3. The Wavelet Transform

Wavelets are all about resolution of signals both in time and in frequency simultaneously.

3.1 Time–frequency atoms

We consider signals \( f \in L^2(\mathbb{R}) \) which are analysed with filters \( \psi \in L^2(\mathbb{R}) \), the time–frequency atoms which satisfy \( \| \psi \| = 1 \) for parameter \( \gamma \in \Gamma \). These filters define a transform

\[
Tf(\gamma) = \int_{\mathbb{R}} \overline{\psi}_\gamma(t) f(t) \, dt = \langle f, \psi_\gamma \rangle.
\]

The Parseval formula gives

\[
Tf(\gamma) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(w) \overline{\hat{\psi}_\gamma}(w) \, dw.
\]

As we have seen, \( \hat{f} \) and \( \hat{\psi}_\gamma \in L^2(\mathbb{R}) \). If, for a given \( \gamma \), \( \psi_\gamma \) is highly concentrated around \( t = u \), then \( Tf(\gamma) \) recovers information of the signal for \( t = u \). If \( \hat{\psi}_\gamma(w) \) is highly concentrated around \( w = \xi \) then \( Tf(\gamma) \) depends on the energy in the spectrum close to \( w = \xi \).

Ideally, one would like \( \gamma \) and \( \psi_\gamma \) such that both these things happen. This is the idea of the time–frequency transform.

Two examples which will be further discussed are...
1. Windowed Fourier transforms, where \( \tau = (s,u) \in \mathbb{R}^2 \) and
\[
\psi_{\tau}(t) = g_{s,u}(t) = e^{is} g(t-u), \\
\text{frequency}
\]
where \( g \) is a "window function" and

2. Wavelets \( \psi = (u,v) \in \mathbb{R} \times \mathbb{R}_+ \) and
\[
\psi_{\tau}(t) = \psi_{s,u}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right), \\
\text{frequency}
\]
The energy for both is typically well localized in time and concentrated in a limited frequency band.

Consider now the time-frequency spread of \( \psi_{\tau} \).

Since
\[
\| \psi_{\tau} \|_2^2 = \int_{\mathbb{R}} |\psi_{\tau}(t)|^2 dt = 1
\]
we interpret \( |\psi_{\tau}| \) as density on \( \mathbb{R} \). We assume that \( |\psi_{\tau}|^2 \) has first and second moments (since it is localized) and let
\[
\mu_{\tau} = \int_{\mathbb{R}} t |\psi_{\tau}(t)|^2 dt,
\]
and
\[
\sigma^2_{t}(\tau) = \int_{\mathbb{R}} (t - \mu_{\tau})^2 |\psi_{\tau}(t)|^2 dt.
\]
By the Plancherel formula one has
\[
\frac{1}{2\pi} \int_{\mathbb{R}} |\hat{\psi}_{\tau}(w)|^2 dw = 1,
\]
and so
\[
\sigma^2_{t}(\tau) = \frac{1}{2\pi} \int_{\mathbb{R}} |\hat{\psi}_{\tau}(w)|^2 dw.
\]
and
\[
\sigma^2_{w}(\tau) = \frac{1}{2\pi} \int_{\mathbb{R}} (w - \mu_{\tau})^2 |\hat{\psi}_{\tau}(w)|^2 dw.
\]
The interpretation is that \( \psi_{\tau} \) is represented by a thin box centered at \((\mu_{\tau}, s_{\tau})\) with

\( \sigma_{t}(\tau) \) and \( \sigma_{w}(\tau) \).
By the Heisenberg uncertainty the \( *\text{area} *\) of the Heisenberg box is bounded by
\[
\sigma_x(p_x) \cdot \sigma_p(x) > \frac{1}{2}
\]
such that the atom \( q_x \) trades off resolution in time for resolution in frequency. (This is the best localization one can get.)

If for every \((u, v) \in \mathbb{R}^2\) there is a unique \( f(u, v) \)
then the corresponding energy of a signal \( f \)
corresponding to \( q_x \) is
\[
P_f(u, v) = |(f, q_x(u, v))|^2 = \int_{\mathbb{R}} |f(t)|^2 q_x(u, v, t) dt.
\]

### 3.2 Windowed Fourier Transform

Gabor introduced "windowed Fourier atoms" to analyze frequency variations of sound. Let the window function \( g(t) \) be real and symmetric, i.e. \( g(t) = g(-t) \) and the atom be
\[
q_{u, v}(t) = e^{i\pi t} g(t - u).
\]
Furthermore, let \( \|g\| = 1 \) and thus \( \|q_{u, v}\| = 1 \) \( \forall u, v \in \mathbb{R} \).

In this case, the time-frequency transform is
\[
(Sf)(u, v) = \langle f, q_{u, v} \rangle = \int_{\mathbb{R}} f(t) g(t - u) e^{i\pi t} dt,
\]
a \( l^2(\mathbb{R}) \).

This is also called the "short-time Fourier transform".
The energy distribution or density for the signal relation to the windowed Fourier transform is

$$Psf(u, \xi) = |Ssf(u, \xi)|^2 = \int f(t) g(t-u) e^{-i \xi t} dt,$$

The Heisenberg box is defined by time & frequency variances. They are for \( g_{u, \xi}(t) \):

$$\sigma_t^2 = \int (t-u)^2 |g_{u, \xi}(t)|^2 dt = \int (t-u)^2 \hat{g}(t) \hat{g}^*(t) dt,$$

and, as \( \hat{g}_{u, \xi}(w) = \hat{g}(w-\xi) \exp[-i w (w-\xi)] \):

$$\sigma_w^2 = \frac{1}{2\pi} \int (w-\xi)^2 |\hat{g}_{u, \xi}(w)|^2 dw = \frac{1}{2\pi} \int w^2 |\hat{\hat{g}}(w)|^2 dw,$$

respectively, where \( \hat{\hat{g}} \) is the Fourier transform of \( \hat{g} \).

Note that \( \sigma_t \) and \( \sigma_w \) are independent of \( u \) and \( \xi \) and only depend on \( g \) (and \( \hat{g} \)).

Thus in this case the Heisenberg boxes are of constant size.

