Recall: Any finite energy signal $f$ can be decomposed in an orthogonal wavelet basis $\{ \psi_{j,k} \}_{j,k \in \mathbb{Z}^2}$

$$f = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \langle f, \psi_{j,k} \rangle \psi_{j,k}$$

The partial sum (at resolution $2^i$)

$$(\text{scale} = \frac{1}{\text{resolution}} = 2^i)$$

$$\sum_{k=-\infty}^{\infty} \langle f, \psi_{j,k} \rangle \psi_{j,k} = d_j$$

can be interpreted as the difference between two approximations of $f$ at scales $2^{-i+1}$ and $2^{-i}$.

d$_j$ is simply the projection of $f$ onto the subspace

$$W_j = \overline{\text{Span} \{ \psi_{j,k} \}}$$

$$d_j = \text{"details of } f \text{ at scale } 2^i\"$$

finer (smaller) scale $\iff$ higher (larger) resolution
Why multiresolution?

Adapting the resolution of our signal analysis allows us to process only the relevant details for a specific task. 

Ex. Burt & Adelson introduced the multiresolution image pyramid that can be used to code or process differently at different resolutions. This is crucial in modern image and video compression. Multiresolution image analysis also facilitates more advanced tasks such as image restoration, segmentation, and object recognition.
Multi-resolution Approximations

\[ f(t) \]

Analysis at scale \( 2^j \)

Compute local averages of \( f \) at positions \( \{k2^j\}_{k \in \mathbb{Z}} \) over intervals of width \( \propto 2^j \) (scale \( 2^j \))

Multi-resolution analysis

Analysis of \( f \) over embedded grids of approximation.

scale \( 2^j \)

\[ 0 \quad 2^j \quad 2^j \quad \cdots \quad t \]

resolution = \( 2^j \)

scale \( 2^{j+1} \)

\[ 0 \quad 2^{j+1} \quad 2^{j+1} \quad \cdots \quad t \]

resolution = \( 2^{j+1} \)

Recall, Haar wavelet analysis

levels of piecewise constant approximation are the local averages of \( f \)
Multiresolution Image Analysis

compute local average over a neighborhood of area proportional to \((2^{-j})^2\)

Multiresolution analysis of d-dimensional object

Compute local average over neighborhood of measure \(\approx 2^{-jd}\)

Formally:

An approximation of a function \(f \in L^2(\mathbb{R})\) at scale \(2^{j}\) is defined as an orthogonal projection of \(f\) onto a subspace \(V_j \subset L^2(\mathbb{R})\) (low resolution)
Definition: Multi-resolution Analysis (MRA)

A sequence \( \{ V_j \}_{j \in \mathbb{Z}} \) of closed subspaces of \( L^2(\mathbb{R}) \) is a multi-resolution analysis (or decomposition) of \( L^2(\mathbb{R}) \) if the following properties hold:

1. \( f \in V_j \iff f(t - 2^j k) \in V_j \quad \forall j, k \)
2. \( V_{j+1} \supset V_j \quad \forall j \)
3. \( f(t) \in V_j \iff f(at) \in V_{j+1} \quad \forall j \)
4. \( V_{-\infty} = \lim_{j \to -\infty} V_j = \{ 0 \} \) \( \iff \) contains only the "zero" function
5. \( V_{\infty} = \lim_{j \to \infty} V_j = L^2(\mathbb{R}) \)
6. There exists a function \( \phi(t) \) such that \( \{ \phi(t-k) \}_{k \in \mathbb{Z}} \) is a Riesz basis of \( V_0 \).
Notes: Here $j$ refers to resolution $2^j$, so $V_{j+1}$ is a finer scale (higher resolution $a^{(j+1)}$) subspace than $V_j$ (resolution $2^j$).

This is the convention of Burrus et al. Others index the subspaces in the inverse manner. That is, they let $j$ denote scale $2^j$, rather than resolution so that $V_{j+1} \subset V_j$ (e.g., Mallat's book).

**Multiresolution Subspaces**

$V_3 \supset V_2 \supset V_1 \supset V_0$

"Nested subspaces"
The fact that the multiresolution subspaces are nested guarantees that an approximation at resolution \(2^j\) (in \(V_j\)) contains all the necessary information to compute an approximation at a lower resolution \(2^{j-1}\) (in \(V_{j-1}\)).

