Subband Coding and Wavelets

National Chiao Tung University
Chun-Jen Tsai
12/04/2014

Concept of Subband Coding

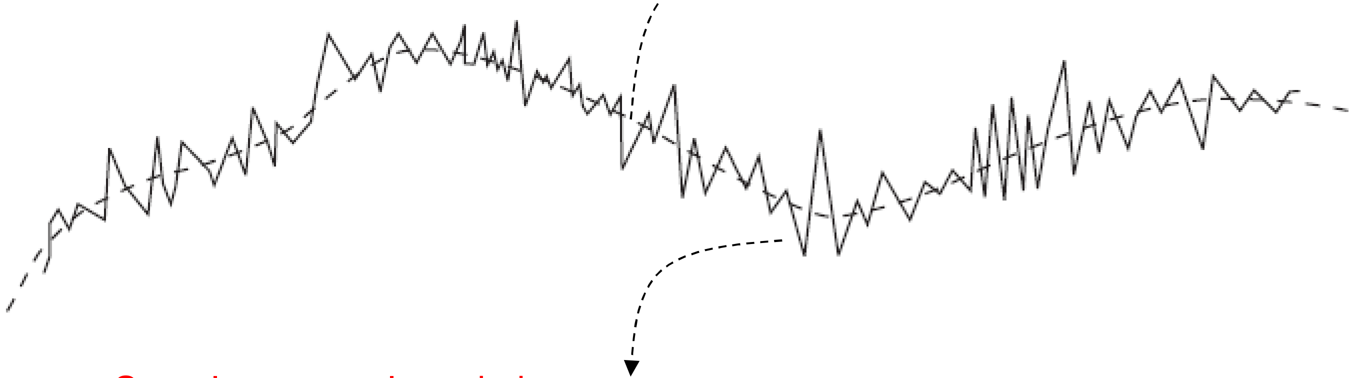
- In transform coding, we use N (or $N \times N$) samples as the data transform unit
  - Transform coefficients are de-correlated data each describing different characteristics of the original data
  - Different coefficients can be quantized differently
- Unfortunately, artificial selection of N causes blocking artifacts in the reconstructed image
  - Lapped Orthogonal Transform could reduce such artifacts
- Key question: can we design a transform that decomposes data without the artificial block size N?

Long-term vs. Short-term Behavior

- Given a sequence $\{x_n\}$, it can be decomposed into two types of behaviors
  - Long-term trend
  - Short-term, sample-to-sample variation

Long-term trend can be extracted by

$$y_n = \frac{x_n + x_{n-1}}{2}.$$

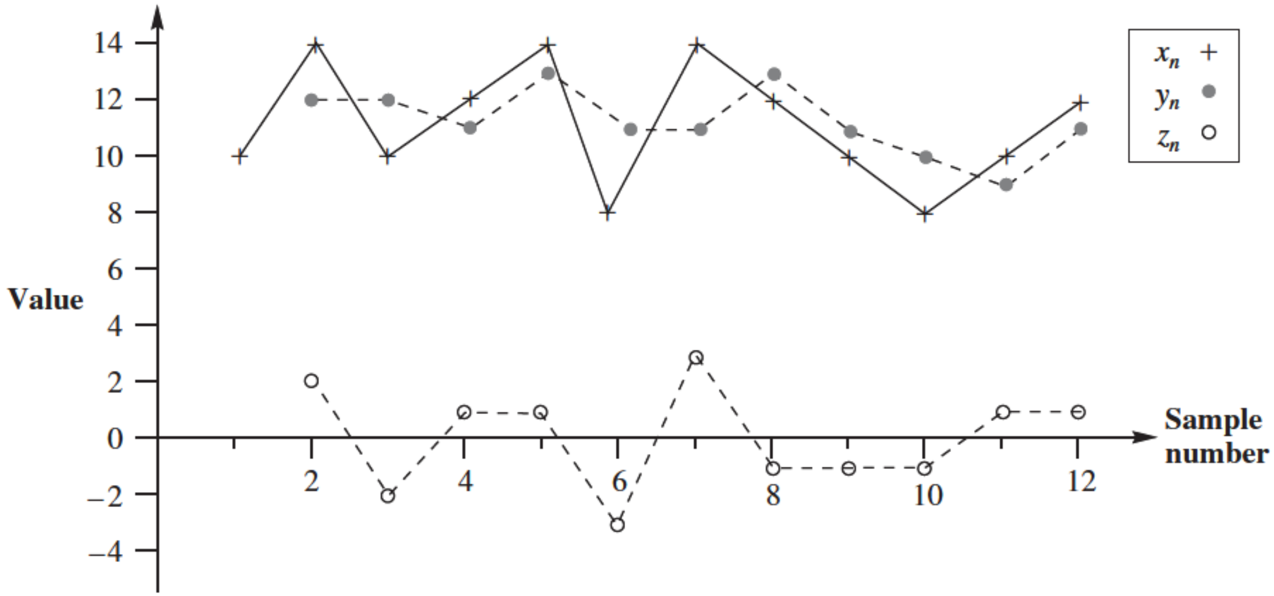


Sample-to-sample variation can be extracted by

$$z_n = x_n - y_n = \frac{x_n - x_{n-1}}{2}.$$

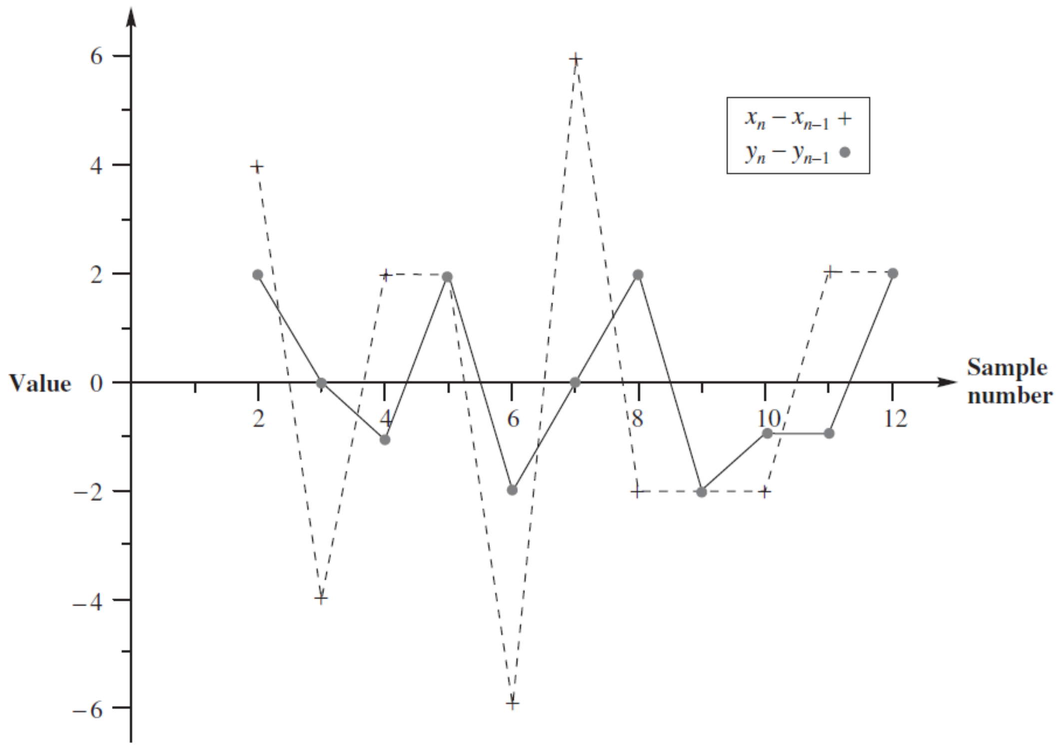
Example: Data Decomposition (1/5)

