From Fourier to time-frequency analyses and perspectives

Pisa PhD Course

S. Meignen

May 2018

Contents

1 Continuous time Fourier transform **5**

1.1 Fourier transform in $L^1(\mathbb{R})$ . . . 5

1.1.1 Density theorems . . . 5

1.1.2 Definition of the Fourier transform in $L^1(\mathbb{R})$ . . . 5

1.1.3 Riemann-Lebesgue Theorem . . . 6

1.1.4 Example . . . 6

1.1.5 Other Properties . . . 7

1.1.6 Inversion of the Fourier transform in $L^1(\mathbb{R})$ . . . 8

1.1.7 Convolution product in $L^1(\mathbb{R})$ . . . 10

1.1.8 Illustration: moving average . . . 11

1.1.9 Convolution and Fourier transform . . . 11

1.2 Fourier transform on $L^2(\mathbb{R})$ . . . 12

1.2.1 The space $L^2(\mathbb{R})$ . . . 12

1.2.2 Convolution in $L^2(\mathbb{R})$ . . . 12

1.2.3 Property of the Fourier Transform in $L^1(\mathbb{R}) \bigcap L^2(\mathbb{R})$ . . . 13

1.2.4 Fourier transform in $L^2(\mathbb{R})$ . . . 14

1.2.5 Property of the Fourier transform in $L^2(\mathbb{R})$ . . . 15

1.3 Exercises . . . 15

2 Fourier Transform of Discrete sequences **19**

2.1 Motivations for the introduction of distributions . . . 19

2.1.1 The space of test functions . . . 19

2.1.2 Definitions of the distribution space . . . 20

2.1.3 Convergence in the distribution space . . . 21

2.1.4 Derivation in the distribution space . . . 22

2.2 Fourier Transform of Distributions . . . 23

2.3 The Schwartz class . . . 23

2.4 The space of tempered distributions $\mathcal{S}'(\mathbb{R})$ . . . 24

2.5 Fourier Transform in $\mathcal{S}'(\mathbb{R})$ . . . 25

2.5.1 Distributions with compact support . . . 27

2.5.2 Convolution $\mathcal{E}'(\mathbb{R}) * \mathcal{D}'(\mathbb{R})$ . . . 28

2.6 Exercises . . . 29

3 Linear Time-Frequency Analysis **31**

3.1 Linear Time-Frequency analysis: the continuous time framework . . . 31

3.1.1 Continuous Time Short Time Fourier Transform . . . 31

3.1.2 Discrete-Time Short-Time Fourier Transform 32
3.1.3 Short-Time Fourier Transform for finite length signal and filter 34
3.2 Applications in Matllab 36
3.2.1 Lab session 36
3.2.2 Files used in Lab session 38

4 Reassignment Techniques and Synchrosqueezing 47
4.1 Introduction 47
4.2 Notation and Multicomponent Signal Definition 48
4.3 Uncertainty Principle for Multicomponent Signals 48
4.4 Time-Frequency Representation Enhancement with Reassignment 49
4.5 Multicomponent Signal Reconstruction with Synchrosqueezing 50
4.5.1 SST in a Nutshell 51
4.5.2 Mathematical Foundations of the SST 52
4.5.3 Denoising Multicomponent Signals using SST 53
4.5.4 Adapting the SST to a Stronger Modulation 54
4.5.5 Computation of the new IF estimate 55
4.6 Applications of SST to Physiological Signals 56

Chapter 1

Continuous time Fourier transform

1.1 Fourier transform in $L^1(\mathbb{R})$

1.1.1 Density theorems

Definition 1 *Let f be a continuons function on a open set Ω of $\mathbb{R}^N$. The support of the function f which we denote by Supp(f) is the complement set in Ω of the largest open set on which f is null.*

$$\mathrm{Supp}\,(f) = \overline{\{x \in \Omega \quad | \quad f(x) = 0\}}$$

Definition 2 *$C_0(\Omega)$ stands for the vector space of the functions continuous on Ω and compactly supported, i.e. :*

$$C_0(\Omega) = \{f \in C(\Omega) \quad | \quad \exists K \, compact \; set \; , K \subset \Omega s.t. \; x \in \Omega \backslash K \;\; f(x) = 0\}$$

Theorem 1 Density theorem

Let $\Omega \subset \mathbb{R}^N$ be an open set. $C_0(\Omega)$ is dense in $L^p(\Omega)$ for $p \in \{1, 2\}$ i.e. :

$$\forall\, p \in \{1,2\},\, \forall\, \varepsilon > 0,\, \forall\, f \in L^p(\Omega),\, \exists g \in C_0(\Omega)\, \|f - g\|_p \leq \varepsilon$$

Remark: This theorem remains true for functions in $C_0^k(\Omega)$, $k \leq \infty$.

1.1.2 Definition of the Fourier transform in $L^1(\mathbb{R})$

Definition 3 (Fourier Transform) *Lett $f \in L^1(\mathbb{R})$, we define the Fourier transform $\hat{f}$ of f as:*

$$\forall \nu \in \mathbb{R},\; \hat{f}(\nu) \stackrel{def}{=} \int_{-\infty}^{+\infty} f(x) e^{-2i\pi\nu x} dx$$

ν is called the frequency (Hz) The application: $\mathcal{F} : f \mapsto \hat{f}$ is called Fourier transform.

1.1.3 Riemann-Lebesgue Theorem

Theorem 2 (Riemann-Lebesgue)

1. $\mathcal{F} : f \mapsto \hat{f}$ is a linear application, continuous from $L^1(\mathbb{R})$ onto $L^\infty(\mathbb{R})$.
2. if $f \in L^1(\mathbb{R})$, then $\hat{f}$ is continuous on $\mathbb{R}$ and $\lim_{\nu \to \pm\infty} \hat{f}(\nu) = 0$.

