Wavelets

Review Fourier Transform and Short Term Fourier Transform (Read the tutorial posted on the web and/or read from either Sayood or Salomon).

The Problems with Fourier Transform

The Fourier transform hides information about time. It gives unambiguous information about the frequencies that the signal contains but it does not say at what times these frequencies occur. As a result, two signals, one stationary and the other no-stationary, containing the same frequencies will give the same frequency spectrum. Every frequency is considered a global characteristic of the signal. A discontinuity in the local part of the signal is translated in the frequency spectrum over the entire time domain – a local characteristic becomes a global characteristic. This does not mean that the information regarding time is totally lost, it becomes embedded in the 'phases' of the components and this is the reason we can reconstruct the original time signal faithfully.

The lack of time information makes Fourier transform error prone. If a signal is received correctly for hours and gets corrupted for only a few second, it totally destroys the signal because the frequencies injected spread over the entire time domain and the errors become global.

A qualitative explanation of why Fourier transform fails to capture time information is the fact that the set of basis functions (sines and cosines) are infinitely long and the transform picks up the frequencies regardless of where it appears in the signal.

Uncertainty Principle

The time and frequency domains are complimentary. If one is local, the other is global. For an impulse signal, which assumes a constant value for a very brief period of time, the frequency spectrum is infinite whereas if a sinusoidal signal extends over infinite time, its frequency spectrum is a single vertical line at the given frequency. We can always localize a signal or a frequency but we cannot do that simultaneously. If the signal has a short duration, its band of frequency is wide and vice versa.

Heisenberg's uncertainty principle was enunciated in the context of quantum physics which stated that the position and the momentum of a particle cannot be precisely determined simultaneously. This principle is also applicable to signal processing where the precise statement is as follows.

Let $g(t)$ be a function with the property $\int_{-\infty}^{\infty} \left| g(t)^2 \right| dt = 1$. Then

$$\left(\int_{-\infty}^{\infty} (t-t_m)^2 \left|g(t)\right|^2 dt\right) \left(\int_{-\infty}^{\infty} (f-f_m)^2 \left|G(f)\right|^2 dt\right) \geq \frac{1}{16\Pi^2}$$

where t_m, f_m denote average values of *t* and *f* ,and *G*(*f*) is the Fourier transform of *g*(*t*).

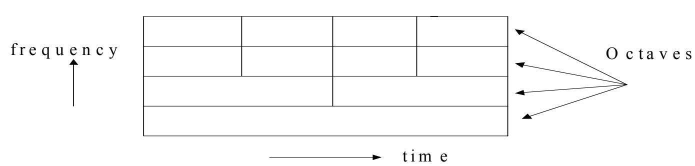
Short-Term Fourier Transform (STFT)

(Read the tutorial posted in the course web page.)

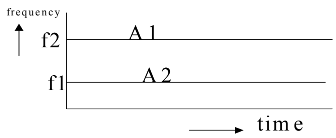
The STFT is an attempt to alleviate the problems with FT. It takes a non-stationary signal and breaks it down into “windows” of signals for a specified short period of time and does FT on the window by considering the signal to consist of repeated windows over time. This gives a better picture of the frequency content of the signal over the segment period but suffers from the disadvantage that the analysis is error prone if the size of the “window” is not suitably chosen.

The following three diagrams depict the differences between FT, STFT and Wavelets For wavelets the time-frequency tiling looks quite different. The frequency spectrum is divided into octaves (scale or pitch). The signals at higher octaves have fine time resolution and the resolution becomes coarser and coarser as the octaves are smaller signifying lower frequencies. The is an ideal scheme for muti-resolution analysis since human eyes are very sensitive to lower frequencies and have good visual perception of the object even if a good fraction of higher frequencies are totally missing from its representation. Of course, for lossless compression we need all the frequency components and enables multi-resolution analysis of the signal easily.

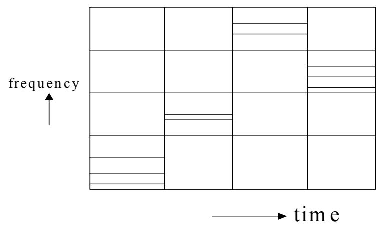
For wavelets the time-frequency tiling looks quite different. The frequency spectrum is divided into octaves (signifying scale or pitch). The signals at higher octaves have fine time resolution and the resolution becomes coarser and coarser as the octaves are smaller signifying lower frequencies.



