EMG and Wavelet Analysis – Part I

Introduction

EMG stands for electromyography; that is, the study of electric signals recorded from muscles (myoelectric signals). Motorneurons in the brain send signals to the motor units (MU) of the muscles, which respond with a motor unit action potential (MUAP). Muscular activity is thus associated with a sum (CMUAP, compound MUAP) of MUAPs (amplitude of the order of 120 mV), which give rise to the signal recorded as the EMG signal, which may be measured using surface electrodes (surface EMG, short SEMG) or needle electrodes. The EMG-signal is of the order from a few µV to a few mV. Differential amplifiers are therefore necessary in order to extract the signal from a noisy environment (bipolar electrodes are typically used). Typical sampling rates used are in the range of 2000 to 600 S/s. With needle electrodes signals up to several kHz can be recorded, but with surface electrodes skin and fatty tissue filter frequencies above 500 Hz.

[Graph showing force and rectified EMG signal]

Isometric contraction: Red curve shows force, while the blue curve shows rectified EMG-signal from the muscle sampled at 1000 S/s.

A number of mathematical methods have been developed for analyzing the EMG-signals of which most are based on the standard Fourier Transform. By dividing a time series into subintervals and calculating the spectrum for every subinterval (giving the so-called spectrogram) one may follow how the spectrum changes with time (Short Time Fourier Transform, STFT).

1 E-mail: borgbros@netti.fi
Isometric test: Spectrum of the EMG-signal.

For instance, it has been found that for a contracting muscle the fatigue is accompanied by a trend of decreasing mean frequency of the EMG-signal, and by a trend of increasing “energy” of the signal (the energy is defined here as the integral of the square of the signal amplitude). Another characteristic property of the EMG-signal is that its mean amplitude correlates positively with the force produced by the contracting muscle; an increasing force yields higher mean EMG-amplitude.

Isometric contraction: Red curve shows a decreasing trend of mean frequency of the EMG-signal. The blue curve is the “energy” which exhibits an increasing trend with time during contraction. Calculations based on the spectrogram for the EMG-signal.
STFT is a very useful mathematical method and has also been generalized in a number ways. Fourier transformation methods are based on the sinus-cosinus functions and are therefore especially well adapted for analyzing periodic signals (the Hartley transform is a variation of the Fourier transform which uses the sum of sinus and cosinus functions as the base functions). The popularity of the Fourier transform is also due to its important use for solving linear differential equations. However, there are many sorts of signals in nature that are non-stationary, non-periodic, “fractal” or seemingly chaotic. It has been realized in recent decades that many processes in nature are “non-linear” (at least when driven off their near equilibrium regions).

Non-linear mathematical models may exhibit very irregular behavior. It is therefore reasonable to suspect that other base functions than the sinus-cosinus functions might serve well for analyzing such “irregular” signals. Thus, instead of decomposing and reconstructing a signal in terms of the sinus-cosinus functions, one might e.g. use saw tooth functions, rectangle waves (Walsh functions), or finite time pulses. Since EMG-signals are typically non-stationary and irregular it could be that such new methods of analysis have a chance of supplementing the traditional methods. The best known of these new tools is the wavelet (fr. ondelette) analysis, which started to develop (by Morlet, Grossmann, Mallat, Daubechies and Meyer) in earnest around the mid-eighties as a method for analyzing seismic waves. During the last decade wavelets have emerged as a major tool in signal processing. (A famous case is FBI using wavelets for compressing fingerprint data).

There are a number of standard wavelet transforms associated with the names of Haar, Daubechies, Meyer, Morlet, Coifman (“Coiflets”), Battle-Lemarié, Chui, etc. The so-called continuous wavelet transforms (CWT) are employed for studying the scalogram of the signal, which generalizes the standard spectrogram. The wavelet has been likened to a “mathematical microscope” that allows us to study the signal at various scales. Scalograms can also be compared to musical transcription, which is based on dividing the frequency range into octaves. By special choices of wavelets one arrives at the so-called multi resolution analysis (MRA), which is linked to the 2-channel filter bank methods. These special wavelets (e.g. Haar and Daubechies) allow signals to be expressed as a sum of a discrete set of scaled versions of the wavelet, thus generalizing the classical Fourier series expansion. Finding a well-adapted wavelet basis makes it possible to express the signal with a high degree of approximation using a small subset of the expansion (wavelet-) coefficients (data compression).

The CWT has found a number of applications in the field of biomedical signals, such as the study of transients and spikes in ECG-signals (cardiac patterns) and EEG-signals (seizure detection). There is also evidence of interesting similarities between the wavelets and biological information processing (cell responses). Thus one may look for optimal wavelets for detecting certain signal patterns (matched filter). One successful application is the detection of so-called QRS complex in the ECG-signals, which can discriminate between normal and abnormal patterns. The Gabor-Morlet wavelet has been used in analyzing arterial sounds. For extracting impulse-like events the so-called B-spline wavelets are quite popular because these wavelets match the potentials quite well. The question is whether there could be similar useful applications to the EMG-signals: Could wavelets provide methods to extract more information from the EMG data? Better timing of transients is an obvious target, but could there also be other useful forms of feature extraction? We may suggest a few lines along which to pursue investigations in this area.

