HST582J/6.555J/16.456J Biomedical Signal and Image Processing Spring 2005

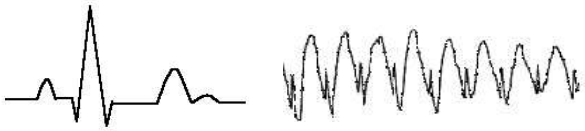
Chapter 16 - Wavelets; Multiscale Activity in Physiological Signals

1 Introduction

Physiological signals are dynamic; they exhibit time-varying statistics (such as value of the mean and variance over a given temporal window) in both the time and frequency domain. Physiological signals also exhibit activity that spans a range of *time scales*. For instance the cycles of an electrocardiogram (ECG) signal contain three components with different time-scales: atrial depolarization represented by the P wave has a duration of 0.1 to 0,15 seconds; ventricular depolarization represented by the QRS complex has a duration around 0.1 seconds; and ventricular repolarization represented by the T-wave has a duration of about 0.2 to 0.4 seconds. However, the inter-beat timing (which when averaged gives the heart rate), changes over much longer time scales; from minutes to hours, to days. Short term variations (activity on the order of seconds and minutes above 0.01 Hz) are due to changes in the sympathetic (*fight-and-flight*) and parasympathetic (*rest-and-digest)* activity of the central nervous system acting on the heart. Longer term variations can be partially attributed to changes in activity and intrinsic circadian controls (that lead to sleep for example). The changes in the heart rate over many scales can provide diagnostic information [12].

The spike-and-slow-wave complex observed in the electroencephalogram (EEG) during the evolution of some epileptic seizures is another example of a physiological signal with multiscale activity [8]. In this case the spike component of the waveform represents the short time-scale event, and the slow-wave component of the waveform represents the long time-scale event.

Classical signal processing tools such as the Fourier transform are not suited for analyzing dynamic, non-stationary signals because implicit in their formulation is an assumption of signal stationarity. Generalizations of the Fourier transform, such as the Short-Time Fourier transform, can be used to analyze signals with time-varying spectral and temporal characteristics. However, the Short-Time Fourier transform cannot be used to simultaneously resolve activity at different time-scale because implicit in its formulation is a selection of a time-scale. This chapter introduces the wavelet transform, a generalization of the Short-Time Fourier transform that can be used to perform multi-scale signal analysis.



ECG Cycle EEG Spike-Wave

Figure 1: Physiological Signals with Multiscale Activity: Each cycle of an ECG signal contains three components: The P-wave, QRS complex, and the T-wave. The spike-and-slow wave complex observed in epileptic EEG has two components with different time-scales: A spike, and a slow wave.

2 Short-Time Fourier Transform

The Fourier transform is well suited for analyzing *stationary signals*; these are signals with time-invariant spectral content. The sum of N sinusoids is an example of a stationary signal because at every point it has the same N frequency components. The chirp signal, which is a sinusoid with linearly or quadratically varying frequency, is an example of a *non-stationary signal*. The Fourier transform cannot capture the spectral evolution of a non-stationary signal for two reasons that are apparent in the Fourier transform analysis/synthesis equation:

$$F(\Omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} \quad x(t) = \int_{-\infty}^{\infty} F(\Omega)e^{j\omega t} \tag{1}$$

1. The Fourier Transform synthesizes $x(t)$ from a linear combination of stationary signals; in particular, the sum of time-invariant, ever-lasting sinusoids. A non-stationary signal cannot be accurately represented using a sum of stationary signals. For example, a chirp signal with a frequency that increases linearly from 0-2 Hz over a two second interval cannot be accurately represented using a linear combination of fixed frequency sinusoids.

2. The Fourier Transform of a signal is a mapping from a function of time $x(t)$ to a function of frequency $F(\Omega)$. The function $F(\Omega)$ tells us the extent to which a signal component with frequency Ω is present in the analyzed signal. It does not indicate how that signal component evolves with time t. For that we need a transform that returns a bivariate function of the form $F(t, \Omega)$.

Figure 2 shows the time and frequency domain plots of a signal composed from the summing of a 10 Hz, 30 Hz, and 50 Hz frequency sinusoids. The Fourier transform reveals the presence of this stationary signal's three frequency components.

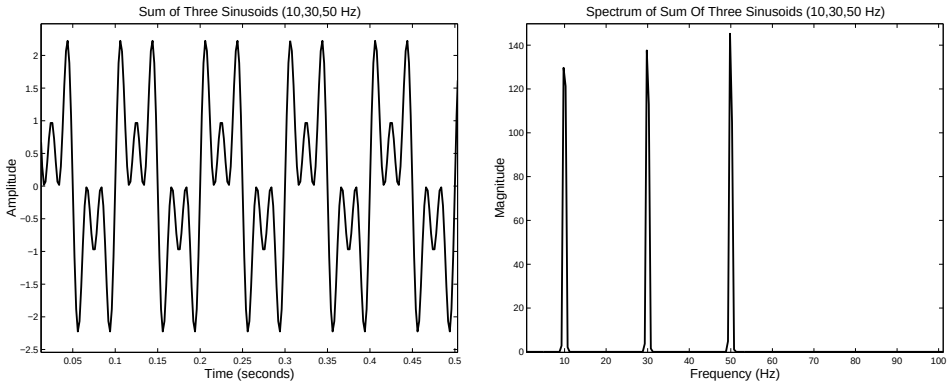


Figure 2: Fourier Analysis of Stationary Signals: The Fourier transform is well-suited for the analysis of stationary signals such as the sum of three sinusoids. In this case, the Fourier transform is used to reveal the presence of three frequency components (10 Hz, 30 Hz, 50 Hz) in the signal.

Figure 3 shows the time and frequency domain representations of a chirp signal with a frequency that increases linearly from 0 to 250 Hz. The Fourier transform of the chirp signal implies the presence of frequencies between 0-250 Hz for the entirety of the signal. In reality, each of these frequency components is only present for a short duration of the chirp signal.

The generalization of the Fourier transform that allows one to study non-stationary signals is the Short-Time Fourier Transform (STFT). The STFT maps $x(t)$ into a bivariate function $F(t, \Omega)$. This function can be used to determine the extent to which a signal component with frequency Ω is present at time $t = t_0$. The construction of $F(t, \Omega)$ involves:

1. A segment of the signal $x(t)$ that begins at time $t = t_0$ is extracted using a window $w(t)$ with duration L. The segment is given by $s(t) = x(t)w(t - t_0)$ and has a duration L.

2. The Fourier Transform of the segment $s(t)$ is computed to give $S(\Omega) = F(t = t_0, \Omega)$. The function $F(t = t_0, \Omega)$ reveals the spectral content of the signal segment in an interval of time beginning at $t = t_0$ and ending at $t = t_0 + L$.

3. The window $w(t)$ is shifted so that it can be used to extract a new signal segment $s(t) = x(t)w(t - t_0 - \Delta t)$. As in the previous step, the Fourier Transform is again used to reveal the spectral content of the new segment.

A plot of the magnitude of the signal segment spectra is known as a *Spectrogram*. The spectrogram has time on the x-axis; frequency on the y-axis; and magnitude of the spectra on the z-axis. The spectrogram illustrates visually how different frequency components evolve over the duration of a signal. Figure 4 is graphical illustration of the process leading to the construction of a Spectrogram.

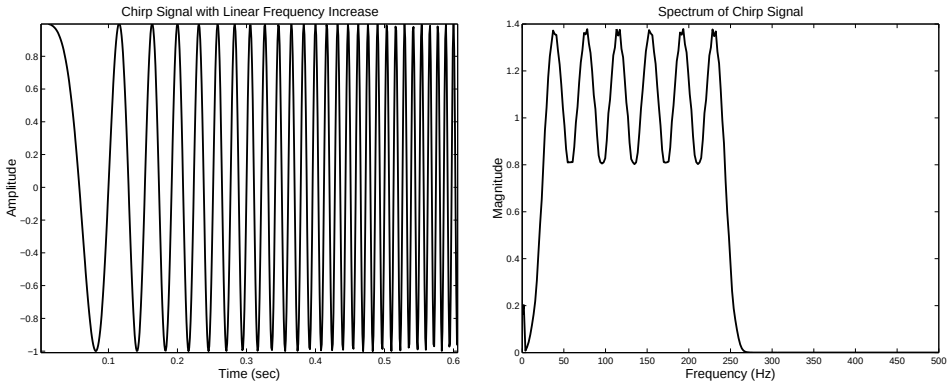


Figure 3: Time and Frequency Plots of Chirp Signal: The Fourier Transform is not well-suited for the analysis of non-stationary signals like the chirp signal. The Fourier Transform in the case of this chirp signal demonstrates the presence of its frequency components for all time, when in actuality each component is present for only a brief duration of time.

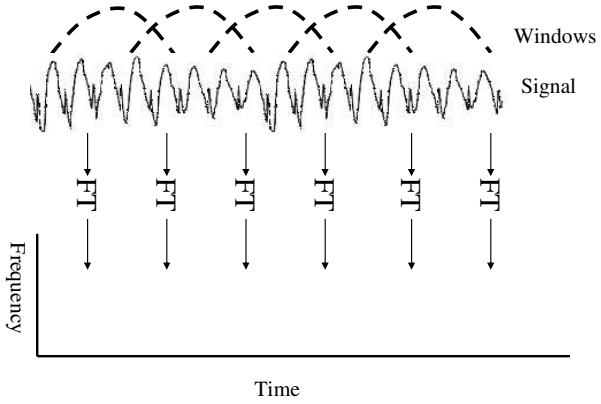


Figure 4: Graphical Construction of the Spectrogram: A spectrogram is constructed by aligning the spectra of adjacent, overlapping signal segments in the time-frequency plane. The spectrogram can be used to study the evolution of a signal's frequency components.

The mathematical expressions for the continuous and discrete short-time Fourier transforms mirror the graphical construction outlined above in Figure 4.

$$F(t,\Omega) = \int x(\tau)w(t-\tau)e^{-j\Omega\tau} \quad F(n,\omega) = \sum x(k)w(n-k)e^{-j\omega k} \tag{2}$$

As an example, consider constructing the spectrograms of two chirp signals. The chirp signal c_1 has linearly increasing frequency, while the chirp signal c_2 has quadratically increasing frequency.

$$c(t)_1 = \sin(2\pi(\alpha + \beta t)t); \tag{3}$$

$$c(t)_2 = \sin(2\pi(\alpha + \beta t + \gamma t^2)t); \tag{4}$$

As expected, the spectrogram of c_1 shows that as time progresses the frequency present in signal segments increases linearly. The spectrogram of c_2 shows a quadratic increase in frequency with time.

