1 Analytic wavelets and sound processing

A complex signal

$$f(t) = u(t) + iv(t)$$

can alternatively be written in polar form as

$$f(t) = A(t)\, e^{i\phi(t)} \tag{1}$$

with amplitude $A(t) \geq 0$ and phase angle $\phi(t)$. If we require that, e.g. $0 \leq \phi(t) < 2\pi$ then the representation (1) is unique. The derivative ϕ' of ϕ is called the *instantaneous frequency*. The representation (1) reflects our interest in the amplitude and phase angle (or instantaneous frequencies of the signal). Unfortunately, the lack of localization of wavelets in both time and frequency as was illustrated by the Heisenberg uncertainty principle makes it difficult to use wavelet analysis in both time and frequency domains. On the other hand, one important application of signal processing is sound processing where amplitudes and frequencies are main issues. In this section we study how approximately analytic wavelets are used to recover amplitudes and frequencies from the transformed signal.

Sound signals are superpositions of sinusoidal waves. A generalized sinusoidal is one of the form

$$f(t) = A(t) \cos \phi(t).$$

where $A(t) \geq 0$ is the amplitude and $\phi(t)$ is the phase. $f(t)$ is the real part of $f_{\mathsf{a}}(t) = A(t)\, e^{i\phi(t)}$. If f_{a} is analytic we call $A(t)$ the analytic amplitude. For example, the signal $f_{\mathsf{a}}(t) = A(t)\, e^{-2\pi i(\omega_0 t + \phi_0)}$ with $\operatorname{supp} \widehat{A} \subset (-\omega_0, \omega_0)$ is analytic since $\widehat{f_{\mathsf{a}}}(\omega) = e^{-2\pi i \phi_0} \widehat{A}(\omega - \omega_0)$. In this case $f(t) = A(t) \cos 2\pi i (\omega_0 t + \phi_0)$.

Speech signals are modeled as

$$f(t) = \sum_{k=1}^{K} A_k(t) \cos \phi_k(t),$$

which is a useful model for pattern recognition and sound processing. Such signals are transmitted by one of two methods: amplitude modulation (changing $A_k(t)$) or frequency modulation (changing $\phi_k'(t)$). Correspondingly there are two methods for sound compression: by sampling without changing ϕ_k' or with changing ϕ_k'.

In the former case, f is synthesized from

$$g(t) = \sum_{k=1}^{K} A_k(\alpha t)\cos\left(\frac{1}{\alpha}\phi_k(\alpha t)\right).$$

At any time t_0, $f(\alpha t_0) = g(t_0)$. If $\alpha > 1$, g is shorter than f but both are perceived as having the same frequency contents. Observe that the instantaneous frequencies for g are $\left(\frac{1}{\alpha}\phi_k(\alpha t)\right)' = \phi_k'(\alpha t)$ which justifies the previous statement.

In the latter case f is synthesized from

$$g(t) = \sum_{k=1}^{K} B_k(t)\cos(\phi_k(\alpha t)),$$

where

$$A_k(t) = F(t, \phi_k'(t)), \quad B_k(t) = F(t, \alpha\phi_k'(t))$$

and $F(t, \omega)$ is a smooth frequency envelop (called the formant in speech processing).

1.1 The ridge algorithm

For the purpose of sound processing, an approximately analytic wavelet is chosen. This is achieved with a symmetric wavelet ψ with compact support in time. The frequency support of the wavelet is then infinite. Since $\widehat{\psi}(0)$ is not exactly zero, the admissibility condition is violated and this means that the reconstruction of the signal is ill posed. In this subsection we discuss an approximate method of reconstruction from the scalogram $P_W f(a,b) = |W_\psi f(a,b)|^2$. The points (a,b) of local maxima of $P_W f(a,b)$ are known as *wavelet ridges*. The ridge algorithm computes the amplitudes and instantaneous frequencies from the wavelet ridges.

Exercise 1 Assume that g is a positive even and continuous function with support in the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$. Show that $\widehat{g}$ is even, real and $\widehat{g}(0) \geq \widehat{g}(\omega)$ for all $\omega \in \mathbb{R}$.