A typical problem is to distinguish between “normal” and “abnormal” states. Thus, one could investigate the possibility of distinguishing signals from healthy and injured muscles (using e.g. filter methods in combination with learning networks); i.e., a diagnostic tool. One problem to address could be to refine the signal processing with the view of getting an improved estimate of the number of motor units activated. A special issue is also the use of matched filter in a noisy environment typical for EMG-signals. Here one may take advantage of works done in related areas (the use of so-called whitening filters). It is also important to advance from the other direction; that is, refining mathematical models for the generation of EMG on which one can devise optimal matching wavelets. - In the next sections some of the mathematical background of wavelets is presented.

**Continuous wavelets**

For a signal $x(t)$ we define its (complex) Fourier transformation as
\( \hat{x}(f) = \int_{-\infty}^{\infty} e^{-i\omega t} x(t) \, dt \)

which is a function of the frequency. Conversely, we get back the signal from its frequency representation through the inverse Fourier transform

\( x(t) = \int_{-\infty}^{\infty} e^{i\omega t} \hat{x}(f) \, df \)

The integral (1) is a measure how well the analyzing function \( e^{i2\pi ft} \) matches the signal. The continuous wavelet transform may use any well-behaved analyzing function (analyzing wavelet) \( g(u) \), which for “admissible wavelets”, however, must satisfy

\[
\int_{-\infty}^{\infty} \left| \hat{g}(\pm f) \right|^2 \, df < \infty, \quad \text{which implies that} \quad \hat{g}(0) = \int_{-\infty}^{\infty} g(u) \, du = 0
\]

The wavelet transform \( W(a, \tau) \) of a the signal \( x(t) \) is defined as (CWT, continuous wavelet transform)

\[
W(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) g^{*} \left( \frac{t-\tau}{a} \right) \, dt
\]

\( g^{*} \) notation the \( g \) introduces by

Here \( a \) is called the scale (or dilatation) factor. If the function \( g \) is concentrated around zero then it follows that for small \( a \) the integral (4) emphasizes the detail of the signal around \( t = \tau \), whereas for large \( a \) it gives a sort of an average value of the signal around \( t = \tau \). It is in this sense that the wavelet transform acts as “mathematical microscope”.

If (3) is satisfied we can reconstruct the signal from the wavelet transform \( W(a, \tau) \) by the formula (applies when the Fourier transform of the wavelet is an even function):

\[
x(t) = \frac{1}{c} \int \left| W(a, \tau) \right|^2 \frac{d\omega}{\omega}, \quad \text{where} \quad c = \int_{0}^{\infty} \left| \hat{g}(f) \right|^2 \, df
\]

The squared modulus \( |W(a, \tau)|^2 \) of the wavelet transform represents the so-called scalogram of the signal. The energy of the signal is defined as

\[
E = \int |x(t)|^2 \, dt = \int |\hat{x}(f)|^2 \, df
\]
where \( |\hat{x}(f)|^2 \) represents the spectrum of the signal. The spectrum and the scalogram can be related by the equation

\[
\int_{a>0} \left| W(a, \tau) \right|^2 \frac{da d\tau}{a^2} = \int |\hat{x}(f)|^2 \frac{d^2}{u} \int |\hat{g}(u)|^2 du
\]

For numerical calculations of \( W(a, \tau) \) one can adopt the Fast Fourier Transform (FFT) to the equation

\[
W(a, \tau) = \sqrt{a} \int e^{2\pi i au} \hat{x}(u) \hat{g}^*(au) du
\]

As an example we may use the “Mexican wavelet” (popular e.g. among the seismologists)

\[
g(x) = (1 - x^2) e^{-x^2/2}
\]

Its Fourier transformation is

\[
\hat{g}(f) = \sqrt{2\pi} (2\pi f)^{1/2} e^{-(2\pi f)^2/2}
\]

for which we have from (5) \( c = \pi \), whence it indeed satisfies the condition (3).

The above figure shows the scalogram – calculated using the “Mexican wavelet” - for the EMG-signal from the isometric test data used before. Time (in seconds) is along the x-axis, whereas the scaling parameter is along the y-axis (logarithmic scale: \( a = 2^y \), thus \( y = 10 \) corresponds to \( a = 1/1024 \) and \( y = 0 \) to \( a = 1 \); we get a finer scale when we move upwards in the figure). The value of \( W(a, \tau) \) is represented by gray scale color; darker means lower value. From the figure it is apparent that the beginning of the contraction is marked by a sharply localized maximum in time. This is a typical feature of scalograms. They are good at localizing transients. Indeed, suppose we have a Dirac pulse signal \( \delta(t-t_0) \), which is peaked at time \( t_0 \), then the transform (4) gives

\[
W(a, \tau) = \frac{1}{\sqrt{a}} g \left( \tau - t_0 \right) \left( \frac{\tau - t_0}{a} \right)
\]

which is sharply peaked at \( \tau = t_0 \) for small \( a \) if \( |g(x)| \) has its maximum at \( x = 0 \).
The next figure above shows the scalogram of a signal which is a sum of two Dirac pulses at \( t = 1 \) and \( t = 2 \); the location of the pulses is easily determined from the scalogram. Since the spectrogram on the other hand uses a fixed “window” (fixed size of the subintervals into which the time series is divided) it does not give a similar sharp location of the transients. For a comparison we give the spectrogram of the same data as contour map.

The same data mapped as contour map. The frequency (Hz) is along the y-axis. The region of contraction is here also visible, and pronounced activity seems to take place around and below 100 Hz.