The original signal can be reconstructed from the time-frequency transform by:

**Theorem** If \( f \in L^2(\mathbb{R}^2) \) then

$$f(t) = \frac{1}{2\pi} \int \int Ssf(u, \xi) \overline{g(t-u)} e^{i \xi t} ds \, du$$

and

$$\int |f(t)|^2 dt = \frac{1}{2\pi} \int \int |Ssf(u, \xi)|^2 ds \, du.$$

(Inversion formula for the windowed Fourier transform.)
Proof. (i) Observe that \( SF \) is essentially a convolution of \( f \) with \( g_\varepsilon(w) = g(w) \exp(\varepsilon iw) \):

\[
SF(w, \varepsilon) = \exp(-iw\varepsilon) \int_{\mathbb{R}} f(t) g(t-w) \exp(iw(t-\varepsilon)) \, dt = \exp(-iw\varepsilon) (f \ast g_\varepsilon)(w)
\]

(ii) Now let

\[
f_\varepsilon(w) = \int_{\mathbb{R}} SF(w, \varepsilon) \exp(-2\pi i w u) \, du
\]

then the convolution theorem gives

\[
\hat{f_\varepsilon}(w) = \hat{f}(w+\varepsilon) \hat{g}_\varepsilon(w+\varepsilon) = \hat{f}(w+\varepsilon) \hat{g}(w)
\]

as \( \hat{g_\varepsilon}(w) = \hat{g}(w-\varepsilon) \).

(iii) Now, one has by the Parseval formula:

\[
\int_{\mathbb{R}} SF(w, \varepsilon) g(t-w) \, du = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f_\varepsilon}(w) \overline{\hat{g}(w)} e^{iw t} \, dw = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(w+\varepsilon) |\hat{g}(w)|^2 e^{iw t} \, dw.
\]

(iv) The reconstruction formula is obtained by integrating and applying of Fubini's

\[
\frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \hat{f}(w+\varepsilon) |\hat{g}(w)|^2 e^{iw t} \, dw \, e^{i\varepsilon t} \, d\varepsilon = \frac{1}{2\pi} \int_{\mathbb{R}} \left( \int_{\mathbb{R}} \hat{f}(w+\varepsilon) \hat{g}(w) \, d\varepsilon \right) |\hat{g}(w)|^2 \, dw = 2\pi \hat{f}(t)
\]

\[= \frac{1}{2\pi} \int_{\mathbb{R}} \left( \int_{\mathbb{R}} \hat{f}(w+\varepsilon) e^{i\varepsilon t} \, d\varepsilon \right) |\hat{g}(w)|^2 \, dw = 2\pi \hat{f}(t)
\]
\[(v) \text{ The Plancherel formula gives:}\]
\[
\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| \hat{f}(u, s) \right|^2 \, du \, ds = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| \hat{f}(w, s) \right|^2 \, dw \, ds
\]

Integrate over $s$ first (Fubini—check!) to get with Plancherel a second time:
\[
\overline{f(t)} = \int_{-\infty}^{\infty} \hat{f}(t, \nu) \, d\nu.
\]

\[
\text{Thus the transform distributes the energy of the signal in the frequency-time plane.}
\]

The reconstruction formula can be restated as:
\[
f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \overline{f(t, \nu)} g(u, \nu) \, du \, d\nu.
\]

The information in $Sf(u, s)$ is redundant and not every $\phi \in L^2(\mathbb{R}^2)$ represents a signal $f \in L^2(\mathbb{R})$.

\[
\text{Proposition: Let } \phi \in L^2(\mathbb{R}^2). \text{ Then}
\]
\[
\mathcal{F} \{ \phi \} = Sf \iff \phi(u_0, s_0) = \frac{1}{2\pi} \int_{\mathbb{R}^2} \phi(u, \nu) K(u_0, u, s_0, 0) \, du
\]

with
\[
K(u_0, u, s_0, s_0) = \langle g(u_0, s_0), g(u, s) \rangle.
\]

\[
\text{Proof: (i) show that this holds for all } \phi = Sf \text{—use the reconstruction formula.}
\]
\[
\text{(ii) let } f = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(u, s) g(t-u) \exp(i\omega t) \, d\omega \, ds
\]

all show that by contr. above one has $Sf = \phi$.

Note that $K$ is a reproducing kernel as it recovers set values.
As \( g(u,t) = e^{i\hat{\tau}g(t-u)} \), one has for the kernel

\[
K(u_0, u, \xi_0, \xi) = \int e^{i(\xi_0 - \xi)u} g(t-u_0) g(t-u) \, dt
\]

\[
= \exp(-i\frac{\xi - \xi_0}{2}(u + u_0)) \cdot \mathcal{A}(\xi) \cdot (u_0 - u, \xi_0 - \xi)
\]

where

\[
\mathcal{A}(\xi, \nu) = \int_{\mathbb{R}} g(\nu + \xi) g(\nu - \xi) e^{-i\nu \nu} \, d\nu
\]

A small assignment: work this out.

Then \( \mathcal{A} \) is the ambiguity function of \( g \). With the

Boas-Parlett formula one gets

\[
\mathcal{A}(\xi, \nu) = \frac{1}{2\pi} \int_{\mathbb{R}} g(\nu + \xi) g(\nu - \xi) e^{-i\nu \nu} \, d\nu
\]

So that the ambiguity set is both represented in the

frequency \& in the time domain by a similar formula.

It provides information about the spread in time

and frequency.
3.3 Wavelet transforms

A wavelet is a $\psi \in L^2(\mathbb{R})$ with zero mean:

$$\int_\mathbb{R} \psi(t) \, dt = 0, \quad \|\psi\| = 1$$

and $\psi$ is centered around $t = 0$.

Time-frequency atoms are obtained by translation and scaling:

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right),$$

Note that $\|\psi_{u,s}\| = 1$.

