Wavelet Transform and its relation to multirate filter banks

Christian Wallinger
ASP Seminar 12th June 2007
Graz University of Technology, Austria

Outline

Short – Time Fourier – Transformation

- Interpretation using Bandpass Filters
- Uniform DFT Bank
- Decimation
- Inverse *STFT* and filter - bank interpretation
- Basis Functions and Orthonormality
- Continuous Time *STFT*

Wavelet – Transformation

- Passing from *STFT* to Wavelets
- General Definition of Wavelets
- Inversion and filter - bank interpretation
- Orthonormal Basis
- Discrete – Time Wavelet Transf.
- Inverse

First, we will develop the short – time Fourier transform (*STFT*) and its relation to filter banks and then the wavelet transform and its relation to multirate filter banks.

Therefore it is much easier to understand, if first the discret time *STFT* and afterwards the continuous time *STFT* will be introduced. Followed by continuous wavelet transform and discret wavelet transform.

• SHORT-Time FOURIER TRANSF.

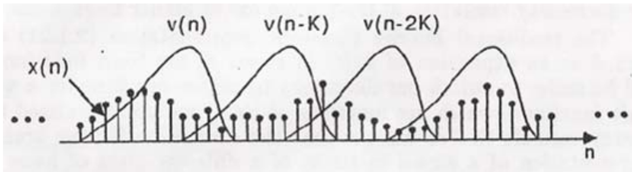


figure 1: *STFT processing in time*

time – frequency plot = Spectogram

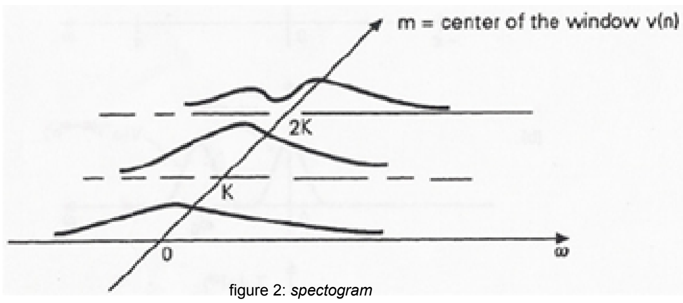


figure 2: *spectogram*

In short – time Fourier transform, a signal *x(n)* is multiplied with a window *v(n)* (typically finite in duration). → The Fourier – transform of the time domain product *x(n)v(n)* is computed, and then the window is shifted in time, and the FT of the new product computed again. (figure 1)

This operation results in a separate FT for each location *m* of the center of the window, which is typically an integer multiple of some fixed integer *K*). (figure 2)

Definition:

$$X_{STFT}\left(e^{j\omega},m\right)=\sum_{n=-\infty}^{\infty}x(n)v(n-m)e^{-j\omega n}$$

m . . . time shift – variable
(typically an integer multiple of some fixed integer K)

ω . . . frequency – variable $-\pi \le \omega < \pi$

From above discussion it is clear that the *STFT* can be written mathematically as shown in the slide, where ω is continuous and takes the usual range between $-\pi$ and $+\pi$.

• *Interpretation using Bandpass Filters*

Traditional Fourier Transform as a Filter Bank

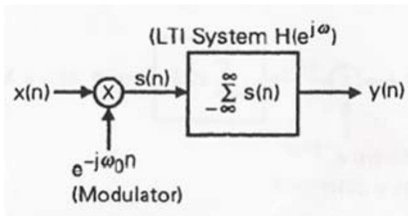


figure 3: Representation of FT in terms of a linear system

1. Modulator $e^{-j\omega_0 n}$ *: → performs a frequency shift*

2. LTI – System $H\left(e^{j\omega}\right)$ *: → ideal lowpass filter*

Before interpreting the *STFT* in terms of filter banks, we will begin by representing a filter bank interpretation for the traditional Fourier – Transform. (figure 3)

→ Figure 3 represents only one channel for one specific frequency ω_0.

Why is $H\left(e^{j\omega}\right)$ an ideal lowpass filter ?

Impulse Response ***h(n) = 1 for all n***

$$\rightarrow \quad H\left(e^{j\omega}\right) = \sum_{n=-\infty}^{\infty} h(n) e^{-j\omega n} = 2\pi\delta_a(\omega) \qquad -\pi \le \omega < \pi$$

→ only zero - frequency passes

→ every other frequency is completely supressed

$$y(n) = X\left(e^{j\omega_0}\right) \qquad \text{for all } n$$

$h(n) = 1$ for all n. This system is evidently unstable, but let us ignore these fine details for the moment.