The spectrogram and the scalogram can be compared through the so-called Wigner-Ville distribution defined for a signal \( x \) by

\[
WV_x(t, f) = \int x(t + \frac{u}{2})x^*(t - \frac{u}{2})e^{-i2\pi fu} du
\]

Apparently, for a harmonic signal of frequency \( f_0 \) the W-V distribution is peaked at this frequency, and for a signal peaked at a time \( t_0 \) the W-V distribution is also peaked at this time. The W-V distribution and the wavelet transform (4) are related by

\[
|W(a, \tau)|^2 = \int WV_x(t, f)WV_x^*(\frac{t - \tau}{a}, af)dt df
\]

For the “Mexican wavelet” we obtain
(14) \[ W(t, f) = \frac{\sqrt{\pi}}{2} e^{-i(2\pi f)^2 - i^2} \left( 3 - 12t^2 - 16\pi^2 f^2 + 4t^4 + 32\pi^2 t^2 f^2 + 64\pi^4 f^4 \right) \]

which attains maximum at \( t = f = 0 \), whence for a small \( a \) the r. h. s. of (13) approaches the W-V distribution. The W-V distribution is sometimes used to decide whether time-frequency or time-scale methods should be used.

The figure above shows the W-V distribution (magnitude) of the EMG-signal from the isometric test. (The W-V distribution was calculated for 32 time points \( t_k \).)

**Multi resolution analysis**

Equation (5) above shows how a signal can be expressed in terms of a wavelet as an integral. The question arises whether a general signal can be expressed as a sum of a discrete subset of the scaled “mother” wavelet, in analogy with the Fourier series. Usually one considers the “dyadic” subset

\[ \psi_{n,k}(t) = 2^{n/2} \psi \left( 2^n t - k \right) \]

\( n, k \in \mathbb{Z} \)

defined by a wavelet \( \psi \) (we adhere to some of the common notations used in the field of MRA); i.e. a “dyadic” scaling. The interesting case is when the functions \( \psi_{n,k} \) form an orthonormal and complete set such that any signal \( x(t) \) may be expressed as sum

\[ x(t) = \sum_{n,k} \beta_k^n \psi_{n,k}(t) = \sum_{n,k} \left\langle x, \psi_{n,k} \right\rangle \psi_{n,k}(t) \]

(16)

\[ \beta_k^n = \int x(t) \psi_{n,k}(t) dt \equiv \left\langle x, \psi_{n,k} \right\rangle \]

(As with the reconstruction formula (5) it also follows from (16) that \( \int x(t) dt = 0 \); thus, for a complete reconstruction covering also functions with non-vanishing integral one has to add basis functions using the so-called “father wavelet” – see below.) The general idea is then: for a given class of signals, find a wavelet such that the essential features of the signals can be captured with a minimum number of coefficients in the wavelet series expansion (16).

The simplest orthonormal wavelet is the Haar wavelet, which was studied already back in 1910 by A Haar as a novel way to expand continuous functions. The Haar wavelet can be defined by a scaling function \( \phi \) (“father wavelet”) which is simply the characteristic function for the interval \([0, 1]\).
The Haar wavelet is then defined by

\[
\psi (x) = \phi (2x) - \phi (2x - 1)
\]

Also the Haar scaling function satisfies a related scaling equation

\[
\phi (x) = \phi (2x) + \phi (2x - 1)
\]

The father wavelet coefficients of a function \(x(t)\) are defined by

\[
\alpha_k^n = \int x(t) 2^{n/2} \phi \left( 2^n t - k \right) dt = \int x(t) \phi_{n,k} (t) dt
\]

The relations, (18) and (19), are generalized later. For the Haar wavelet it is not difficult to see that the set (15) forms an orthonormal set. As pointed out above this is not a complete set but it can be completed by adding the functions \(\phi(x-k), k = \ldots, -2, -1, 0, 1, 2, \ldots\). Instead of (16) we then have

\[
x(t) = \sum_{n,k} \beta^R_n \psi_{n,k} (t) + \sum_k a_k \phi(x-k)
\]

Suppose we have numerical data \(x(t_k)\) sampled at regular intervals \(t_k = k/N\) for \(N = 2^R\) and \(k = 0, \ldots, N-1\). We may think of \(x(t)\) as a function constant on the intervals \((k/N, (k+1)/N)\) with value \(x_k = x(t_k)\) there, and we may write

\[
x(t) = \sum_{0 \leq k \leq 2^R - 1} \alpha_{R,k} \Phi_{R,k} (t) \quad \text{with}
\]

\[
\alpha_{R,k}^n = x(t_k) 2^{-R/2}
\]

Also, the wavelet coefficients \(\beta^R_n\) defined by (16) will vanish for \(n \geq R\), since \(x(t)\) in (21) will be orthogonal to \(\psi_{n,k}\) for \(n \geq R\). Using the equations (18) and (19) one can show that the \(\alpha, \beta\)-coefficients satisfy

\[
\alpha_{R,k}^n = \frac{1}{\sqrt{2}} \alpha_{2k}^{n+1} + \frac{1}{\sqrt{2}} \alpha_{2k+1}^{n+1}
\]

\[
\beta_{R,k}^n = \frac{1}{\sqrt{2}} \alpha_{2k}^{n+1} - \frac{1}{\sqrt{2}} \alpha_{2k+1}^{n+1}
\]
Thus, if we know the \( a\)-coefficients for a certain level \( n \) we can calculate the coefficients for all sub-levels. All this is translated into the following algorithm (Fast Haar Transform) for calculating the coefficients for the data \( x_k, k = 0, \ldots, 2^R - 1 \). From it we produce two sub-sequences \( a', b' \) of length \( N/2 \) (\( N = 2^R \)) by

\[
\begin{align*}
a'_k &= \frac{1}{\sqrt{2}} x_{2k} + \frac{1}{\sqrt{2}} x_{2k+1} \\
b'_k &= \frac{1}{\sqrt{2}} x_{2k} - \frac{1}{\sqrt{2}} x_{2k+1}
\end{align*}
\]

(22*)

\[ k = 0, \ldots, \frac{N}{2} - 1 \]

Thus, the sequence \( x \) is decomposed into \( (a' | b') \). Applying the same procedure to \( a' \) we get \( (a^2 | b^2) \) of length \( N/4 \) and \( N/4 \). Then continue with \( a' \), etc, till finally \( x \) is decomposed into \( (a^R | b^R | b^{R-1} | \ldots | b') \).