The wavelet transform of $f \in L^2(\mathbb{R})$ is

$$W_f(u,s) = \langle f, \psi_{u,s} \rangle = \int_\mathbb{R} f(t) \frac{1}{\sqrt{s}} \psi^*\left(\frac{t-u}{s}\right) \, dt,$$

With

$$\hat{\psi}_s(t) = \frac{1}{\sqrt{s}} \psi^*(-t/s),$$

this is a convolution:

$$W_f(u,s) = f \ast \hat{\psi}_s(u) \rightarrow \psi$$

filter bank

and:

$$f \ast \hat{\psi}_s(u) = \hat{f}(u) \cdot \hat{\psi}_s(u)$$

where

$$\hat{\psi}_s(u) = \sqrt{s} \hat{\psi}(su).$$

2 main types:

- complex analytic wavelet = windowed F.T.
- real wavelet → to detect signal transitions.
3.3.1 Real wavelets

- detecting transients & analysing fractals

Examples: "Mexican hat" 2nd derivative of Gaussian

\[
\begin{align*}
\psi(t) &= \frac{2}{\pi \sqrt{30}} \left( \frac{t^2}{\sigma^2} - 1 \right) \exp \left( -\frac{t^2}{2\sigma^2} \right) \\
\hat{\psi}(\omega) &= -\sqrt{8 \sigma^{3/2}} \frac{1}{\sqrt{\pi \sigma}} \omega^2 \exp \left( -\frac{\sigma^2 \omega^2}{2} \right)
\end{align*}
\]

Use in computer vision - detect multiscale edges.

Reconstruction formula: (includes energy conservation)

**Theorem (Calderon, Grossmann, Morlet)**

Let \( \psi \in L^2(\mathbb{R}) \) be a real wavelet such that

\[
C_\psi = \int_0^\infty \frac{1}{\omega} |\hat{\psi}(\omega)|^2 d\omega < \infty.
\]

Then

\[
s(t) = \frac{1}{C_\psi} \int_0^\infty \int_\mathbb{R} |W_F(u,s)|^2 \psi \left( \frac{t-u}{s} \right) du \cdot ds
\]

and

\[
\|s(t)\|_2^2 = \frac{1}{C_\psi} \int_0^\infty \int_\mathbb{R} |W_F(u,s)|^2 du \cdot ds.
\]

(Inversion formula for wavelet transform.)
The regularity \( C_T = \int_0^\infty \frac{\left| \hat{\psi}(\omega) \right|^2}{\omega} \, d\omega < \infty \) is called admissibility condition. If follows that \( \hat{\psi}(0) = 0 \), which, together with \( \hat{\psi} \in C'(\mathbb{R}) \) is sufficient for admissibility. This occurs if
\[
\int_{\mathbb{R}} (1 + |t|) \cdot |\psi(t)| \, dt < \infty.
\]

**Proof of the reconstruction:**

(i) Let \( \psi_s(t) = \frac{1}{\sqrt{s}} \psi(-t/s) \) such that
\[
Wf(u,s) = f * \psi_s^*(u) \quad \text{and} \quad \overline{\psi}_s(t) = \psi_s(-t).
\]
One has
\[
\hat{\psi}_s(\omega) = \sqrt{s} \hat{\psi}(-s\omega) \quad \text{and} \quad \hat{\psi}_s(\omega) = -\sqrt{s} \hat{\psi}^*(-s\omega).
\]

(ii) By Plancherel's convolution theorem one has
\[
\psi_s(t) = \int_{\mathbb{R}} Wf(u,s) \overline{\psi}_s(t-u) \, du = f * \psi_s^* * \overline{\psi}_s(t)
\]
and, by the Fourier inversion theorem \( \hat{\psi}(-\omega) = \hat{\psi}(\omega) \) (as \( \psi \) real):
\[
= \frac{i}{2\pi} \int_{\mathbb{R}} \hat{\psi}(\omega) \left| \frac{\psi(s\omega)}{s} \right|^2 \, e^{i\omega t} \, d\omega
\]
(\*)

(iii) Now \( (\ast) / s^2 \) is in \( L^1(\mathbb{R} \times \mathbb{R}_+) \) and, with Fubini,
\[
\int_0^\infty \psi_s(t)/s^2 \, ds = \frac{1}{2\pi} \int_\mathbb{R} \hat{\psi}(\omega) e^{i\omega t} \left( \int_0^\infty |\hat{\psi}(s\omega)| \, ds \right) \, d\omega = C_T \quad \text{s.t. s \to ws}
\]
\[
= f(t) \cdot C_T \quad \text{from which we get (iv).}
\]

(iv) The energy conservation obtained in similar way, left to reader.
Like for windowed Fourier, the range of \( W(E^c(u^c)) \) is a subset of \( L^2(\mathbb{R} \times \mathbb{R}^*_+) \). As before, one has

\[
Wf(u_0, s_0) = \frac{1}{c_4 \pi} \int_{\mathbb{R}} K(u, u_0, s, s_0) Wf(u, s) \, du \, ds
\]

where

\[
K(u_0, u, s, s_0) = \langle \Psi(u, s), \Psi(u_0, s_0) \rangle
\]

is the reproducing kernel. This follows by combining the reconstruction formula with the transform “in the opposite way”:

\[
K = W^* R^* - \text{where } R \text{ is the co-reciprocal, i.e., } RW = I
\]

\[
\implies KW = WRW = W, \text{ i.e., } R \text{ is the generalized inverse.}
\]

If one is given the information about the small scales, the scaling function summarises the effect of the large scales. It is defined through the Fourier transform (up to an arbitrary phase):

\[
\|\hat{\phi}(w)\|^2 = \int_0^{\infty} |\hat{\psi}(sw)|^2 \, ds = \int_0^{\infty} |\hat{\psi}(s)|^2 \, ds
\]

Note: \( |\hat{\phi}(w)|^2 \xrightarrow{w \to 0} C \) and let \( \hat{\psi}(s) = \frac{1}{\sqrt{s}} \hat{\psi}(\sqrt{s}) \) & \( \hat{\psi}(t) = \hat{\psi}^*(t) \).