Wavelet: At low frequency, we use a larger time window. With increasing frequency, the time window gets smaller according to the uncertainty principle explained earlier.



Fourier Transform: The signal is depicted to have only two frequencies with amplitudes A1 and A2 extending over the entire time domain.



STFT: The time-freqquency plane is broken up into smaller boxes, each box denoting certain frequency and time spread.

Wavelet Fundamentals (Read Sayood Section 14.3, pp.459-462, Section 14.4, pp.462-468. Salomon Section 5.5, pp.467-474, Section 5.6 , pp. 474-484, Section 5.6.3, pp. 490-491)

CWT – Continuous Wavelet Transfomation

- $\psi(t)$ – A "mother" wavelet function.
- A scaled and translated mother wavelet: $\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi(\frac{t-b}{a})$

where a = scaling factor, b = translation to right.

- $f(t)$: an arbitrary function of time
- Wavelet transform

$$w_{a,b} = <\psi_{a,b}(t), f(t)>$$
$$= \int_{-\infty}^{\infty} \psi_{a,b}(t) f(t) dt$$
$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{a}}\psi(\frac{t-b}{a}) dt$$

- The inverse wavelet transform

$$w_{a,b} = \frac{1}{C_\psi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w_{a,b}\psi_{a,b}(t) da db$$

Two conditions:

(i) Wave condition: $\int_{-\infty}^{\infty} \psi_{a,b}(t) = 0$

(ii) Admissibility condition: $\int_{-\infty}^{\infty} |\psi_{a,b}(t)|^2 \; dt < \infty$

For any pair of values $(a,b) \in R^2$, there is $w_{a,b}$; The entire *real* plane is the *support* of $w_{a,b}$.

DWT – Discrete Wavelet Transform

If *(a,b)* take discrete value in R^2, we get DWT. A popular approach to select *(a,b)* is

$$a = \frac{1}{a_0^m} \qquad a_0 = 2 \;\rightarrow\; a = \frac{1}{2^m} = <1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots>. \quad m\text{: integer}$$

$$b = \frac{nb_0}{a_0^m} \qquad a_0 = 2,\; b_0 = 1, \quad b = \frac{n}{2^m} \qquad n, m\text{: integer}$$

$$\therefore \quad \psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) = 2^{m/2}\psi\left(\frac{t-\frac{n}{2^m}}{\frac{1}{2^m}}\right) = \psi_{m,n}(t) = 2^{m/2}\psi\left(2^m t - n\right)$$

Wavelet Transform:

$$w_{m,n} = < f(t), \psi_{m,n}(t) > = 2^{m/2}\int f(t)\psi(2^m t - n)dt$$

Inverse Wavelet Transform

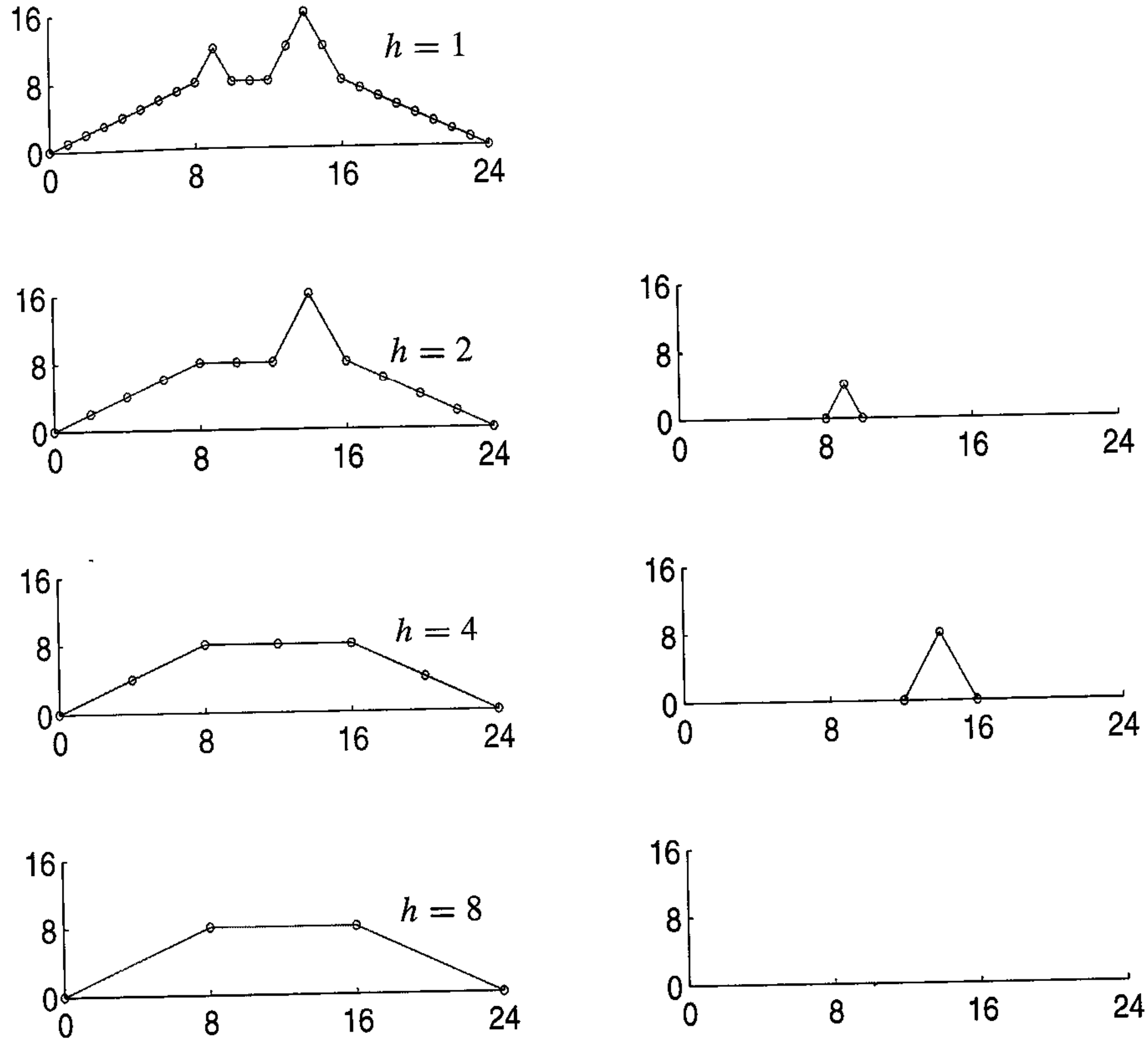
$$f(t) = \sum_m \sum_n w_{m,n}\psi_{m,n}(t)$$

If *f(t)* is continuous while $w_{a,b}$ consists of discrete values, the series is called the discrete time wavelet transform (DTWT). If *f(t)* is sampled (that is, discrete in time, as well as $w_{a,b}$ are discrete, we get discrete wavelet transform (DTWT).

Multi-resolution Analysis (MRA)

Let's say , we want to represent a real number N=87/7=12. 4285714… by a series of successive approximation. Depending on the desired accuracy, we can approximate N successively as sequence of "round-off" values 10, 12, 12.4,12.42, …etc. The successive difference between two consecutive round-off values (2, 0.4, 0.02, 0.008, …) is called the "detail" part. The round off values are sometimes also called the "averages". In multi-resolution analysis, we use a *scaling function* to represent the round off values and a *wavelet function* to represent the detail values. The further we descend the level of details, the more accurate is the approximation. In the other direction, if we "stretch" the scaling function more and more, we end up seeing nothing at as if we trying to approximate 87/7 by 100's and at that point all information is in the details: 10+2+0.4+0.02+0.008. This numerical example gives only a conceptual idea of multi-resolution analysis.

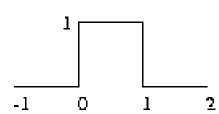
Let's take another example. A function is to be viewed at various levels of approximation and resolutions. The idea is to divide a complicated function into several simple functions and study them separately. If a function has both a slowly varying as well as a rapidly varying component, we have to discretize it using step size (*h*) determined by the rapidly varying segment. This needs a large number of data point. The coarsest approximation of the function together with the details at every level completely represents the original function. Note with every level (scale) the step size is doubled.



Study MRA equations and applications (Section 14.5, 14.6 and 14.7) from Sayood.