The wavelet coefficients are then given by

\[
\beta^n_k = b_k^{R-n} 2^{-R/2} \quad (= 0 \text{ if } n \geq R)
\]

\[ k = 0, \ldots, 2^n - 1 \]

For data with \( 2^R \) points we therefore obtain \( 1 + 2 + 2^2 + \ldots + 2^{R-1} = 2^R - 1 \) wavelet coefficients. If one add to this the average \( a^R \) of the data one gets in all \( 2^R \) constants from which the signal can be reconstructed. The algorithm is easily implemented in computer code. Conversely, from the decomposition \( (a^R | b^R | b^{R-1} | \ldots | b') \) we can reconstruct the data \( a^0 \) through successive steps

\[
\begin{align*}
a'^{R-1}(k) &= a'(k) + b'(k) \\
a'^R(k) &= \sum_i \alpha^{R-1}_{i} \phi_{R-j_i}(t_k) \\
b'^R(k) &= \sum_i \beta^{R-1}_{i} \psi_{R-j_i}(t_k)
\end{align*}
\]

From (22) and (16*) it is evident that the “energy” is conserved for the transformation

\[
\sum x_k^2 = \sum (\beta^n_k)^2 + \sum a_k^2
\]

Furthermore, (22*) implies that if the signal is constant the \( \beta \)-coefficients will vanish. If the signal is linear (as a function of time) the \( R-1 \) level \( \beta \)-coefficients are constant, whence the next level of \( \beta \)-coefficients will vanish. For a quadratic signal the \( R-3 \) level \( \beta \)-coefficients will vanish, and so on. Thus, the \( \beta \)-coefficients will be large where the signal changes abruptly. This can be viewed directly if we map the \( \beta \)-coefficients on the time axis such that \( \beta^n_k \) will mapped to time \( 2^{-n} k \). The next figure shows the EMG-signal of the isometric test along with its Haar wavelet coefficients (the upper curve in the figure – it has been displaced vertically in order to distinguish the curves).
We can also display the $\beta$–coefficients for the different scales in sort of a discrete version of the scalogram.

In this figure there are 12 levels. The 0th level corresponds to just one $\beta$–coefficient, on the next level we have two $\beta$–coefficients, and finally on the 11th level we have $2^{11} = N/2$ $\beta$–coefficients. Dark color means low (negative) value, white is for high value. The next figure shows the absolute values of the $\beta$–coefficients of the same data and with more colors.
In the diagram the region of muscle contraction is clearly enhanced, as the resting signal seems to be suppressed. The diagram can be read as a “musical transcription” of the EMG-signal. The 11th level corresponds to high pitch around 500 Hz, 10th level to 250 Hz, and so on. Apparently the signal for the contraction seems to be concentrated around the 7th and 9th level (around 30 – 125 Hz).

The Haar transform is a special instance of a so-called sub-band coding filter bank, which can be pictured as follows

The information flows from left to right. In the first step data (a0) is put through a high pass filter H1 which produces the first set of “detail part” b1 of the signal, whereas the low pass produces a sort of “moving average” a1 of the data (compare (22*)). The next step repeats the procedure with a1 as input data, and so on. In equations this can be written

\[ b^n[k] = \sum_j h_1[2k - j]a^{n-1}[j] = \sum_j h_1[j]a^{n-1}[2k-j] \]
\[ a^n[k] = \sum_j h_0[2k - j]a^{n-1}[j] \]

where \( h_0[m] \) and \( h_1[m] \) denote the low and high pass filter coefficients. In (25) r. h. s. we have included the effect of down-sampling (denoted by downward arrows in the diagram) by using \( 2k \) instead of \( k \) in the argument of filter coefficients. Thus, down-sampling a sequence \( \{r_k\} \) means we throw every second element obtaining a sequence \( \{s_k\} \) with \( s_k = r_{2k} \). Up-sampling on the other hand means that we enlarge a sequence \( \{r_k\} \) by inserting zeros between consecutive elements obtaining a sequence \( \{s_k\} \) with \( s_{2k} = r_k \) and \( s_{2k+1} = 0 \). Then the synthesis of the signal \( (a0) \) from its decomposition can be described by the following diagram (showing the two final steps – up-sampling is denoted by an upward arrow).