Let \(P_{V_j}\) denote the projection operator projecting \(f \in L^2\) to \(f \in V_j\):

\[
V_\infty = \{0\} \implies \lim_{j \to \infty} \|P_{V_j} f\| = 0 \quad \forall f \in L^2(\mathbb{R})
\]

implying that as the resolution \(2^j \to 0\) we lose all the details of \(f\):

\[
V_\infty = L^2(\mathbb{R}) \implies \lim_{j \to \infty} \|P_{V_j} f - f\| = 0
\]

showing that the signal approx \(f_j = P_{V_j} f\) converges to the true signal as resolution \(2^j \to \infty\).

(rate at which \(\|P_{V_j} f - f\| \to 0\) depends on regularity of \(f\))
Riesz Bases

Less restrictive than insisting on an orthogonal basis.

A sequence is a Riesz basis of $L^2(\mathbb{R})$ if it is linearly independent and if there exist $A > 0$, $B > 0$ such that for any $f \in L^2(\mathbb{R})$ we can find a representation

$$f = \sum_{n} c_n e_n$$

that satisfies

$$\frac{1}{B} \| f \|_2^2 \leq \sum_n c_n^2 \leq \frac{1}{A} \| f \|_2^2$$

This "Parseval-like" energy relationship guarantees the stability of representations (convergence in norm) using Riesz basis.
Proposition

A family \( \{ \Omega(t-k) \}_{k \in \mathbb{Z}} \) is a Riesz basis of the space \( V_0 = \text{Span} \{ \Omega(t-k) \} \) if and only if there exist \( A > 0, B > 0 \) such that

\[
\frac{1}{B} \leq \sum_{n=\infty}^{\infty} | \Phi(\omega - 2n\pi) |^2 \leq \frac{1}{A} \quad \forall \omega \in [-\pi, \pi]
\]

where \( \Phi \) denotes the FT of \( \Omega(t) \).
Examples

Ex. Piecewise Constant Approximations

\[ V_j \] is the set of all \( f \in L^2(\mathbb{R}) \)
such that \( f(t) \) is constant on
\[ t \in [k 2^{-j}, (k+1) 2^{-j}) \]
for each \( \text{ne } \mathbb{Z} \).

\[ V_0 = \text{Span} \left\{ \phi(t-k) \right\} \]

\[ \phi(t) = \begin{cases} 1 & t \in [0,1) \\ 0 & \text{o.w.} \end{cases} \]

(Hasar scaling function)
Ex. Shannon Approximations
(Sinc Wavelet)

$V_j$ is defined to be the set of $f \in L^2(\mathbb{R})$ whose FT has support in the band $[-2^j \pi, 2^j \pi]$

$V_j = \{ f \in L^2(\mathbb{R}) : F(w) = 0 \text{ for } |w| > 2^j \pi \}$

$V_0 = \text{span} \left\{ \frac{\sin \pi(t-k)}{\pi(t-k)} \right\}$

$P_{V_j}$ is a "lowpass" approx. of $f$.

$P_{V_j}$ is an ideal lowpass filter.
**Ex. Beyond piecewise constant approximations**

**Spline approximation**

**Box spline of degree m**

Convolve \(1_{[0,1]}\) box function with itself \(m\) times.

- Degree \(m\) box spline is \(m-1\) times continuously differentiable.
- For all \(m \geq 0\), box spline basis \(\{\phi(t-k)\}\) is a Riesz basis for \(V_0\).
The Scaling Function

The family

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

is a Riesz basis of \( V_j \) for all \( j \in \mathbb{Z} \).

\( \Phi(t) \) is called the scaling function of the associated multi-resolution analysis (MRA):

\[ V_j = \text{span} \{ 2^{j/2} \Phi(2^j t - k) \} \]
The nesting of \( V_j \supset V_{j-1} \) implies that \( \phi(t) = \phi_{0,0}(t) \in V_0 \) can be expressed as a linear combination of \( \sum_{k} \phi_{i,k}(t) \).

\[
\phi(t) = \sum_{n=-\infty}^{\infty} a(n) \phi(2t-n)
\]

for some sequence \( \{a(n)\} \in \mathbb{Z} \).