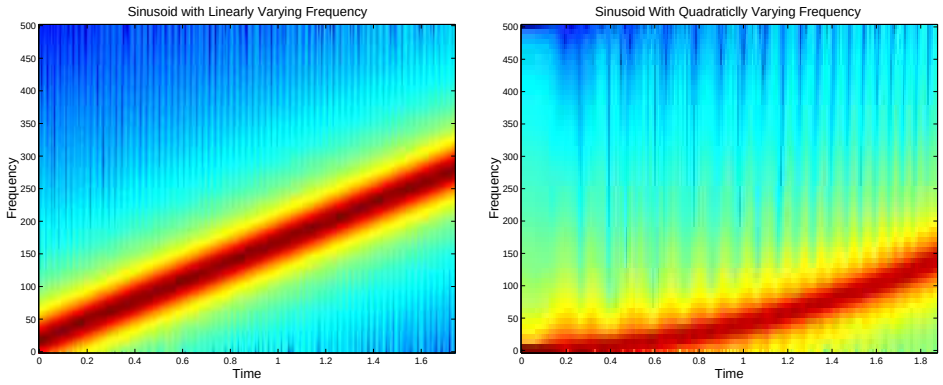


Figure 5: Spectrogram of Chirp Signals With Linearly and Quadratically Increasing Frequency

The window $w(t)$ determines the spectral and temporal resolution of the Short-Time Fourier Transform. Temporal resolution refers to the smallest time-separation below which two temporal events cannot be distinguished on the Spectrogram. Similarly, spectral resolution refers to the smallest frequency-separation below which two spectral events cannot be distinguished on the Spectrogram. A long window $w(t)$ results in poor temporal resolution and good frequency resolution. Conversely, a short window $w(t)$ results in good temporal resolution and poor spectral resolution. Time and frequency resolution can only be traded for one another, they cannot both be improved simultaneously. Figure 6 demonstrates how

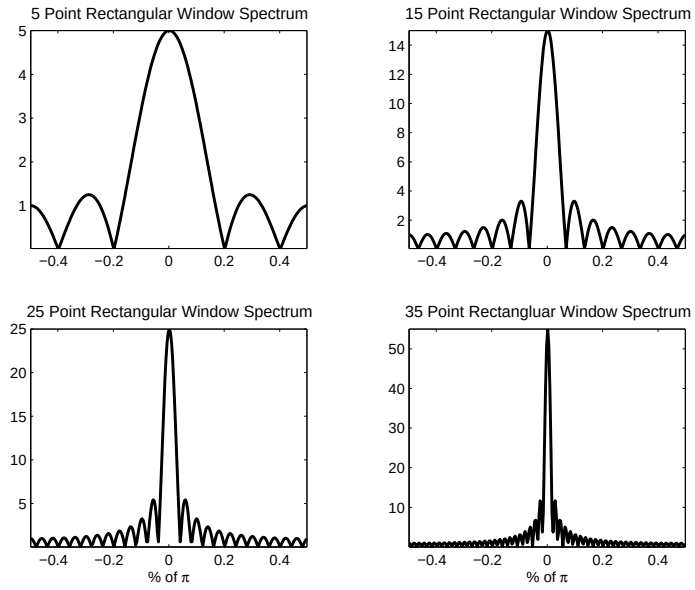


Figure 6: Spectral and Temporal Resolution of the Short-Time Fourier Transform: A short analysis window results in good temporal resolution and poor spectral resolution. As the analysis window length increases its frequency resolution increases and its temporal resolution decreases.

the frequency resolution of a window improves as the window's length increases (temporal resolution decreases).

From the perspective of the time-frequency plane, the analysis window length limits our knowledge of signal activity to a two-dimensional cell. The dimension of the cell along the time axis indicates the limits of temporal resolution; temporal events with time-separation smaller than this dimension cannot be differentiated using a spectrogram. The dimension of the cell along the frequency axis indicates the limits of frequency resolution; spectral events with frequency-separation smaller than this dimension cannot be distinguished using a spectrogram. The dimensions of the cell can be altered so as to favor either temporal or spectral resolution, but the area of the cell remains constant.

As an example of trading between temporal and spectral resolution, consider computing two spectrograms for a sinusoid with a sudden change in frequency. The spectrogram in the second panel of Figure 8 emphasizes temporal resolution over frequency resolution, which is why the break in the signal is evident, but frequencies on each side of the break are poorly resolved. The spectrogram in the third panel of Figure 8 emphasizes spectral resolution over temporal resolution, which is why we can resolve the signal frequencies but cannot determine when the transition from one frequency to the other occurs.

The trade-off between spectral and temporal resolution forced by the Short-Time Fourier Transform makes it unsuitable for the analysis of signals with multiscale activity. Consider the spike-and-slow wave signal in Figure 9. If we analyze this signal using an STFT biased

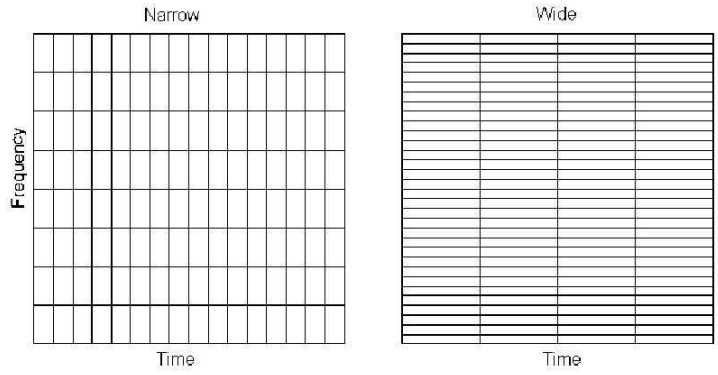


Figure 7: Time-Frequency Plane Division using the Short-Time Fourier Transform: The Short-Time Fourier Transform limits our knowledge of signal activity to a two-dimensional cell in the time-frequency plane. The dimensions of the cell can be changed to favor either spectral or temporal resolution. The same temporal and spectral resolution applies to low and high frequency activity.

towards temporal resolution, the onset of the spike component emerges clearly, but resolution of the frequency of the sinusoidal component is lost. In the spectrogram, the spike component is represented by a periodic columnar band with energy across all frequencies. Each of these columns is a result of shifting the analysis window so that it is centered over a single spike(an impulse-like signal), and then applying the Fourier transform. Recall that the Fourier transform of an impulse (a single spike) has equal energy across all frequencies. In summary, the STFT with high temporal resolution clearly demonstrates the onset of the spike component (short time-scale event), but poorly resolves the sinusoidal component of the signal (long time-scale events).

If we analyze the same signal using an STFT biased towards spectral resolution, the frequency of each sinusoidal component in the signal is resolved clearly. In the spectrogram, the sinusoidal components are represented by rows with energy across time. The darkest row near 2 Hz represents the visible sinusoidal component in the spike-and-slow wave signal. The remaining rows (sinusoidal components) arise from Fourier transformation of the multiple spikes that fall within each long analysis window (a signal resembling a periodic impulse train). Recall that a periodic impulse train is equivalent to an infinite sum of harmonically related sinusoids; each of these sinusoids is represented by a row in the spectrogram. In summary, the STFT with high spectral resolution clearly resolves the spectral components (long time-scale event) present in the signal, but poorly resolves the onset of the spike components (short time-scale event).

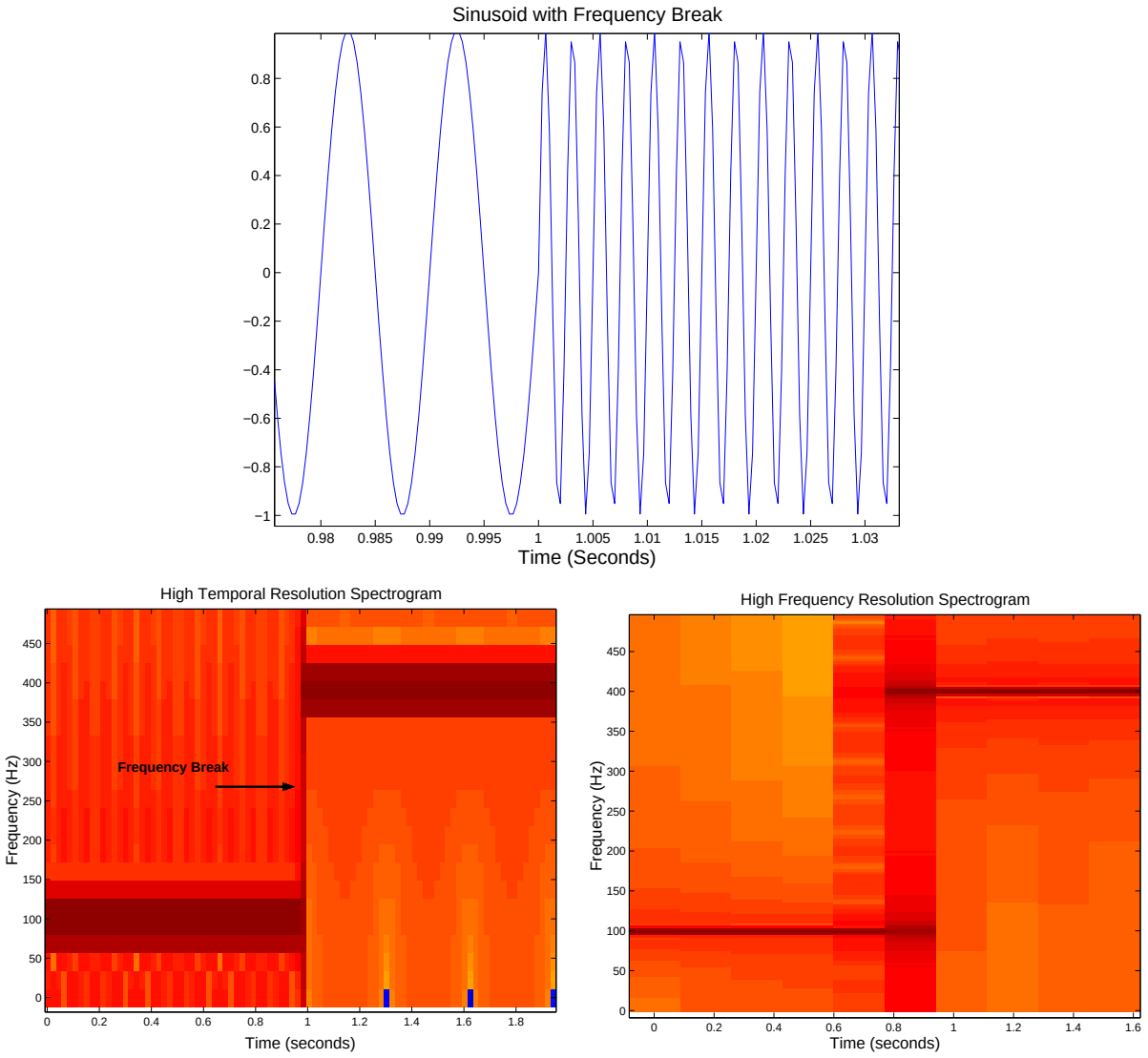


Figure 8: Spectrograms of Sinusoid with Frequency Break: A spectrogram biased towards temporal resolution accurately shows the time at which the sinusoid's frequency changes; it does not clearly show the value of the frequency components on either side of the breakdown. A spectrogram biased towards spectral resolution accurately shows the sinusoid's two frequency components; it does not clearly show the termination of one component and the onset of the second.