- Let $\{x_n\}$: 10, 14, 10, 12, 14, 8, 14, 12, 10, 8, 10, 12
  - $\{x_n - x_{n-1}\}$: 10, 4, –4, 2, 2, –6, 6, –2, –2, –2, 2, 2
- If we use an m-bit uniform quantizer to code $\{x_n - x_{n-1}\}$:
  - $M = 2^m$, $\Delta = 12/M$, maximum error = $\Delta/2 = 6/M$
- If we compute y_n and z_n, we have



Example: Data Decomposition (2/5)

- $\{y_n\}$: 10, 12, 12, 11, 13, 11, 11, 13, 11, 10, 9, 11
  - $\{y_n - y_{n-1}\}$: 10, 2, 0, –1, 2, –2, 0, 2, –2, –1, –1, 2
  - If we use an m-bit uniform quantizer to code $\{y_n - y_{n-1}\}$:
    - $M = 2^m$, $\Delta = 4/M$, maximum error = $2/M$



Example: Data Decomposition (3/5)

- $\{z_n\}$: 0, 2, –2, 1, 1, –3, 3, –1, –1, –1, 1, 1
  - Differential coding of $\{z_n\}$ does not make sense
  - If we use an m-bit uniform quantizer to code $\{z_n\}$:
    - $\Delta = 6/M$, maximum QE = $3/M$
- If FLC is used to code the quantizer outputs, the rate of $\{y_n\}$ plus $\{z_n\}$ is twice of the rate of $\{x_n\}$
  - On the other hand, the maximum quantization error is smaller ($5/M$ vs. $6/M$)
  - We can reduce the rate by only sending every other values of y_n and z_n.

Example: Data Decomposition (4/5)

- We can divide $\{y_n\}$ into subsequences $\{y_1, y_3, \ldots\}$ and $\{y_2, y_4, \ldots\}$; and $\{z_n\}$ into subsequences $\{z_1, z_3, \ldots\}$ and $\{z_2, z_4, \ldots\}$. Note that

$$y_{2n} = \frac{x_{2n} + x_{2n-1}}{2}, \quad z_{2n} = \frac{x_{2n} - x_{2n-1}}{2}.$$

- The sequence $\{x_n\}$ can be reconstructed by using either the even sequences, $\{y_{2n}\}$ and $\{z_{2n}\}$, or the odd sequences, $\{y_{2n-1}\}$ and $\{z_{2n-1}\}$ as follows:

$$y_{2n} + z_{2n} = x_{2n}, \quad y_{2n} - z_{2n} = x_{2n-1}.$$

Example: Data Decomposition (5/5)

- By transmitting only the even or odd sequences of $\{y_n\}$ and $\{z_n\}$, the rate is the same as that of $\{x_n\}$. But do we still have smaller quantization error?
  - Quantization error is only affected by the dynamic range of the sequences
  - The dynamic range of a subsequence will be smaller or equal to the original sequence

Remarks on Data Decomposition

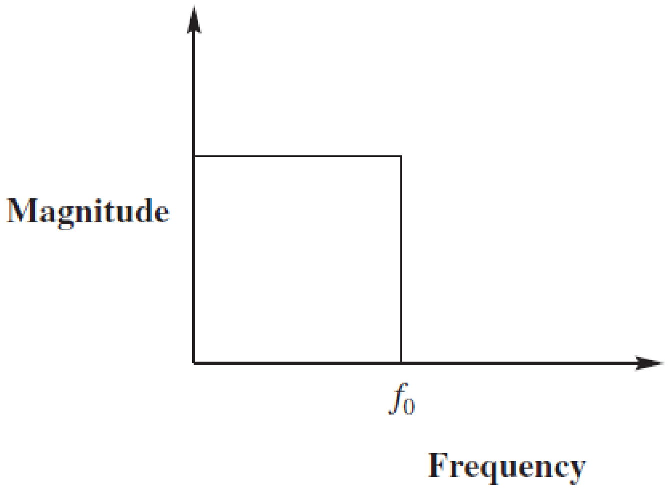
- Decomposing the $\{x_n\}$ sequence into subsequences may not result in any increase in rate
- The two subsequences had distinctly different characteristics, which should be encoded using different techniques
  - If we had not split the $\{x_n\}$ sequence, same approach will be used to compress both subsequences
- It is possible to further decomposing subsequences into subsequences, and so on
  - Data decomposition is also known as bank-filtering

Filters

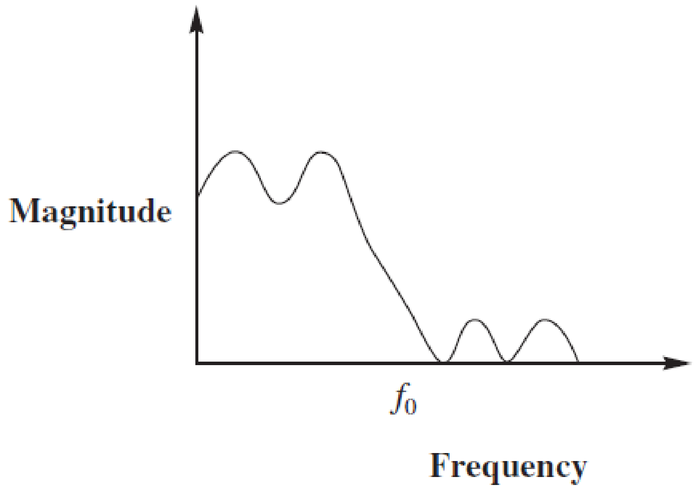
- Filters are popular tool for data decomposition. A filter isolates certain frequency components from others.
- Low-pass filters:
  - Filters that only let through components below a certain frequency f_0. f_0 is called the cut-off frequency
- High-pass filters:
  - Filters that block all frequency components below a certain value f_0 are called high-pass filters
- Band-pass filters:
  - Filters that let through components that have frequency content above some frequency f_1 but below frequency f_2