**Ex. Piecewise Constant Haar Scaling function**

\[
\begin{align*}
\phi(t) &= \phi(2t) + \phi(2t-1) \\
\{a(n)\} &= \{0,0,\ldots,0,1,1,0,0,\ldots\}
\end{align*}
\]
Ex. First order Box Spline

\[ \Phi(t) = \frac{1}{2} \Phi(2t) + \Phi(2t-1) + \frac{1}{2} \Phi(2t-2) \]

\[ \{a(n)\} = \{ ... , 0, \frac{1}{2}, 1, \frac{1}{2}, 0, ... \} \]

\[ n = 0, 1, 2 \]

It is usual to express the relationship between \( \Phi(t) \) and \( \{\Phi(2t-n)\} \) as

\[ \Phi(t) = \sum_{n} h(n) \sqrt{2} \Phi(2t-n) \]

since \( \sqrt{2} \Phi(2t-n) \) is normalized.

The coefficients \( \{h(n)\} \) are called the scaling filter.

Note:

\[ h(n) = \langle \Phi(t), \sqrt{2} \Phi(2t-n) \rangle \]
So, we see that the scaling filter \( \{ h(n) \} \) is intimately related to the MRA.

Under what conditions does a filter \( \{ h(n) \} \) correspond to a valid MRA?

**Theorem (Mallat, Meyer)**

Let \( \phi \in L^2(\mathbb{R}) \) be an integrable, orthogonal scaling function. Then the Fourier series of \( h[n] = \langle \phi(t), \sqrt{2} \phi(2t-n) \rangle \) satisfies

1. \[ |H(\omega)|^2 + |H(\omega+\pi)|^2 = 2 \quad \forall \omega \in [-\pi, \pi] \]

and

2. \[ H(0) = \sqrt{2} \]

Conversely, if \( H(\omega) \) is 2\pi-periodic, continuously differentiable at \( \omega = 0 \), satisfies (1) and (2), and if

\[ \inf_{\omega \in [\frac{\pi}{2}, \frac{\pi}{2}]} |H(\omega)| > 0 \]

then

\[ \phi(\omega) = \prod_{p=1}^{\infty} \frac{H(\omega/2^p)}{\sqrt{2}} \]

is the FT of an orthonormal scaling function \( \phi \in L^2(\mathbb{R}) \)

(See "A Wavelet Tour of Signal Processing," S. Mallat (1998))
This theorem shows that discrete filters satisfying certain conditions completely characterize a MRA.

Filters that satisfy \(①\) are called quadrature (or conjugate) mirror filters. These filters make it possible to carry out MRA directly on discrete (sampled) signals and images.

There are many variations and generalizations to this theorem that we will look at later.
MRA and Wavelets

Suppose we have a scaling function \( \phi(t) \) and the associated MRA
\[
\{ \cdots \subset V_1 \subset V_0 \subset V_V \subset \cdots \subset L^2(\mathbb{R}) \}
\]
Furthermore, assume that
\[
\left\{ \phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k) \right\}
\]
is an orthogonal basis for
\[
V_j = \text{span} \left\{ 2^{j/2} \phi(2^j t - k) \right\}
\]

Aside: If we start with a Riesz basis that is not orthonormal, we can find a relating scaling function that does generate an orthonormal basis.

Suppose \( \theta(t) \) is a scaling function and that \( \left\{ \theta_{j,k}(t) = 2^{j/2} \theta(2^j t - k) \right\} \) is a Riesz basis for \( V_j \). Then the scaling func \( \phi(t) \) whose FT is given by
\[
\hat{\phi}(\omega) = \frac{\Theta(\omega)}{\left( \sum_{k \in \mathbb{Z}} |\Theta(\omega + 2\pi k)|^2 \right)^{1/2}}
\]
generates the following orthonormal basis for \( V_j \):
\[
\left\{ \phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k) \right\}
\]
Wavelet Subspaces

Let \( W_j \) denote the orthogonal complement of \( V_j \) in \( V_{j+1} \).

That is, \( W_j \) is the closed subspace consisting of functions in \( V_{j+1} \) that are orthogonal to \( V_j \).

\[
W_j = \{ f \in V_{j+1} : P_{V_j} f = 0 \} = \text{"wavelet subspace at resolution } 2^j \text{"}
\]

With this definition \( V_{j+1} \) can be expressed as the direct sum of \( V_j \) and \( W_j \):

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

(\( \oplus \) denote direct sum)

Aside: A vector space \( X \) is the direct sum of two subspaces \( Y \) and \( Z \) if every vector in \( x \in X \) has a unique representation of the form \( x = y + z \), where \( y \in Y \), \( z \in Z \).