By the canonical Huang-Fedderi identity:

\[
\int_0^{\infty} Wf(u, s) \ast \psi_s(t) \, ds = \frac{1}{2\pi} \int_{\mathbb{R}} \int_0^{\infty} \hat{f}(w) \, \hat{\psi}(sw) \, |w|^2 \, dw \, ds
\]

\[
= \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(w) e^{iwt} \int_0^{\infty} \hat{\psi}(sw) \, |w|^2 \, dw \, ds = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(w) \int_0^{\infty} \hat{\phi}(sw) \, |w|^2 \, dw \, ds
\]

and, with \( f \ast \overline{\psi}_{s_0} = Lf(-s_0), \) one gets

\[
f(t) + \frac{1}{c_4 \pi} \int_{\mathbb{R}} Wf(u, s) \ast \psi_s(t) \, du \, ds + \frac{1}{c_4 \pi} Lf(-s_0) \ast \overline{\psi}_{s_0}(t)
\]

small-scale parts

large-scale "trends"
3.3.2 Analytic wavelets

Recall that analytic functions are of the form $f(z) = \sum_{k=0}^{\infty} a_k z^k$ and, with $z = r e^{i\theta}$, this gives the Fourier series

$$f(re^{i\theta}) = \sum_{k=0}^{\infty} a_k r^k e^{i k \theta}$$

(no $k<0$ occurs).

Similarly, a signal $f \in L^2(\mathbb{R})$ is analytic if

$$\hat{f}_a(w) = 0 \quad \text{for} \quad w < 0.$$ 

While $f_a$ is not real, it is determined by its real part

$$f = \text{Real}(f_a) \Rightarrow \hat{f}(w) = \hat{f}_a(w) + \hat{f}_a^*(-w)$$

from which

$$\hat{f}_a(w) = \int 2\hat{f}(w), \quad w > 0 \quad (*) -$$

Furthermore, the analytic part $f_a$ of a general signal $f$ is defined by $(*)$.

An analytic wavelet $\psi$ satisfies $\hat{\psi}(w) = 0$ for $w < 0$.

The analytic wavelet transform is then

$$Wf(w,s) = \langle f, \psi_{w,s} \rangle = \int f(t) \frac{1}{\sqrt{s}} \psi^*(\frac{t-s}{s}) \, dt.$$ 

Assume that $\int_{\mathbb{R}} |\psi(t)|^2 \, dt = \infty$ and define $\sigma^2 = \int_{\mathbb{R}} |\psi(t)|^2 \, dt < \infty$.

Thus $u = \int_{\mathbb{R}} |\psi_{w,0}(t)|^2 \, dt$ and

$$\int_{\mathbb{R}} (|\psi_{w,0}(t)|^2)^2 \, dt = s \sigma^2.$$
In the frequency domain, one has the mean frequency as
\[ \mu = \frac{1}{2\pi} \int_0^{\infty} \omega |\psi^2(\omega)| d\omega \]
and, as
\[ \psi_{\text{gs}}(\omega) = \sqrt{s} \psi(8\omega) \exp(-i\omega) \]
the mean freq. of \( \psi_{\text{gs}} \) is \( 1/s \) and the energy spread is
\[ \frac{1}{2\pi} \int_0^{\infty} (\omega - \mu)^2 |\psi_{\text{gs}}(\omega)|^2 d\omega = \frac{s^2}{s^2} \]
where \( \mu_s = \frac{1}{\sqrt{s}} \int_0^{\infty} (\omega - \mu)^2 |\psi^2(\omega)| d\omega \).

So the Heisenberg box is \( \eta w \)

\[ \begin{array}{c}
| & 1/\sqrt{s} & \mu_w/s \\
1/\sqrt{s} & \mu_w/s & 1/\sqrt{s} \\
\mu_w/s & 1/\sqrt{s} & \eta w \\
\end{array} \]

and one observes that for larges (low frequencies) one has small frequency spreads and high time spreads etc. The signal energy belonging to that box is given by the density (in \( R^* \times R^* \)):
\[ P_{\text{WF}}(u, \nu) = |W_f(u, s)|^2 = |W_f(u, \nu/2)|^2 \]
this is the scalogram.
Again, we get a reconstruction formula by

**Theorem:** For an analytic wavelet test, one has

\[
WF(u, s) = \frac{1}{2} Wf_a(u, s), \quad f \in L^2(\mathbb{R})
\]

and, if \( f \) is real and \( C \psi = \int_{\mathbb{R}} \frac{|\psi'(w)|^2}{w} < \infty \) then

\[
f(t) = \frac{2}{C \psi} \text{Real} \left( \int_{-\infty}^{\infty} \int_{0}^{\infty} WF(u, s) \psi_s(t-u) \, du \, \frac{ds}{s^2} \right)
\]

and

\[
\|f\|^2 = \frac{2}{C \psi} \int_{\mathbb{R}} \left| \int WF(u, s) \right|^2 \, du \, \frac{ds}{s^2}
\]

**Proof:**

(i) Let \( f_s(u) = WF(u, s) = f \ast \overline{\psi_s}(u) \)

then \( \hat{f}_s(w) = \hat{f}(w) \overline{\hat{\psi}_s}(sw) \)

and, as \( \hat{\psi}(w) = 0 \), \( w < 0 \) & \( \hat{\psi}_{\infty}(w) = 2 \hat{f}(w), w > 0 \)

\( \Rightarrow \hat{f}_s(w) = \frac{1}{2} \hat{f}_a(w) - \hat{f}_s \psi^*(sw) \Rightarrow (\star) \).

(ii) As for the previous (real) transform one gets the reconstruction.

(iii) Energy conservation from the Plancherel formula.

In frequency, one has:

\[
\|f\|^2 = \frac{2}{C \psi} \int_{\mathbb{R}} \int WF(u, s) \, du \, ds
\]

(\( \Rightarrow \) justification for the \( \frac{1}{2} \) past \( dS \) in eq. \( \frac{1}{2} \) instead of \( \frac{1}{2} ds \))
Let \( g(t) \) be a real & symmetric window such that 
\( \hat{g}(\omega) = 0 \) for \( |\omega| > \eta > 0 \) (n low pass). Then 
\[
\psi(t) = g(t) \exp(i\eta t)
\]
has Fourier transform 
\[
\hat{\psi}(\omega) = \hat{g}(\omega - \eta)
\]
and thus \( \psi \) is an analytic wavelet, with centre \( \eta \) and 
\[
|\hat{\psi}(\eta)| = \sup_{\omega \in \mathbb{R}} |\hat{\psi}(\omega)| = |\hat{g}(0)|.
\]
An approximate analytic wavelet is the Gabor wavelet with 
\[
\psi(t) = \frac{1}{(\sigma^2_\eta)^{3/4}} \exp\left(\frac{-t^2}{2\sigma^2_\eta}\right), \quad \text{as, for } \sigma^2_\eta \gg 1
\]
one has \( \hat{\psi}(\omega) \approx 0 \) for \( |\omega| > \eta \).