The synthesis low and high pass filters are denoted by G0 and G1 with filter coefficients \( g_0[m] \) and \( g_1[m] \). For the synthesis equation we obtain

\[ a^{n-1}[k] = \sum_j g_0[k-2j]a^n[j] + \sum_j g_1[k-2j]b^n[j] \]

Consider the two-step case consisting of analysis and synthesis.
A signal $x$ is analyzed and re-synthesized as a signal $x_r$. The process can be conveniently described using $z$-transforms. Given a sequence $\{x_k\}$ then its $z$-transformation, denoted $X(z)$, is defined as

$$X(z) = \sum_{k} x_k z^{-k}$$

for a complex number $z$. The convolution of two sequences $\{h_k\}$ and $\{x_k\}$ is defined by

$$h \ast x[k] = \sum_{j} h[k - j]x[j]$$

with the important property that its $z$-transform is the product $H(z)X(z)$. In terms of $z$-transforms the up-sampling and down-sampling are described as

$$X(z) \downarrow = \frac{1}{2} \left( X\left(\sqrt{z}\right) + X\left(-\sqrt{z}\right) \right)$$

$$X(z) \uparrow = X\left(z^2\right)$$

Using these properties the process of the above diagram can be written in the $z$-domain as

$$Xr(z) = \frac{1}{2} \left( G_0(z)H_0(z) + G_1(z)H_1(z) \right)X(z) +$$

$$\frac{1}{2} \left( G_0(z)H_0(-z) + G_1(z)H_1(-z) \right)X(-z)$$

From this it follows that $x_r$ is a perfect reconstruction of $x$ if we have

$$G_0(z)H_0(z) + G_1(z)H_1(z) = 2z^{-K}$$

$$G_0(z)H_0(-z) + G_1(z)H_1(-z) = 0$$

This means that $x_r$ is identical with $x$ except for a time delay, $x_r[n] = x[n-K]$.

In terms of these new concepts the Haar transform is represented by the following filters

$$G_0(z) = \frac{1 + z^{-1}}{\sqrt{2}} \quad G_1(z) = \frac{1 - z^{-1}}{\sqrt{2}}$$

$$H_0(z) = \frac{1 + z}{\sqrt{2}} \quad H_1(z) = \frac{1 - z}{\sqrt{2}}$$
(We observe that for the Haar filters we have \( G_0(z) = G_1(−z) \) and similarly for the analysis filter, which is characteristic of the so-called Quadrature Mirror Filters, QMF.) If we set \( z = e^{j2\pi f} \) in \( H(z) \) we obtain the discrete Fourier transformation \( H(f) \).

\[
H(f) = \sum_{k} h[k] e^{-j2\pi kf}
\]

(32)

Using the relations (assuming real-valued filter coefficients in (33.b))

\[
H(−z) = H(f + \frac{1}{2})
\]

(33)

\[
H(z^{-1}) = H^*(f)
\]

(\( z = e^{j2\pi f} \))

and (30) one can show that the Haar filter function satisfies

\[
|G_0(f)|^2 + |G_0(f + \frac{1}{2})|^2 = 2
\]

(34)

This equality is in fact necessary for the corresponding wavelet to be orthogonal. If we can solve (30) for \( G \) we obtain

\[
G_0(z) = \frac{2z^{-K}}{D(z)} H_1(−z)
\]

(35)

\[
G_1(z) = -\frac{2z^{-K}}{D(z)} H_0(−z)
\]

if \( D(z) = H_0(z)H_1(−z) - H_0(−z)H_1(z) \neq 0 \)

(we have for the Haar filter \( D(z) = 2z \) and \( K = -1 \)). The so-called conjugate quadrature filters (CQF) are obtained by setting \( D(z) = 2z^k \); i.e.,

\[
G_0(z) = H_1(−z)
\]

(36)

\[
G_1(z) = -H_0(−z)
\]

plus postulating a special relation between high pass and low pass filter (the exponent in (37) must be odd since \( D(z) \) (35) is an odd function)

\[
H_1(z) = -z^{-2k+1}H_0(−z^{-1})
\]

(37)

From (37), (36) and (30) we obtain

\[
H_0(z)H_0(z^{-1}) + H_0(−z)H_0(−z^{-1}) = 2z^{2k+1-K}
\]

(38)

which in term of the discrete Fourier transform becomes

\[
|H_0(f)|^2 + |H_0(f + \frac{1}{2})|^2 = 2z^{2k+1-K}
\]

(39)

Since the l. h. s. of (39) is real-valued we must have \( K = 2k + 1 \) on the r. h. s.; i.e., the CQF filters satisfy the orthogonality condition (34). Equation (38) can now be written as
One approach to filter design is to find a polynomial \( P(z) \) satisfying (40.a) and then try to factorize it on the form (40.b) obtaining the filter \( H \). As an example, for the Haar filter we have

\[
P(z) = \frac{1}{2} \left( z + 2 + z^{-1} \right) = \left( \frac{1 + z}{2} \right) \left( \frac{1 + z^{-1}}{2} \right)
\]

which was generalized by Daubechies (1988) as

\[
P(z) = \left( \frac{1 + z}{2} \right)^N \left( \frac{1 + z^{-1}}{2} \right)^N Q_N(z)
\]

where \( Q_N \) is a “symmetrical” polynomial of degree \( N - 1 \); i.e., it satisfies

\[
Q_N(z) = Q_N(z^{-1})
\]

\[
Q_N(z) = q_{N-1} z^{-N-1} + q_{N-2} z^{-N+2} + \ldots + q_0 + \ldots + q_{N-1} z^{N-1}
\]