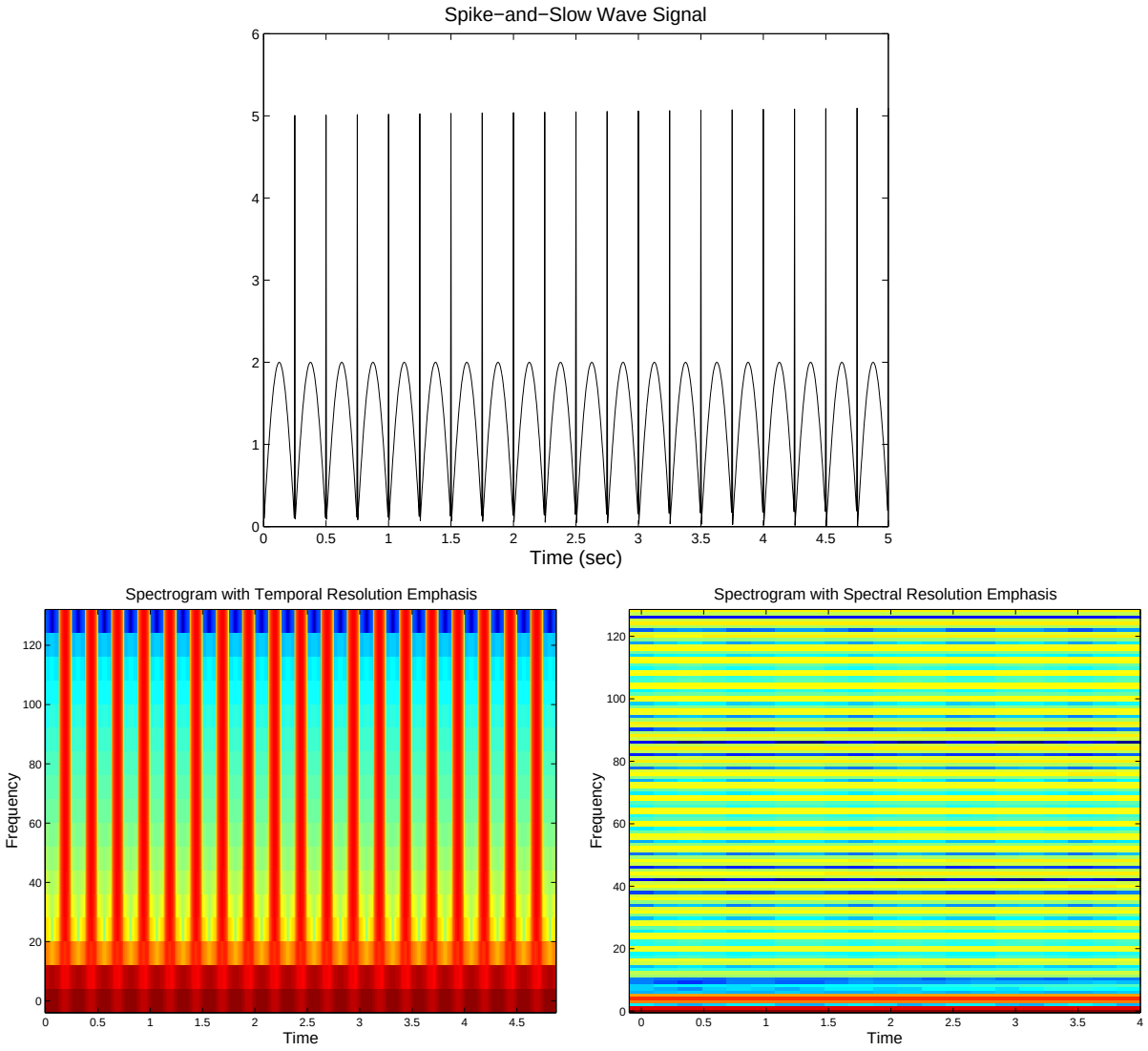


Figure 9: Spectrograms of Spike-And-Slow Wave Signal: A spectrogram biased towards temporal resolution accurately shows the onset of the spike component (short time-scale event); it does not clearly resolve the frequency of the sinusoidal component (long time-scale event). A spectrogram biased towards spectral resolution accurately shows the frequency of the sinusoidal component; it does not clearly show the onset of spike component.

3 The Continuous Wavelet Transform

The continuous wavelet transform (CWT) is a generalization of the Short-Time Fourier Transform that allows for the analysis of non-stationary signals at multiple scales. Similar to the STFT, the CWT makes use of an analysis window to extract signal segments; in this case the window is called a *wavelet*. Unlike the STFT, the analysis window or wavelet is not only translated, but *dilated* and *contracted* depending on the scale of activity under study. Wavelet dilation increases the CWT's sensitivity to long time-scale events, and wavelet contraction increases its sensitivity to short time-scale events.

$$C(a,\tau) = \int \frac{1}{\sqrt{a}} \Psi(\frac{t-\tau}{a}) x(t) dt \tag{5}$$

The mathematical expression for the continuous wavelet transform is shown above. The equation shows that a wavelet $\Psi(t)$ is shifted by τ and dilated or contracted by a factor a prior to computing its *correlation* with the signal $x(t)$. The correlation between the signal and the wavelet is defined as the integral of their product. The CWT maps $x(t)$ into a bivariate function $C(a,\tau)$ that can be used to determine the similarity between $x(t)$ and a wavelet scaled by a at given time τ. The correlation is localized in time, it is computed over an interval beginning at $t = tau$ and ending $t = \tau + L$ where L is the duration of the wavelet. A time plot of the correlation between the signal and the scaled wavelets is called a *Scalogram*. The steps for constructing a scalogram are visualized in figure 10.

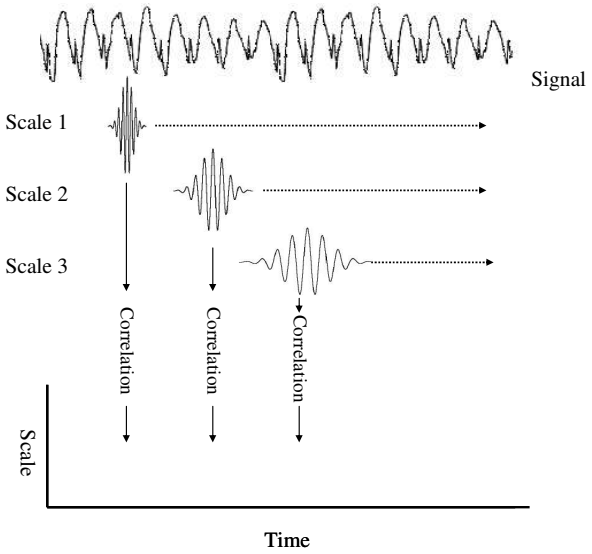


Figure 10: Constructing a Scalogram: A scalogram illustrates how signal activity within a range of time-scales evolves over time. The scalogram is constructed by evaluating the correlation between a signal and wavelets with different scales, and then plotting how the correlation with each wavelet varies over time.

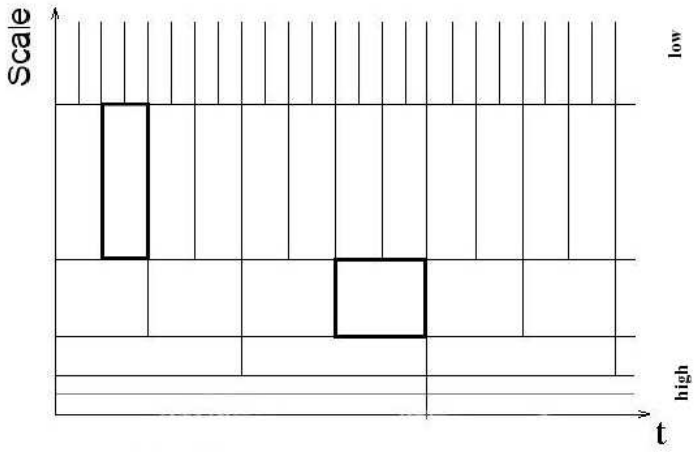


Figure 11: Time-Frequency Plane Tiling of the Continuous Wavelet Transform: The Continuous Wavelet Transform automatically adjusts its time and frequency resolution depending on the scale of activity of interest by dilating or contracting the analysis window. This allows the transform divide the time-frequency plane into regions that highlight either short or long time-scale events.

When the wavelet is contracted ($a < 1$), the wavelet offers high temporal resolution and is well-suited for determining the onset of short-time events such as a spikes and transients. When the wavelet is dilated ($a > 1$) the wavelet offers high spectral resolution and is well-suited for determining the frequency of sustained, long-term events such as baseline oscillations. This time-frequency trade-off is practical since we are often more interested in knowing with accuracy the onset of impulse-like transients as opposed to details of their broad frequency structure. Similarly, knowing the frequency of long-term, sustained activity is often more important than knowledge of the exact onset of the change since it is gradual.

Figure 11 illustrates how wavelet analysis limits our knowledge of signal activity to variable size, two-dimensional cells. For small values of a (upper portion of scale axis) we see that we have high temporal resolution (short dimension along time-axis of the time-frequency cell) and poor frequency resolution (large dimension along scale axis of the time-frequency cell). It is in this portion of the time-frequency plane that spike and transients in the signal are highlighted. For large values of a (lower portion of scale axis) we have high spectral resolution (short dimension along the scale axis of the time frequency cell). It is in this portion of the time-frequency plane that sustained oscillations and other long-term events are highlighted. The ability of the continuous wavelet transform to separate short time-scale events into one portion of the time-frequency plane and long time-scale events into another allows one to simultaneously study signal activity at multiple scale. Recall that this was not possible using the STFT since an implicit selection of a single scale was made through the choice of the the analysis window length.

We will now examine the scalogram of several signals, including a spike-and-slow wave signal

to illustrate the ability of the CWT to simultaneously reveal activity at various scales. Figure 12 shows the scalogram of a stationary sine wave. The scalogram demonstrates that energy (bright columnar bands) predominates in the higher scales; there is a bright band for each peak and trough of the sine wave. The dark bands between the brighter bands are a result of the correlation integral evaluating to a small value due to overlap of the wavelet with both positive and negative portions of the sine wave. The distance between the bright bands can be used to determine the sine wave frequency.