### 3.4 Quadratic time-frequency energy - Wigner-Ville distribution

The time-frequency densities obtained so far all depended on the wavelets or windows used and their resolution was not only limited by the Heisenberg uncertainty but also by the actual wavelet or window. Here a distribution is introduced from which the spectrogram, the scalogram and all squared time-frequency decompositions can be derived.

The idea is to use the signal itself as the "wavelet" or filters. (related to self-similarity!)

The Wigner-Ville distribution is
\[
\text{WV} f(u, s) = \int_{\mathbb{R}} f(u + \frac{t}{2}) f^*(u - \frac{t}{2}) e^{-i\omega s} \, dt
\]
The Parseval formula shows that time & frequency are

\[ \mathcal{P}_v f(u, \xi) = \frac{1}{2\pi} \int \hat{f}(\xi + \frac{\tau}{2}) \hat{f}(\xi - \frac{\tau}{2}) e^{i\tau u} d\tau. \]

Check that \( \mathcal{P}_v f \in L^2 \).

If \( f \) is concentrated in time/frequency then so is \( \mathcal{P}_v f \).

**Proposition:**

(a) \( \text{supp} f = \left[ u_0 - T/2, u_0 + T/2 \right] \Rightarrow \)
\[ \text{supp} \mathcal{P}_v f(\cdot, \xi) \subseteq \left[ u_0 - T/2, u_0 + T/2 \right] \ni \xi \]

(b) \( \text{supp} \hat{f} = \left[ \xi_0 - \Delta/2, \xi_0 + \Delta/2 \right] \Rightarrow \)
\[ \text{supp} \mathcal{P}_v f(u, \cdot) \subseteq \left[ \xi_0 - \Delta/2, \xi_0 + \Delta/2 \right] \ni u. \]

**Proof:** (a) \( \Rightarrow \) (b) by symmetry, (a): check supports of integrand \( \square \)

**Examples:**

- \( f(t) = \delta(t - u_0) \Rightarrow \mathcal{P}_v f(u, \xi) = \delta(u - u_0) \)
- \( f(t) = \exp(i\xi_0 t) \Rightarrow \mathcal{P}_v f(u, \xi) = \frac{1}{2\pi} \delta(\xi - \xi_0) \)
- \( f(t) = \frac{\exp(-t^2/2\sigma^2)}{(\sigma^2\pi)^{1/4}} \Rightarrow \mathcal{P}_v f(u, \xi) = \frac{1}{\sqrt{\pi}} \exp(-u^2/(\sigma^2 - \xi^2)) \)

\[ \mathcal{P}_v f(u, \xi) = |f(u)|^2 |\hat{f}(\xi)|^2 \]

in \( L^2 \) case!
Properties of the Wigner-Ville distribution:

- \( f(t) = e^{i\phi} g(t) \Rightarrow \text{WV} f(u,s) = \text{WV} g(u,s) \)
- \( f(t) = g(t - u_0) \Rightarrow \text{WV} f(u,s) = \text{WV} g(u - u_0, s) \)
- \( f(t) = \exp(i\pi s t) g(t) \Rightarrow \text{WV} f(u,s) = \text{WV} g(u, s - 2\pi) \)
- \( f(t) = \frac{1}{\sqrt{\pi}} g(t/\sqrt{s}) \Rightarrow \text{WV} f(u,s) = \text{WV} g(u/s, s/\pi) \)

Proposition (Instantaneous Frequency):

Let \( a(t) \) be analytic part of \( f \) and

\[ f_a(t) = a(t) \exp[i\phi(t)] \quad (a(t) \neq 0) \]

Let the instantaneous frequency be

\[ \omega(t) = \phi'(t). \]

Then

\[ \omega(t) = \sqrt{\text{WV} f_a(t,s) ds \over \text{WV} f_a(t,s) ds} \]

(Proof: use Fourier transform.)

For example, a chirp has

\[ \text{WV} f(u,s) = \delta(s - 2au) \]

\[ f(t) = \exp(iat^2) \]
Now for a signal \( f(t) \) we can consider \( |f(t)|^2 \) and 
\( \frac{1}{2\pi} |\hat{f}(w)|^2 \) to be time & frequency densities, respectively. One has:
\[
\|f\|_2^2 = \int |f(t)|^2 dt = \frac{1}{2\pi} \int |\hat{f}(w)|^2 dw.
\]

Furthermore, these densities are the marginals of the Wigner-Ville distribution:

Proposition. For \( f \in L^2(\mathbb{R}) \):
\[
\int_{\mathbb{R}} P_{\hat{f}}(u,s) du = |f(s)|^2
\]
and
\[
\frac{1}{2\pi} \int_{\mathbb{R}} P_{\hat{f}}(u,s) ds = |f(u)|^2.
\]

Proof: \( g_s(u) = P_{\hat{f}}(u,s) \Rightarrow \hat{g}_s(\omega) = \hat{f}(\omega + s) \hat{f}^*(\omega - s) \).

Now \( |\hat{f}(s)|^2 = \hat{g}_s(0) = \int g_s(u) du = \int P_{\hat{f}}(u,s) du \)

and similarly for \( f(u) \).

Problems: All the above comments make the Wigner-Ville distribution the ideal time-frequency distribution but there are 2 problems:

1. \( P_{\hat{f}}(u,s) \) can be negative
2. Components of the signal can produce counterintuitive interferences.
Interferences: \[ f = f_1 + f_2 \]

\[ \Rightarrow P_{vf} = P_{vf_1} + P_{vf_2} + P_{vf_1,f_2} + P_{vf_2,f_1} \]

Where the "cross Wigner-Ville distribution" of \( f_1 \) is

\[ P_{vf_1,g}(u,s) = \int_R \hat{f}(u + \frac{v}{2}) \hat{g}^*(u - \frac{v}{2}) e^{-jvs} dv \]

The interference term is

\[ I[f_1,f_2] = P_{vf_1,f_2} + P_{vf_2,f_1} \]

These terms can be complicated, oscillatory, and contribute nothing to the marginal distributions if the signals do not have overlapping time-frequency components.