Proof

1. - $\mathcal{F}$ is linear (linearity of $\int$).

   - To show the continuity of $\mathcal{F}$ it suffices to prove the result in 0:

$$\forall f \in L^1(\mathbb{R}),\ \|\mathcal{F}(f)\|_{L^\infty(\mathbb{R})} \leq C\|f\|_{L^1(\mathbb{R})}$$

$$\forall \nu,\ \left|\hat{f}(\nu)\right| = \left|\int_{-\infty}^{+\infty} f(x)e^{-2i\pi\nu x}\right| \leq \int_{-\infty}^{+\infty} |f(x)|\,dx = \|f\|_{L^1(\mathbb{R})}$$

   so $\hat{f} \in L^\infty(\mathbb{R})$ et $\|\hat{f}\|_{L^\infty(\mathbb{R})} \leq \|f\|_{L^1(\mathbb{R})}$.

2. Let $g \in C_0^1(\mathbb{R})$, then

$$\begin{aligned} \hat{g}(\nu) &= \int_{\mathbb{R}} g(x)e^{-2i\pi\nu x}dx = \left[g(x)\frac{e^{-2i\pi\nu x}}{-2i\pi\nu}\right]_{-\infty}^{+\infty} - \int_{\mathbb{R}} g'(x)\frac{e^{-2i\pi\nu x}}{-2i\pi\nu}dx \\ |\hat{g}(\nu)| &\leq \frac{1}{2\pi\,|\nu|}\int_{\mathbb{R}} \left|g'(x)\right| dx \xrightarrow[\nu\to\pm\infty]{} 0 \text{ since } \left\|g'\right\|_{L^1(\mathbb{R})} < +\infty. \end{aligned}$$

   But $C_0^1(\mathbb{R})$ is dense in $L^1(\mathbb{R})$,so : $\forall f \in L^1(\mathbb{R}),\ \forall \varepsilon > 0,\ \exists g \in C_0^1\mathbb{R},\ \|f - g\|_{L^1(\mathbb{R})} < \varepsilon$.
   Then, as

$$|\hat{f}(\nu)| \leq \|f - g\|_{L^1(\mathbb{R})} + |\hat{g}(\nu)|,$$

   we get $\lim_{\nu \to \pm\infty} \hat{f}(\nu) = 0$.

1.1.4 Example

$\Pi = \mathbb{1}_{]-\frac{1}{2};\frac{1}{2}[}$, $\hat{\Pi}(\nu) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-2i\pi\nu x}dx = \frac{\sin(\pi\nu)}{\pi\nu}$: cardinal sine function.

1.1.5 Other Properties

Proposition 1 (delay)

$f \in L^1(\mathbb{R})$, $\tau \in \mathbb{R}$
$\forall x \in \mathbb{R}$, $g(x) = f(x-\tau)$ for $g \in L^1(\mathbb{R})$, then $\forall \nu \in \mathbb{R}$, $F(g)(\nu) = \hat{g}(\nu) = e^{-2i\pi\nu\tau}\hat{f}(\nu)$

Proposition 2

$f \in L^1(\mathbb{R})$, $a > 0$.
$\forall x \in \mathbb{R}$, $g(x) = f(ax)$ with $g \in L^1(\mathbb{R})$. $\forall \nu \in \mathbb{R}$, $\mathcal{F}(g(\nu)) = \frac{1}{a} f(\frac{\nu}{a})$.

Theorem 3

1. if $x \to x^k f(x)$ is in $L^1(\mathbb{R})$ for $k \in \{0, \cdots, n\}$ then $\hat{f}$ is n times differentiable, and one has:
$$\hat{f}^{(k)}(\nu) = \widehat{g_k}(\nu) \quad \forall \nu \in \mathbb{R}$$
where $g_k(x) = (-2i\pi x)^k f(x)$
2. If $f \in L^1(\mathbb{R}) \bigcap C^n(\mathbb{R})$ and if $f^{(k)} \in L^1(\mathbb{R})$ then for all $k \in \{1, \cdots, n\}$ one has:
$$\widehat{f^{(k)}}(\nu) = (2i\pi\nu)^k \hat{f}(\nu) \quad \forall \nu \in \mathbb{R}$$
3. If $f \in L^1(\mathbb{R})$ and if supp(f) is bounded, then $\hat{f} \in C^\infty$.

Proof

1. For all $k \leq n$, $\frac{\partial^k f(x) e^{-2i\pi\nu x}}{\partial^k \nu}$ is continuous for all ν and almost all x. Furthermore, $|\frac{\partial^k f(x) e^{-2i\pi\nu x}}{\partial^k \nu}| = |(-2i\pi x)^k f(x)|$ belongs to $L^1(\mathbb{R})$ and $\hat{f}$ belongs to C^k and then one applies the theorem on the differentiation of an integral dependent on a parameter.

2. Let us compute $\widehat{f'}$. By integrating by parts, we get that:
$$\widehat{f'}(\nu) = [f(x)e^{-2i\pi\nu x}]_{-\infty}^{\infty} + \int_{\mathbb{R}} f(x)(2i\pi\nu)e^{-2i\pi\nu x} dx.$$
Here we need to remark that if f is integrable and belongs to C^1, and is such that f' is also integrable then
$$f(x) = f(a) + \int_a^x f'(t)dt.$$
As f' is integrable, the integral has a limit when x tends to $\pm\infty$, so $f(x)$ has a limit when x tends to infinity. Moreover, this limit is necessarily null since f is integrable. We thus get $\widehat{f'}(\nu) = (2i\pi\nu)\hat{f}(\nu)$. Reasoning by induction, we get the expected result.