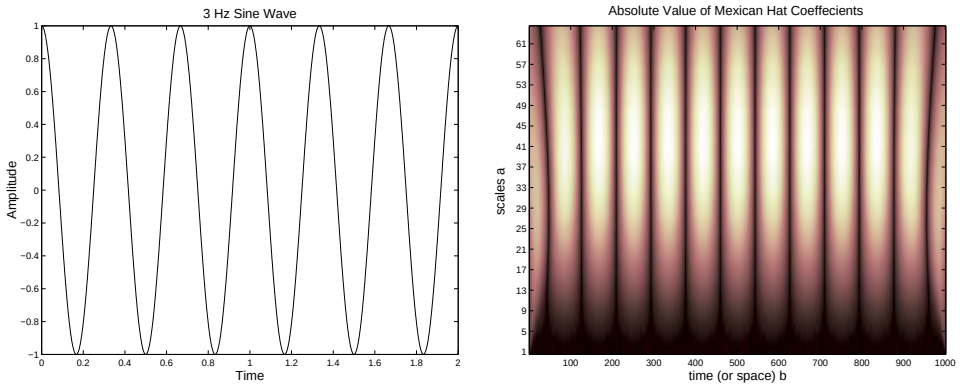


Figure 12: CWT of Sinusoid with Fixed Frequency

Figure 13 shows the scalogram of a sine wave with an abrupt change in frequency. We noted that that the STFT can be biased to emphasize the frequency break using a short window, or the constant oscillations on each side of the break using a long window. The wavelet transform accomplishes both simultaneously. Note that for small scales (small a) the frequency break is very clearly defined in time. Furthermore, the range of available scales allow us to determine with accuracy the frequency of the sinusoids by noting the periodicity of the bright bands.

The scalogram of a sinusoid with linearly increasing frequency is shown in Figure 14. The scalogram highlights the signal's increasing frequency content by the presence of energy in increasingly smaller scales.

Finally we examine the scalogram of the spike-and-slow wave signal. At the lower scales $(1-17)$ the scalogram shows narrow, bright columnar bands that lie in between the thicker columnar bands; these narrow bands represent the spikes in the signal. At the higher scales $(17-81)$ the scalogram shows thick, bright columnar bands which represent the wave component of the signals. The CWT is able to simultaneously capture the multiscale activity within this signal.

The continuous wavelet transform is redundant because it varies the wavelet scaling parameter a continuously. Typically, not much more information is gained by analyzing a signal at $a = 20$ and $a = 20.5$; in practice a discrete set of scales is chosen. The most commonly chosen set of scales is known as the *dyadic scale*, it includes all scales such that $a = 2^i$ for

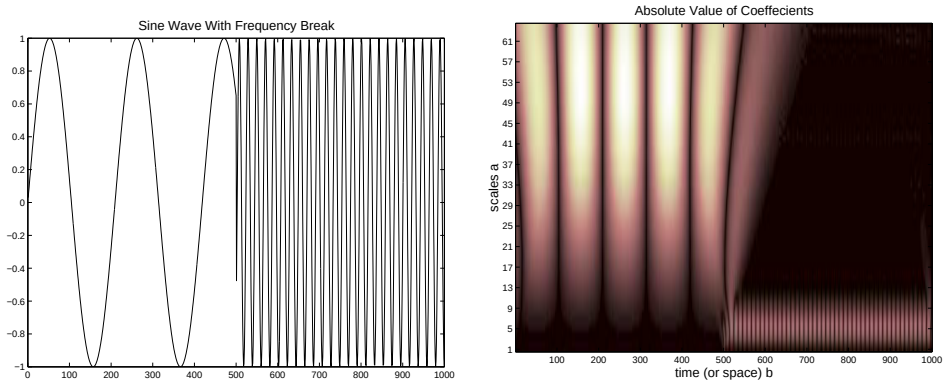


Figure 13: CWT of Sinusoid With Frequency Breakdown

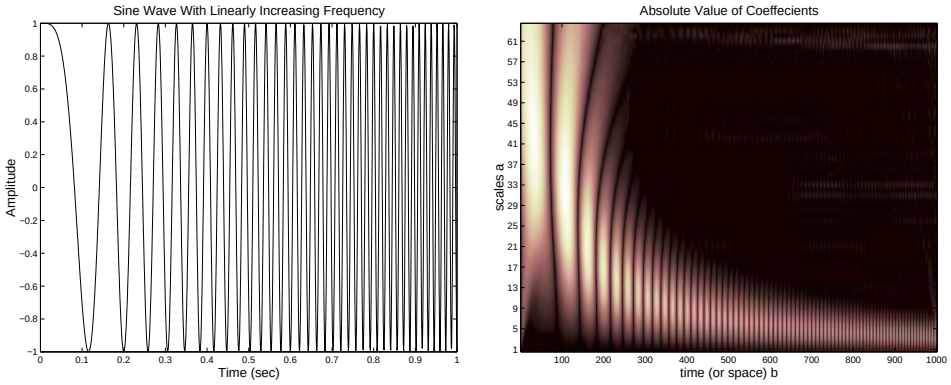


Figure 14: CWT of Sinusoid With Linearly Increasing Frequency

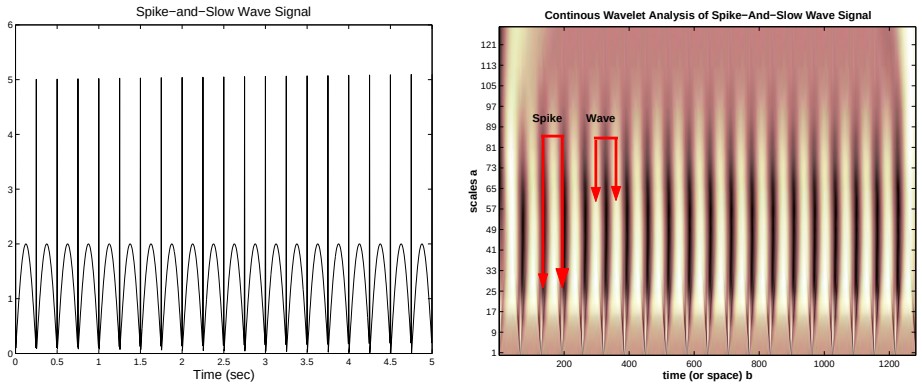


Figure 15: CWT of Spike-And-Slow-Wave Signal

$i = 1, \ldots, N$. There is no loss of information in this process of subsampling the parameter a, the signal can be perfectly reconstructed from knowledge of the continuous wavelet transform over the dyadic scale. Even better is that computing the continuous wavelet transform over the dyadic scales leads to an efficient filterbank implementation of this transform. When the continuous wavelet transform is computed over the the dyadic scale it is more commonly called the Discrete Wavelet Transform (DWT); the word discrete refers to the discrete nature of the scale parameter a.

4 Filterbanks

A filter bank is a collection of filters that decomposes a signal into a set of frequency bands. This decomposition allows one to selectively examine or modify the content of a signal within the chosen bands for the purpose of compression, filtering, or signal classification. The Short-Time Fourier Transform and the Continuous Wavelet Transform can be computed efficiently using filterbanks. Furthermore, the filterbank formulation makes the application of the STFT and CWT in the setting of signal compression, filtering, or classification very natural.

The Short-Time Fourier Transform can be computed using a filterbank known as an *M-channel filterbank*. An M-channel filterbank is shown in Figure 16; it consists of M parallel filters all with equal bandwidths but different center frequencies.

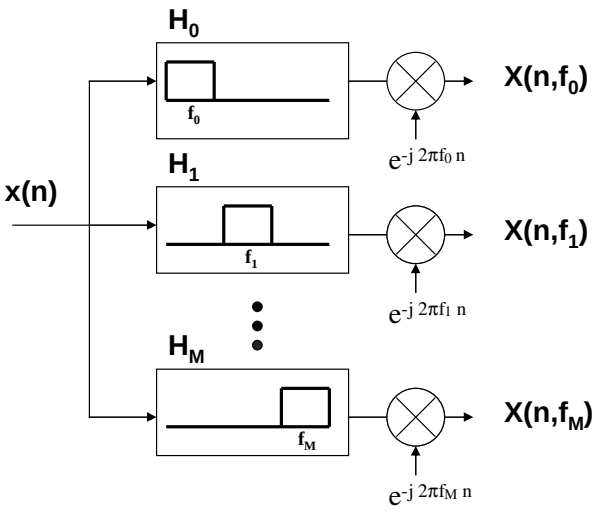


Figure 16: Short-Time Fourier Transform Filter Bank: The Short-Time Fourier Transform at times n and frequencies f_k can be computed using and M-channel filterbank. The filters in this structure have a frequency response consisting of the spectrum of the analysis window $w(n)$ modulated to the frequency of interest f_k.

To see this note the following manipulation of the STFT analysis equation

$$X(n, f_k) = \sum_{m=-\infty}^{m=+\infty} x(m)w(n-m)e^{-j2\pi f_k m}$$

$$X(n, f_k) = e^{-j2\pi f_k n}(\sum_{m=-\infty}^{m=+\infty} x(m)w(n-m)e^{j2\pi f_k (n-m)}) \quad (6)$$

$$X(n, f_k) = e^{-j2\pi f_k n}(x(n) * w(n)e^{j2\pi f_k n})$$

The equation and figure show that the value of the input signal's transform at time n and frequencies in a band centered around f_k is given by filtering $x(n)$ using filters with impulse responses $w(n)e^{j2\pi f_k n}$. These filters have frequency responses with the spectrum of the analysis window $w(n)$ modulated to the center frequency f_k. Typically the frequencies f_k are uniformly sampled over the range $0to1$, so $f_k = k/N$ for $k = 0, \cdots, N-1$.