Interference can also occur for real signals \( f(t) \) with a single instantaneous frequency, \( f(t) = \text{Real}(f_1(t)) \) and \( f(t) = \text{Real}(f_2(t)) \). Then \( P_{vf} \) has energy concentrated at \( s_1 = \phi'(u) \), \( s_2 = -\phi'(u) \) and an interference at around \( s = 0 \). In this case should use \( P_{vf} \) not \( P_{vf} \).

Finally, the positivity:

Negative values of \( P_{vf} \) occur in the interference terms. These terms can be suppressed by smoothing:

\[ P_{vf}(u,s) = \int_R P_{vf}(u,s') \delta(u,s,s') ds' \]
where $\Theta$ is smooth enough to impose
$$\Theta f(u, s) > 0 \quad \text{for} \quad (u, s) \in T^2.$$ Examples include the spectrogram & scalogram.

More generally, let $\{\psi_k\}$ be time-frequency atoms. The time-frequency density of these atoms is
$$\Theta f(u, s) = \int f(u, s) \psi(u, s) \, du \, ds.$$ Then one gets
$$\Theta f(u, s) = \int \Theta f(u', s) \psi(u, u', s, s') \, du' \, ds'$$
with
$$\psi(u, u', s, s') = \frac{1}{2\pi} P_v \Theta \psi(u, s) (u', u', s, s').$$
by the Moyal formula which will be shown on next page.

**Examples are:**

- **Spectrogram:** $\psi(u, s) (t) = g(t-u) e^{ist}$

  $$\Rightarrow \psi(u, u', s, s') = \frac{1}{2\pi} P_v \Theta \psi(u, s) (u'-u, s'-s).$$

  Note: If $g$ Gaussian then $\Theta$ Gaussian.

- **Scalogram:** $\psi$ analytic, wavelet, centered at $u$

  $$\Theta f(u, s) = \int f(u, s) \, du \, ds,$$

  Then

  $$\psi(u, u', s, s') = \frac{1}{2\pi} P_v \Theta \psi(u, s) (u'-u, s'-s).$$

  positive dist. remove interference at the cost of some resolution.
Theorem [Hoyland]: \( f, g \in L^2(\mathbb{R}) \)
\[
\left| \int_{\mathbb{R}} f(t)g^{*}(t)dt \right|^2 = \frac{1}{2\pi} \int_{\mathbb{R}^2} \hat{f}(u,\bar{\tau})\hat{g}(u,\bar{\tau})d\bar{\tau}d\bar{\tau}
\]

Proof: (i) \( F_{\tau}(\tau) = \hat{f}(u+\tau/2)f^{*}(u-\tau/2) \)
\( G_{\bar{\tau}}(\bar{\tau}) = g(u+\bar{\tau}/2)g^{*}(u-\bar{\tau}/2) \) by definition.

then
\[
P_{\tau}f(u,\bar{\tau}) = \hat{F}_{\tau}(\bar{\tau}), \quad P_{\bar{\tau}}g(u,\bar{\tau}) = \hat{G}_{\bar{\tau}}(\bar{\tau})
\]
(iii) \( \int \int P_{\tau}f(u,\bar{\tau})P_{\bar{\tau}}g(u,\bar{\tau})d\bar{\tau}d\bar{\tau} = \int \hat{F}_{\tau}(\bar{\tau})\hat{G}_{\bar{\tau}}(\bar{\tau})d\bar{\tau}d\bar{\tau} \)

by Plancherel

(iii) \( I = \int \int P_{\tau}f(u,\bar{\tau})P_{\bar{\tau}}g(u,\bar{\tau})d\bar{\tau}d\bar{\tau} \)
\( \) (insert (i)):
\[
= \frac{1}{2\pi} \int \int f(u+\tau/2)f^{*}(u-\tau/2)g(u+\bar{\tau}/2)g^{*}(u-\bar{\tau}/2)d\tau du
\]

\( \) (iv) \( t = u+\tau/2 \quad t' = u-\tau/2 \quad d\tau du = dtdt' \) change of variables
\[
= \frac{1}{2\pi} \int \int f(t)g^{*}(t)f^{*}(t')g(t')dt dt'
\]

\( \) (v) \( = \frac{1}{2\pi} \int \int f(t)g^{*}(t)f^{*}(t')g(t')dt dt'
\]

\( \) \( = 2\pi |\langle f, g \rangle|^2 \) \( \) \( \) \( \)
So while we can remove negative parts and interferences, it comes at a cost and, in particular, the marginals are perturbed:

**Theorem (Wigner)** There is no positive quadratic energy distribution $Pf$ with

\[
\sqrt{Pf(\omega,dz)} = 2\pi |f(\omega)|^2 \quad \& \quad \sqrt{Pf(\omega,du)} = |f(\omega)|^2
\]

**Proof**: Uses the fact that correctly supported signals cannot vanish on intervals. ---

---
4. Wavelet Tools and Lipschitz Regularity

So far, we have seen how wavelet transforms do provide some insight in the distribution of the signals. For every $n$, the time-frequency domain for this wavelet will be used for further analysis of functions. In many ways, wavelets are the modern microscope to analyze real functions.

4.1 Lipschitz regularity and Fourier analysis

If $f$ is sufficiently smooth in $x$ it can be represented by its Taylor polynomials:

$$p_x(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x)}{k!} (t-x)^k$$

The error $e_x(t) = f(t) - p_x(t)$ is bounded by

$$|e_x(t)| \leq \left|\frac{t-x}{n!}\right|^n \sup_{x \in [x-h, x+h]} |f^{(n)}(\xi)|, \quad t \in [x-h, x+h]$$

if $f \in C^n[x-h, x+h]$,

The differentiability is generalized by Lipschitz regularity, whose $n$ is replaced by Lipschitz or H"older exponents $(\alpha > 0)$

**Definition (Lipschitz):** (i) $f$ is **Lipschitz** regular at $x$, if there exist $K > 0$ and $\alpha$ such that:

$$|f(t) - p_x(t)| \leq K|t-x|^\alpha, \quad t \in \mathbb{R}, \quad K = K(x)$$

(ii) $f$ is uniformly Lipschitz $\alpha$ over $[a,b]$ if $f \in L^\alpha$, where:

$$|f(t) - f(x)| \leq K|t-x|^\alpha, \quad t \in \mathbb{R}, \quad x \in [a,b]$$

(iii) Lipschitz regularity of $f$ at $x$ over $[x-h, x+h]$ is the supremum and thus if $f$ is Lipschitz $\alpha$. 