Compute the following Fourier transform

- Let $-\infty < a < b < +\infty$ et $f = \chi_{[a,b]}$.
- We denote $u(t)$ the Heavydide function (equal to 1 if $t > 0$ et 0 otherwise). $sgn(t)$ is the sign function. Let α be a complex number with positive real part. Compute the following Fourier transforms:

$$\begin{array}{ll} i) f(t) = e^{-\alpha t}u(t) & ii) f(t) = e^{\alpha t}u(-t) \\ iii) f(t) = e^{-\alpha|t|} & iv) f(t) = \frac{t^k}{k!}e^{-\alpha t}u(t) \\ v) f(t) = \frac{t^k}{k!}e^{\alpha t}u(-t) & vi) f(t) = sign(t)e^{-\alpha|t|} \end{array}$$

- Compute $\mathcal{F}(f)$ with $f : x \mapsto e^{-\pi x^2}$

Proposition 3

Let $f, g \in L^1(\mathbb{R})$, then $f\hat{g}$ et $\hat{f}g$ both belong to $L^1(\mathbb{R})$ and one has:

$$\int_{\mathbb{R}} f\hat{g} = \int_{\mathbb{R}} \hat{f}g$$

1.1.6 Inversion of the Fourier transform in $L^1(\mathbb{R})$

Definition 4 *For any function f belonging to $L^1(\mathbb{R})$ let us write:*

$$\overline{\mathcal{F}}(f)(\nu) = \int_{\mathbb{R}} f(x)e^{2i\pi\nu x}dx.$$

One then have the following inversion theorem:

Theorem 4

1. Let $f \in L^1(\mathbb{R})$. Let us assume f is continuous at $x \in \mathbb{R}$ and that $\hat{f} \in L^1(\mathbb{R})$. Then,
$$\overline{\mathcal{F}}\hat{f}(x) = f(x)$$
2. Let $f \in L^1(\mathbb{R})$ and $\hat{f} \in L^1(\mathbb{R})$ then
$$\overline{\mathcal{F}}\hat{f}(x) = f(x) \text{ for almost all } x$$

Proof 1) Let us first prove the first point. For $n \in \mathbb{N}^*$, let us define $g_n(x) = e^{-\frac{2\pi}{n}|x|}$, for which we get $\widehat{g_n}(\nu) = \frac{1}{\pi}\frac{n}{1+n^2\nu^2}$. Since g_n is in $L^1(\mathbb{R})$, we can write:

$$\int_{\mathbb{R}} \hat{f}(\nu)g_n(\nu)e^{2i\pi x\nu}d\nu = \int_{\mathbb{R}} f(\nu)\widehat{g_n}(\nu - x)d\nu$$

The term on the left hand side tends to $\overline{\mathcal{F}}\hat{f}(x)$ using the dominated convergence theorem. Let us show that the term on the right hand side tends to $f(x)$. As $\int_{\mathbb{R}} \widehat{g_n}(\nu)d\nu = 1$, one may write

$$\int_{\mathbb{R}} f(\nu)\widehat{g_n}(\nu - x)d\nu - f(x) = \int_{\mathbb{R}} (f(\nu + x) - f(x))\widehat{g_n}(\nu)d\nu.$$

Let $\epsilon > 0$, there exists $\eta = \eta(\epsilon, x)$ such that $|y - x| \leq \eta \Rightarrow |f(y) - f(x)| \leq \epsilon$ (f continuous at x). One can then write:

$$\int_{\mathbb{R}} (f(x+\nu) - f(x))\widehat{g_n}(\nu)d\nu = \int_{|\nu|\leq\eta} (f(x+\nu) - f(x))\widehat{g_n}(\nu) + \int_{|\nu|\geq\eta} (f(x+\nu) - f(x))\widehat{g_n}(\nu).$$

For all $n \in \mathbb{N}^*$, one has:

$$\int_{|\nu|\leq\eta} |f(x+\nu) - f(x)|\widehat{g_n}(\nu)d\nu \leq \epsilon \int_{\mathbb{R}} \widehat{g_n}(\nu)d\nu = \epsilon.$$

Furthermore,

$$|\int_{|\nu|\geq\eta} f(x)\widehat{g_n}(\nu)|d\nu \leq |f(x)|(1 - \frac{2}{\pi}atan(\eta n)),$$

which tends to 0 when n tends to infinity. Furthermore, as $\widehat{g_n}$ is even and decreasing over $\mathbb{R}^+$

$$|\int_{|\nu|\geq\eta} f(x+\nu)\widehat{g_n}(\nu)| \leq \widehat{g_n}(\eta)\|f\|_1,$$

this expression tends to 0 when n tends to infinity. This proves the theorem.
2) Let us now show point 2. We multiply the function to be integrated by $\hat{h}_\epsilon(\nu) = e^{-\pi\epsilon^2\nu^2}$:

$$I_\epsilon = \int_{\mathbb{R}} (\int_{\mathbb{R}} f(u)e^{-\pi\epsilon^2\nu^2}e^{2i\pi\nu(x-u)}du)d\nu.$$

Then, we have $(u, \nu) \to \phi(u, \nu) = f(u)e^{-\pi\epsilon^2\nu^2}e^{2i\pi\nu(x-u)} \quad \in L^1(\mathbb{R}^2)$. By applying Fubini theorem, we get two different expressions of I_ϵ:

i) By integrating with respect to u, one gets $I_\epsilon = \int_{\mathbb{R}} \hat{f}(\nu)e^{-\pi\epsilon^2\nu^2}e^{2i\pi\nu x}d\nu$.
But since $|\hat{f}(\nu)e^{-\pi\epsilon^2\nu^2}e^{2i\pi\nu x}| \leq |\hat{f}(\nu)|$ which belongs to $L^1(\mathbb{R})$, and since $\lim_{\epsilon\to 0} e^{-\pi\epsilon^2\nu^2} = 1$, by applying the dominated convergence theorem, we get that $\lim_{\epsilon\to 0} I_\epsilon = \int_{\mathbb{R}} \hat{f}(\nu)e^{2i\pi\nu x}dx$.