$\delta_a(\omega)$ is the Dirac delta function.

Summarizing, the process of evaluating $y(n) = X\left(e^{j\omega_0}\right)$ can be looked upon as a linear system, which takes the input $x(n)$ and produces a *constant* output $y(n)$.

Therefore, the FT operator is a bank of modulators followed by filters. This system has an uncountably infinite number of channels.

STFT as a Bank of Filters

→ *Expansion* of *Definiton* for further insight!

$$X_{STFT}\left(e^{j\omega}, m\right) = e^{-j\omega m} \sum_{n=-\infty}^{\infty} x(n) v(n-m) e^{j\omega(m-n)}$$

with:

$$v(n-m)e^{j\omega(m-n)} = v(-(m-n))e^{j\omega(m-n)}$$

→ Convolution of x(n) with the impulse response of the LTI – System $v(-n)e^{j\omega n}$

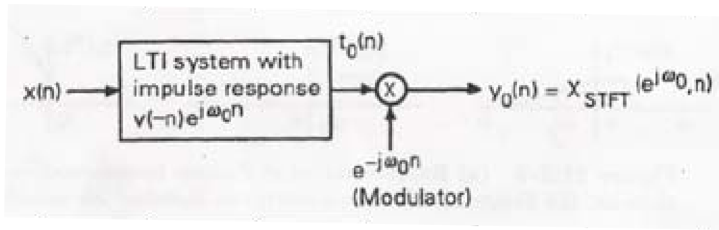


figure 4: Representation of *STFT* in terms of a linear system

In most applications, $v(n)$ has a lowpass transform $V(e^{j\omega})$.

→

$$v(-n) \circ\!\!-\!\!\bullet \; V(e^{-j\omega})$$

$$v(-n)e^{j\omega_0 n} \circ\!\!-\!\!\bullet \; V(e^{-j(\omega-\omega_0)})$$

Figure 4 shows the interpretation of the *STFT* in terms of a filter bank. (Again, only one channel can be seen).

The first is an LTI filter followed by the modulator.

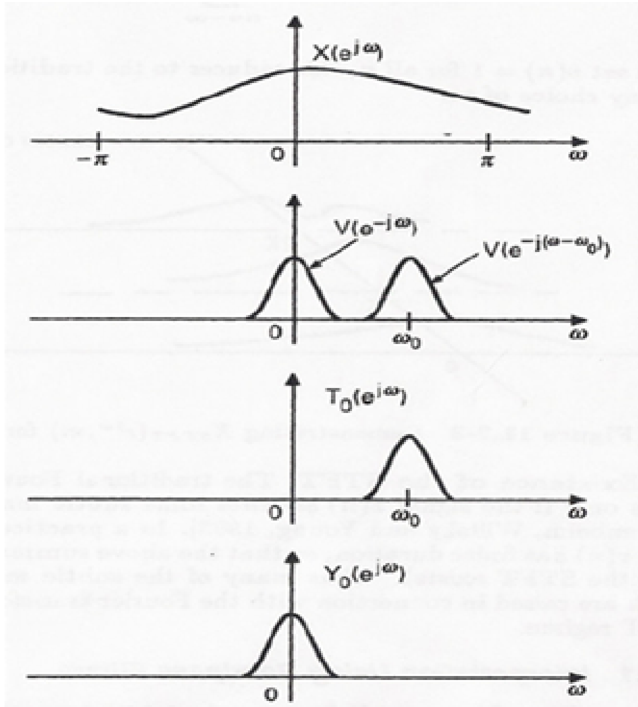


figure 5: Demonstration of how *STFT* works

Figure 5 demonstrates how the *STFT* works.

(a) FT of an arbitrary choosen input signal *x(n)*

(b) the window – transform and its shifted version

(c) output of LTI filter

(d) traditional Fourier transform of $X_{STFT}\left(e^{j\omega_0}, n\right)$

Hence, the *STFT* can be looked upon as a filter bank, with *infinite* number of filters (one per frequency) !