In our case above, each \( f \in V_{j+1} \) is uniquely represented by

\[
f = P_{V_j} f + P_{W_j} f
\]
The projection $P_{W_j}f$ provides the "details" of the signal that appear at resolution $2^j$, but which disappear at lower resolution $2^{j'}$ (coarser scale $2^{-j'}$).

Applying this relationship

$$V_1 = V_0 \oplus W_0$$
$$V_2 = V_1 \oplus W_1 = V_0 \oplus W_0 \oplus W_1$$

... 

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

Note that the initial space is arbitrary; we also have, for example,

$$L^2(\mathbb{R}) = V_6 \oplus W_6 \oplus W_7 \oplus \cdots$$

$$L^2(\mathbb{R}) = V_{-5} \oplus W_{-5} \oplus W_{-4} \oplus \cdots$$
or even

\[ L^2(\mathbb{R}) = \bigoplus_{j=-\infty}^{\infty} W_j \]

eliminating the subspaces \( \{V_j\} \) altogether. This shows that all the information provided by the MRA is completely contained in the wavelet subspaces \( \{W_j\} \).

In practice, we usually start with a scale subspace \( V_j \), with \( j \) chosen to represent the finest details of interest in a signal.

For example, if we sample a signal, then we have a fundamental limit on the resolution that is meaningful.
Scale and Wavelet Subspaces

Note:

(i) \( W_j \perp V_j \quad \forall j \quad (\text{That is, if } w \in W_j \text{ and } v \in V_j, \text{ then } \langle w, v \rangle = 0) \)

(ii) \( W_i \perp W_j \quad \forall i \neq j \)

(iii) \( W_j \subseteq V_{j+1} \)
Recall that the translates and dilates of the scaling function provide orthonormal bases for the scale spaces \( \{V_j\} \).

Similarly, we would like a set of orthonormal bases for the wavelet subspaces \( \{W_j\} \).

**Notes**

\[ W_j \perp V_j \quad \text{if we } W_j, v \in V_j \quad \text{then } \langle w, v \rangle = 0, \]

\[ W_i \perp W_j \quad i \neq j \]

\[ W_j \subset V_{j+1} \]

Let \( \{\psi_{j,k}\} \) be an orthonormal basis for \( W_j \), \( j \in \mathbb{Z} \). Then

\[ \langle \psi_{j,k}, \psi_{i,l} \rangle = 0 \quad \forall k, l \in \mathbb{Z}, \ i \neq j \]
Also note
\[ \psi_{j,k} \in V_{j+1} \]

In particular,
\[ \psi(t) = \psi_{0,0}(t) \in V_0 \]

\[ \Rightarrow \quad \psi(t) = \sum_n \alpha_n \sqrt{2} \Phi(2^{j}t - n) \]

for some \( \{ \alpha_n \} \)

We will see later that taking
\[ \alpha_n = (-1)^n h(1-n) \]

where \( \{ h(n) \} \) is the scaling filter associated with \( \Phi \) generates a \( \psi(t) \) such that
\[ \{ \psi_{j,k}(t) = 2^{j/2} \psi(2^{j}t - k) \} \quad k \in \mathbb{Z} \]

is an orthogonal basis for \( W_j \)

for all \( j \).
The function

\[ \psi(t) = \sum_{n} (-1)^n h(1-n) \sqrt{2} \ \phi(2t-n) \]

is called the prototype or mother wavelet associated with the MRA \( \cdots V_{-1} \subset V_0 \subset V_1 \subset \cdots \).