ii) Integrating with respect to v:

$$I_\epsilon = \int_{\mathbb{R}} f(u) \left(\int_{\mathbb{R}} e^{-\pi\epsilon^2\nu^2}e^{2i\nu(x-u)}d\nu \right) du = \int_{\mathbb{R}} f(u)\frac{1}{\epsilon}e^{-\pi(\frac{x-u}{\epsilon})^2}du,$$

using the properties of the Fourier transforms of Gaussian functions and the dilation formula. Furthermore, we know that the function $h_\epsilon(x) = \frac{1}{\epsilon}e^{-\pi(\frac{x}{\epsilon})^2}$ has its integral equal to 1. One then deduce that:

$$\begin{aligned}
\int_{\mathbb{R}} |I_\epsilon(x) - f(x)| &= \int_{\mathbb{R}}\int_{\mathbb{R}} |(f(x-u) - f(x))h_\epsilon(u)|du \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} |f(x - \epsilon u) - f(x)|h_\epsilon(u)du \\
&\leq \int_{\mathbb{R}} \|f(x - \epsilon u) - f(x)\|_1 h(u)du.
\end{aligned}$$

In $L^1(\mathbb{R})$, one has the following property:

Proposition 4

Let $f \in L^1(\mathbb{R})$, $h \in \mathbb{R}$, and define $\tau_h f(x) = f(x-h)$. Then $\tau_h h \in L^1(\mathbb{R})$ et $\lim_{h\to 0} \|\tau_h f - f\|_1 = 0$.

Proof The theorem uses the density of $C_0(\mathbb{R})$ in $L^1(\mathbb{R})$. Indeed, let g_n be a sequence in $C_0(\mathbb{R})$ tending to f in $L^1(\mathbb{R})$, i.e.:

$$\forall \epsilon > 0 \ \exists N \ \forall n \geq N \ \|f - g_n\|_1 \leq \epsilon$$

One may then write:

$$\int_{\mathbb{R}} |f(x+\eta) - f(x)| \leq \int_{\mathbb{R}} |f(x+\eta) - g_n(x+\eta)| + \int_{\mathbb{R}} |g_n(x+\eta) - g_n(x)| + \int_{\mathbb{R}} |g_n(x) - f(x)|$$

Let N be such that $|f - g_N| \leq \frac{\epsilon}{3}$ and choose η such that $\|g_N(x+\eta) - g_N(x)\|_1 \leq \frac{\epsilon}{3}$ (dominated convergence theorem), hence the result.

Since $\|f(x-\epsilon u) - f(x)\|_1 |h(u)| \leq 2\|f\|_1 |h(u)|$ which belongs to $L^1(\mathbb{R})$, applying the dominated convergence theorem, we deduce that: $\lim_{\epsilon\to 0} \|I_\epsilon - f\|_1 = 0$.

So I_ϵ tends to f in $L^1(\mathbb{R})$ so there exists a sub-sequence $I_{\phi(\epsilon)}$ converging to f almost everywhere, hence the result.

1.1.7 Convolution product in $L^1(\mathbb{R})$

Theorem 5 (and definition)

$f \in L^1(\mathbb{R})$, $g \in L^1(\mathbb{R})$. Let us define : $\forall x \in \mathbb{R}$, $(f \star g)(x) = \int_{\mathbb{R}} f(y) g(x-y) dy$.
Then $(f \star g)$ is defined almost everywhere, integrable and $\|f \star g\|_{L^1(\mathbb{R})} \leq \|f\|_{L^1(\mathbb{R})} \|g\|_{L^1(\mathbb{R})}$.

Proof Using Fubini theorem:

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(y) g(x-y)| \, dy \right) dx = \int_{\mathbb{R}} |f(y)| \left(\int_{\mathbb{R}} |g(x-y)| \, dx \right) dy \tag{1.1}$$

and by changing variables $u = x - y$, we obtain:

$$(1.1) \quad = \quad \int_{\mathbb{R}} |f(y)| \left(\int_{\mathbb{R}} |g(u)| \, du \right) dy = \|f\|_{L^1(\mathbb{R})} \|g\|_{L^1(\mathbb{R})} < +\infty,$$

since $\int_{\mathbb{R}} |g(u)| \, du = \|g\|_{L^1(\mathbb{R})}$. So $x \mapsto \int_{\mathbb{R}} |f(y) g(x-y)| \, dy$ is integrable and thus finite almost everywhere. Consequently $(f \star g)$ is defined almost everywhere, integrable and:

$$\int_{\mathbb{R}} |(f \star g)(x)| \, dx \leq \|f\|_{L^1(\mathbb{R})} \|g\|_{L^1(\mathbb{R})}$$

Proposition 5

Let $f, g, h \in L^1(\mathbb{R})$.

- $f \star g = g \star f$
- $(f \star g) \star h = f \star (g \star h)$
- $(f + g) \star h = f \star h + g \star h$

1.1.8 Illustration: moving average

At each point $x \in \mathbb{R}$, one replaces $f(x)$ by its average $\bar{f}(x)$ over an interval of length τ :

$$\bar{f}(x) = \frac{1}{\tau} \int_{x-\frac{\tau}{2}}^{x+\frac{\tau}{2}} f(t)dt = \frac{1}{\tau} \int_{\mathbb{R}} \chi_{[x-\frac{\tau}{2};x+\frac{\tau}{2}]}(t) f(t)dt = \int_{\mathbb{R}} h(x-t) f(t)dt$$

where $h : u \mapsto \frac{1}{\tau} \mathbb{1}_{[-\frac{\tau}{2};\frac{\tau}{2}]}$.