Perfect reconstruction of the input signal is possible following analysis by an M-channel filterbank. The process involves reconstituting the signal's spectrum by adding the frequency content extracted into channel as shown in Figure 17

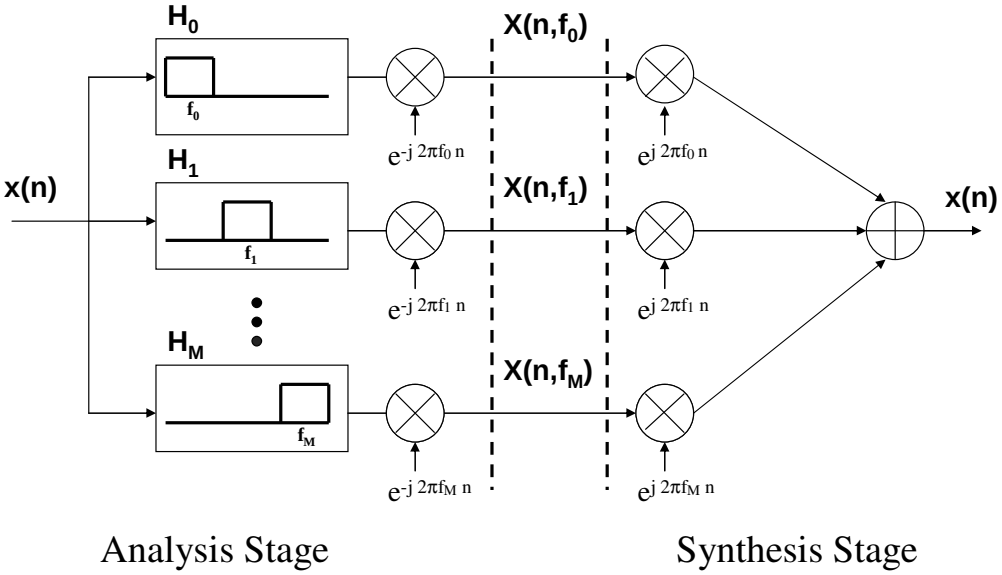


Figure 17: Signal Reconstruction From Short-Time Fourier Transform: The input signal can be reconstructed following analysis by an M-channel filterbank. The process involves addition of the frequency content extracted into each channel by the analysis stage of the filter.

The structure of the M-channel filterbank offers another perspective on the time-frequency plane tiling associated with the Short-time Fourier Transform. For any given time n_0, which corresponds to a column in the time-frequency plane, the impulse response of all channels are of equal length. This implies that the STFT offers the same temporal resolution across all frequencies. Similarly the bandwidth of all the channels are equal, which implies the STFT offers the same spectral resolution across all frequencies. Fixed spectral and temporal resolution across all frequencies leads to the uniform tiling of the time-frequency plane.

The *tree-structured filter bank* in Figure 18, can be used to compute the wavelet coefficients $C(a, \tau)$ of the continuous wavelet transform; however, only over a dyadic scale of dilations and contractions. A tree-structured filterbank splits an incoming signal into a low-pass channel using the filter $H_0(z)$, and a high-pass channel using the filter $H_1(z)$. The low-pass channel can be recursively split N times using the same two filters. Signals extracted from the filterbank at higher iteration levels contain increasingly longer time-scale activity, while those extracted from lower levels contain shorter time-scale activity. The mathematical derivation linking the tree-structured filterbank to the CWT is presented in the appendix. We motivate the connection by showing a qualitative equivalence between their time-frequency decomposition of the signal.

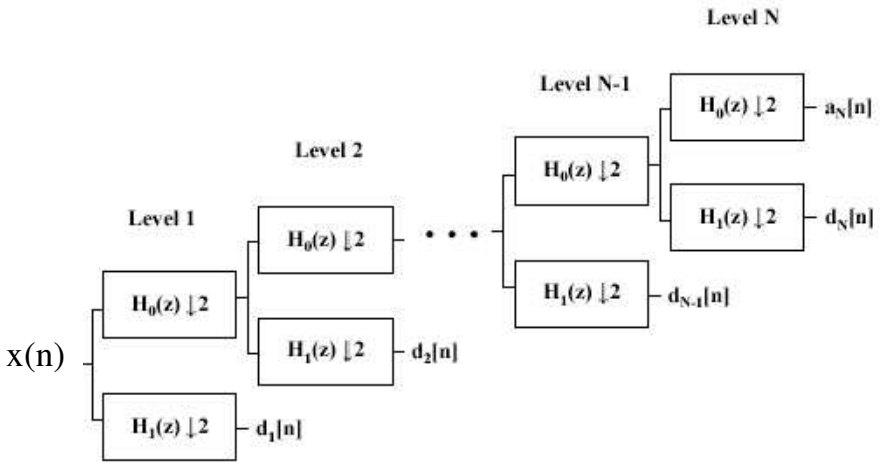


Figure 18: Tree-Structured Filterbank: This filterbank architecture can be used to compute the coefficients of the continuous wavelet transform. The coefficients are only computed over the dyadic scale.

Consider deriving the input to output transfer functions of the 3-level filterbank shown in Figure 19; to do this we make use of the *Noble Identity*:$\downarrow 2\, H(z) = H(z^2) \downarrow 2$. Outputs extracted from the third-level (large scale a) are generated using a transfer function with a narrow bandwidth, which implies good frequency resolution but poor temporal resolution. At the other extreme, the output extracted form the first-level (small scale a) is generated using a transfer function that has a large bandwidth, which suggests poor frequency resolution but good temporal resolution. A decrease in frequency resolution and an increase in temporal resolution as the scale a decreases (activity in signal is at higher frequency, or shorter time-scale) is similar to the time-frequency trade-off offered by the CWT. Note that the bandwidth of the transfer function increases by a factor of two for each level of the filterbank. This is a consequence of each level of the filterbank extracting wavelet coefficients that are greater than those of the previous level by a factor of two.

Using only a dyadic scale of wavelet coefficients one can perfectly reconstruct the input signal; this possibility highlights the redundancy of varying the scale parameter a continuously in the

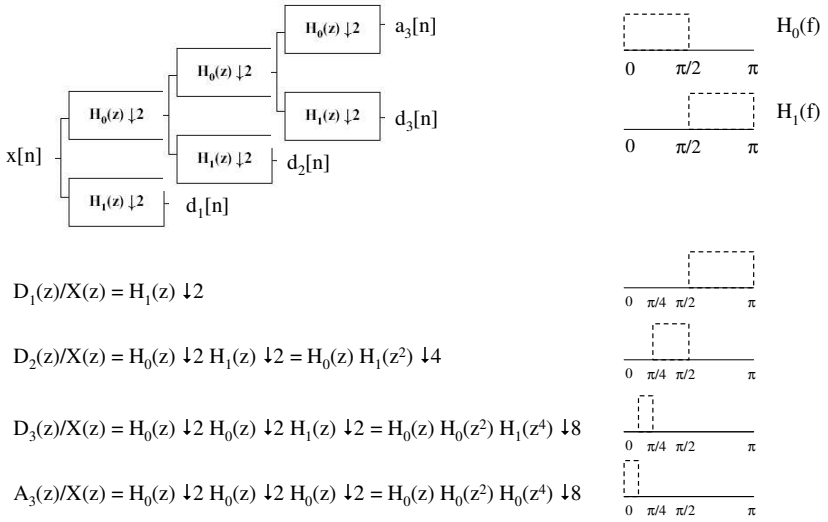


Figure 19: Tree-Structured Filterbank: The filterbank uses high-frequency resolution and poor temporal resolution to extract long time-scale activity. On the other hand, the filterbank uses poor-frequency resolution and high temporal resolution to extract short time-scale activity. This time-frequency trade-off mirrors that offered by the Continuous Wavelet Transform.

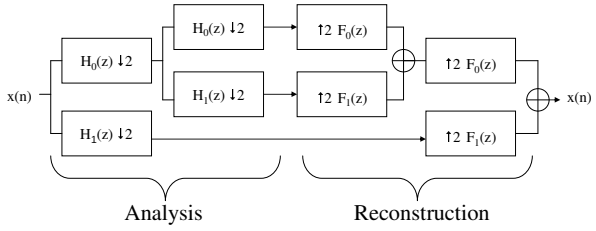


Figure 20: Iterated Filterbank Inverse Filter: The filter that reconstructs the input signal from the outputs of an iterated filterbank is the mirror image of the analysis filter. The reconstruction filter uses upsampling as opposed to the the downsampling found in the analysis filter.

Continuous Wavelet Transform. The reconstruction, or synthesis filterbank is a mirror image of the analysis filterbank as shown in Figure 20. The analysis filters $H_0(z)$ and $H_1(z)$ and the reconstruction filters $F_0(z)$ and $F_1(z)$ must be carefully chosen such that the decomposed signal can be perfectly reconstructed. The analysis and reconstruction filters have to satisfy an *anti-alias* and *zero-distortion* conditions that we will derive below.

Consider the two-channel iterated filterbank in Figure 21. We will trace the signal through the filterbank channels, and then note which conditions on the analysis and reconstruction filters guarantee perfect reconstruction. Since the analysis and reconstruction filters are causal, we expect the output to be a delayed version of the input in the case of perfect reconstruction. First we derive an expression for the intermediate signals w_0 and w_1.

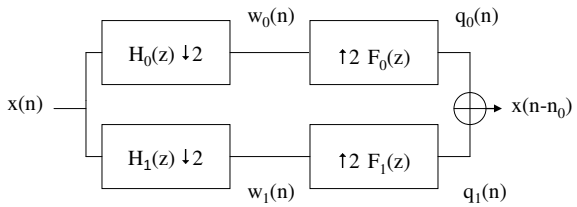


Figure 21: Two Channel Analysis and Synthesis Filterbanks.

$$(\downarrow M)H(e^{jw}) = \frac{1}{M}\sum_{i=0}^{M-1} H(e^{j(w/M-2\pi i/M)})$$

$$\begin{aligned} W_0(e^{jw}) &= \tfrac{1}{2}(X(e^{jw/2})H_0(e^{jw/2}) + X(e^{j(w/2-\pi)})H_0(e^{j(w/2-\pi)})) \\ W_1(e^{jw}) &= \tfrac{1}{2}(X(e^{jw/2})H_1(j^{jw/2}) + X(e^{j(w/2-\pi)}H_1(e^{j(w/2-\pi)}))) \end{aligned} \tag{7}$$

The terms containing $X(e^{j(w/2-\pi)}$ are aliased copies of the original signal spectrum; they are introduced by downsampling the output of each filterbank channel. To achieve perfect reconstruction these terms will have to be eliminated by the synthesis filter. The terms containing $X(e^{jw/2})$ are distorted versions of the original spectrum; distortion is introduced by frequency-axis scaling (downsampling) and magnitude scaling ($H_0(z)$ and $H_1(z)$). To achieve perfect reconstruction the synthesis filterbank must correct for this distortion by upsampling and filtering ($F_0(z)$ and $F_1(z)$). Now we derive an expression for the intermediate signals q_0 and q_1.

$$(\uparrow M)H(e^{jw}) = H(e^{jwM})$$

$$\begin{aligned} Q_0(e^{jw}) &= W_0(e^{j2w})F_0(e^{jw}) = \tfrac{1}{2}(X(e^{jw})H_0(e^{jw})F_0(e^{jw}) + X(e^{j(w-\pi)})H_0(e^{j(w-\pi)})F_0(e^{jw})) \\ Q_1(e^{jw}) &= W_1(e^{j2w})F_1(e^{jw}) = \tfrac{1}{2}(X(e^{jw})H_1(e^{jw})F_1(e^{jw}) + X(e^{j(w-\pi)})H_1(e^{j(w-\pi)})F_1(e^{jw})) \end{aligned} \tag{8}$$