In practice, we are interested in computing the Fourier transform at a discrete set of frequencies

→ $0 \leq \omega_0 < \omega_1 < \ldots < \omega_{M-1} < 2\pi$

Therefore the STFT reduces to a filter bank with M bandpass filters

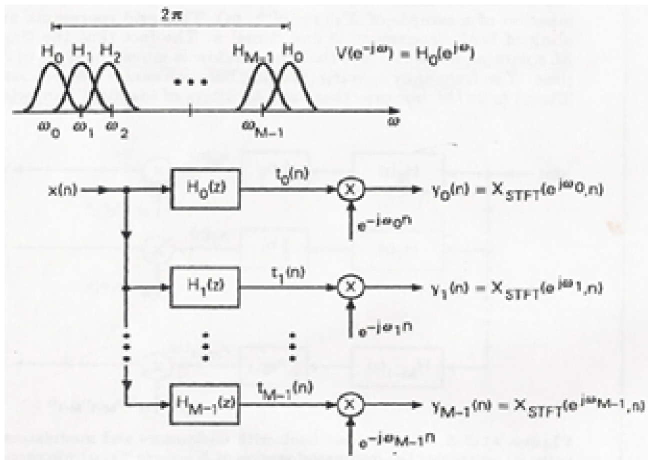


figure 6: *STFT* viewed as a filter bank

Uniform DFT bank

If the frequencies ω_k are uniformly spaced, then the system becomes the uniform DFT bank.

The *M* filters are related as in the following manner

$$H_k(z) = H_0\left(zW^k\right) \qquad 0 \le k \le M-1 \qquad W = e^{-j\frac{2\pi}{M}}$$

$$\rightarrow \qquad H_k(e^{j\omega}) = H_0\left(e^{j(\omega - \frac{2\pi}{M}k)}\right) \qquad H_0(e^{j\omega}) = V(e^{-j\omega})$$

→ The *uniform DFT bank* is a device to compute the *STFT* at uniformely spaced frequencies.

→ The frequency responses $H_k\left(e^{j\omega}\right)$ are uniformly shifted versions of $H_0\left(e^{j\omega}\right)$

Decimation

if passband width of $V(e^{j\omega})$ is narrow

→ output signals $y_k(n)$ *are narrowband lowpass signals*

this means, that *yk(n)* varies slowly with time

According to this variying nature, one can exploit that to decimate the output.

Decimation Ratio of M = moving the window *v(k)* by M samples at a time

if filters have *equal* bandwidth

→ $n_k = M$

→ *maximally decimated* analyses bank

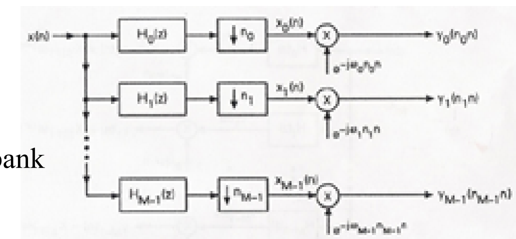


figure 7: Analysis bank with decimators

Figure 7 shows a decimated *STFT* system, where the modulators have been moved past the decimators.

In a more general system n_k could be different for different k, and moreover $H_k(z)$ may not be derived from one prototype by modulation. Such a system, however, does not represent the *STFT* obtainable by moving a single window across the data *x(n)*. → this systems will be admitted in the wavelet transform.

Time – Frequency Grid

Uniform sampling of both, ‘*time’ n* and ‘*frequency’* ω

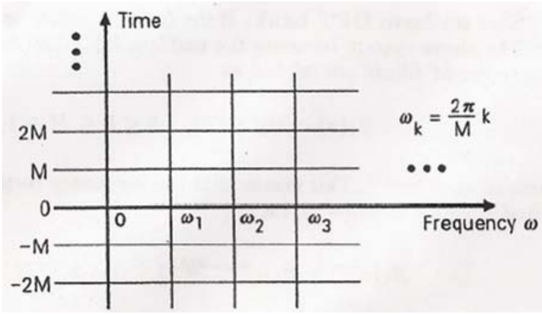


figure 8: time – frequency grid

Time spacing *M* corresponds to moving the window *M* units (= samples) at a *time*.

$$\text{frequency spacing of adjacent filters} = \frac{2\pi}{M}$$

Inversion of the STFT

From traditional Fourier – viewpoint

$X_{STFT}\left(e^{j\omega},m\right)$ is the FT. from the time domain product

$x(n)v(n-m)$

→

$$x(n)v(n-m)=\frac{1}{2\pi}\int_0^{2\pi}X_{STFT}\left(e^{j\omega},m\right)e^{j\omega n}d\omega$$

For example, if we set n = m we obtain the *STFT* inversion formula for *x(m)* as long as *v(0)* exists. If it does not, we can pick some other value of *m*.