Ex. Haar (piecewise constant) MRA

Recall

\[ \phi(t) = \begin{cases} 
1 & \text{te}[0,1) \\
0 & \text{o.w.}
\end{cases} \]

\[ h(n) = \{ 0, \ldots, 0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, \ldots \} \]

\[ h_{-1} = 0 \quad h_{0} = 1 \]

\[ \Rightarrow \]

\[ \psi(t) = \frac{1}{\sqrt{2}} \cdot \sqrt{2} \ \phi(2t) - \frac{1}{\sqrt{2}} \sqrt{2} \ \phi(2t-1) \]

\[ = \phi(2t) - \phi(2t-1) \]
Ex. First order Box spline

$$\Phi(t)$$

$$\{ h(n) \} = \{ \cdots, 0, \frac{1}{2\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{2\sqrt{2}}, 0, \cdots \}$$

$$n = 0 \quad n = 1 \quad n = 2$$

$$\Rightarrow$$

$$\Psi(t) = \frac{1}{2} \Phi(2t) - \Phi(2t-1) + \frac{1}{2} \Phi(2t-2)$$

Caution: This is not an orthogonal system.
Wavelet Representations

\[ V_0 \oplus W_0 \oplus W_1 \oplus \ldots = L^2(\mathbb{R}) \]

Orthogonality

\[ \Rightarrow \]

\[ f = P_{V_0} f + P_{W_0} f + P_{W_1} f + \ldots \]

If \( \{ \phi(t-k) \}_{k} \) is an o.n. basis for \( V_0 \)

and \( \{ 2^{j/2} \phi (2^j t-k) \}_{k} \) is an o.n. basis for \( W_j \)

\( j \geq 0 \)

then we can write

\[ f(t) = \sum_{k=-\infty}^{\infty} c(k) \phi(t-k) + \sum_{j \geq 0} \sum_{k=-\infty}^{\infty} d_j(k) 2^{j/2} \phi (2^j t-k) \]

where

\[ c(k) = \langle f, \phi(t-k) \rangle \]

\[ d_j(k) = \langle f, 2^{j/2} \phi (2^j t-k) \rangle \]
More generally, we can start at any scale $2^{j_0}$ and write

$$f(t) = \sum_{k} c_{j_0}(k) 2^{j_0/2} \phi(2^{j_0} t - k) + \sum_{k} \sum_{j = j_0}^\infty d_{j}(k) 2^{j/2} \psi(2^{j} t - k)$$

The coefficients $\{c_{j_0}(k)\}_k$ applied to the scaling basis functions $\{2^{j_0/2} \phi(2^{j_0} t - k)\}_k$ produce the low resolution (coarse scale) approximation of $f$, corresponding to the projection of $f$ onto the subspace $V_{j_0}$.

$\{d_{j}(k)\}_{j,k}$ and $\{2^{j/2} \psi(2^{j} t - k)\}_{k,j \geq j_0}$ provide the high resolution details of the signal.
Geometrically,

\[ V_{j_0} \rightarrow L^2(\mathbb{R}) \]

scale space (resolution \( 2^{j_0} \))

\[ P_{V_{j_0}} f \]

\[ f - P_{V_{j_0}} f \]

Wavelet subspaces \( W_{j_0} \oplus W_{j_0+1} \oplus \cdots \)

The coefficients

\[ \{ c_{j_0}(k) \}_{k} , \{ d_{j}(k) \}_{k, j \geq j_0} \]

are called the

discrete wavelet transform (DWT)

of \( f \).
Connection to DSP

At high resolutions, the scaling functions are similar to Dirac delta functions (assuming \( \phi(t) \) is localized and well-behaved).

That is,

\[ 2^{j/2} \phi(2^j t) \rightarrow \delta(t) \quad \text{as} \quad j \rightarrow \infty \]

Ex. Haar scaling function

\[ \phi(t) \]

\[ 2^{1/2} \phi(2^j t) \]

[Diagram of Haar scaling function]
So, for $j$ sufficiently large we have

$$c_j (k) = \langle f, \Phi_{j, k} \rangle$$

$$= \int_{-\infty}^{\infty} f(t) \ 2^{j/2} \ \phi (2^j t - k) \ dt$$

$$= \int_{-\infty}^{\infty} f(t) \ \delta (t - k 2^{-j}) \ dt$$

$$= f (k 2^{-j})$$

In other words, the scaling coefficients are approximately equal to signal samples at a sampling rate of $T = 2^{-j}$ (sampling frequency $\omega_s = 2\pi 2^{j}$).
Computing the DWT

Mallat's Fast Wavelet Transform (FWT) algorithm

1. Assume an initial set of scaling coefficients \( \{ c_j(n) \}_n \) representing an approximation \( f_j = \mathcal{P}_{v_j} f \) to a signal \( f \) at a certain scale related to the sampling period \( T = 2^{-j} \).