The polynomial \( Q_N \) is uniquely determined by the equation (32). Indeed, this equation means that the even terms in \( P(z) \) must vanish (except for the constant term), and since \( P(z) \) is “symmetrical” of degree \( 2N - 1 \) this leaves \( N \) conditions for the \( N \) coefficients \( q_k \). Having obtained \( Q_N \), one has yet to factorize it according to (40.1). One possible factorization in case of \( N = 2 \) is the famous Daubechies wavelet (of 4th order and usually referred to as Daub4) with the low pass analysis filter

\[
H_0(z) = \frac{1}{4 \sqrt{2}} \left\{ 1 + \sqrt{3} \right\} z^3 + \left\{ 3 + \sqrt{3} \right\} z^2 + \left\{ 3 - \sqrt{3} \right\} z + \left\{ 1 - \sqrt{3} \right\}
\]

The high pass filter is then given applying (37) as

\[
h_1[n] = (-1)^n h_0[3 - n]
\]

Finally we can obtain the scaling function and the wavelet from the filters as follows. From (20) we see that using a Dirac impulse \( \delta(t - t_m) \) as the input signal \( x(t) \) we obtain (assuming real-valued functions)

\[
a^{R-n}[k] = \int \delta(t - t_m) \phi_{n,k}(t) dt = 2^{n/2} \phi(2^n t_m - k)
\]

\[
b^{R-n}[k] = \int \delta(t - t_m) \psi_{n,k}(t) dt = 2^{n/2} \psi(2^n t_m - k)
\]

Using \( R \) steps in the filtering we can determine the scaling functions at points \( 2^R \); thus, with an increasing number of steps the scaling function (and similarly the wavelet) will be determined at a dense set of points. From (25) and (46) we find that the functions \( \phi \) and \( \psi \) must satisfy the following scaling relations

\[
\phi(t) = \sqrt{2} \sum_j h_j [-j] \phi(2t - j)
\]

\[
\psi(t) = \sqrt{2} \sum_j h_j [-j] \phi(2t - j)
\]
The equations in (47) generalize the Haar case (18) and (19). Applying the Fourier transformations to (47) we can derive the following formal expression for the Fourier transform of the scaling function

\[ \hat{\phi}(f) = \prod_{k=1}^{\infty} M_0 \left( \frac{k}{2^k} \right) \]

(48)

\[ M_0(f) = \frac{1}{\sqrt{2}} H_0(-f) \]

Similarly we obtain for the Fourier transform of the wavelet

\[ \hat{\psi}(f) = M_1 \left( \frac{k}{2^k} \right) \]

(49)

\[ M_1(f) = \frac{1}{\sqrt{2}} H_1(-f) \]

One can also use the equations in (47) in order to implement an iteration procedure

\[ \phi^{(i+1)}(t) = \sqrt{2} \sum_j h_0[-j] \phi^{(i)}(2t - j) \]

(50)

\[ \phi^{(0)}(t) = \chi_{[0,1]}(t) \]

With the initial function taken as the characteristic function for the unit interval [0,1] it follows that the limit function of (50) will be zero outside [0,3], which is also true of the corresponding wavelet defined by (47.b); the wavelet is said to have compact support.
The first figure above shows the Daub4 scaling function computed using the iteration procedure in six steps (it is normalized so that its integral is equal to 1). The second figure shows the Daub4 wavelet obtained by using (46) in the form of 
\[ \beta_{r-2}^{R} = 2\psi(4t_{m}) \] employing as input signals the discrete versions \( x^{m}[k] = \delta_{m}^{k}, \quad m = 0, \ldots, 2^{R} - 1, \) of the Dirac impulse \( \delta(t - t_{m}). \) The wavelet coefficients were calculated using the filters (44–45). Of course we could have obtained it also from the scaling function through (47.b).

As the picture of the Daub4 wavelet shows, it is a bit more regular than the Haar wavelet but perhaps not too similar to the myoelectric potentials we wish to match. Smoother version can be obtained by going to the higher order Daubechies wavelets. An interesting class of wavelets in this regard is the ones using the so-called B-splines as the scaling function. The \( 0^{th} \) case is the Haar wavelet using the characteristic function for the unit interval as the scaling function. The \( N:th \) order B-spline scaling function is obtained by convoluting the \( (N - 1):th \) order with the characteristic function for the unit interval

\[
\phi_{N}(t) = \int \phi_{0}(t-u)\phi_{N-1}(u)du \\
\phi_{0}(t) = \chi_{I_{0.1}}(t)
\]

If we take the Fourier transform of (51) we obtain

\[
\hat{\phi}_{N} = \left( \hat{\phi}_{0} \right)^{N+1} \\
\hat{\phi}_{0}(f) = -e^{-m\pi f} \left( \frac{\sin(\pi f)}{\pi f} \right)
\]

Using (48) we can derive the relation

\[
\hat{\phi}(2f) = \hat{M}_{0}(f)\hat{\phi}(f)
\]

That is, knowing the Fourier transform of the scaling function we can calculate the filter \( H_{0} \) (from \( M_{0} \)). Applying this method to (52) we obtain
This gives for the low pass filter coefficients in case of the $N$:th order B-spline

$$H_0(f) = \sqrt{2}M_0(-f) = \sqrt{2} e^{i(N+1)\pi f} \left( \cos(\pi f) \right)^{N+1} = \frac{\sqrt{2}}{2^{N+1}} (1 + e^{i2\pi f})^{N+1}$$

From this we can calculate the high pass filter coefficients (cmp (45)) and finally the wavelet function from (47.b). The quadratic case ($N = 2$) is graphed below.