Another inversion formula is given by:

$$x(n) = \frac{1}{2\pi} \int_0^{2\pi} \left(\sum_{m=-\infty}^{\infty} X_{STFT}\left(e^{j\omega}, m\right) v^*(n-m) \right) e^{j\omega n} d\omega$$

which is provided by $\sum_m |v(m)|^2 = 1$

if $\sum_m |v(m)|^2 \neq 1$ but finite→ divide right side of the formula by $\sum_m |v(m)|^2$

but if window energy is infinite → one cannot apply this formulation

Filter Bank Interpretation of the Inverse

With $F_k(z)$ as synthesis - filter

Reconstruction can be done by the following synthesis bank:

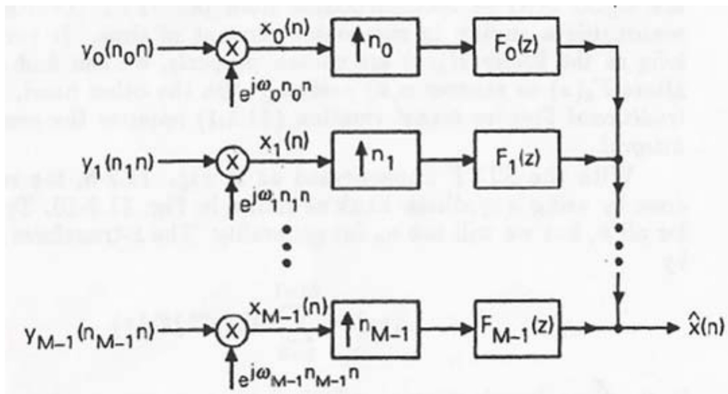


figure 9: synthesis – bank used to reconstruct x(n)

typically $n_k = M$ for all *k*

The *z – Transformation* of $\hat{x}(n)$ is given by

$$\hat{X}(z) = \sum_{k=0}^{M-1} X_k\left(z^{n_k}\right)F_k(z)$$

in *time* – domain

$$\hat{x}(n) = \sum_{k=0}^{M-1}\sum_{m=-\infty}^{\infty} x_k(m) f_k(n - n_k m)$$

$$= \sum_{k=0}^{M-1}\sum_{m=-\infty}^{\infty} y_k(n_k m) e^{j\omega_k (n_k m)} f_k(n - n_k m)$$

$y_k(n_k m)\ldots STFT\ - Coefficien\ ts$

Reconstruction is stable, if the filters $F_k(z)$ are stable!

Perfect reconstruction will be obtained, if $\hat{x}(n) = x(n)$

Basis Functions and Orthonormality

Functions of interest

$$\eta_{km}(n) \hat{=} f_k(n - n_k m) \ldots basis\ functions$$

For these double indexed functions (*basis functions* $\{\eta_{km}(n)\}$),
the orthonormality property means that

$$\sum_{n=-\infty}^{\infty} f_{k1}^*(n - n_{k1}m_1) f_{k2}(n - n_{k2}m_2) = \delta(k_1 - k_2)\delta(m_1 - m_2)$$

should be zero, except for those cases where $k_1 = k_2 \ and \ m_1 = m_2$

Remember: k . . . filter number m . . . time shift

How should we design the filters $F_k(z)$ in order to ensure this orthonormality property ?

Therefore, the paraunitary property of the polyphase matrix is sufficient!

The Continuous - Time Case

Main points:

$$X_{STFT}(j\Omega,\tau)=\int_{-\infty}^{\infty}x(t)v(t-\tau)e^{-j\Omega t}\,dt \qquad (STFT)$$

$$x(t)v(t-\tau)=\frac{1}{2\pi}\int_{-\infty}^{\infty}X_{STFT}(j\Omega,\tau)e^{j\Omega t}\,d\Omega \qquad (inv.\ STFT)$$

$$x(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}X_{STFT}(j\Omega,\tau)v^{*}(t-\tau)\,d\tau\right)e^{j\Omega t}\,d\Omega \qquad (inv.\ STFT)$$

Because of the close resemblance to the discrete – time case, we only summarize the main points for the continuous – time case.