2. The wavelet and scaling coefficients at coarser scales, \( j < J \), are then computed recursively using the (lowpass) scaling filter \( \{ h(n) \}_n \) and (highpass) wavelet filter \( \{ h_w(n) = (-1)^n h(1-n) \}_n \).
\[ C_j(k) = \langle f, 2^{j/2} \phi(2^j t - k) \rangle \]
\[ = \sum_n h(n - 2k) c_{j+1}(n) \]
\[ d_j(k) = \langle f, 2^{j/2} \psi(2^j t - k) \rangle \]
\[ = \sum_n h_1(n - 2k) c_{j+1}(n) \]

for \( j = J-1, J-2, \ldots \).

\( a \) lowpass filter \( c_{j+1} \) to obtain lower resolution approximation

\( b \) highpass filter \( c_{j+1} \) to obtain details in \( c_{j+1} \) but not in \( c_j \).

Decimation: The filters are shifted by \( 2k \) (rather than \( k \)) so that only even indexed terms (at filter outputs) are retained. This eliminates redundant information in full sequence outputs \( \sum_n h(n-k)c_{j+1}(n) \) and \( \sum_n h,(n-k)c_{j+1}(n) \).
With these coefficients (computed using simple digital filters!) we have the representation

\[ P_{v_j}f = \sum_{k=-\infty}^{\infty} c_{j_0}(k) \, 2^{j_0/2} \, \phi(2^{j_0}t - k) \]
\[ + \sum_{k=\infty}^{\infty} \sum_{j=j_0}^{j-1} d_j(k) \, 2^{j/2} \, \psi(2^j t - k) \]

\[ \uparrow \]
finite sum

Moreover, if \( f \) is of finite time duration, then the sums over \( k \) are also finite and hence \( P_{v_j}f \) is completely determined by a handful of numbers (scaling and wavelet coefficients).

Note: This opens the door to a whole new world of DSP. Instead of processing signal samples, we can analyze and process a signal using its DWT!
Ex: Haar Analysis

\[ h(n) = \{ \ldots, 0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, \ldots \} \]
\[ h_i(n) = \{ \ldots, 0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0, \ldots \} \]

Begin with
\[ \{ c_j(k) \}_k \]

for \( j < J \),
\[ c_j(k) = \frac{1}{\sqrt{2}} \ c_{j+1}(2k) + \frac{1}{\sqrt{2}} \ c_{j+1}(2k+1) \]
\[ d_j(k) = \frac{1}{\sqrt{2}} \ c_{j+1}(2k) - \frac{1}{\sqrt{2}} \ c_{j+1}(2k+1) \]

\[ \{ c_j(k) \}_j \] = local sums of \( \{ c_{j+1}(k) \}_j \)
\[ \{ d_j(k) \}_j \] = local differences of \( \{ c_{j+1}(k) \}_j \)

Note: In this special case (Haar) we have
\[ c_{j+1}(2k) = \frac{1}{\sqrt{2}} \ c_j(k) + \frac{1}{\sqrt{2}} \ d_j(k) \]
\[ c_{j+1}(2k+1) = \frac{1}{\sqrt{2}} \ c_j(k) - \frac{1}{\sqrt{2}} \ d_j(k) \]

Which gives us a simple reconstruction (synthesis) procedure!
Associated Scaling and Wavelet Functions

Take $j_0 = 0$, for simplicity.

$$P_{\phi} f = \sum_k c_0(k) \phi(t-k) + \sum_{\ell} \sum_{j=0}^{J-1} d_{\ell j}(k) 2^{\ell/2} \psi(2^{\ell} t - k)$$
Multiresolution Analysis of $L^2(\mathbb{R})$

Main Idea: Decompose $L^2(\mathbb{R})$ into a sequence of nested subspaces, each consisting of functions with successively finer "detail" and structure (i.e., higher resolution).