Example: Low-Pass Filter

- A filter is defined by their magnitude transfer function:
  - The ratio of the magnitude of the output and input of the filter as a function of frequency



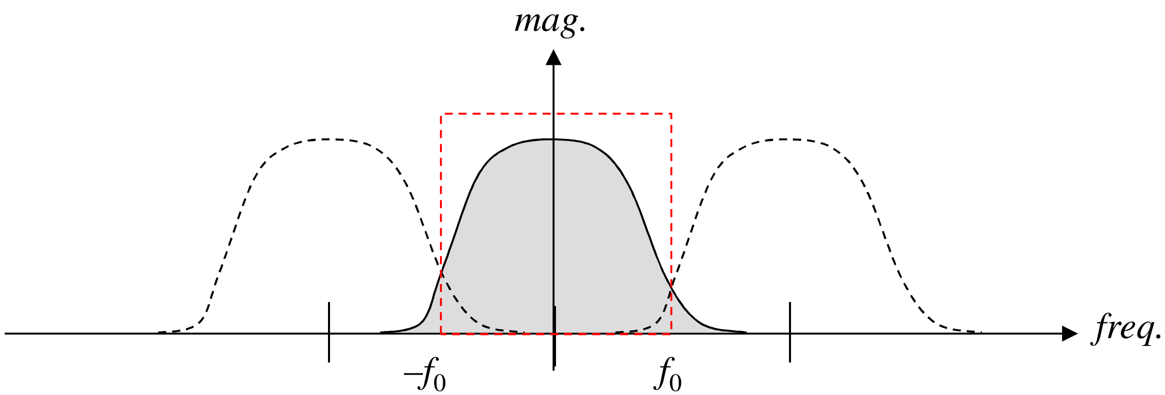
Ideal low-pass filter



Realistic low-pass filter

Sampling and Aliasing

- Nyquist theorem for digitization of analog signals:
  - If the highest frequency component of a continuous-time signal is f_0, then we need to sample the signal at more than $2f_0$ times per second in order to fully reconstruct the original signal from the samples
- Aliasing effect: if the sampling rate is less than $2f_0$, frequency component higher than f_0 will not be distinguishable from low frequency components:



Digital Filters

- Digital filtering involves taking a weighted sum of current and past inputs to the filter and, in some cases, the past outputs of the filter
- The general form of the input-output relationships of the filter is given by

$$y_n = \sum_{i=0}^{N} a_i x_{n-i} + \sum_{i=1}^{M} b_i y_{n-i},$$

where the sequence $\{x_n\}$ is the input to the filter, the sequence $\{y_n\}$ is the output from the filter, and the values $\{a_i\}$ and $\{b_i\}$ are called the filter coefficients

Example: Filter for Decomposition

- In previous example, $\{x_n\}$ is decomposed into $\{y_n\}$ and $\{z_n\}$. It is easy to show that the filter we use for extracting $\{y_n\}$ and $\{z_n\}$ out of $\{x_n\}$ are as follows:
  - Filter for $\{y_n\}$

$$h_n = \begin{cases} 0.5 & n = 0 \\ 0.5 & n = 1 \\ 0 & \text{otherwise} \end{cases}.$$

  - Filter for $\{z_n\}$

$$h_n = \begin{cases} 0.5 & n = 0 \\ -0.5 & n = 1 \\ 0 & \text{otherwise} \end{cases}.$$

Filter Terminology

- If the input sequence is a single 1 followed by all 0s, the output sequence is called the impulse response of the filter
- The number N is often called the number of *taps* in the filter
- If the b_i are all 0, then the impulse response will die out after N samples. These filters are called finite impulse response (FIR) filters
- If any of the b_i have nonzero values, the impulse response can, in theory, continue forever. Filters with nonzero values for some of the b_i are called infinite impulse response (IIR) filters.

Example: A Simple Two-Tap Filter

- If $a_0 = 1.25$, $a_1 = 0.5$, all other a_i and b_i are zeros; x_n is an impulse function:

$$x_n = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases},$$

then the output y_n is

$$y_0 = a_0 x_0 + a_1 x_{-1} = 1.25$$

$$y_1 = a_0 x_1 + a_1 x_0 = 0.5$$

$$y_n = 0, \ \ n < 0 \text{ or } n > 1,$$

$\{y_n\}$ is the impulse response (often denoted by $\{h_n\}$)

Example: An IIR Filter

- If $a_0 = 1$, $b_1 = 0.9$, all other a_i and b_i are zeros; x_n is the impulse function, then the output y_n is

$$y_0 = a_0 x_0 + b_1 y_{-1} = 1 \cdot 1 + 0.9 \cdot 0 = 1$$

$$y_1 = a_0 x_1 + b_1 y_0 = 1 \cdot 0 + 0.9 \cdot 1 = 0.9$$

$$y_2 = a_0 x_2 + b_1 y_1 = 1 \cdot 0 + 0.9 \cdot 0.9 = 0.9^2$$

$$\cdots$$

$$y_n = (0.9)^n.$$

Thus, the impulse response $\{h_n\}$ is

$$h_n = \begin{cases} 0, & n < 0 \\ (0.9)^n, & n \geq 0 \end{cases}.$$

Convolution Operation

- If $\{x_n\}$ and $\{y_n\}$ are the input and output, respectively, of a filter with impulse response $\{h_n\}_{n=0\ldots M}$, then $\{y_n\}$ can be obtained by the convolution of $\{x_n\}$ and $\{h_n\}$:

$$y_n = \sum_{k=0}^{M} h_k x_{n-k},$$

where M is finite for an FIR filter and infinite for an IIR filter.

Stability of a Filter

- A filter is stable if any bounded inputs will produce bounded outputs
- Because FIR filters are simply weighted averages, they are always stable
- For IIR filters, it is possible to have unbounded output even when the input is bounded
  - Although IIR filters can become unstable, they can also provide better performance, in terms of sharper cutoffs and less ripple in the passband and stopband for a fewer number of coefficients