Historically, the *STFT* was first developed for the continuous – time case by Dennis Gabor.

Choice of "Best Window"

***R**oot **M**ean **S**quare duration of window function v(t) in*

time domain D_t

$$D_t^2 = \frac{1}{E}\int_{-\infty}^{\infty} t^2 v^2(t)\,dt$$

frequency domain D_f

$$D_f^2 = \frac{1}{2\pi E}\int_{-\infty}^{\infty} \Omega^2 \left|V(j\Omega)\right|^2 d\Omega$$

with:
E . . . window energy $E = \int v^2(t)\,dt$

Uncertainty principle:

$$D_t D_f \geq 0.5$$

Iff Gaussian – window, this inequality becomes an equality !

D_t is the rms time domain duration and D_f the rms frequency domain duration of the window.

Filter Bank Interpretation

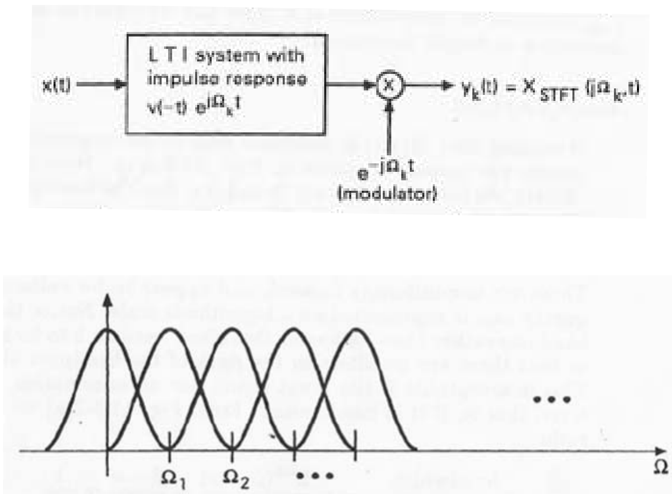


figure 10: continuous – STFT

Figure 10 shows again the filtering interpretation for the continuous – time *STFT*.

• THE WAVELET TRANSFORM

Disadvantage of STFT

uniform time – frequency box $\left(D_t = const., D_f = const.\right)$

→ The accuracy of the estimate of the Fourier transform is poor at low frequencies, and improves as the frequency increases.

⇨ *Expected properties for a new function:*

- window width should adjust itself with 'frequency'
- as the window gets wider in time, also the step sizes for moving the window should become wider.

These goals are nicely accomplished by the ***wavelet transform*****.**

Passing from STFT to Wavelets

Step 1:

Giving up the *STFT* modulation scheme and obtain filters

$$h_k(t) = a^{-k/2} h(a^{-k} t) \qquad a > 1 \ldots scaling\ factor, k = \text{int}eger$$

in the frequency domain:

$$H_k(j\Omega) = a^{k/2} H(ja^k \Omega)$$

→ all reponses are obtained by *frequency – scaling* of a prototype response $H(j\Omega)$

This is unlike the case of *STFT*, where all filters were obtained by *frequency – shift* of a prototype.

The scale factor $a^{-k/2}$ is meant to ensure that the energy $\int_{-\infty}^{\infty} |h_k(t)|^2 dt$ is independent of *k*.

Example:

Assuming $H(j\Omega)$ is a bandpass with cutoff frequencies α and β.

Also $a = 2,\ \beta = 2\alpha$ and the center frequency should be the geometrical mean of the two cutoff edges

$$\Omega_k = 2^{-k}\sqrt{\alpha\beta} = \alpha\, 2^{-k}\sqrt{2}$$

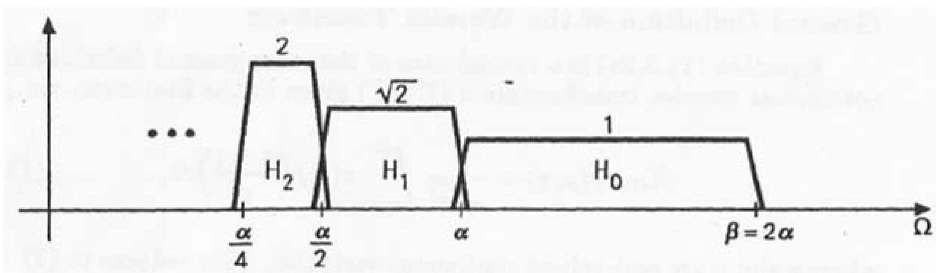


figure 11: frequency – response obtained by scaling process

Ratio:

$$\frac{bandwidth}{center-frequency\ \Omega_k} = \frac{2^{-k}(\beta-\alpha)}{2^{-k}\sqrt{\alpha\beta}} = \frac{1}{\sqrt{2}}$$

is independent of integer *k*

In *electrical filter theory* such a system is often said to be a *'constant Q'* system!