As can be seen from the graph, the quadratic B-spline wavelet is much more like an “action pulse” than the Daubechies wavelet. Unlike the Haar and Daubechies wavelets the B-spline wavelets of order $N > 0$ are not orthogonal. There is a mathematically straightforward way, however, to orthogonalize the wavelets. Define a new scaling function $\phi_{orth}$ by

$$\hat{\phi}_{orth}(f) = \frac{\hat{\phi}(f)}{\sqrt{\sum_k |\hat{\phi}(f+k)|^2}}$$

which will automatically satisfy the orthogonality condition

$$\sum_k |\hat{\phi}(f+k)|^2 = 1$$

The drawback with the orthogonal B-spline wavelets obtained in this way is that they have infinite filters. One way to circumvent this problem are the so-called bi-orthogonal wavelets, which use different analysis and synthesis wavelets, called primary and dual wavelets. The dual and primary wavelets are orthogonal vis-a-vis each other but not within themselves (dual wavelet marked with a tilde)

$$\langle \psi_{n,k}, \tilde{\psi}_{m,l} \rangle = \delta_n^m \delta_k^l$$
Instead of (16) we get

\[ x(t) \approx \sum_{n,k} \langle x, \psi_{n,k} \rangle \psi_{n,k}(t) \]  

(59)

Bi-orthogonal B-spline filters are obtained by assuming the relations

\[ G_0(z) = z^{-1}H_1(-z) \]
\[ G_1(z) = -z^{-1}H_0(-z) \]  

(60)

between synthesis and analysis filters (this is a special case of (35)). The perfect reconstruction condition (30.a) can thus be written as

\[ P(z) + P(-z) = 2z^{-K} \]
\[ P(z) = G_0(z)H_0(z) \]  

(61)

The bi-orthogonal B-spline wavelets consist of the solutions (see the Appendix for proof)

\[ H_0(z) = \sqrt{2} \left( \frac{1 + z^{-1}}{2} \right)^N z^{N/2} \]

(62)

\[ G_0(z) = \sqrt{2} \left( \frac{1 + z}{2} \right)^N \sum_{j=0}^{L-1} \left( \frac{L-1+j}{L} \right) \sin^2(\pi f \frac{N}{L}) \]
\[ L = \frac{N + \tilde{N}}{2}, \quad z = e^{2\pi i f}, \quad N + \tilde{N} \text{ must be even.} \]

The analysis and synthesis high pass filters are then obtained using (60).

**Appendix**

A. The orthogonality condition

Assume the translates of the scaling function are orthogonal

\[ \int \phi(x-n)\phi(x)dx = \delta_n^0 \]  

(A.1)

In the frequency domain this condition reads as

\[ \delta_n^0 = \int e^{2\pi i nf} \left| \Phi(f) \right|^2 df = \]

\[ \int_0^1 \sum_k \left| \Phi(f + k) \right|^2 df \]

(A.2)

where in the last equality we used the periodicity of the exponential function. Condition (A.2) is equivalent with
(A.3) \[ \sum_k \hat{\phi}(f + k)^2 = 1 \]

which was referred to above in equ (57). Suppose we have

(A.4) \[ \hat{\phi}(f) = M_0 \left( \frac{f}{2} \right)^\frac{L}{2} \]

with

\[ M_0(f + 1) = M_0(f) \]

then (A.3) leads to the following form of the orthogonality condition

(A.5) \[ |M_0(f)|^2 + |M_0(f + \frac{1}{2})|^2 = 1 \]

B. The B-spline solution

We prove that (62) satisfies (61.a) with \( K = 0 \). For \( P(z) \) we obtain the following expression ( \( z = e^{j2\pi f} \))

\[ P(z) = 2 \cos(\pi f)^N \sqrt{\Delta} Q_L \left( (\sin(\pi f))^2 \right) \]

\[ Q_L(y) = \sum_{j=0}^{L-1} \binom{L-1 + j}{j} y^j \]

Equation (61.a) becomes using \( y = (\sin(\pi f))^2 \)

(B.2) \[ (1 - y)^L Q_L(y) + y^L Q_L(1 - y) = 1 \]

Equation (B.2) can be proved by induction on \( L \). It is obviously true for \( L = 1 \) when \( Q_1 = 1 \). Using the property

(B.3) \[ \binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1} \]

for the binomial coefficients we deduce, directly from the definition of \( Q \), that

(B.4) \[ (1 - y)L(y) = Q_{L-1}(y) + y^{L-1} \left( \binom{2L-3}{L-1} - y \binom{2L-2}{L-1} \right) \]

\[ = Q_{L-1}(y) + \binom{2L-3}{L-1} y^{L-1} (1 - 2y) \]

Using (B.4) we get (interchanging \( y \) and \( 1 - y \))

(B.5) \[ (1 - y)^L Q_L(y) + y^L Q_L(1 - y) = (1 - y)^L Q_{L-1}(y) + y^{L-1} Q_{L-1}(1 - y) \]

which completes the induction.
C. Frames

For an admissible wavelet, like the Mexican wavelet (9) above, the signal can be reconstructed from the wavelet transform of the signal via the inverse CWT (5). For special wavelets, like the Haar wavelets and the Daubechies wavelets, the signal can be reconstructed using a discrete subset of the dilated and scaled versions of the wavelet (equ (16*)). Does this work also e.g. for the Mexican wavelet? That question is addressed by the concept of frames. The square integrable functions on \( \mathbb{R} \) define an infinite dimensional vector space (a so-called Hilbert space), \( L^2(\mathbb{R}) \), with the inner product

\[
\langle f, g \rangle \equiv \int f(x) g^*(x) dx
\]

\[
\|f\| \equiv \sqrt{\langle f, f \rangle}
\]