Example: Unstable IIR Filter

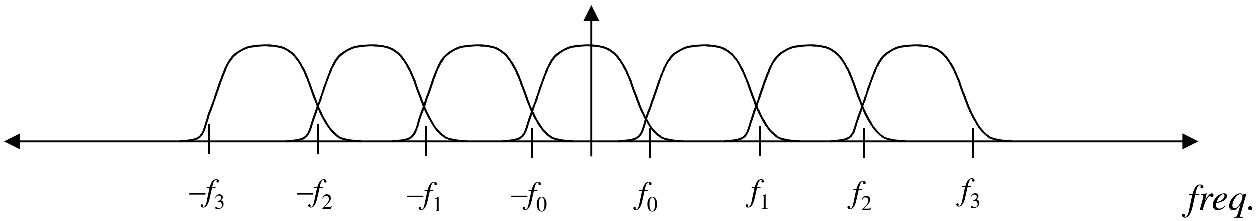
- Consider a filter with $a_0 = 1$ and $b_1 = 2$. Suppose the input is a single 1 followed by 0's, the output is

$$
\begin{aligned}
y_0 &= a_0 x_0 + b_1 y_{-1} = 1\cdot 1 + 2\cdot 0 = 1 \\
y_1 &= a_0 x_1 + b_1 y_0 = 1\cdot 0 + 2\cdot 1 = 2 \\
y_2 &= a_0 x_2 + b_1 y_1 = 1\cdot 0 + 2\cdot 2 = 2^2 \\
&\cdots \\
y_n &= 2^n.
\end{aligned}
$$

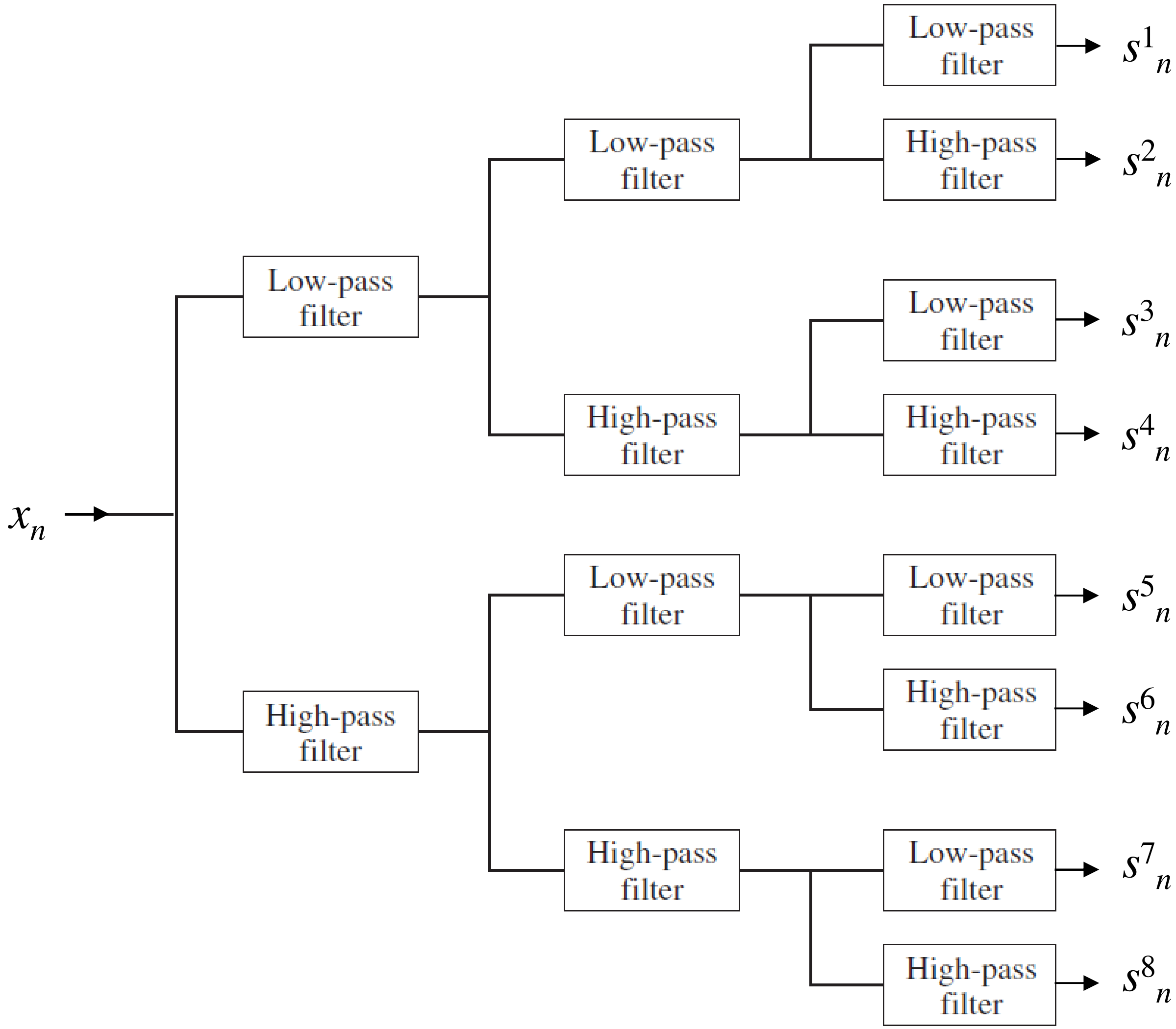
Even though the input contains a single 1, the output at time $n = 30$ is 2^{30}!

Filter Banks

- In multimedia compression, we often have to decompose input data into multiple subsequences (i.e. frequency bands)
  - We need more than two filters to do the JOB
  - An array of filters is often called a filter bank
- In practice, we can cascade multiple use of a pair of low-pass and high-pass filters to decompose data into multiple frequency bands



Example: An 8-Band Filter Bank



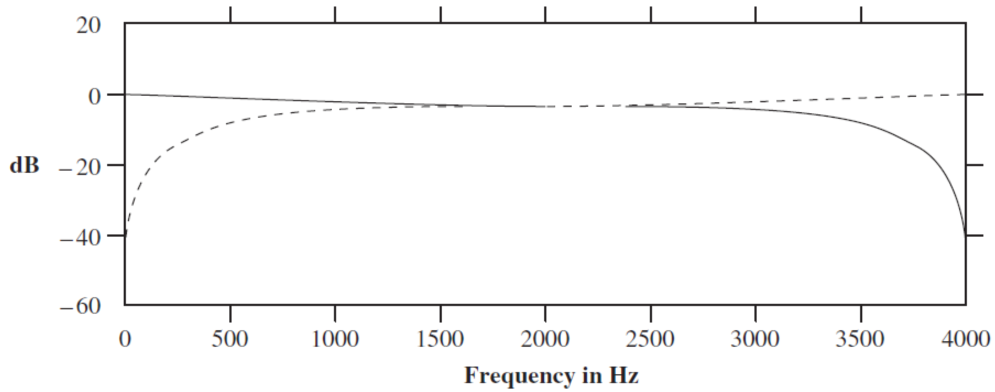
Design of a Pair of Low-High Filters

- For cascading implementation of filter banks, we need a pair of low-pass and high-pass filters
- The most popular filter pairs are the quadrature mirror filters (QMF), which were first proposed by Crosier, Esteban, and Galand in 1976
  - These filters have the property that if the impulse response of the low-pass filter is given by $\{h_n\}$, then the high-pass impulse response is given by $\{(-1)^n h_{N-1-n}\}$, that is,