For perfect reconstruction we would like the reconstructed signal to be a delayed version of the original.

$$\begin{gathered} q_0(n) + q_1(n) = x(n - n_0) \\ Q_0(e^{jw}) + Q_1(e^{jw}) = e^{-jwn_0} X(e^{jw}) \end{gathered} \tag{9}$$

Collect the terms with $X(e^{jw})$ into one set and the terms with $X(e^{j(w-\pi)})$ into another set. The collection of terms with $X(e^{jw})$ are the only ones that can give rise to an equality with the term $e^{-jwn_0}X(e^{jw})$. The collection of terms with $X(e^{j(w-\pi)})$, which are aliased copies of the original spectra, must be annihilated. When these two conditions are met by the four filters $H_0(z)\, H_1(z)\, F_0(z)\, F_1(z)$ perfect reconstruction is achieved.

$$\begin{gathered} \text{Zero-Distortion Condition:} \\ \tfrac{1}{2} X(e^{jw})(H_0(e^{jw})F_0(e^{jw}) + H_1(e^{jw})F_1(e^{jw}) = e^{-jwn_0} X(e^{jw}) \\ H_0(e^{jw})F_0(e^{jw}) + H_1(e^{jw})F_1(e^{jw} = 2e^{-jwn_0} \\ \text{Anti-Alias Condition:} \\ X(e^{j(w-\pi)})(H_0(e^{j(w-\pi)})F_0(e^{jw}) + H_1(e^{j(w-\pi)})F_1(e^{jw})) = 0 \\ H_0(e^{j(w-\pi)})F_0(e^{jw}) + H_1(e^{j(w-\pi)})F_1(e^{jw}) = 0 \end{gathered} \tag{10}$$

The discrete-time analysis and synthesis filters that satisfy the above conditions are not equivalent to the wavelet functions $\Psi(t)$ used in the Continuous Wavelet Transform; they are derived from the wavelet function. The link is established in the appendix. In summary, performing wavelet analysis involves:

1. Select a wavelet appropriate for analyzing the signal of interest. The wavelet should have morphological features that match those to be extracted, highlighted, or detected in the input signal.

2. Derive the filters $H_0(z)\, H_1(z)$ so that an efficient filterbank implementation can be used to compute the wavelet coefficients.

3. Derive the filters $F_0(z)\, F_1(z)$ so that an efficient inverse filterbank can be used to reconstruct a new version of the signal from the modified wavelet coefficients.

4. Fortunately, the filters $H_0(z)\, H_1(z)\, F_0(z)\, F_1$ have already been computed for a large number of wavelet functions. The filters can be immediately used to study signals of interest.

If all the wavelet coefficients produced by the analysis filterbank are preserved and the signal is reconstructed, the synthesized signal will exactly equal the input signal. If some coefficients are selectively preserved, then we are effectively *filtering* in the *scale-domain* as opposed to the conventional *frequency-domain*.

5 Biomedical signal processing examples

Transient (nonstationary) changes in a signal may be isolated by combining the time-domain and frequency-domain analysis of a signal. In this way we can take advantage of both these paradigms and allow filtering of both persistent signal sources within the observation, and short transient sources of noise. Joint time-frequency analysis (JTFA) is then essentially a transformation of an N-point M-dimensional signal (usually where $M = 1$ for the ECG) into a $M + 1$-dimensional signal.

Consider the noisy sine wave below and its 5-level decomposition. If we zero out the details d_1 d_2 d_3 d_4, and save the approximation a_5 we will be eliminating the coefficients that are most sensitive to the noise. If we reconstruct only using the coefficients a_5, we obtain the result in figure 22.

Now consider using a wavelet decomposition to separate components in a signal occurring over different time scales. For example, suppose we would like to separate the spike and wave components of the signal shown below. The wave component of the signal is primarily contained in the approximation coefficients $a5$, and the spike is mostly contained in the detail coefficients d_{1-4}. If we reconstruct the signal only using the approximation coefficients we will recover the wave component. If we reconstruct the signal only using the detail coefficients we will recover the spike component. (See figure 23.)

Another example of visualizing a nonstationary signal is given in Fig. 24. Here we can see one beat from a normal an ECG (upper plot) and the corresponding scalogram (lower plot) produced by the CWT of this segment. Note that the lighter regions of the scalogram which correspond to the higher energy regions such as the QRS complex and the T-wave, and are more defined at shorter scales.

In practice, the CWT provides a vast amount of redundancy in the representation (with more than an order of magnitude more wavelet values than original signal components) and therefore effects a decompression rather than a data reduction. In order to extract information from a wavelet decomposition and remove much of the redundancy, we can consider only the local maxima and minima of the transform. These include wavelet *ridges* and the wavelet *modulus maxima*. Wavelet ridges are used to determine instantaneous frequencies and amplitudes and are define by

$$\frac{dS(a,\tau)}{da} = 0 \tag{11}$$

where $S(a,\tau) = |C(a,\tau)|^2/a$ is the rescaled scalogram. The wavelet modulus maxima are

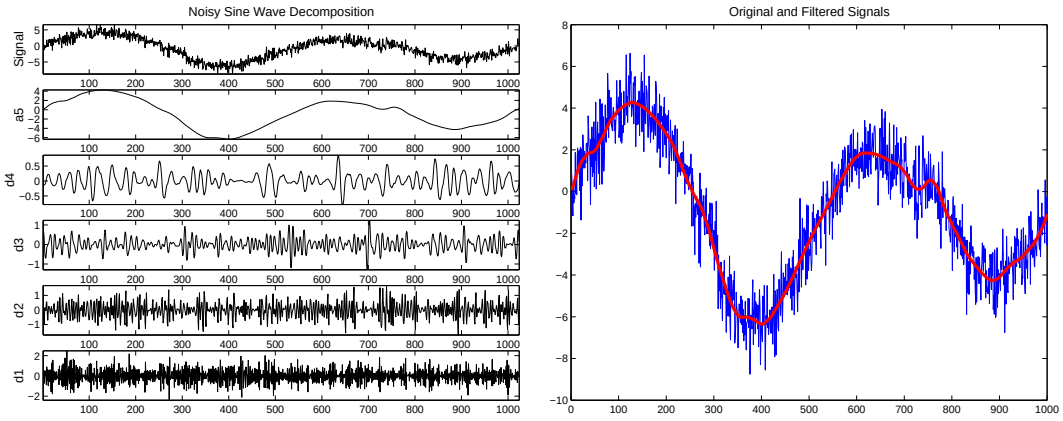


Figure 22: Wavelet Noise Suppression: Filtering using the wavelet transform involves zeroing coefficients at scales sensitive to the noise, and then reconstructing a signal from only using coefficients insensitive to the noise.

used for locating and characterizing singularities in the signal and are given by

$$\frac{dS(a,\tau)}{d\tau} = 0. \tag{12}$$

Another effective way to produce a data reduction is through the Discrete Wavelet Transform (DWT).

5.0.1 An ECG denoising example; wavelet choice

Fig. 25 illustrates a selection of biorthogonal wavelets denoted *biorJ.K*, where J and K refer to the number of vanishing moments in the LP and HP filters respectively. Note that in most literature, J refers to the length of the lowpass filter for and K to the length of the highpass filter, therefore MATLAB'S *bior4.4* has 4 vanishing moments[1], with 9 LP and 7 HP coefficients (or 'taps') in each of the filters.

Fig. 26 illustrates the effect of using different mother wavelets to filter a section of clean ('zero-noise') ECG, using only the first approximation of each wavelet decomposition. The clean (upper) ECG is created by averaging 1228 R-peak aligned, 1s long segments of a healthy ECG. Gaussian pink noise is then added with an SNR of 20dB. The RMS error

[1]If the Fourier transform of the wavelet is J continuously differentiable, then the wavelet has J vanishing moments. Type $waveinfo('bior')$ at theMATLAB prompt for more information. Viewing the filters using $[lp_{decon}, hp_{decon}, lp_{recon}, hp_{recon}] = wfilters('bior4.4')$ in MATLAB reveals one zero coefficient in each of the LP decomposition and HP reconstruction filters, and three zeros in the LP reconstruction and HP decomposition filters. Note that these zeros are simply padded, and do not count when calculating the filter size.

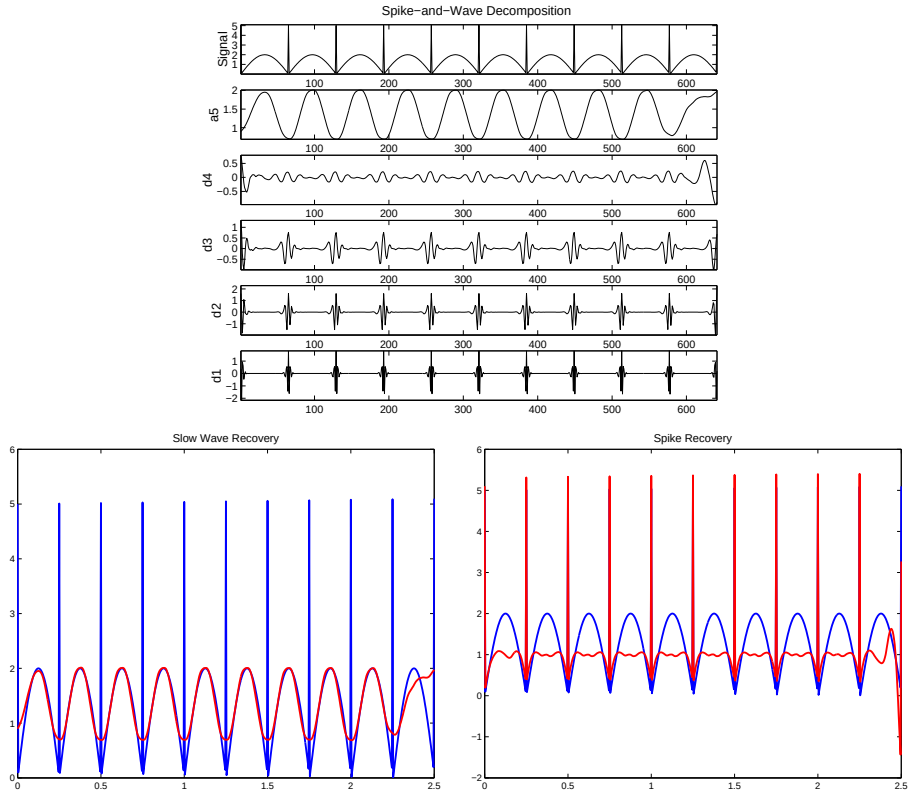


Figure 23: Wavelet Separation of Activity at Different Scales: Wavelets are effective tools for separating activity at different time scales. To recover short time-scale activity a signal is reconstructed only using wavelet coefficients sensitive to the short time-scale activity. To recover long time-scale activity a signal is reconstructed only using coefficients sensitive to long time-scale activity. The definition of short and long time-scale activity is application and signal dependent.