(Q ... Quality factor $Q = \dfrac{center - frequency}{bandwidth}$)

→ filter ouputs can be obtained by:

$$a^{-k/2} \, e^{-j\Omega_k \tau} \int_{-\infty}^{\infty} x(t) h\left(a^{-k}(\tau - t)\right) dt$$

Step 2:

$$k \uparrow \quad \rightarrow \quad bandwidth\ of\ H_k(j\Omega) \downarrow \quad \rightarrow \quad Samplerate \downarrow$$

or in *time* domain

$$k \uparrow \quad \rightarrow \quad window\ length \uparrow \quad \rightarrow \quad step\ size \uparrow$$

Since the bandwidth of $H_k(j\Omega)$ is smaller for larger k, we can sample its output at a correspondingly lower rate. → Viewed in time domain, the width of $h_k(t)$ is larger so that we can afford to move the window by a larger step size!

Therefore:

$$\tau = na^k T \qquad n \ldots \text{integer}, \quad a^k T \ldots \text{step size}$$

hence:

$$h\left(a^{-k}\left(na^k T - t\right)\right) = h\left(nT - a^{-k} t\right)$$

Summarizing, we are computing:

$$X_{DWT}(k,n) = a^{-k/2} \int_{-\infty}^{\infty} x(t) h\left(nT - a^{-k} t\right) dt$$

→

$$X_{DWT}(k,n) = \int_{-\infty}^{\infty} x(t) h_k\left(na^k T - t\right) dt$$

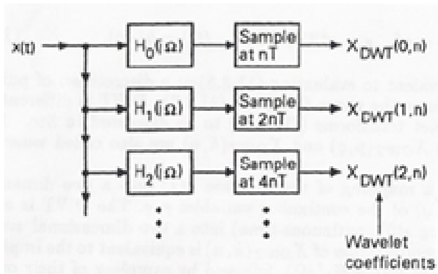


figure 12: Analysis bank of DWT

DWT...Discrete Wavelet Transform

This can be done by replacing the continuous variable τ as shown in the slide.

The modulation factor $e^{-j\Omega_k \tau}$ has been omitted.

What we can see is, that the above integral represents the convolution between *x(t)* and $h_k(t)$, evaluated at a discrete set of points na^kT. In other words, the output of the convolution is sampled with spacing a^kT. (figure 12 is a schematic of this for a = 2).

Time Frequency Grid

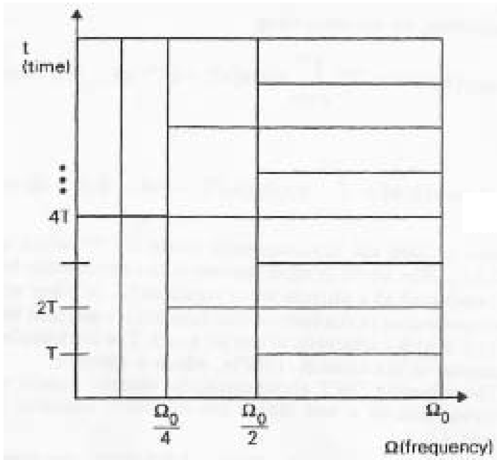


figure 13: time – frequency grid

$$D_t D_f = const.$$

Frequency spacing is smaller at low frequencies, and the corresponding time – spacing is larger.