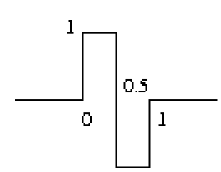
Define a scaling function $\phi(t)$.

$$\phi(t) = \begin{cases} 1 & 0 \le t < 1 \\ 0 & \text{otherwise} \end{cases}$$

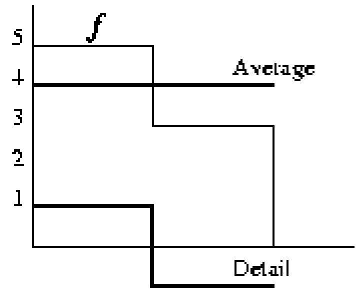


Harr Mother wavelet

$$\psi(t) = \begin{cases} 1 & 0 \le t < \frac{1}{2} \\ -1 & \frac{1}{2} \le t < 1 \end{cases}$$



Ex.1. Now consider, a function



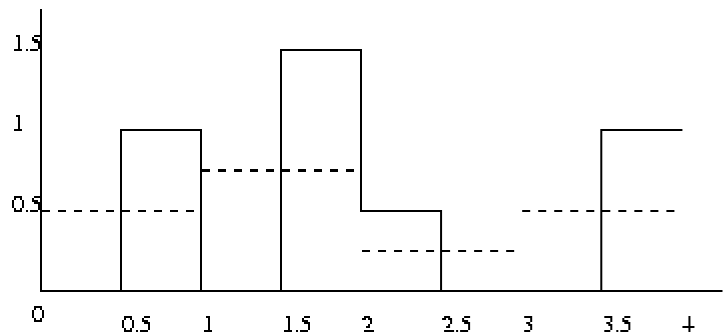
Do you agree?

$$f(t) = 4 \bullet \phi(t) + 1 \bullet \psi(t)$$

average detail

The coefficient for $\phi(t)$ is the average = (5+3)/2 = 4.
The coefficient for $\psi(t)$ is the difference avg. = (5-3)/2 = 1.

Ex.2.



$$\begin{aligned} f(t) = {} & 0.5\phi(t) + 0.75\phi(t-1) + 0.25\phi(t-2) + 0.5\phi(t-1) \\ & - 0.5\psi(t) + (-0.75)\psi(t-1) + 0.25\psi(t-2) + (-0.5)\psi(t-3) \end{aligned}$$

Haar Transform

$$f(t) = \sum_{k=-\infty}^{\infty} c_k \phi(t-k) + \sum_{k=-\infty}^{\infty} \sum_{j=0}^{\infty} d_{j,k} \psi(2^j t - k)$$

Interpretation of Computing Haar transform as averages and details taking O(N) time.
Examples of Haar Trasnforms (pp. 479-488, Salomon)).

Matrix multiplication interpretation of Haar transform (Salomom, pp.490-491)

Subband Coding

Subband coding is a technique of decomposing the source signal into constitutent parts and coding the parts separately.

One drawback of transform coding is that it artificially divides the source output into "blocks", which might give rise to blocking artifacts. In subband coding the source signal output is decomposed into "constituent" frequencies. The constituents might have characteristics that relates to specific features of the image, compressability, R(D) values, quantization rates etc.

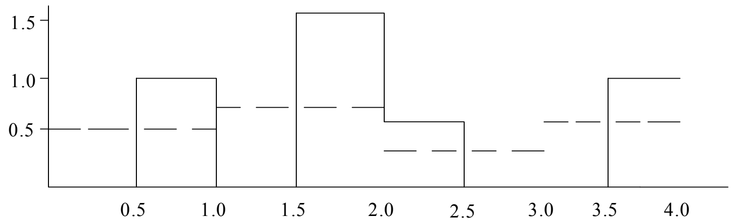
Read pp. Sayood, Chapter 13 Sections 13.2 and 13.3, pp.405-416

Nyquist's Law : If the highest frequency component of a signal is f_0 , then we need to sample the signal at more than or equal to 2 f_0 times per second in order to achieve accurate reconstruction of the signal. If the signal has frequency components between f_1 and f_2 , then we need to sample the signal at more than or equal to $2(f_2 - f_1)$ times per second to have faithful reconstruction. If this law is violated, a kind distortion called *aliasing* take place which introduces unwanted signals with frequencies higher than half the sampling rate at lower frequencies. To prevent this kind of distortion, the system is equipped with *anti-aliasing filter* which restricts the input to the sampler to be less than half the sampling frequency.

