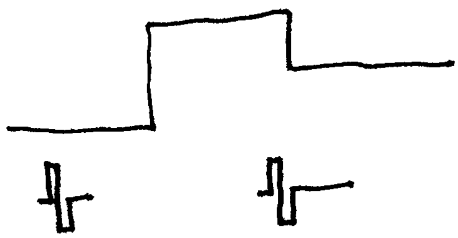
or piecewise polynomial



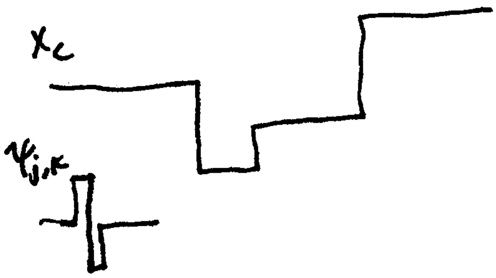
Haar works well for piecewise constant functions since

① it has small support

→ only "touches" edges in a few locations



② $\int \psi_{j,k}(t)\, dt = 0$



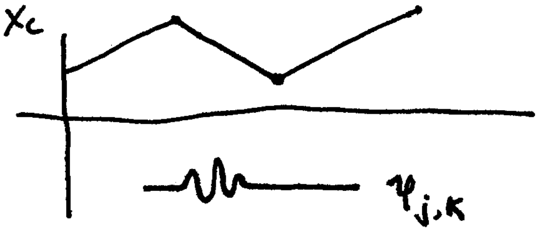
→ for wavelets lying in flat regions

$$\langle x_c, \psi_{j,k} \rangle = 0$$

What would we need for

$$\langle x_c, \psi_{j,k} \rangle$$

to be zero inside linear regions?



We would need

$$\int t\, \psi_{j,k}(t)\, dt = 0$$ and ② above

Say $x_c(t)$ is polynomial over the support of degree $\leq p$
of $\psi_{j,k}(t)$.

If

(*) $$\int t^m \psi_{j,k}(t)\, dt = 0 \qquad m = 0, \dots, p$$

then it is easy to see that

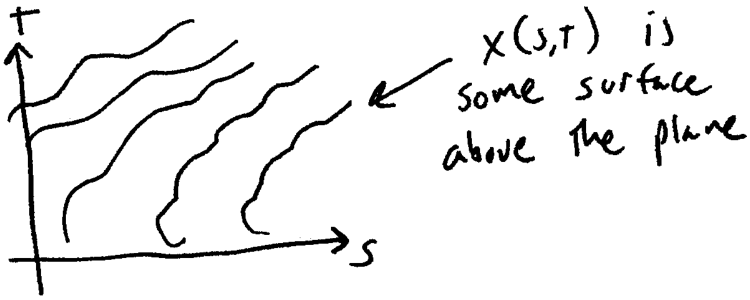
$$\langle x_c, \psi_{j,k} \rangle = 0$$

If (*) is true, we say $\psi_{j,k}$ has m <u>vanishing</u> <u>moments</u>.

This is a desirable property, as it allows us to extend our approximation results w/ Haar and piecewise constants to m^{th} order piecewise polynomials.

2D Wavelets

Many times we are interested in processing signals that are functions of two spatial variables ("images")

$X(s,t)$ = 

We can easily move from 1D to 2D ($x(t) \in L_2(\mathbb{R})$ to $x(s,t) \in L_2(\mathbb{R}^2)$) using the notion of separability.

Def: We call a function $f(s,t)$ separable if

$$f(s,t) = g(s) \cdot h(t)$$

for some $g(\cdot)$ and $h(\cdot)$.

Ex: The 2D Fourier basis functions

$$e^{j(\omega_1 s + \omega_2 t)}$$

are separable.

$g(s) =$

$h(t) =$

104

Exercise: Suppose g and h are 1D functions with $g \perp h$ $\left(\langle g,h\rangle = 0, \text{ i.e. } \int g(t)\overline{h(t)}\,dt = 0\right)$

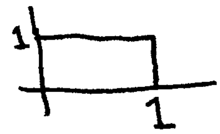
Show that the four functions

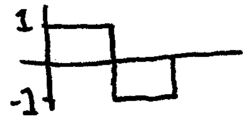
$$f_1(s,t) = g(s)\,g(t)$$
$$f_2(s,t) = g(s)\,h(t)$$
$$f_3(s,t) = h(s)\,g(t)$$
$$f_4(s,t) = h(s)\,h(t)$$

are orthogonal.

105

Exercise:

Let $\phi(t) =$  (Haar scaling function and wavelet)

$\psi(t) =$ 

Sketch

$$\phi^{LL}(s,t) = \phi(s)\cdot\phi(t)$$

$$\psi^{LH}(s,t) = \phi(s)\,\psi(t)$$

$$\psi^{HL}(s,t) = \psi(s)\cdot\phi(t)$$

$$\psi^{HH}(s,t) = \psi(s)\cdot\psi(t)$$

on $[0,1]^2$.

Given a 1D wavelet basis for $L_2(\mathbb{R})$:

$$\{\phi_{0,k}\}_{k\in\mathbb{Z}} \cup \{\psi_{j,k}\}_{j\geq 0,\, k\in\mathbb{Z}}$$

(or more generally

$$\{\phi_{J,k}\}_{k\in\mathbb{Z}} \cup \{\psi_{j,k}\}_{j\geq J,\, k\in\mathbb{Z}} \quad)$$

where

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

We can create a 2D separable wavelet basis for $L_2(\mathbb{R}^2)$ with

$$\{\phi^{LL}_{0,k_1,k_2}\}_{k_1,k_2\in\mathbb{Z}} \cup \{\psi^{LH}_{j,k_1,k_2}\}_{j\geq 0,\, k_1,k_2\in\mathbb{Z}}$$
$$\cup \{\psi^{HL}_{j,k_1,k_2}\}_{j\geq 0,\, k_1,k_2} \cup \{\psi^{HH}_{j,k_1,k_2}\}_{j\geq 0,\, k_1,k_2}$$

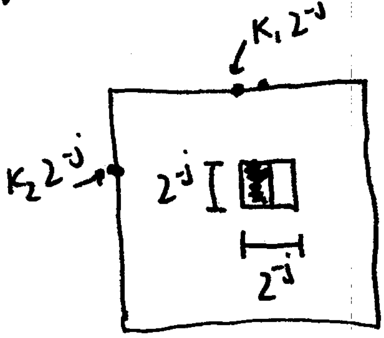
where

$$\phi^{LL}_{j,k_1,k_2}(s,t) = 2^j\phi(2^j s - k_1)\phi(2^j t - k_2)$$
$$\psi^{LH}_{j,k_1,k_2}(s,t) = 2^j\phi(2^j s - k_1)\psi(2^j t - k_2)$$
$$\psi^{HL}_{j,k_1,k_2}(s,t) = 2^j\psi(2^j s - k_1)\phi(2^j t - k_2)$$
$$\psi^{HH}_{j,k_1,k_2}(s,t) = 2^j\psi(2^j s - k_1)\psi(2^j t - k_2)$$

Locality:

At a scale j and a shift (k_1,k_2) the
ψ^{LH}_{j,k_1,k_2}, ψ^{HL}_{j,k_1,k_2}, ψ^{HH}_{j,k_1,k_2}

are centered at $\approx (k_1 2^{-j}, k_2 2^{-j})$
and concentrated on a square of width $\sim 2^{-j}$



As a general rule, wavelet coefficients will be small if the image is smooth over this region, and will be "active" if there is an edge or texture in this region.