General Definition of the Wavelet Transform

$$X_{CWT}(p,q) = \frac{1}{\sqrt{|p|}} \int_{-\infty}^{\infty} x(t) f\left(\frac{t-q}{p}\right) dt$$

p,q ... real – valued continuous variables

According to former definition:

$$p = a^k \qquad q = a^k T n \qquad f(t) = h(-t)$$

$$X_{CWT}(p,q) \quad \text{and} \quad X_{DWT}(k,n) \quad \ldots\ldots\ldots \quad \text{wavelet coefficients}$$

Inversion of Wavelet Transform

$$x(t) = \sum_k \sum_n X_{DWT}(k,n) \psi_{kn}(t)$$

where $\psi_{kn}(t)$ are the basis functions

Filter Bank Interpretation of Inversion

Reconstruction of *x(t)* as a designing problem of the following synthesis filter bank

$X_{DWT}(k,n)$ … *sequence*

$F_k(j\Omega)$ … *continuous in time*

→ output of synthesis filter bank :

$$\hat{x}(t) = \sum_k \sum_n X_{DWT}(k,n) f_k\left(t - a^k nT\right)$$

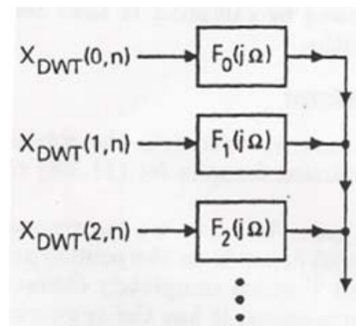


figure 14: synthesis bank

Figure 14 shows the synthesis filter bank.

We have to be careful with the interpretation of this figure. Since $X_{DWT}(k,n)$ is a *sequence*, the signal which is input to the *continuous –time* filter $F_k(j\Omega)$ is actually an impulse train.

All synthesis filters are again generated from a fixed prototype synthesis filter *f(t) (→ mother wavelet)*

$$f_k(t) = a^{-k/2} f\left(a^{-k} t\right)$$

Substituting this in the preceding equation and assuming perfect reconstruction, we get

$$x(t) = \sum_k \sum_n X_{DWT}(k,n) a^{-k/2} f\left(a^{-k} t - nT\right)$$

with:

$$\psi(t) = f(t) \quad \rightarrow \quad \psi_{kn}(t) = a^{-k/2} \psi\left(a^{-k} t - nT\right) = a^{-k/2} \psi\left[a^{-k}\left(t - na^k T\right)\right] \ldots set\ of\ basis\ functions$$

using this, we can express each basis function in terms of the filter $f_k(t)$

$$\psi_{kn}(t) = f_k\left(t - na^k T\right)$$

Orthonormal Basis

Of particular interest is the case where $\{\psi_{kn}(t)\}$ is a set of orthonormal functions

Therefore, we expect:

$$\int_{-\infty}^{\infty} \psi^*_{kn}(t)\psi_{lm}(t)dt = \delta(k-l)\delta(n-m)$$

using Parseval's theorem, this becomes

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \Psi^*_{kn}(j\Omega)\Psi_{lm}(j\Omega)d\Omega = \delta(k-l)\delta(n-m)$$

and get :

$$X_{DWT}(k,n) = \int_{-\infty}^{\infty} x(t)\psi^*_{kn}(t)dt$$

Comparing these results, we can conclude:

$$\psi_{kn}(t) = h^*_k(a^k nT - t)$$

And in particular for *k = 0 and n = 0:*

$\psi_{00}(t) = \psi(t) = h^*(-t)$ → for the *orthonormal case* → $f_k(t) = h^*{}_k(-t)$

Discrete – Time Wavelet Transform

Starting with the frequency domain relation and a scaling factor *a* = 2

$$H_k(e^{j\omega}) = H(e^{j2^k\omega}) \quad \ldots \quad k \text{ is a nonnegative integer}$$

for highpass $H(e^{j\omega})$ and *k = 1, k = 2*

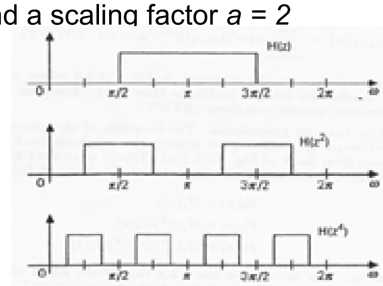


figure 15: Magnitude responses

Let *G(z)* be a lowpass with response

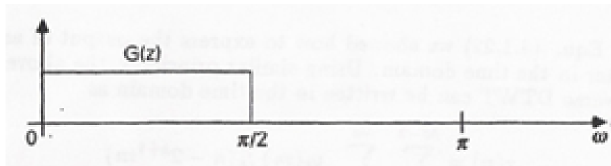


figure 16: Magnitude – response of G(z)