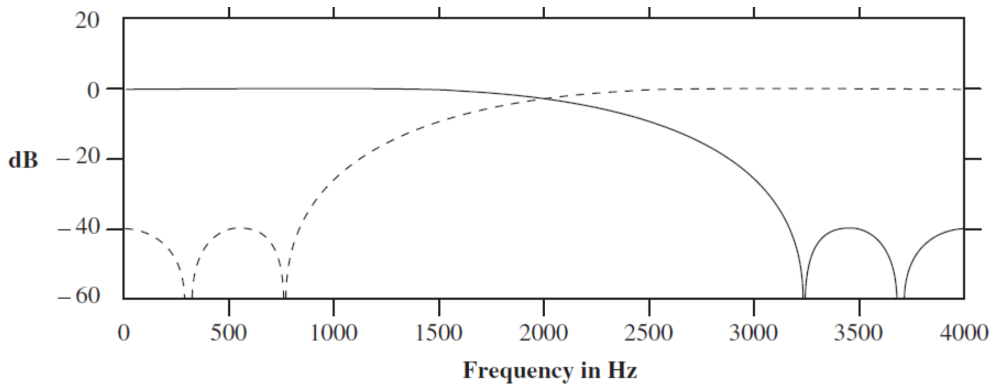
$$h_{N-1-n} = h_n, \qquad n = 0,1,\ldots,\tfrac{N}{2}-1.$$