In pratice

- choice of a more regular window.
- choice for τ depends on the scale of the phenomena one wants to highlight.

1.1.9 Convolution and Fourier transform

Theorem 6 (Convolution and Fourier transform)

i) Let $f \in L^1(\mathbb{R})$, $h \in L^1(\mathbb{R})$. Then $\forall \nu \in \mathbb{R}$, $\mathcal{F}(f \star h)(\nu) = \hat{f}(\nu)\hat{h}(\nu)$.

ii) Let $f \in L^1(\mathbb{R})$, $h \in L^1(\mathbb{R})$ such that $\hat{f}$ and $\hat{h}$ are also in $L^1(\mathbb{R})$, then for almost all ν, one has: $\hat{f} \star \hat{h} = \mathcal{F}(fh)$.

Example 1 $\mathcal{F}(\bar{f})(\nu) = \hat{h}(\nu)\hat{f}(\nu) = \frac{\sin(\pi\nu\tau)}{\pi\nu\tau}\hat{f}(\nu)$. *$\hat{h}$ is called transfer function. One can then adapt $\frac{1}{\tau}$ to the frequencies of interest in signal f.*

Proof i) Applying Tonelli's theorem: $\int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(y)g(x-y)|\, dy \right) \left| e^{-2i\pi\nu x} \right| dx = \|f\|_{L^1(\mathbb{R})} \|g\|_{L^1(\mathbb{R})} < +\infty$

since $\left|e^{-2i\pi\nu x}\right| = 1$. Then, from Fubini's theorem:

$$\begin{aligned}
\mathcal{F}(f \star g)(\nu) &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(y) g(x-y) dy \right) e^{-2i\pi\nu x} dx \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(y) g(x-y) dy \right) e^{-2i\pi\nu(x-y+y)} dx \\
&= \int_{\mathbb{R}} f(y) e^{-2i\pi\nu y} \left(\int_{\mathbb{R}} g(x-y) e^{-2i\pi\nu(x-y)} dx \right) dy \\
&= \int_{\mathbb{R}} f(y) e^{-2i\pi\nu y} \left(\int_{\mathbb{R}} g(u) e^{-2i\pi\nu u} du \right) dy = \hat{f}(\nu)\hat{g}(\nu)
\end{aligned}$$

ii) Since $\hat{f}$ and $\hat{g}$ are both in $L^1(\mathbb{R})$, we get, remarking that $\overline{\mathcal{F}}$ has the same properties as $\mathcal{F}$:

$$\begin{aligned}
\overline{\mathcal{F}}(\hat{f} \star \hat{g}) &= \overline{\mathcal{F}}(\hat{f})\overline{\mathcal{F}}(\hat{g}) \\
&= fg \text{ almost everywhere}
\end{aligned}$$

Finally, since $f = \overline{\mathcal{F}}(\hat{f})$, f is bounded one can compute the Fourier transform of fg to obtain: $\hat{f} \star \hat{g} = \mathcal{F}(fg)$.

1.2 Fourier transform on $L^2(\mathbb{R})$

One of the main drawback with considering the Fourier transform in $L^1(\mathbb{R})$, is its non invertibility in general. In what follows, we are going to see how to define the Fourier transform on $L^2(\mathbb{R})$ as a bijective application from $L^2(\mathbb{R})$ onto $L^2(\mathbb{R})$.

1.2.1 The space $L^2(\mathbb{R})$

Let $f, g \in L^2(\mathbb{R})$, we recall that $L^2(\mathbb{R})$ is equipped with the inner product $\langle f, g \rangle = \int_{\mathbb{R}} f(x)\overline{g(x)}dx$ and that the norm on $L^2(\mathbb{R})$ is defined by:$\|f\|_2 = \sqrt{\langle f, f \rangle}$. $L^2(\mathbb{R})$ is an Hilbert space for which one has the Cauchy-Schwarz theorem:

Theorem 7 (Cauchy-Schwarz)

Let f and g belong to $L^2(\mathbb{R})$, we then have the following property:

$$\left| \int_{\mathbb{R}} f(t)\bar{g}(t) \right| \leq \sqrt{\int_{\mathbb{R}} |f|^2(t)dt} \sqrt{\int_{\mathbb{R}} |g|^2(t)dt}$$

1.2.2 Convolution in $L^2(\mathbb{R})$

Convolution in $L^2(\mathbb{R})$ is defined for f and g in $L^2(\mathbb{R})$ by $F(x) = \int_{\mathbb{R}} f(x-y)g(y)dy$, satisfying :

Theorem 8

i) $F \in C^0(\mathbb{R}) \bigcap L^\infty(\mathbb{R})$

ii) F is continuous

Proof i) direct application of Cauchy-Schwarz theorem
ii) We have

$$\begin{aligned} |F(x+\eta) - F(x)| &= |\int_{\mathbb{R}} (f(x+\eta-y) - f(x-y))\, g(y)|dy \\ &\leq \int_{\mathbb{R}} |f(x+\eta-y) - f(x-y)|^2 dy \ \|g\|_2. \end{aligned}$$

The term depending on η tends to 0 with η (to prove it we use the density of $C_0(\mathbb{R})$ in $L^2(\mathbb{R})$). So, f is continuous at x.

Example : the correlation in $L^2(\mathbb{R})$ is defined by:

$$G(x) = \int_{\mathbb{R}} f(x+t)\bar{f}(t)dt = \check{f} \star \bar{f}(-x),$$

which is continuous.