The remainder of the reading assignment in Sayood gives a basic introduction to signal processing and explains concepts like *finite impulse response* (FIR), *convolution* and *quadrature mirror filter* and filter banks for analysis and synthesis. Although not required for your final exam, this material is fundamental for your understanding of multi-resolution analysis using wavelets. (You may read this material from any standard text book on signal processing).

Multi-Resolution Analysis

Let's now take a hypothetical signal as shown in Ex 2 and repeated below.



$$f = 0.5\phi(t) + 0.75\phi(t-1) + 0.25\phi(t-2) + 0.5\phi(t-3)$$
$$-0.5\psi(t) + -0.75\psi(t-1) + 0.25\psi(t-2) - 0.5\psi(t-3)$$

Note, if the function has a resolution at a frequency f. Then both the scaling and the wavelet function has a resolution at half the frequency f. We have used the Haar scaling and wavelet functions to express the given function. In general we can write

$$f(t) = \sum_{k=0}^{\infty} c_k \phi(t-k) + \sum_{k}^{\infty} \sum_{j=0}^{\infty} d_{j,k} \psi(2^j t - k)$$

The multi-resolution can be performed if the signal satisfies the following conditions:

1) The scaling function must be orthogonal to its translates by integers.

2) The signal at any resolution contains all the information of the signal at coarses resolution. This is mathematically expressed by talking about the space V_j.

   V_0= {space generated by the scaling function and its translates}

   V_1= { space generated by scaling function compressed by a factor of two and translated by half integers} and so on for V_j. We also know that

   $V_0 \subset V_1 \subset V_2 \subset V_3 \subset$........

3) $\bigcap V_j = \{0\}$

3) Any signal can be approximated with arbitrary precision.

If the above four conditions are satisfied, then there exists a wavelet that, with its translates, by integers and dilates by a factor of two can encode the difference of information between the signal seen at two successive resolutions. In mathematical terms, as we have seen earlier,

$$W_j \otimes V_j = V_{j+1}$$

Multi-resolution theory gives a simple and fast method for decompressing a signal into its components at different scales. We progressively drain the signal of its information . At each step, we encode the details as wavelet coefficients and then work at the next level with the signal seen at half its previous resolution.

In the language of wavelet theory, the scaling function is *dilated* to make an image of the signal at half resolution. In the language of signal processing, a low pass filter is applied to the signal and the result is sub-sampled, taking only half as many samples.

Filter Banks

The Haar transform can be looked upon as a bank of **filters, lowpass and highpass filters.** A tree structured architecture of low and high pass filters in tandem with **down-sampling or decimation** can also be interpreted as sub-band decomposition corresponding to multi-resolution analysis**.**

A filter is a linear operator in terms of filter coefficients h(0), h(1), h(2), …h(t). The number t is called the "tap" of the filter. Thus if t=2, it is called a 2-tap filter. See a few more examples of higher tap filters in p. 513, Sayood. The filter is used to transform the input vector x(n) to an output vector y(n) by the **Convolution** operation, as defined earlier:

$$y(n) = \sum_k h(k) \bullet x(n-k)$$

for all values of n=1,2, …. For the sampled input sequence x(n).

Matrix Formulation of the Filter Bank Operations

In practice, digital filters perform these scaling (smoothing the signal) and wavelet (picking the differences) operations using simple arithmetic of multiplication and addition via the ***convolution operation***.

The k th term of the convolution of two sequences *a* and *b* is defined as

$$(a \oplus b)_k = \sum_{j=-\infty}^{\infty} a_j b_{k-j}$$

If a_j and b_j are non-zero for j greater than or equal to 0, then

$$(a \oplus b)_k = \sum_{j=0}^{k} a_j b_{k-j}$$

For example , $$(a \oplus b)_2 = \sum_{j=0}^{2} a_j b_{2-j} = a_0 b_2 + a_1 b_1 + a_2 b_0$$

If we interpret *a* and *b* to be represent a radix 10 number, the the above expansion gives how many hundreds are there in the product of the numbers *a* and *b*.