Suppose now that we have a function g with the properties stated in Exercise 1 and suppose we choose $\eta > 0$ sufficiently large that $\widehat{g}(\omega) \approx 0$ for $|\omega| > \eta$. Define the "wavelet" ψ by

$$\psi(t) = g(t)\, e^{i\eta t}. \tag{2}$$

Then $\widehat{\psi}(0) \approx 0$ and ψ is approximately analytic. Here

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} g\left(\frac{t-b}{a}\right) e^{i\eta \frac{t-b}{a}}$$

is supported in the interval $\left[b - \frac{a}{2}, b + \frac{a}{2}\right]$.

Exercise 2 Consider the following choices of g (restricted to the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$).

Name	g
Recangle	1
Hamming	$0.54 + 0.46\cos(2\pi t)$
Gaussian	e^{-18t^2}
Hanning	$\cos^2(\pi t)$
Blackman	$0.42 + 0.5\cos(2\pi t) + .08\cos(4\pi t)$

1. Compute and plot $\widehat{g}(\omega)$ (Hint: It is the same as the Fourier series).
2. Experiment with the choice of η and make plots of the real and imaginary parts $\psi(t)$ as constructed in (2).

Theorem 1 *(estimation of wavelet coefficients)*
Suppose $f(t) = A(t)e^{i\phi(t)}$, ψ *is defined by (2) and* $\|g\| = 1$. *Then*

$$\langle f, \psi_{a,b}\rangle = \sqrt{a} f(b) \widehat{g}\left(\frac{\eta - a\phi'(b)}{2\pi}\right) + \epsilon(a,b) \tag{3}$$

$$= \sqrt{a} A(b) e^{i\phi(b)} \widehat{g}\left(\frac{\eta - a\phi'(b)}{2\pi}\right) + \epsilon(a,b), \tag{4}$$

where

$$|\epsilon(a,b)| \leq \frac{a^{3/2}}{\sqrt{12}}\left(\sup_{t\in[b-a/2,b+a/2]} |A'(t)| + 2A(b) \sup_{t\in[b-a/2,b+a/2]} |\phi'(t)|\right) \tag{5}$$

$$\leq \frac{a^{3/2}}{\sqrt{3}} \sup_{t\in[b-a/2,b+a/2]} \max\{|A'(t)|, 2A(b)|\phi'(t)|\}. \tag{6}$$

Proof. Write

$$f(t) = A(t)e^{i\phi(t)} = \widetilde{f}(t) + h(t)$$

with

$$\begin{aligned}\widetilde{f}(t) &= A(b)\, e^{i(\phi(b)+(t-b)\phi'(b))} \\ &= f(b)\, e^{i(t-b)\phi'(b)}.\end{aligned}$$

Then

$$\langle f, \psi_{a,b}\rangle = \left\langle \widetilde{f}, \psi_{a,b}\right\rangle + \langle h, \psi_{a,b}\rangle .$$

The first term gives

$$\begin{aligned}\left\langle \widetilde{f}, \psi_{a,b}\right\rangle &= f(b)\int_{-\infty}^{\infty} e^{-i(t-b)(\eta/a-\phi'(b))}\frac{1}{\sqrt{a}}g\left(\frac{t-b}{a}\right)dt \\ &= f(b)\int_{-\infty}^{\infty} e^{-it(\eta/a-\phi'(b))}D_a g(t)\,dt \\ &= f(b)\,D_{1/a}\widehat{g}\left(\frac{\eta/a-\phi'(b)}{2\pi}\right).\end{aligned}$$

To estimate the second term we write

$$\begin{aligned}h(t) &= A(t)\,e^{i\phi(t)} - A(b)\,e^{i(\phi(b)+(t-b)\phi'(b))} \\ &= (A(t)-A(b))\,e^{i\phi(t)} + A(b)\left(e^{i\phi(t)} - e^{i(\phi(b)+(t-b)\phi'(b))}\right) \\ &= (A(t)-A(b))\,e^{i\phi(t)} + A(b)\left(e^{i\phi(t)} - e^{i\phi(b)}\right) + f(b)\left(1-e^{i(t-b)\phi'(b)}\right) \\ &= T_1(t) + T_2(t) + T_3(t).\end{aligned}$$