1.2.3 Property of the Fourier Transform in $L^1(\mathbb{R}) \bigcap L^2(\mathbb{R})$

Theorem 9 (Plancherel-Parseval)

Let f and h belonging to $L^2(\mathbb{R}) \bigcap L^1(\mathbb{R})$, then one has:

$$\int_{\mathbb{R}} f(t)\overline{h(t)}dt = \int_{\mathbb{R}} \hat{f}(\nu)\overline{\hat{h}(\nu)}d\nu$$

If $f = h$, one has the following property : $\int_{\mathbb{R}} |f|^2 = \int_{\mathbb{R}} |\hat{f}|^2$

Proof We first prove that the Fourier transform of a function in $L^1(\mathbb{R}) \bigcap L^2(\mathbb{R})$ is in $L^2(\mathbb{R})$ showing that $\|f\|_2^2 = \|\hat{f}\|_2^2$.
Let us first consider $g_\alpha(x) = e^{-\alpha x^2}$, whose Fourier transform is $\hat{g}_\alpha(x) = \sqrt{\frac{\pi}{\alpha}} e^{-\frac{\pi^2 x^2}{\alpha}}$. Applying the monotone convergence theorem, one gets:

$$\int_{\mathbb{R}} g_\alpha(x)|\hat{f}(x)|^2 \underset{\alpha \to 0}{\to} \int_{\mathbb{R}} |\hat{f}|^2 \leq +\infty$$

since $g_\alpha(x)|\hat{f}(x)|^2$ is positive, belongs to $L^1(\mathbb{R})$ and is increasing when α decreases.

Moreover, as the function $(x,u,y) \to f(y)\overline{f(u)}e^{i2\pi x(u-y)}g_\alpha(x)$ is in $L^1(\mathbb{R}^3)$ (applying Tonnelli's theorem).

$$\int_{\mathbb{R}} g_\alpha(x)|\hat{f}(x)|^2 = \int_{\mathbb{R}} f(y)\int_{\mathbb{R}}\overline{f(u)}\int_{\mathbb{R}} e^{-i2\pi x(y-u)}g_\alpha(x)dxdydu$$
$$= \int_{\mathbb{R}} f(y)\int_{\mathbb{R}}\overline{f(u)}\hat{g}_\alpha(y-u)dydu = \int_{\mathbb{R}}\int_{\mathbb{R}} f(y+u)\overline{f(u)}du\ \hat{g}_\alpha(y)dy$$
$$= \int_{\mathbb{R}} G(y)\hat{g}_\alpha(y)dy = \int_{\mathbb{R}} G(\sqrt{\frac{\alpha}{\pi}}y)e^{-\pi y^2}dy$$
$$\underset{\alpha\to 0}{\to} G(0) = \|f\|_2^2,$$

the limit being obtained applying the dominated convergence theorem, this means that the Fourier transform of a function in $L^1(\mathbb{R})\bigcap L^2(\mathbb{R})$ is in $L^2(\mathbb{R})$.
We are now going to show Plancherel formula. Let f and h be in $L^1(\mathbb{R})\bigcap L^2(\mathbb{R})$. $\hat{f}\overline{\hat{h}}$ belongs to $L^1(\mathbb{R})$ as a product of functions in $L^2(\mathbb{R})$. Moreover, defining $\check{h}(t) = \bar{h}(-t)$, one has $\mathcal{F}(f * \check{h}) = \hat{f}\overline{\hat{h}}$ which belongs to $L^1(\mathbb{R})$. So, from the inversion theorem of the Fourier transform, and as $f * \check{h}$ is continuous being, the convolution of functions in $L^2(\mathbb{R})$, one has $f * \check{h}(x) = \overline{\mathcal{F}}(\hat{f}\overline{\hat{h}})(x)$, for all x. Considering its value at $x = 0$, one gets Plancherel inequality.

1.2.4 Fourier transform in $L^2(\mathbb{R})$

Proposition 6

$L^1(\mathbb{R})\bigcap L^2(\mathbb{R})$ is dense in $L^2(\mathbb{R})$.

Proof Let us define $f_N(x) = \chi_{[-N,N]}(x)f(x)$ which belongs to $L^1(\mathbb{R})\bigcap L^2(\mathbb{R})$, one checks that f_N tends to f in $L^2(\mathbb{R})$.

Let f_N be a sequence of functions in $L^1(\mathbb{R})\bigcap L^2(\mathbb{R})$ converging to f in $L^2(\mathbb{R})$. We have seen that $\hat{f}_N$ belongs to $L^2(\mathbb{R})$, furthermore $\hat{f}_N$ is a Cauchy sequence in $L^2(\mathbb{R})$ since

$$\|\hat{f}_N - \hat{f}_P\|_2^2 = \int_{N\geq x\geq P} |\hat{f}_N - \hat{f}_P|^2 \to 0 \text{ when P tends to infinity.}$$

We then define $\hat{f}_\infty$ the limit in $L^2(\mathbb{R})$ of $\hat{f}_N$.
It remains to show that this limit is independent of the choice of sequence f_N tending to f. It is easing to see that this arises from Parseval equality. Indeed, let f_N and $\tilde{f}_N$ in $L^1(\mathbb{R})\bigcap L^2(\mathbb{R})$ tending to f in $L^2(\mathbb{R})$, then:

$$\|f_N - \tilde{f}_N\|_2 = \|\hat{f}_N - \widehat{\tilde{f}_N}\|_2 \to 0,$$

meaning the Fourier transforms have the same limit in $L^2(\mathbb{R})$.
We then have the following definition:

Definition 5 *The Fourier transform of a function $f \in L^2(\mathbb{R})$ is defined as the limit in $L^2(\mathbb{R})$ of the Fourier transform of any $f_N \in L^1(\mathbb{R})\bigcap L^2(\mathbb{R})$ tending to f in $L^2(\mathbb{R})$.*

In the sequel, we will note $\mathcal{F}(f)$ the Fourier transform of f when the latter is in $L^2(\mathbb{R})$.
Remark: for the sake of simplicity, one takes $f_N = \chi_{[-N,N]}f$.