A set of functions \( \{e_k\} \) is called a frame if there are constants \( B > A > 0 \) such that

\[
A\|f\| \leq \left| \sum_k \langle f, e_k \rangle e_k \right| \leq B\|f\|
\]

for every \( f \) in \( L^2(\mathbb{R}) \). The property (C.2) ensures that the function \( f \) can be reconstructed from its “Fourier coefficients” \( \langle f, e_k \rangle \). Indeed, define the map \( T \) which sends functions \( f \) in \( L^2(\mathbb{R}) \) to sequences \( \{c_n\} \) in \( l^2 \) (the space of square summable sequences) given by \( c_k = \langle f, e_k \rangle \). The dual map \( T^* \) mapping sequences \( \{c_n\} \) in \( l^2 \) on functions \( f \) in \( L^2(\mathbb{R}) \) is defined by \( f = \sum_k c_k e_k \). It follows then from (C.2) that

\[
(AI) \leq T^*T \leq BI
\]

(\( I \) denotes the identity map), which ensures that \( T^*T \) has an inverse. It is therefore possible to define a dual set of vectors through

\[
\tilde{e}_k = (T^*T)^{-1} e_k
\]

It follows that the function \( f \) can be represented as \( \sum_k \langle f, e_k \rangle \tilde{e}_k \) since

\[
\sum_k \langle f, e_k \rangle \tilde{e}_k = \sum_k \langle f, e_k \rangle (T^*T)^{-1} e_k
\]

\[
= (T^*T)^{-1} T^*T(f) = f
\]

Daubechies (1992) has given criteria for the functions

\[
\psi_{m,n}(t) = a^{m/2} \psi(a^{m/2} - nb)
\]

to constitute a frame for a given wavelet \( \psi \). Especially, for the Mexican wavelet

\[
\psi(x) = \frac{2}{\sqrt{3}} \pi^{-\frac{1}{4}} (1 - x^2)^{\frac{\gamma}{2}}
\]

(the constants in (C.6) make its square integral equal to 1) we have the frame bounds in case \( a = 2 \):
If \( A \) is close to \( B \) we can use the following formula (only a few terms are usually needed for a good approximation) to calculate the dual elements

\[
\hat{e}_k = \frac{2}{A + B} \sum_{j=0}^{n} \left( I - \frac{2}{A + B} T^* T \right)^j e_k
\]

Equation (C.7) follows from (C.4) and the operator equivalent of the formula \( \frac{1}{1-x} = 1 + x + x^2 + \cdots \).

**D. The Morelet “wavelet”**

The Morelet “wavelet”

\[
\psi_{f_o, \alpha}(t) = \left( \frac{\alpha}{\pi} \right)^{\frac{1}{4}} e^{\frac{2\pi i f_o - \alpha}{2} t^2}
\]

is widely used in signal processing, though it is not a proper wavelet. Its Fourier transform is given by

\[
\hat{\psi}_{f_o, \alpha}(f) = \sqrt{2} \left( \frac{\pi}{\alpha} \right)^{\frac{1}{4}} e^{-\frac{2\pi^2}{\alpha} (f - f_o)^2}
\]

which does not vanish for \( f = 0 \), but for typical values of the constants it can still be quite small at the point \( f = 0 \). If we define the time and frequency resolution by

\[
\Delta f = \sqrt{\int (f - f_o)^2 \left| \psi_{f_o, \alpha}(f) \right|^2 df}
\]

\[
\Delta t = \sqrt{\int t^2 \left| \psi_{f_o, \alpha}(t) \right|^2 dt}
\]

we get for the Morelet wavelet

\[
\Delta f = \frac{\sqrt{2\alpha}}{4\pi}
\]

\[
\Delta t = \frac{1}{\sqrt{2\alpha}}
\]

Thus, the time and frequency resolutions are inversely related, \( \Delta f \Delta t = \frac{1}{4\pi} \) and this is the best bound that can be achieved for any distribution (Heisenberg relation). This is why the Morelet wavelet has been considered as kind of an optimal tool for time-frequency analysis. Conventionally one defines the “quality factor” for the Morelet wavelet (D.1) by
\( (D.5) \quad Q = \frac{\pi f_0}{\sqrt{2} \alpha} \)

which is proportional to \( f_0/\Delta f \). A typical procedure is to calculate the transforms of a signal,

\( (D.6) \quad x_{f,a}(t) = \int \psi_{f,a}(t-u)x(u)du \)

for a sequence of a frequencies determined by

\[ f_s = q^s f_0 \]

\( (D.7) \quad q = \frac{1 + \frac{1}{2Q}}{1 - \frac{1}{2Q}} \)

This means that the overlap between two adjacent frequency bands is of the order of \( e^{-2} \) if \( Q \gg \frac{1}{2} \) (i.e. about \(-8.7\) dB).

The above figure shows the (logarithmic) absolute value of \((D.6)\) as a function of frequency (vertical axis, ranging from 50 Hz to 300 Hz on a logarithmic scale) and time for the EMG-signal of the isometric test. The parameter values are \( \alpha = 200, f_0 = 50 \text{ Hz}, q = 1.136 \) and \( Q = 7.85 \). These correspond, according to \((D.4)\), to a time resolution \( \Delta t \) of 0.05 s and a frequency resolution \( \Delta f \) of 1.59 Hz.

**Literature**

(Hint: many valuable papers can be accessed on internet via ResearchIndex at http://citeseer.nj.nec.com/.)

1. Antenucci J: “Notes on spectral analysis”,