Then

$$\begin{aligned}\left|\langle T_1, \psi_{a,b}\rangle\right| &\leq \int_{-\infty}^{\infty}|A(t)-A(b)|\,D_a g(t-b)\,dt \\ &= \int_{-\infty}^{\infty}|A(t+b)-A(b)|\,D_a g(t)\,dt \\ &= \int_{-a/2}^{a/2}|A(t+b)-A(b)|\,D_a g(t)\,dt \\ &\leq \int_{-a/2}^{a/2}|A'(\xi(t))\,t|\,D_a g(t)\,dt,\end{aligned}$$

where $\xi(t)$ is between t and $t+b$. Hence,

$$\left|\langle T_1, \psi_{a,b}\rangle\right| \leq \sup_{t\in[b-a/2,b+a/2]}|A'(t)|\int_{-a/2}^{a/2}|t|\,D_a g(t)\,dt.$$

Furthermore,

$$\begin{aligned}
&\int_{-a/2}^{a/2} |t|\, D_a g(t)\, dt \\
=\;& \frac{1}{\sqrt{a}} \int_{-1/2}^{1/2} |at|\, g(t)\, a dt \\
=\;& a^{3/2} \int_{-1/2}^{1/2} |t|\, g(t)\, dt \le a^{3/2} \|g\| \sqrt{\int_{-1/2}^{1/2} t^2 dt} \\
=\;& \frac{a^{3/2}}{\sqrt{12}}.
\end{aligned}$$

Therefore,

$$\left|\left\langle T_1, \psi_{a,b} \right\rangle\right| \le \frac{a^{3/2}}{\sqrt{12}} \sup_{t\in[b-a/2,b+a/2]} |A'(t)|\,.$$

Similarly

$$\begin{aligned}
\left|\left\langle T_2, \psi_{a,b} \right\rangle\right| &\le A(b) \int_{-a/2}^{a/2} \left|e^{i\phi(t+b)} - e^{i\phi(b)}\right| D_a g(t)\, dt \\
&\le a^{3/2} A(b) \sup_{t\in[b-a/2,b+a/2]} |\phi'(t)| \int_{-1/2}^{1/2} |t|\, g(t)\, dt \\
&\le \frac{a^{3/2}}{\sqrt{12}} A(b) \sup_{t\in[b-a/2,b+a/2]} |\phi'(t)|
\end{aligned}$$

and

$$\begin{aligned}
\left|\left\langle T_3, \psi_{a,b} \right\rangle\right| &\le |f(b)| \int_{-a/2}^{a/2} \left|1 - e^{i\phi'(b)t}\right| D_a g(t)\, dt \\
&\le a^{3/2} A(b)\, |\phi'(b)| \int_{-1/2}^{1/2} |t|\, g(t)\, dt \\
&\le \frac{a^{3/2}}{\sqrt{12}} A(b) \sup_{t\in[b-a/2,b+a/2]} |\phi'(t)|\,.
\end{aligned}$$

■

The smallness of the corrective term $\epsilon(a,b)$ can be seen as follows. For small scales, the CWT emphasizes the local rapid change of f where $|A'(t)|, |\phi'(t)|$

are large. In this case $\epsilon(a,b)$ is small because of the coefficient $a^{3/2}$. For large scales, the CWT emphasizes the slow local change of f where $|A'(t)|, |\phi'(t)|$ are naturally small, which means that $\epsilon(a,b)$ is also small.

Real Signals Equation (3) and the estimates (5), (6) do not change if f is a real signal of the form

$$f(t) = A(t)\cos\phi(t).$$

To see this we put $F(t) = A(t)\, e^{i\phi(t)}$. Then

$$f(t) = \frac{F(t) + \overline{F}(t)}{2}$$

and

$$\begin{aligned}\langle f, \psi_{a,b}\rangle &= \frac{1}{2}\langle F, \psi_{a,b}\rangle + \frac{1}{2}\langle \overline{F}, \psi_{a,b}\rangle \\ &= \frac{\sqrt{a}}{2} F(b)\,\widehat{g}\left(\frac{\eta - a\phi'(b)}{2\pi}\right) + \frac{\sqrt{a}}{2}\overline{F}(b)\,\widehat{g}\left(\frac{\eta + a\phi'(b)}{2\pi}\right) \\ &\quad + \frac{1}{2}\epsilon_1(a,b) + \frac{1}{2}\epsilon_2(a,b).\end{aligned}$$