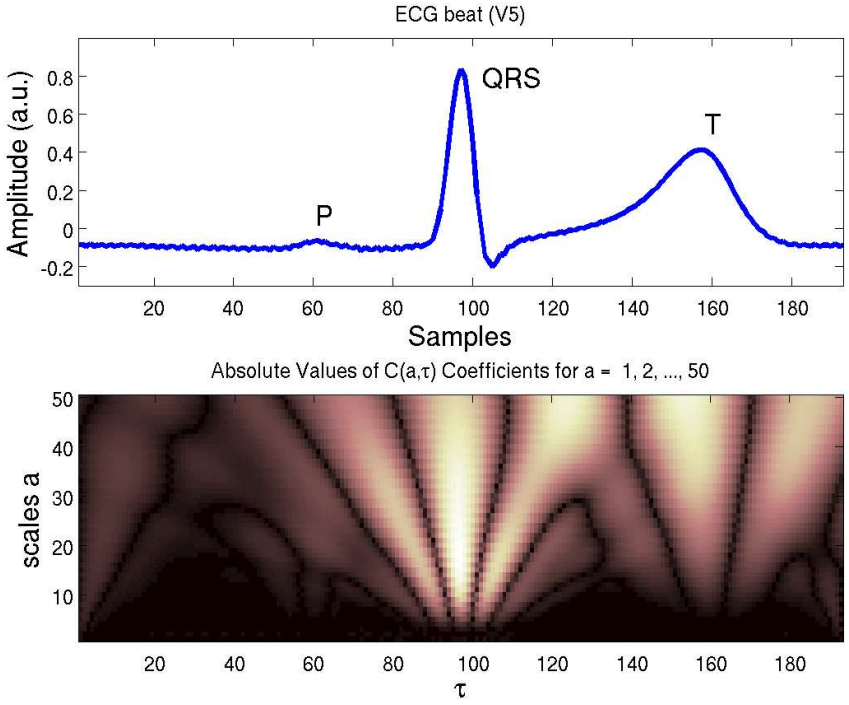


Figure 24: A relatively clean 0.75s segment of lead V5 ECG recorded at 256 Hz and its corresponding scalogram form the CWT for scales $0 < a \leq 50$.

Wavelet Family	Family member	RMS error
Original ECG	N/A	0
ECG with pink noise	N/A	0.3190
Biorthogonal *'bior'*	*bior3.3*	0.0296
Discrete Meyer *'dmey'*	*dmey*	0.0296
Coiflets *'coif'*	*coif2*	0.0297
Symlets *'sym'*	*sym3*	0.0312
Symlets *'sym'*	*sym2*	0.0312
Daubechies *'db'*	*db2*	0.0312
Reverse biorthogonal *'rbio'*	*rbio3.3*	0.0322
Reverse biorthogonal *'rbio'*	*rbio2.2*	0.0356
Haar *'haar'*	*harr*	0.0462
Biorthogonal *'bior'*	*bior1.3*	0.0472

Table 1: Signals displayed in Fig. 26 (from top to bottom) with RMS error between clean and wavelet filtered ECG with 20dB additive Gaussian pink noise. N/A indicates 'not applicable'.

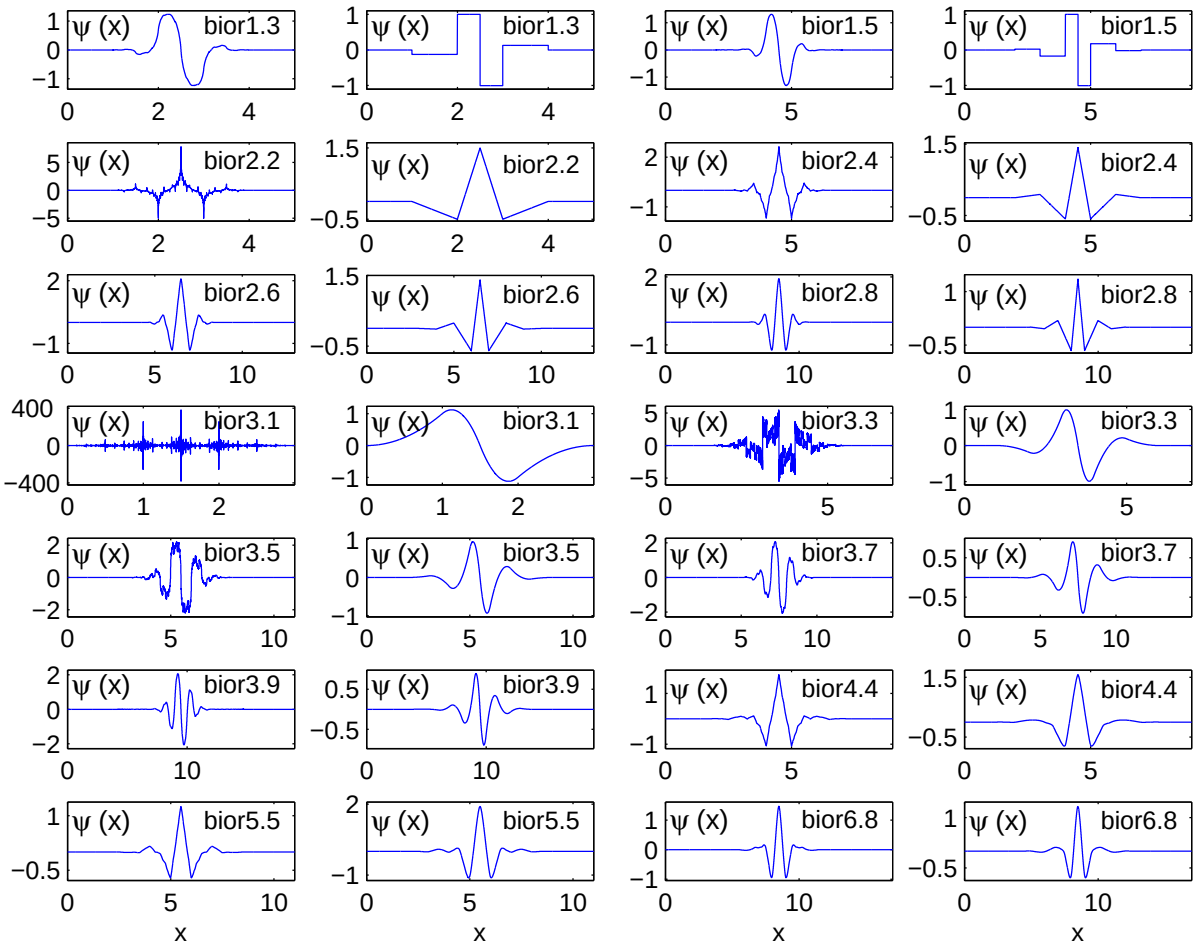


Figure 25: Biorthogonal Wavelets labeled by their MATLAB nomenclature. for each filter, two wavelets are shown; one for signal decomposition (on the left side) and one for signal reconstruction (on the right side). Type *waveinfo('bior')* in MATLAB for more information. Note how increasing the order of the filter leads to increasing similarity between the mother wavelet and typical ECG morphologies.

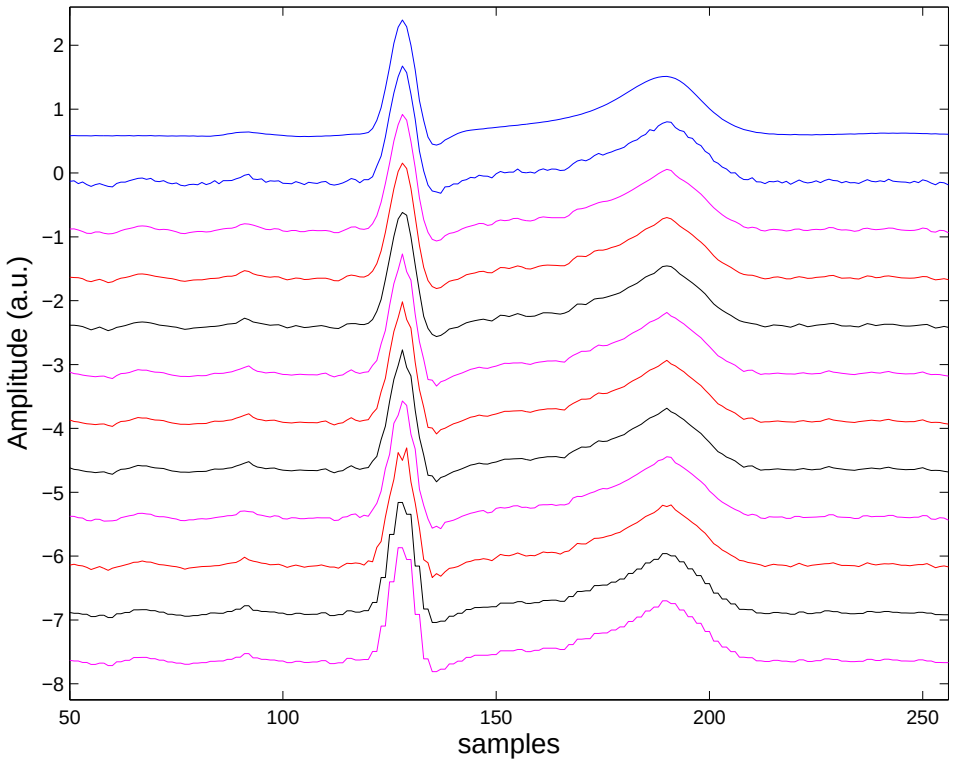


Figure 26: The effect of a selection of different wavelets for filtering a section of ECG (using the first approximation only) contaminated by Gaussian pink noise (SNR=20dB). From top to bottom; original (clean) ECG, noisy ECG, biorthogonal (8,4) filtered, discrete Meyer filtered, Coiflet filtered, symlet (6,6) filtered, symlet filtered (4,4), Daubechies (4,4) filtered, reverse biorthogonal (3,5), reverse biorthogonal (4,8), Haar filtered and finally, biorthogonal (6,2) filtered. The 'zero-noise' clean ECG is created by averaging 1228 R-peak aligned, 1s long segments of a healthy ECG. RMS error performance of each filter is listed in table 1.

between the filtered waveform and the original clean ECG for each wavelet is given in table 1. Note that the biorthogonal wavelets with $J,K \geq 8,4$, the discrete Meyer wavelet and the Coiflets appear to produce the best filtering performance in this circumstance. The RMS results agree with visual inspection, where significant morphological distortions can be seen for the other filtered signals. In general, increasing the number of taps in the filter produces a lower error filter.