To transform a signal into wavelets, we *convolve* the sampled signals with the *a*-coefficients (the low pass filter coefficients associated with the scaling function) and the *b*-coefficients (so denoted as *d*-coefficients (the high-pass filter coefficient associated with the wavelet corresponding to the scaling function). A mathematical proof of this statement is given in Sayood and other references on wavelets. For our running example,

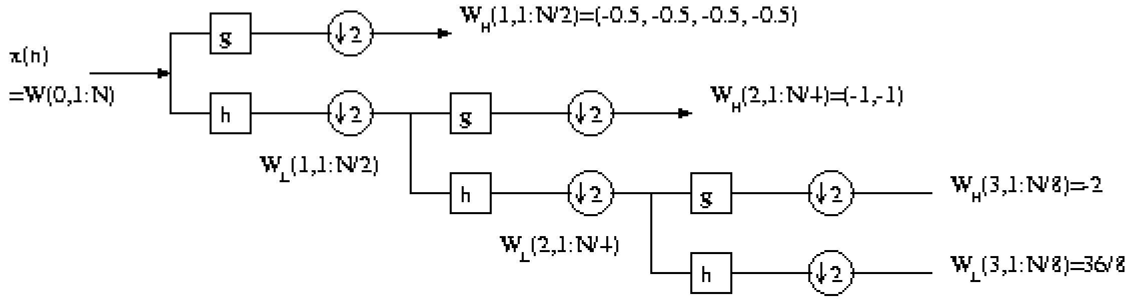
Haar scaling function has: $a_{-1}=0.5$, $a_0=0.5$ (average computation) and the corresponding Haar wavewlet co-efficients are $d_{-1}=-0.5$, $d_0=0.5$ (difference) [ignoring a square root of 2 factor]. The signal has values are depicted in the last figure. Now, convolve the signal values with the sequence *a*. The values at time t=0,2,4 6 are .5, .75, .25 and .5. These correspond to averages. If we convolve the signal with the *d*-sequence, we get .5, -.75, .25, -.5 which are the same as we observed earlier.

A more practical way to compute this transformation is given below. For the average and the detail coefficients, two different filters are used. Since the output is computed for every input, this will produce all average and details for consecutive inputs pairs. Since we need this output for every consecutive disjoint pairs, a down sampling step is needed.

Let W_L and W_H denote the low-pass and the high-pass coefficients. $j = 1,2,3,\ldots,\log N$, $N = 2^j$, j is called the **octave** value. The 0^{th} row of the 2-dimensional matrix W is initialized with the input values $X = [x(1), x(2),\ldots, x(n)]$.
Example:
Let the Harr filter be denoted as:

	m=0	m=1
h(m)	0.5	0.5
g(m)	-0.5	0.5

Let the input sampled signal values be $X = (1,2,3,4,5,6,7,8)$



Imagine the computation to proceed on a matrix of $\log_2 N+1$ rows. At the 0th row, there are 8 elements (for illustration, assume N=8). The first octave produces two groups: $N/2$ low resolution values and $N/2$ high resolution (detail) values. In the next octave computation, only the first $N/2$ values are modified to produce their averages and detail values. This process is recursed until at the $\log_2 N$ -th octave we only have one average and one detail value. This is showed in the following matrix:

This can be expressed mathematically as:

$$W_L(j,n)=\sum_{m=0}^{m=1} W_L(j-1,2n-m)h(m) \qquad n=1,2,\ldots,\frac{N}{2^j}$$

$$W_L(j,n)=\sum_{m=0}^{m=k} W_H(j-1,2n-m)g(m) \qquad n=\frac{N}{2^j}+1,\ldots,\frac{N}{2^{j-1}}$$

Let's do one row:

j=1,   h(0)=0.5,   h(1)=0.5, n=1,2,…,8

$W_L(1,1)=2\times(0.5)+1\times(0.5)=1.5$

$W_L(1,2)=4\times(0.5)+3\times(0.5)=3.5$ …,etc.

$W_H(1,5)=2\times(-0.5)+1\times(0.5)=-0.5$

$W_H(1,6)=4\times(-0.5)+3\times(0.5)=-0.5$ …,etc.

	j\n	1	2	3	4	5	6	7	8	
$W_L(0,N)$	0	1	2	3	4	5	6	7	8	
$W_L(1,1:\frac{N}{2})$	1	1.5	3.5	5.5	7.5	-0.5	-0.5	-0.5	-0.5	$W_H(1,\frac{N}{2}+1:N)$
$W_L(2,1:\frac{N}{4})$	2	2.5	6.5	-1	-1					$W_H(2,\frac{N}{4}+1:\frac{N}{2})$
$W_L(3,1:\frac{N}{8})$	3	4.5	-2							$W_H(3,\frac{N}{8}+\ :\frac{N}{4})$