Using QMF – banks or its equivalent

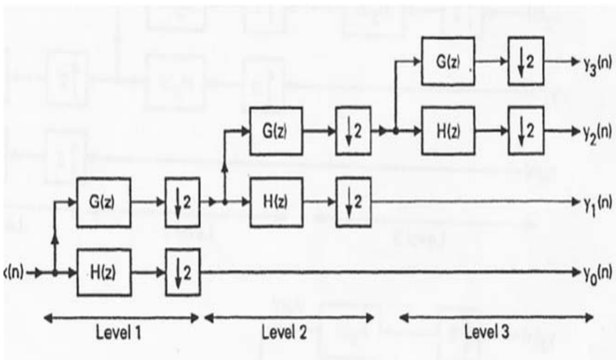


figure 17: 3 level binary tree-structured QMF

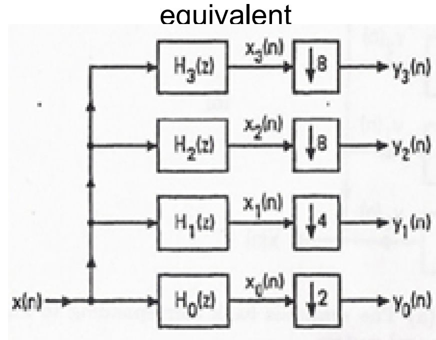


figure 18: equivalent 4-channel system

Responses of the filters $H(z), G(z)H(z^2), G(z)G(z^2)H(z^4), \ldots\ldots$

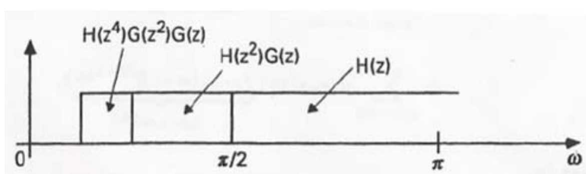


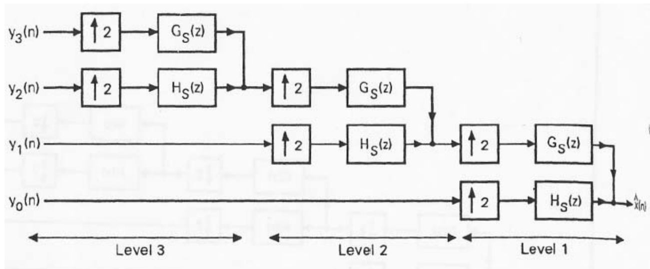
figure 19: combinations of H(z) and G(z)

Defining the Discrete –Time Wavelet Transform

$$y_k(n) = \sum_{m=-\infty}^{\infty} x(m) h_k\left(2^{k+1} n - m\right), \qquad 0 \le k \le M-2$$

$$y_{M-1}(n) = \sum_{m=-\infty}^{\infty} x(m) h_{M-1}\left(2^{M-1} n - m\right), \qquad \left(D_{iscrete} T_{ime} WT\right)$$

Inverse Transform



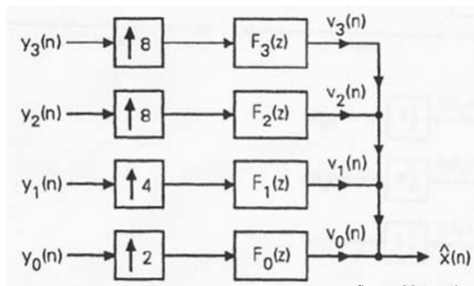


figure 20: synthesis filters

$$F_0(z) = H_s(z), \quad F_1(z) = H_s(z^2)G_s(z), \quad \ldots$$

For perfect reconstruction $\hat{x}(n) = x(n)$ we can express

$$X(z) = F_0(z)Y_0(z^2) + F_1(z)Y_1(z^4) + \ldots + F_{M-2}(z)Y_{M-2}\left(z^{2^{M-1}}\right) + F_{M-1}(z)Y_{M-1}\left(z^{2^{M-1}}\right)$$

and in time domain:

$$x(n) = \sum_{k=0}^{M-2} \sum_{m=-\infty}^{\infty} y_k(m) f_k\left(n - 2^{2k+1} m\right) + \sum_{m=-\infty}^{\infty} y_{M-1}(m) f_{M-1}\left(n - 2^{M-1} m\right)$$

Main References

Multirate Systems and Filter Banks

(Prentice Hall Signal Processing Series)

by *P. P. Vaidyanathan*