1. If \( f \in C^m, m=\lfloor \alpha \rfloor \Rightarrow f \Delta \Gamma \) is a Taylor expansion.
2. Lipschitz exponents can vary.
3. Multi-valued field with non-isolated singularities, where \( \mathbb{R} \) is different at each point.
4. Uniform Lipschitz exponents were required since \( \mathbb{R} \) is non-isolated.

\[
|f(t) - f(v)| \leq K |t - v|^\alpha
\]

\( \alpha \)-Lipschitz \( \alpha \)-continuously.

- Bounded & discontinuous \( \Rightarrow \) Lipschitz 0.
- \( \alpha < 1 \Rightarrow \) not differentiable (like \( \sqrt{t} \)) in \( C^\alpha \) sense.

Characteistics with Fourier theory:

**Theorem:** If

\[
\int |\hat{f}(w)| (1 + |w|^\alpha) \, dw < \infty \quad \text{for some } \alpha > 0
\]

then \( f \) is bounded and uniformly Lipschitz \( \alpha \) over \( \mathbb{R} \).

**Proof:**

(i) boundedness by inverse: \( |f(t)| \leq \int |\hat{f}(w)| \, dw < \infty \)

(ii) case \( 0 \leq \alpha < 1 \). \( \rho(f)(t) = f(v) \), to show \( |f(t) - f(v)| (|t - v|^\alpha) \leq K \).

Fourier inversion:

\[
|f(t) - f(v)| \leq \frac{1}{2\pi} \int |\hat{f}(w)| \cdot \frac{\text{exp}(iwt) - \text{exp}(ivt)}{|t - v|^\alpha} \, dw
\]

New bound (**):

\[
|f(t) - f(v)| \leq \frac{1}{2\pi} \int |\hat{f}(w)| \cdot \frac{2 |w|^\alpha}{|t - v|^\alpha} \, dw = K < \infty \quad \text{by cond.}
\]

(iii) Case \( \alpha > 1 \), easier reg. thus \( \Rightarrow f \in C^m(\mathbb{R}), m = \lfloor \alpha \rfloor \).

\[ f \text{ with bounded \( \alpha \)-order; } f \in C^m \text{ uniformly Lipschitz } \Rightarrow f \text{ uniform Lip } \alpha \text{ over } \mathbb{R} \]

\( \Rightarrow \) use (ii) and \( \varphi^{(m)}(w) = (iwt)^m \varphi(w) \).

\[ \text{complete the proof. (p. 106)} \]

- In Haile.
### 4.2 Wavelets with vanishing moments

A wavelet $\psi$ is said to have $n$ vanishing moments if

$$\int_{\mathbb{R}} t^k \psi(t) \, dt = 0 \quad \text{for} \quad 0 \leq k < n.$$  

Such a wavelet is orthogonal to polynomials of degree $n-1$. Now let, as before,

$$f(t) = p_x(t) + \varepsilon_n(t) \quad \text{with} \quad |\varepsilon_n(t)| < K |t|^n$$  

and $n \geq 1$. Then one can verify that for $W$ defined by $\psi$ one has

$$Wp_x(u,s) = \int_{\mathbb{R}} p_x(t) \frac{1}{t^s} \psi(t^s - u) \, dt = 0$$  

and so

$$Wf(u,s) = WE_n(u,s),$$  

i.e., the wavelet transform does only analyze the "non-polynomial part" and is thus very similar to a derivative. This is made more concise in the following theorem.

First let the wavelet $\psi$ have fast decay if for any $m \in \mathbb{N}$

$$|\psi(t)| \leq \frac{C}{1 + |t|^m}, \quad t \in \mathbb{R}.$$  

The wavelet transform (using $\psi$) can then be interpreted as the derivative of another wavelet transform (with $\varepsilon$). This is called a 

"multiscale differential operator"
One then has:

**Theorem:** A wavelet \( \psi \) will fast decay has \( n \) vanishing moments if \( \exists \tilde{\psi} \) will fast decay such that
\[
\psi(t) = (-1)^n \frac{d^n \theta(t)}{dt^n}
\]

As a consequence
\[
Wf(u,s) = s^n \frac{1}{du^n} \left( f \ast \overline{\tilde{\psi}_s} \right)(u)
\]

where
\[
\tilde{\psi}_s(t) = s^{1/2} \tilde{\psi}(-t/s)
\]

Moreover, \( \psi \) has no more than \( n \) vanishing moments if and only if \( \int \psi(t) dt \neq 0 \).

**Proof.**

(i) A \( \psi \) has a fast decay \( \Rightarrow \) Its Fourier transform is \( \hat{\psi} \in C^\infty \) by the regularity results of the Fourier transform:
\[
0 = \int \hat{\psi}(t) e^{it\omega} dt = (i\omega)^n \hat{\psi}^{(n)}(\omega)
\]
and so bounded \( \hat{\psi}(\omega) = \tilde{\psi}(\omega) = (i\omega)^n \hat{\psi}(\omega) \).

(iii) It follows that \( \psi(t) = (-1)^n \tilde{\psi}^{(n)}(t) \). Integrating \( n \) times provides the fast decay of \( \tilde{\psi} \) (by induction).