1.2.5 Property of the Fourier transform in $L^2(\mathbb{R})$

Theorem 10

The Fourier transform (resp $\overline{\mathcal{F}}$) can be extended into an isometry from $L^2(\mathbb{R})$ onto $L^2(\mathbb{R})$. Let us denote $\mathcal{F}$ et $\overline{\mathcal{F}}$, these extensions, one then gets:

- $\forall f \in L^2(\mathbb{R})\ \mathcal{F}\overline{\mathcal{F}}(f) = \overline{\mathcal{F}}\mathcal{F}(f) = f$ almost everywhere.
- $\forall f, g \in L^2(\mathbb{R})\ \int_{\mathbb{R}} f(x)\overline{g(x)}dx = \int_{\mathbb{R}} \mathcal{F}(f)\overline{\mathcal{F}}(g)d\xi$
- $\forall f \in L^2(\mathbb{R})\ \|f\|_2 = \|\mathcal{F}(f)\|_2$

Proof The proof stems from the density theorem of functions of $L^1(\mathbb{R}) \bigcap L^2(\mathbb{R})$ in $L^2(\mathbb{R})$ (the equalities being true in $L^1(\mathbb{R}) \bigcap L^2(\mathbb{R})$, they are also true in $L^2(\mathbb{R})$).

1.3 Exercises

Exercise 1 *Properties of the Fourier transform F*
Show the following properties:

1. $F(f + \lambda g) = \hat{f} + \lambda \hat{g}$, $\forall f, g \in L^1(\mathbb{R})$, $\forall \lambda \in \mathbb{R}$.
2. $F[f(ax)](\nu) = \frac{1}{|a|}\hat{f}(\frac{\nu}{a})$, $\forall f \in L^1(\mathbb{R})$, $\forall a \in \mathbb{R}^*$.
3. $F[f(x - \tau)](\nu) = e^{-2i\pi\nu\tau}\hat{f}(\nu)$, $\forall f \in L^1(\mathbb{R})$, $\forall \tau \in \mathbb{R}$.
4. $F[f'](\nu) = 2i\pi\nu\hat{f}(\nu)$, $\quad \forall f \in L^1(\mathbb{R}) \bigcap C^1(\mathbb{R})$, such that $f' \in L^1(\mathbb{R})$.
5. Compute $F[xf(x)](\nu)$ as a function of $\hat{f}(\nu)$ (hypotheses on f?).

Exercise 2 *Computation of simple Fourier transforms*

1. Compute Fourier transform of $f(x) = e^{-|x|}$, that of $g(x) = U(x)f(x)$, U Heavyside function.
2. Compute Fourier transform of $\rho_n(x) = n\Pi(nx)$ (Π indicator function of $[-1/2, 1/2]$).
3. Plot ρ_n and $\hat{\rho}_n$. What happens when $n \to +\infty$?
4. *Modulation* :Compute $F[\cos(2\pi\nu_0 x)f(x)]$. Example : $f(x) = \chi_{[-a,a]}(x)$.

Exercise 3 *Computation of the Fourier transform of* $f(x) = e^{-\pi x^2}$.

1. Check that $f \in L^1(\mathbb{R})$
2. Show that f is solution to the following differential equation

$$y' + 2\pi xy = 0 \tag{1.2}$$

3. Compute Fourier transform of (1.2) and deduce differential equation satified by $\hat{f}$.
4. Deduce the computation of $\hat{f}$.

Exercise 4 *Door function*

1. Compute $\hat{\Pi}$. Check that $\lim_{\nu\to\infty}\hat{\Pi}(\nu)=0$.
2. Deduce the value of the integral:

$$\int_{-\infty}^{+\infty}\frac{\sin\pi\nu}{\pi\nu}e^{2i\pi\nu x}dx$$

3. Deduce that (difficult question):

$$\int_{-\infty}^{+\infty}\frac{\sin t}{t}dt=\pi$$

4. Compute $\int_{-\infty}^{+\infty}\Pi^2(x)dx$
5. Deduce that :

$$\int_{-\infty}^{+\infty}\left(\frac{\sin t}{t}\right)^2dt=\pi$$

Exercise 5 *Hat function*
Let Λ be the piecewise affine function, equal to 0 on $]-\infty,-1]$ and $[1,+\infty[$, with value 1 at $x=0$.
1. Give the expression of $\Lambda(x)$.
2. Show that $\Lambda'(x)=\Pi(x+1/2)-\Pi(x-1/2)$.
3. Compute the Fourier transform of Λ'. Deduce that of Λ.

Exercise 6 *On the relation between Fourier transform and Fourier coefficients*
f_0 a function of $L^1(\mathbb{R})$, null outside the interval $[0,T]$. f T-periodic extension of f:

$$f(x)=\sum_{n\in\mathbb{Z}}f_0(x+nT)$$

1. Since f periodic function integrable on $[0,T]$, show that Fourier coefficients of f, $c_n(f)$ satisfy:

$$c_n(f)=\frac{1}{T}\hat{f}_0(\frac{n}{T})$$

where $\hat{f}_0$ is the Fourier transform of function f_0.

Exercise 7 Let us define

$$f(x)=\int_0^{\infty}\frac{1}{\sqrt{t}}e^{-\frac{x^2}{2t}-\frac{t}{2}}dt$$

1. Show that $f\in L^1(\mathbb{R}^+)$.
2. Compute $\hat{f}$. Deduce that $f(x)=\sqrt{2\pi}e^{-|x|}$.

Exercise 7 Let $f(x)=\frac{\sin x}{|x|}$ and $f_\lambda(x)=e^{-\lambda|x|}\frac{\sin x}{|x|}$ $(\lambda>0)$.
1. Show that f and f_λ belong to $L^2(\mathbb{R})$, and that if λ tends to 0, f_λ converges to f in $L^2(\mathbb{R})$.
2. Compute of a fixed ξ, $\frac{\partial}{\partial\lambda}\hat{f}_\lambda(\xi)$. Deduce $\hat{f}_\lambda(\xi)$ and then $\hat{f}(\xi)$.