Characteristics of Filters

- Johnston 8-tap QMF filter

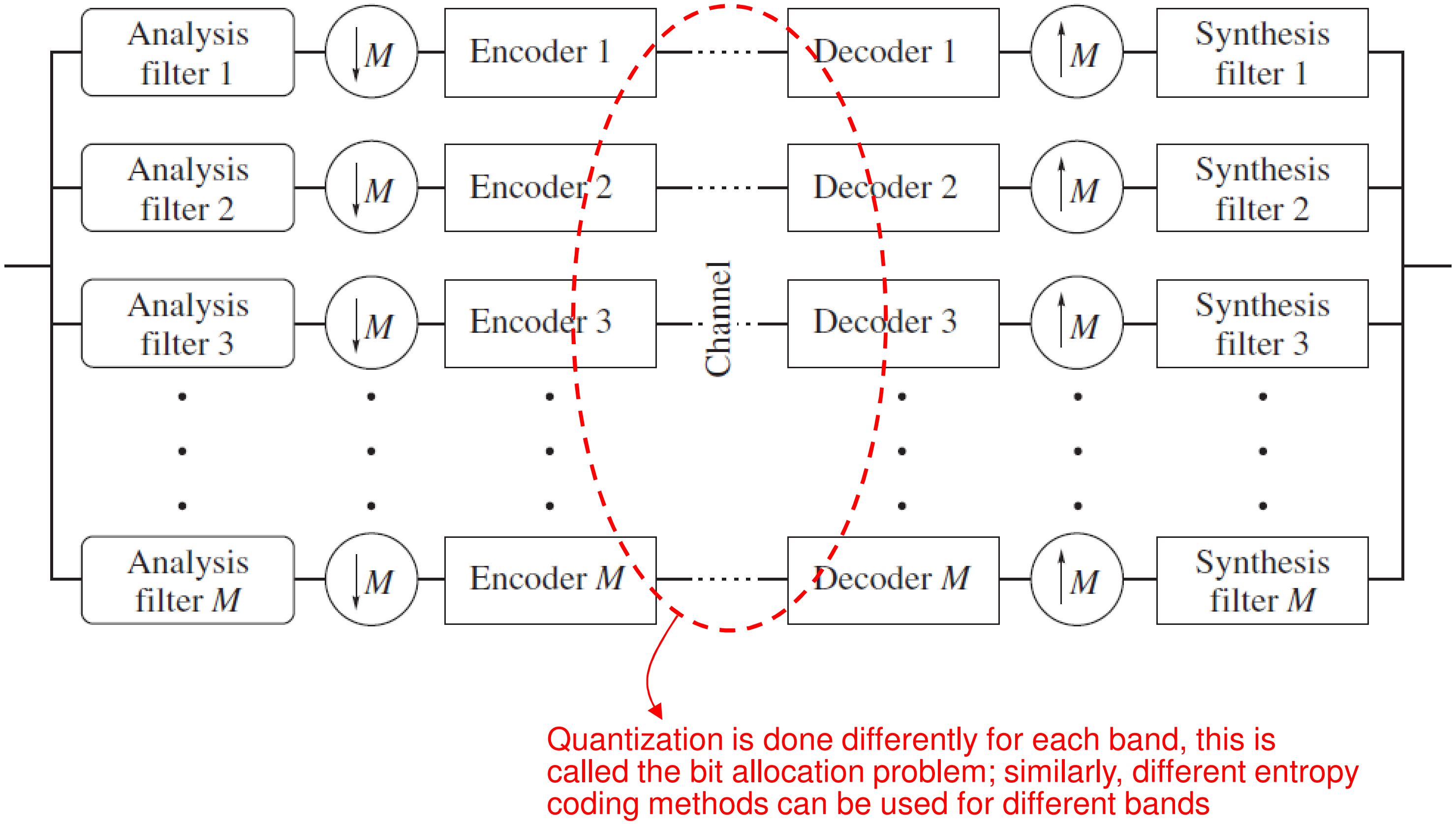


- Smith-Barnwell 8-tap QMF filter



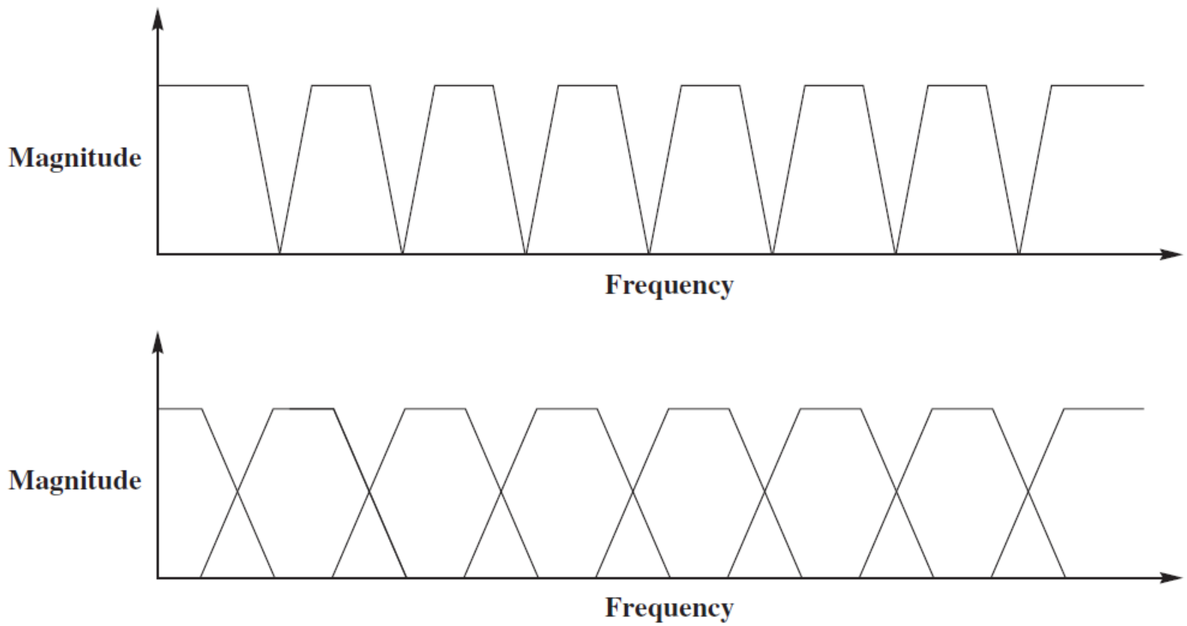
Basic Subband Coding Algorithm

- Block diagram of the subband coding system:



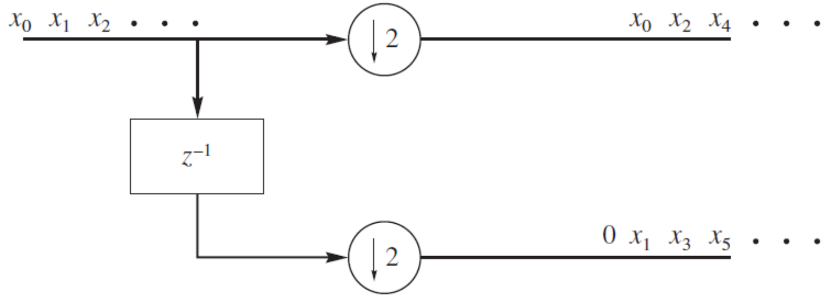
Analysis Filters

- The source output is passed through a bank of filters, called the analysis filter bank. The filters can be non-overlapping or overlapping.

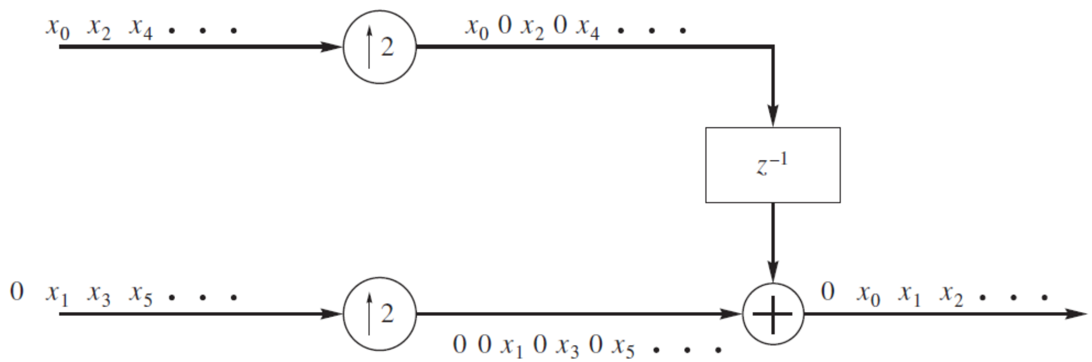


Downsampling and Upsampling

- Since each filter output has smaller bandwidth, we can discard some samples without losing information; this is called down-sampling or decimation