Since $\widehat{g}$ is even, $\widehat{g}\left(\frac{\eta + a\phi'(b)}{2\pi}\right) = \widehat{g}\left(\frac{\eta - a\phi'(b)}{2\pi}\right)$. Therefore,

$$\begin{aligned}\langle f, \psi_{a,b}\rangle &= \frac{\sqrt{a}}{2}\left(F(b) + \overline{F}(b)\right)\widehat{g}\left(\frac{\eta - a\phi'(b)}{2\pi}\right) + \epsilon(a,b) \\ &= \sqrt{a} f(b)\,\widehat{g}\left(\frac{\eta - a\phi'(b)}{2\pi}\right) + \epsilon(a,b),\end{aligned}$$

where $\epsilon(a,b) = \frac{1}{2}\epsilon_1(a,b) + \frac{1}{2}\epsilon_2(a,b)$. Furthermore, since F and $\overline{F}$ have the same amplitudes and (up to a sign change) phases, the estimates (5), (6) are the same for ϵ_1, ϵ_2 and, hence, for ϵ.

Recovery of the Amplitude and Phase Equation (3) gives the approximate value of the normalized scalogram as

$$\frac{1}{a}\left|W_\psi f(a,b)\right|^2 \approx A^2(b)\left|\widehat{g}\left(a\frac{\eta - \phi'(b)}{2\pi}\right)\right|^2.$$

Since $\widehat{g}$ is maximum at $\omega = 0$, the expression on the right is maximum when $\frac{\eta - a\phi'(b)}{2\pi} = 0$, giving

$$\phi'(b) = \frac{\eta}{a} \tag{7}$$

. The analytic amplitude is given by

$$A(b) = \frac{\sqrt{\frac{1}{a} P_W f(a,b)}}{\widehat{g}(0)}. \tag{8}$$

The ridge algorithm then works as follows: at a given scale a one finds the time instant b at which $\frac{1}{a} P_W f(a,b)$ achieves its maximum. In other words, (a,b) is the ridge point at scale a. The instantaneous frequency is computed from (7) and the amplitude is calculated from (8). The phase $\phi(b)$ is calculated as the phase of $W_\psi f(a,b)$ at the ridge point. The signal $f(t)$ is thus synthesized as

$$f(t) \approx A(b)\, e^{i(\phi(b)+\phi'(b)(t-b))}.$$

In practice the ridge points are found by saving the normalized scalogram in a matrix W whose rows correspond to scales and whose columns correspond to time instants and then, for each row, finding the column where the maximum occurs.

Exercise 3 How are the amplitude and phase angles recovered for a real signal?

Exercise 4 Write a program that computes and plots the ridges of the hyperbolic chirp

$$f(t) = \cos\left(\frac{\alpha}{t-\beta}\right)$$

for various values of α, β using a wavelet with $g(t) = \chi_{[-1/2,1/2]}(t)$.

Multiple Frequencies Suppose the signal f has the form

$$f(t) = A_1(t)\, e^{i\phi_1(t)} + A_2(t)\, e^{i\phi_2(t)}.$$

Then

$$\begin{aligned} \langle f, \psi_{a,b} \rangle &= \sqrt{a} A_1(b)\, e^{i\phi_1(b)} \widehat{g}\left(\frac{\eta - a\phi_1'(b)}{2\pi}\right) \\ &\quad + \sqrt{a} A_2(b)\, e^{i\phi_2(b)} \widehat{g}\left(\frac{\eta - a\phi_2'(b)}{2\pi}\right) + \epsilon(a,b). \end{aligned}$$

The two components are discriminated if

$$\widehat{g}\left(\frac{a\left|\phi_1'(b)-\phi_2'(b)\right|}{2\pi}\right)\approx 0$$

for all b. This means that

$$\left|\phi_1'(b)-\phi_2'(b)\right|\geq\frac{2\pi\eta}{a}.$$

In this case $\eta = a\phi_1'(b)$ creates a ridge point from which we can calculate $A_1(b)$ and $\phi_1'(b)$ and $\eta = a\phi_2'(b)$ generates a second ridge point from which to calculate $A_2(b)$ and $\phi_2'(b)$.

Assignment 1 Write a program to synthesize a signal of one or two frequencies where the amplitudes and phases are calculated from a wavelet transform. Experiment with several choices of wavelets and compare the calculated values with the original ones.