Basic Properties:

1. $f \in V_j \iff f(t - 2^{-j}k) \in V_j$ for all $j, k$ integers.
Let \( f \in V_j \) and let \( V_j \) be Haar subspace. The \( f \) is piecewise constant on intervals \( [k2^{-j}, (k+1)2^{-j}) \), \( k \)

\[ \Rightarrow f \text{ is constant on intervals } [k2^{-(j+1)}, (k+1)2^{-(j+1)}], k \in \mathbb{Z} \]

\[ \Rightarrow f \in V_{j+1} \]

Another view:

\[ V_j \equiv \text{Span} \left( \{ \phi(2^j t - k) \}_{k \in \mathbb{Z}} \right) \]

where

\[ \phi(2^j t) \quad \xrightarrow{\text{i.o., there exists a func}} \quad \phi(t) \text{ such that } \{ \phi(2^j t - k) \} \text{ is a basis for } V_j \]

\[ f \in V_j \quad \Rightarrow \quad f = \sum_{k=-\infty}^{\infty} \alpha_k \phi(2^j t - k) \]

\[ V_{j+1} \equiv \text{Span} \left( \{ \phi(2^{j+1} t - k) \}_{k \in \mathbb{Z}} \right) \]
In particular,

\[ \phi(z^{j+1}t - k) \in V_{j+1} \]

\[ \phi(z^j t) = \phi(z^{j+1} t) + \phi(z^{j+1} t - 1) \]

\[ \Rightarrow \quad V_j \subset V_{j+1} \]

**Beyond Haar:**

Suppose \( \phi(t) \) is 1st order box spline

\[ \phi(z^j t) = \frac{1}{2} \phi(z^{j+1} t) + \phi(z^{j+1} t - 1) + \frac{1}{2} \phi(z^{j+1} t - 2) \]

\[ \Rightarrow \quad V_j = \text{span}( \{ \phi(z^{j+1} t - k) \} ) \subset V_{j+1} = \text{span}( \{ \phi(z^{j+1} t - k) \} ) \]
\( f \in V_j \iff f(z \cdot t) \in V_{j+1} \) for all \( j \)

\[
f \in V_j \Rightarrow f(t) = \sum_k \alpha_k \phi(2^j t - k)
\]

\[
\Rightarrow f(zt) = \sum_k \alpha_k \phi(2^j (zt) - k)
\]

\[
= \sum_k \alpha_k \phi(2^{j+1} t - k)
\]

\[
\Rightarrow f(zt) \in V_{j+1}.
\]

4) \( V_{-\infty} = \lim_{j \to \infty} V_{-j} = \{0\} \)

5) \( V_{\infty} = \lim_{j \to \infty} V_j = L^2(\mathbb{R}) \)

\[ V_j \to \mathbb{L}^2(\mathbb{R}) \]

\( V_{-j} \to \text{empty set (except for zero-function)} \)
Key Equation:

\[ V_j \subset V_{j+1} \Rightarrow \alpha^{2j} \phi(2^j t) = \sum_{n=-\infty}^{\infty} h(n) 2^{j+\frac{1}{2}} \phi(2^{j+1} t - n) \]

There exists a sequence \( \{ h(n) \} \) relating basis functions for \( V_j \) to basis for \( V_{j+1} \).

Note that if the basis functions are orthogonal (e.g., Haar), then

\[ h(n) = \left\langle \alpha^{2j} \phi(2^j t), 2^{j+\frac{1}{2}} \phi(2^{j+1} t - n) \right\rangle \]
**Key Theorem:**

Let \( \phi \in L^2(\mathbb{R}) \), \( \{ \phi(t-n) \}_{n \in \mathbb{Z}} \) an orthogonal basis for \( V_0 \subset L^2(\mathbb{R}) \). Then, to ensure a valid multi-resolution analysis, the sequence

\[
h(n) = \langle \phi(t), \sqrt{2} \phi(2t-n) \rangle
\]

must satisfy

1. \( |H(w)|^2 + |H(w+\pi)|^2 = 2 \) for \( w \in [0, 2\pi] \)
2. \( H(0) = \sqrt{2} \),

where \( H(w) = \sum_n h(n) e^{-jwt} \).
Conversely, if $H(w)$ is continuously differentiable at $w = 0$, satisfies (1) and (2), and

$$\inf_{w \in [-\pi, \pi]} |H(w)| > 0$$

then

$$\phi(w) = \prod_{p=1}^{\infty} \frac{H(\frac{w}{2^p})}{\sqrt{2}}$$

produces an orthonormal scaling function

$$\phi(t) = \int_{-\infty}^{\infty} \phi(w) e^{jwt} \, dw$$

Proof: see Mallat
pp. 229 - 234