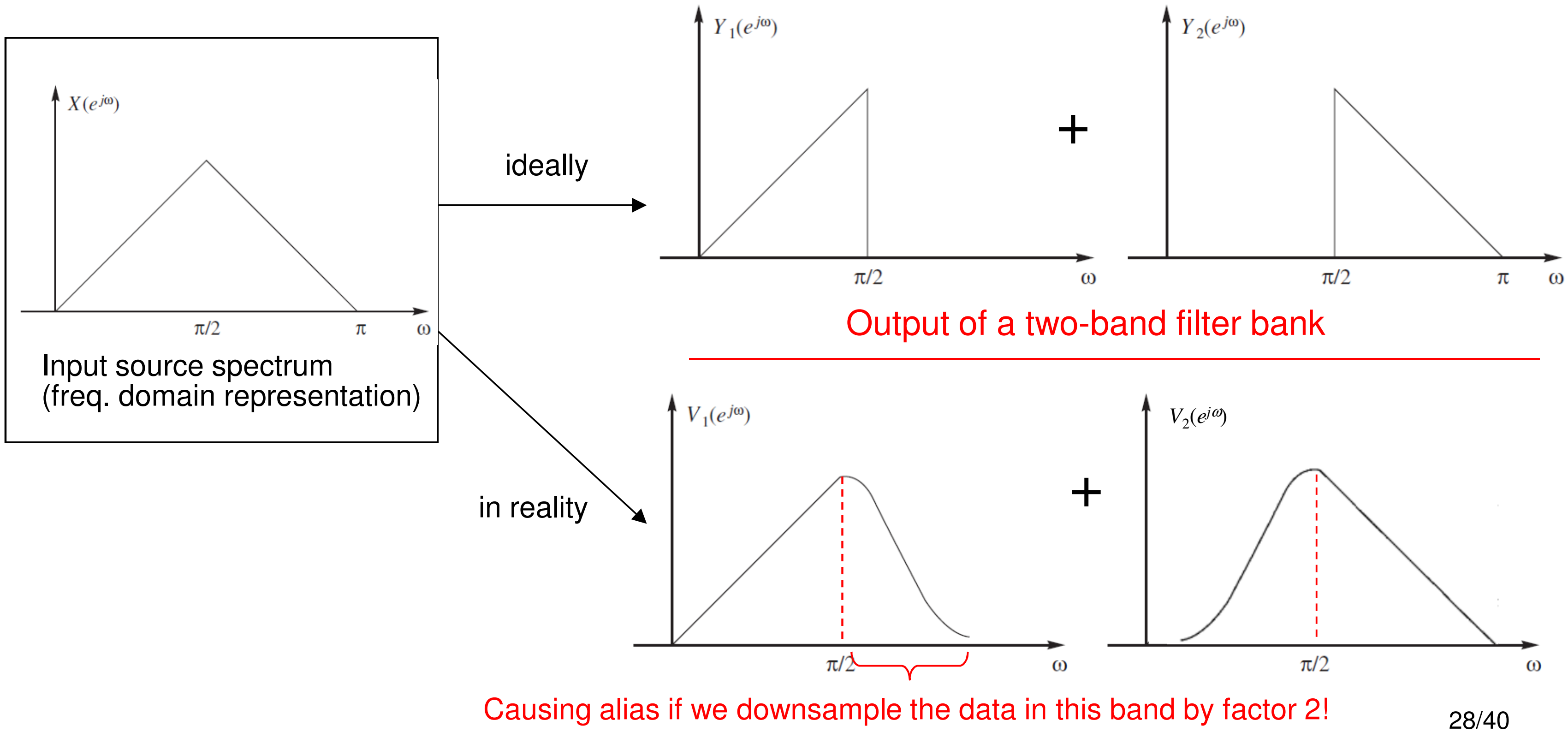


- For reconstruction, we perform upsampling



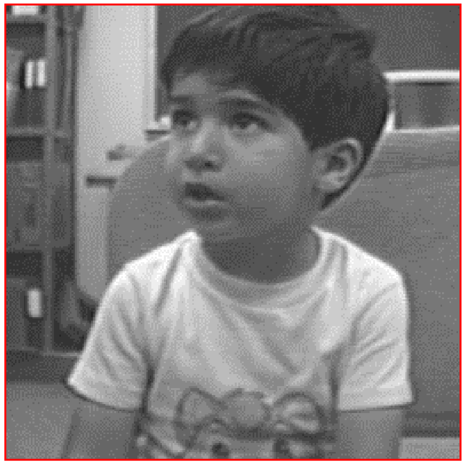
Non-Perfect Reconstruction

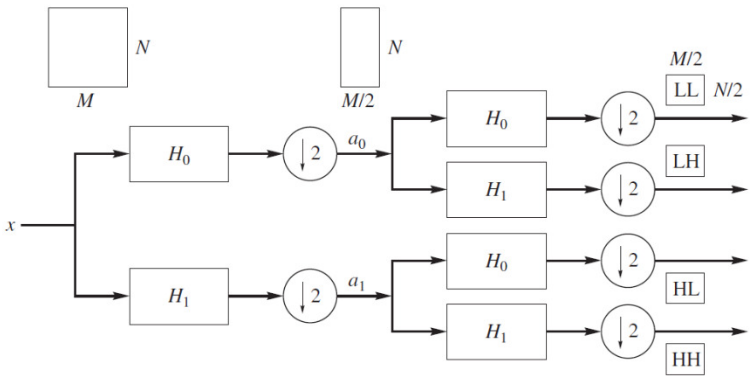
- Although theoretically there is no data loss for subband decomposition, in reality, this is not trivial



Applications of Subband Coding

- Most audio codecs today uses subband coding
  - Human ears can be modeled by a filter bank of 25 overlapping bands
- Some researchers try to apply subband coding on images and videos, but not very successful
  - Key issue: cascaded 2-D decomposition using separable 1-D filters is not very meaningful





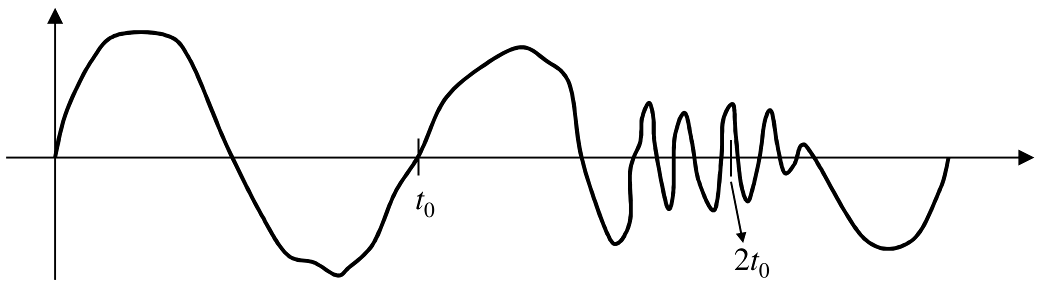
Problems with Time-Freq. Analysis

- Given a non-stationary signal, it is difficult to tell which frequency component happened at what time
- We can use short-term Fourier transform (STFT):

$$F(\omega,\tau)=\int_{-\infty}^{\infty} f(t)\,g^{*}(t-\tau)\,e^{j\omega t}\,dt.$$

Location of the frequency analysis

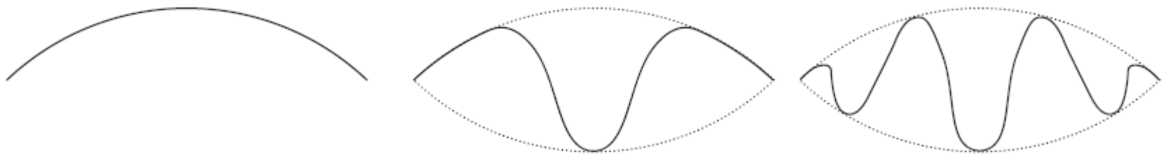
Window function that selects the signal segment for frequency analysis while reducing boundary effects



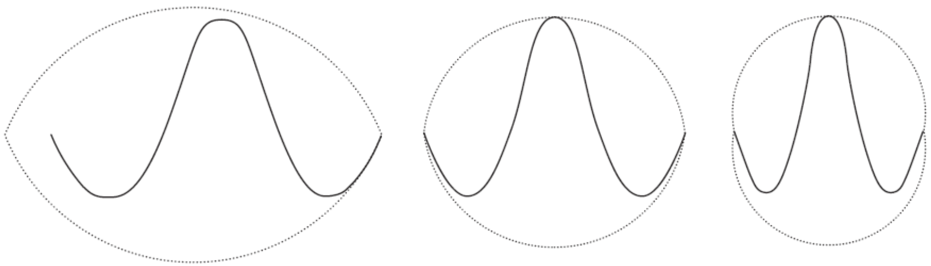
Key question: how do we pick the window size of $g(t)$?

Concept of Wavelets

- Once the window function $g(t)$ is fixed, the basis functions, $g(t)e^{jm\omega_0 t}$, $m = 0, 1, \ldots,$ of STFT are:



- A different ideas is to adapt the window size to the basis function such that it contains one cycle only:



Such basis functions are called wavelets

Wavelet Basis Functions

- We can start with a single basis function, called mother wavelet, then translate and scale this function to create other basis functions
  - Scaling: $f(t/a)$ is the scaling of $f(t)$ by a constant a
  - Translation: $f(t-b)$ is the translation of $f(t)$ by a constant b

- Note that scaling will change the norm of a function:

$$\left\| f\left(\tfrac{t}{a}\right) \right\|^2 = \int_{-\infty}^{\infty} f^2(\tfrac{t}{a})dt = a\int_{-\infty}^{\infty} f^2(x)dx = a\left\| f(t) \right\|^2.$$

To have the same norm, we must multiply $f(t/a)$ by $1/\sqrt{a}$.