(iv) The wavelet transform is
\[
Wf(u,s) = f \ast \overline{\tilde{\psi}_s}(u) = \frac{1}{s} \tilde{\psi}_s(t) \text{ and } \frac{d^n}{dt^n}
\]
\[
Wf(u,s) = s^n \frac{1}{du^n} \left( f \ast \overline{\tilde{\psi}_s} \right)(u) = s^n \frac{1}{du^n} \left( f \ast \overline{\tilde{\psi}} \right)(u)
\]

(v) As
\[
\int t^n \psi(t) dt = \langle (-1)^n \tilde{\psi}(0) \rangle = \langle (-1)^n \tilde{\psi}(0) \rangle
\]
the \( n \)-th moment is only zero if \( \tilde{\psi}(0) = \int \theta(t) dt \neq 0 \)
At small scales, the wavelet transform converges to the derivative \( \text{wth} s \to 0 \) as:

\[
\lim_{s \to 0} \frac{1}{s^\frac{1}{5}} \tilde{\psi}_s (u) = K \cdot \psi (u).
\]

in the sense of weak convergence in \( \mathcal{D}' \), i.e.,

\[
\lim_{s \to 0} \psi * \frac{1}{s^\frac{1}{5}} \tilde{\psi}_s (u) = K \cdot \psi (u).
\]

Furthermore, if \( f \in C^n [u-r, u+r] \) then

\[
\lim_{s \to 0} \frac{Wf(u,s)}{s^{n+\frac{1}{5}}} = \lim_{s \to 0} f^{(n)} \ast \frac{1}{s^\frac{1}{5}} \tilde{\psi}_s (u) = K f^{(n)} (u),
\]

and if the \( n \)-th derivative is bounded then \( Wf(u,s) = O(s^{n+\frac{1}{5}}) \).

4.3. Measuring the regularity of signals with wavelets

It was seen earlier how wavelets allow the extraction of instantaneous frequencies and corresponding transient components of signals. Here we show that wavelets can also reveal the regularity. This is important as irregularities do contain the main information. Examples — Signal switched off, Sources of fields — breaking waves — edges in images.

Consider first the uniform regularity:

**Theorem.** Let the wavelet \( \psi \) have \( n \) vanishing moments and \( \psi^{(n)} \) have fast decay, i.e., \( \psi^{(n)} \to 0 \) for \( s \to \infty \) and \( n \in \mathbb{N} \). Then \( \psi^{(n)} \leq C n! s^{-\frac{n}{5}} \).

(i) If furthermore \( f \in L^2 (\mathbb{R}) \) is uniformly Lipschitz \( \alpha \) on \( [a,b] \), then \( F \in A^2 [a,b] \).

\( (*) \quad |Wf(u,s)| \leq A s^{\frac{n}{5} - \frac{1}{2}} \), \( (u,s) \in \mathbb{R} \times \mathbb{R}^+ \).

(ii) If conversely \( f \) is bounded and \( (*) \) holds for \( \alpha < n/5 \), then \( f \) is uniformly Lipschitz \( \alpha \) on \( [a,e, b-e] \), \( e > 0 \).

Proof uses pointwise regularity result.
$N.B. \ (*) \text{ is a condition for } s = 0 \text{ as by Cauchy-Schwarz.}$

\[ |Wf(u, s)| \leq \|f\|_{\ell^1} \|\psi_{us}\| \leq \|f\|_{\ell^1} \|\psi\|_{L^\infty}. \]

- Similar to Fourier of sound with respect.

For the following, let $\psi$ again have $n$ vanishing moments and $n$ derivatives have fast decay. Then one has

**Theorem (Jaffard 1991):** If $f \in L^1(\mathbb{R})$ is Lipschitz $\alpha < n$ at $x = 0$ then $f \in A^\alpha.$

\[ |Wf(u, s)| \leq As^{\alpha + \frac{1}{2}} \left( 1 + \frac{|u| - \sqrt{s}}{s} \right). \]

Conversely, if $\alpha < n$ and $A \in A^\alpha$ then $\psi$ is Lipschitz $\alpha$ at $x = 0.$

\[ |Wf(u, s)| \leq As^{\alpha + \frac{1}{2}} \left( 1 + \frac{|u - \sqrt{s}|}{s} \right). \]

Then $f$ is Lipschitz $\alpha$ at $x = 0.$ (use bilinearly)

**Proof:** (only first part, second is more difficult)

(i) If $f$ is Lipschitz $\alpha$ at $x = 0 \Rightarrow p(x, dy) = \rho(x) \delta_{\rho(x)} \leq K \rho(x) \delta_{\rho(x)}$ \( \leq K \mathcal{L}^{-1}(\rho). \)

(ii) If $n$ vanishing moments $\Rightarrow Wf(u, s) = 0.$

\[ |Wf(u, s)| = \left| \int f(t) \psi(t - u) \frac{1}{t} \left( \frac{t - u}{s} \right)^{1/2} dt \right|. \]

\[ x = \frac{t - u}{s} \Rightarrow \int K \mathcal{L}^{-1}(\rho) \frac{1}{s} \left( \frac{t - u}{s} \right)^{1/2} \frac{1}{\sqrt{s}} \left( \frac{t - u}{s} \right) |d\psi(t)| \]

\[ \leq \sqrt{s} \int K \mathcal{L}^{-1}(\rho) \frac{1}{\sqrt{s}} |d\psi(t)| dx \]

Since $\mathcal{L}^{-1}(\rho) \leq 2 \mathcal{L}^{-1}(\rho)$

\[ |Wf(u, s)| \leq K 2^{1/2} \int |\mathcal{L}^{-1}(\rho)(x)| dx \]

\[ \leq K 2^{1/2} \left( \int \left\{ 1 + (x^2 + 1) \mathcal{L}^{-1}(d\psi(x)) \right\} dx \right)^{1/2} \]
The bounds in the theorem are not quite Lipschitz for the wavelet transform but can be seen to be Lipschitz if \( \phi \) is compactly supported, say, \( \text{supp } \phi = [-C, +C] \). Define then the cone of influence of \( \nu \in \mathbb{R} \) in the scale-space:

\[
\begin{pmatrix} x \\ u \end{pmatrix} \in \mathbb{R}^2, \quad \text{for } |u - \nu| < C \cdot s
\]

The value \( \phi^\alpha f(\nu) \) will only have an impact on \( Wf(u, s) \) if \( (u, s) \) are in the cone of influence. In this case one gets from the previous bound:

\[
|Wf(u, s)| \leq A(1 + C^\alpha) \cdot s^\alpha + \nu^\alpha
\]

which is a Lipschitz condition again. Note that the largest wavelet coefficients are in the cone of influence of the singular points.