The wavelet transform can be considered either as a spectral filtering over many time scales or viewed as a linear time filter $\Psi[(t-\tau)/a]$ centered at a time τ with scale a that is convolved with the time series, $x(t)$. Therefore convolving the filters with a shape more commensurate with that of the ECG produces a better filter. Fig. 25 illustrates this point. Note that as we increase the number of taps in the filter, the mother wavelet begins to resemble the ECG's P-QRS-T morphology more closely.

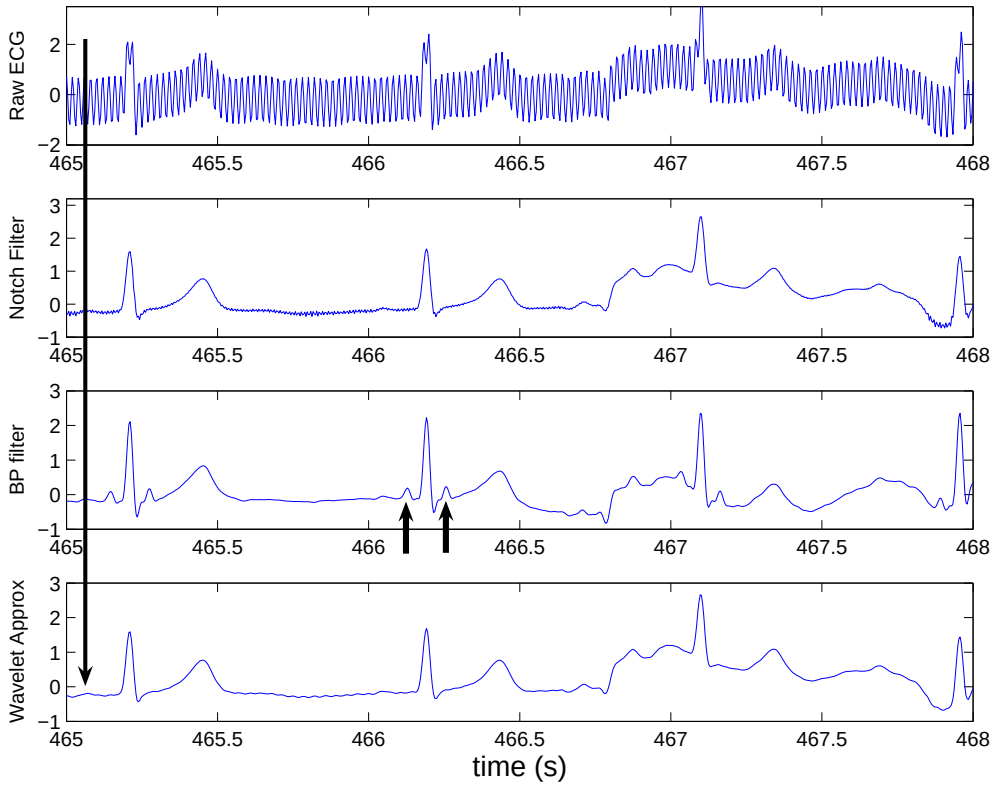


Figure 27: Raw ECG with 50 Hz mains noise, IIR 50 Hz notch filtered ECG, 0.1-45 Hz FIR band-pass filtered ECG and *bior3.3* wavelet filtered ECG. The left-most arrow indicates the low amplitude P-wave. Central arrows indicate Gibbs oscillations in the FIR filter causing a distortion larger then the P-wave.

The biorthogonal wavelet family are FIR filters and therefore possess a linear phase response, which is an important characteristic for signal and image reconstruction. In general, biorthogonal spline wavelets allow exact reconstruction of the decomposed signal. This is not possible using orthogonal wavelets (except for the Haar wavelet). Therefore, *bior*3.3 is a good choice for a general ECG filter. It should be noted that the filtering performance of each wavelet will be different for different types of noise, and an adaptive wavelet-switching procedure may be appropriate. As with all filters, the wavelet performance may also be application specific, and a sensitivity analysis on the ECG feature of interest is appropriate (e.g. QT-interval or ST-level) before selecting a particular wavelet.

As a practical example of comparing different common filtering types to the ECG, observe Fig. 27. The upper trace illustrates an unfiltered recording of a V5 ECG lead from a healthy adult in his 30s undergoing an exercise test. Note the high amplitude 50 Hz (mains) noise[2]. A 3-tap IIR 50 Hz notch-filter is then applied to reveal the underlying ECG. Note some baseline wander disturbance from electrode motion around t=467s, and the difficult in discerning the P-wave (indicated by a large arrow at the far left). The third trace is a band-pass (0.1-45 Hz) FIR filtered version of the upper trace. Note the baseline wander is reduced significantly, but a Gibbs[3] ringing phenomena is introduced into the Q- and S-waves (illustrated by the small arrows), which manifests as distortions with an amplitude larger than the P-wave itself. A good demonstration of the Gibbs phenomenon can be found at [6] and [11]. This *ringing* can lead to significant problems for a QRS detector (looking for Q-wave onset) or any technique for analyzing at QT intervals or ST changes. The lower trace is the first approximation of a biorthogonal wavelet decomposition (*bior3.3*) of the notch-filtered ECG. Note that the P-wave is now discernible from the background noise and the Gibbs oscillations are not present.

6 Postscript

The wavelet transform (WT) is a popular technique for performing joint time-frequency analysis (JTFA) and belongs to a family of JTFA techniques that include the STFT, the Wigner Ville transform (WVT), the Zhao-Atlas-Marks distribution and the Hilbert-Huang transform[4]. Unfortunately, all but the WT suffer from significant cross-terms which reduce their ability to locate events in the time-frequency plane. Reduced Interference Distribution (RID) techniques such as the exponential or Choi-Williams distribution, the (pseudo) WVT, and the Margenau-Hill distribution, have been developed to suppress the cross terms to some extent, but in general, they do not provide the same degree of (time or frequency) resolution as the WT [13]. Furthermore, the WT, unlike other fixed resolution JTFA techniques allows

[2]60 Hz mains noise is encountered in North and Central America, Western Japan, South Korea, Taiwan, Liberia, Saudi Arabia, and parts of the Carribean, South America and some South Pacific Islands

[3]The existence of the ripples with amplitudes independent of the filter length. Increasing the filter length narrows the transition width but does not affect the ripple. One technique to reduce the ripples is to multiply the impulse response of an ideal filter by a tapered window.

[4]All the JTFA techniques have been unified by Cohen [4]

a variable resolution and facilitates better time resolution of high frequencies and better frequency resolution of lower frequencies. It should be noted that although wavelet analysis has often been quoted as the panacea for analyzing nonstationary signals (and thereby overcoming the problem of the Fourier transform, which assumes stationarity), it is sometimes important to segment data at non-stationarities because model assumptions may no longer hold. Of course, JTFA may help with this segmentation.

The number of articles concerning wavelets applied to biomedical signals, and the ECG in particular, is enormous and an excellent overview of many of the key publications concerning multiscale ECG analysis can be found in Addison [1]. Chapter 4 in Akay *et al.* [5] on late potentials and relevant publications by Pablo Laguna [10, 9]. It should be noted however, that there has been much discussion of the use of wavelets in heart rate variability (HRV) analysis since long range beat-to-beat fluctuations are obviously non-stationary. Unfortunately, very little attention has been paid to the unevenly-sampled nature of the RR-interval time series and this can lead to serious errors (see Chapter 3 in [3]). Techniques for wavelet analysis of unevenly sampled data do exist [2, 7], but it is not clear how a discrete filter bank formulation with up-down sampling could avoid the inherent problems of resampling an unevenly sampled signal.

It should also be noted that wavelet filtering is a lossless supervised filtering method where the basis functions are chosen *a priori*, much like the case of a Fourier-based filter (although some of the wavelets do not have orthogonal basis functions). Unfortunately, because the CWT and DWT are signal separation methods that effectively occur in the frequency domain[5], it is difficult to remove *in-band* noise (biomedical signals and associated noises often have a significant overlap in the frequency domain). In later chapters we will look at techniques which *discover* the basis functions in the data, based either on the statistics of the signal's distributions, or with reference to a known signal model. The basis functions may overlap in the frequency domain and therefore we may separate out in-band noise.

7 Appendix

The appendix will contain the following derivations

1. Derivation linking Continuous Wavelet Transform to iterated filterbank
2. Derivation linking wavelet function to the iterated filterbank filters $H_0(z)\, H_1(z)\, F_0(z)\, F_1(z)$

References

[1] Paul S Addison. Wavelet transforms and the ECG: a review. *Physiological Measurement*, 26(5):R155–R199, 2005.

[5] the wavelet is convolved with the signal

[2] A. Antoniadis and J. Fan. interaction. *Journal of the American Statistical Association*, 96(455):939–967, 2001.

[3] G. D. Clifford, F. Azuaje, and P. E. McSharry. *Advanced Methods and Tools for ECG Analysis*. Artech House, Norwood, MA, USA, October 2006.

[4] R. Cohen, A. and.Ryan. *Wavelets and Multiscale Signal Processing*. Chapman and Hall, London, 1995.

[5] H. Dickhaus and H. Heinrich. *Analysis of ECG Late Potentials Using Time-Frequency Methods; in Time Frequency and Wavelets in Biomedical Signal Processing*, chapter 4. Wiley-IEEE Press, 1997.

[6] P. Grinfeld. The Gibbs Pheonomenon. http://www.math.drexel.edu/˜pg/fb//java/la_applets/Gibbs/index.html.

[7] Godfrey K.R. Chappell M.J. Cayton R.M. Hilton M.F., Bates R.A. Evaluation of frequency and time-frequency spectral analysis of heart rate variability as a diagnostic marker of the sleep apnoea syndrome. *Med Biol Eng Comput.*, 37(6):760–769, Nov 1999.

[8] J.Gotman and P.Gloor. Automatic recognition and quantification of interictal epileptic activity in the human scalp eeg. *Electroencephalography and Clinical Neurophysiology*, 41:513–529, 1976.

[9] P. Laguna. Home page. http://diec.unizar.es/˜laguna.

[10] J. P. Martínez, R. Almeida, S. Olmos, A. P. Rocha, and P. Laguna. A wavelet-based ECG delineator: Evaluation on standard database. *IEEE Transactions on Biomedical Engineering*, 51(4):570–58, 2004.

[11] MIT. Design of FIR Filters by Windowing. http://web.mit.edu/6.555/www/fir.html.

[12] C.-K. Peng, Shlomo Havlin, H. Eugene Stanley, and Ary L. Goldberger. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 5(1):82–87, 1995.

[13] W. Williams. *Recent Advances in Time-Frequency Representations: Some Theoretical Foundation; in Time Frequency and Wavelets in Biomedical Signal Processing*, chapter 1. Wiley-IEEE Press, 1997.