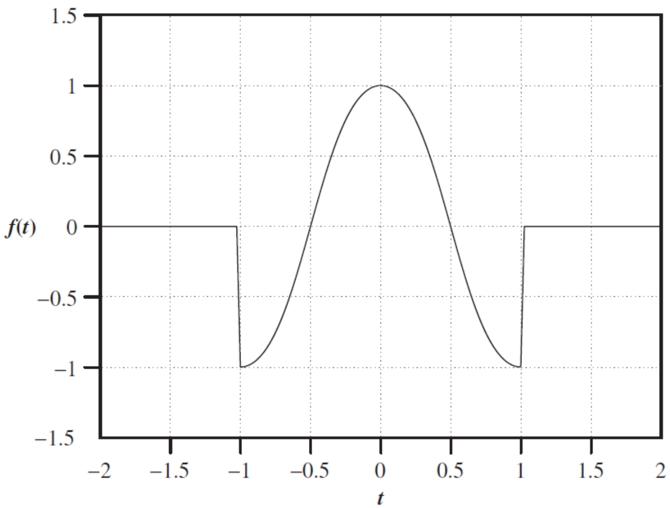
Example: Scaling

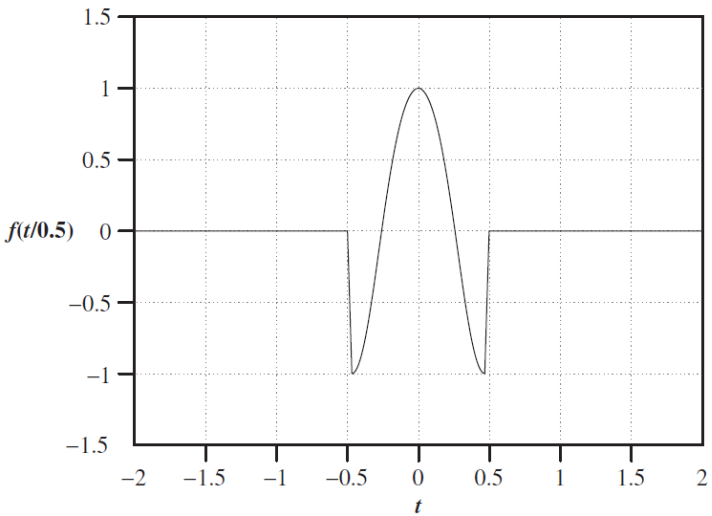
- Given the function

$$f(t)=\begin{cases}\cos(\pi t) & -1\le t\le 1\\ 0 & otherwise\end{cases}.$$

scale it by 0.5:

$$f(\frac{t}{0.5})=\begin{cases}\cos(2\pi t) & -\frac{1}{2}\le t\le \frac{1}{2}\\ 0 & otherwise\end{cases}.$$

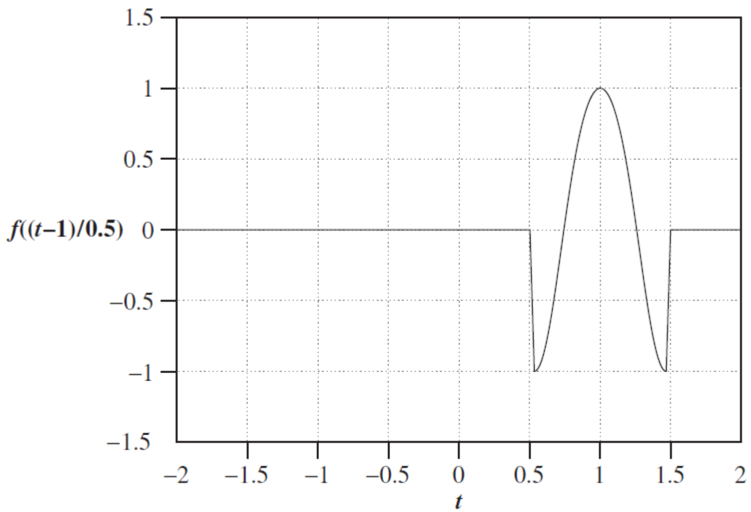




Example: Translation

- Translate the function $f(t/0.5)$ to the right

$$f\left(\frac{t-1}{0.5}\right)=\begin{cases}\cos(2\pi(t-1)) & \frac{1}{2}\le t\le\frac{3}{2}\\ 0 & otherwise\end{cases}.$$



Basis Function Generation

- Given a mother wavelet $\psi(t)$, the remaining functions are obtained as

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right).$$

  Note that the frequency domain representation of $\psi(t)$ is $\Psi_{a,b}(\omega) = \mathscr{F}[\psi_{a,b}(t)]$, where $\mathscr{F}[\cdot]$ is the Fourier transform.

- If a and b are continuous, then $\psi_{a,b}(t)$ are the basis functions for continuous wavelet transform (CWT).

Wavelet Transforms

- Forward transform:
  - the coefficient w.r.t. each wavelet basis function are:

$$w_{a,b} = \left\langle \psi_{a,b}(t), f(t) \right\rangle = \int_{-\infty}^{\infty} \psi_{a,b}(t) f(t) dt.$$

- Inverse transform
  - $f(t)$ can be recovered from $w_{a,b}$ by

$$f(t) = \frac{1}{C_\Psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w_{a,b} \psi_{a,b}(t) \frac{da\,db}{a^2},$$

where

$$C_\Psi = \int_0^{\infty} \frac{\left|\Psi(\omega)\right|^2}{\omega} d\omega.$$

Selection of Mother Wavelet

- Note that C_Ψ must be finite for the inverse transform to exist $\rightarrow \Psi(0) = 0$
  - The mother wavelet $\psi(t)$ must have zero mean
- The wavelets bases should have finite energy as well, i.e.,

$$\int_{-\infty}^{\infty} \left|\Psi(\omega)\right|^2 d\omega < \infty.$$

  - Since $\Psi(\omega)$ only has energy distributed in some frequency band, it is effectively a band-pass filter

Selection of a and b

- If the scaling and translating parameters a and b are discrete values, they must be set properly so that no input data are missed in the action
  - Small a and large b causes "gaps" in the analysis process
- A popular approach is to select a and b according to

$$a = a_0^{-m},\ b = nb_0 a_0^{-m},$$

where m and n are integers, $a_0 = 2$, $b_0 = 1$. That is,

$$\Psi_{m,n}(t) = a_0^{m/2}\Psi(a_0^m t - nb_0),\quad m, n \in Z.$$

Example: Harr Wavelet

- The Harr wavelet is given by

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

  By translating and scaling $\psi(t)$, we can synthesize a variety of functions.

- Harr wavelet is effectively a simple high-pass filter that analyze the signal at various resolutions and locations

Subband Coding vs. Wavelet Trans.

- In subband coding, a signal is decomposed into different frequency subbands for analysis
  - The dimensions of time- and freq.-domain are the same
- In wavelet analysis, a signal can be decomposed into a (possibly) higher dimensional space for analysis
  - Each subspace can represent different characteristics of the original signals (beyond frequency)