Exercise 8 Let a and b two real numbers such that $a,b>0$ et $a\neq b$.

1. Compute the Fourier transform of $e^{-a|x|}$.
2. Deduce the values of the following convolution products: $\frac{1}{a^2+x^2} * \frac{1}{b^2+x^2}$ and $e^{-a|x|} * e^{-b|x|}$.

Exercise 9 *Heat equation de la chaleur*
Let us consider the following partial derivatives equation:

$$\begin{cases} \frac{\partial^2 f}{\partial x^2} = \frac{\partial f}{\partial t} \\ f(x,0) = \varphi(x) \end{cases} \tag{1.3}$$

where φ belongs to $C_c^\infty(\mathbb{R})$. Let us define:

$$F(\nu,t) = \int_{-\infty}^{+\infty} f(x,t)\ e^{-2i\pi\nu x} dx$$

1. Let us assume $f \in L^1(\mathbb{R})$. Check that F satisfies:

$$\frac{\partial F}{\partial t} + 4\pi^2\nu^2 F = 0$$

2. Deduce F, and then f.

Chapter 2

Fourier Transform of Discrete sequences

The Fourier transform of Discrete sequences makes use of the theory of distributions of which we give a brief introduction.

2.1 Motivations for the introduction of distributions

The theory of distributions has been introduced to extend the notions of functions and that of derivation. It is the basis to the unification of discrete and continuous phenomena, and are widely used in mechanical physics, electronic, and probabilities.
To model impulses, the physician P. Dirac had the idea, around 1920 to use a pseudo-fonction, already introduced by par O. Heaviside, now known as the Dirac distribution and assumed to satisfy:

$$\delta_a(x) = \begin{cases} +\infty \text{ if } & x = a \\ 0 \text{ otherwise} & \end{cases}$$

and, for any continuous function ϕ

$$\int_{-\infty}^{+\infty} \delta_a(x)\varphi(x)dx = \varphi(a).$$

δ_a is definitely not a function but one had to wait until the years 1945-1950, and the work by L. Schwartz, for a proper mathematical definition of this object. This is the main motivation to the introduction of distribution theory.

2.1.1 The space of test functions

The distributions are going to be defined as applications on a function space which is called the *space of test functions.*

Definition 1 *One defines $\mathcal{D}(\Omega)$ (also denoted $C_0^\infty(\Omega)$) the set of smooth functions (admitting derivatives of any orders) defined on Ω, with values in $\mathbb{C}$, and compactly supported in Ω.*

$\mathcal{D}(\Omega)$ is a vector space.

Remark 1 *Let* $\varphi \in \mathcal{D}(\Omega)$. $\mathrm{Supp}\,(\varphi)$ *is a compact set and* $\mathrm{Supp}\,(\varphi) \subset \Omega$. *If* $\tilde{\varphi}(x) = \begin{cases} 0 & si\ x \in \mathbb{R} \setminus \mathrm{Supp}\,(\varphi) \\ \varphi(x) & si\ x \in \mathrm{Supp}\,(\varphi) \end{cases}$ *then* $\tilde{\varphi} \in \mathcal{D}(\mathbb{R})$.

Example 1 $\varphi(x) = \begin{cases} \exp \frac{-1}{1-\|x\|^2} & si\ \|x\| < 1 \\ 0 & si\ \|x\| \geq 1 \end{cases}$

$\varphi \in \mathcal{D}(\mathbb{R})$, $\mathrm{Supp}\,(\varphi) = \overline{B(0,1)}$

Definition 2 *Convergence in* $\mathcal{D}(\Omega)$ *Let* φ_n *and* $\varphi \in \mathcal{D}(\Omega)$. *$\varphi_n$ converges to φ in $\mathcal{D}(\Omega)$ if:*

- $\exists K$ *a compact set ,* $K \subset \Omega$ *such that* $\forall\, n$, $\mathrm{Supp}\,(\varphi_n) \subset K$
- $\forall\, \alpha \in \mathbb{N}^N$, $\partial^\alpha \varphi_n \longrightarrow \partial^\alpha \varphi$ *uniformly.*

Theorem 1

$\mathcal{D}(\Omega)$ is dense in $L^p(\Omega)$, $1 \leq p < +\infty$.

2.1.2 Definitions of the distribution space

Definition 3 *A distribution T on Ω is a linear form continuous on $\mathcal{D}(\Omega)$, i.e.*

(i) $\forall\, \varphi_1, \varphi_2 \in \mathcal{D}(\Omega)$, $\forall\, \lambda \in \mathbb{C}$, $T(\varphi_1 + \lambda\varphi_2) = T(\varphi_1) + \lambda T(\varphi_2)$

(ii) If $\varphi_n \longrightarrow \varphi$ *in* $\mathcal{D}(\Omega)$, *then* $T(\varphi_n) \longrightarrow T(\varphi)$ *in* $\mathbb{C}$

One notes $\langle T, \varphi \rangle$ or $T(\varphi)$.

Remark 2 *Point ii) is equivalent to showing: for any compact set* $K \subset \Omega$, *there exists* $C_k > 0$ *and* $k \in \mathbb{N}$ *such that for all* $\varphi \in \mathcal{D}(\Omega)$ *with* $\mathrm{Supp}\,(\varphi) \subset K$, *one has* $\langle T, \varphi \rangle \leq C_k \|\varphi\|_{C^k(K)}$, *with* $\|\varphi\|_{C^k(K)} = \max_{\alpha \leq k} \|\varphi^{(\alpha)}\|_{\infty, K}$.

One denotes $\mathcal{D}'(\Omega)$ the set of distributions on Ω, which is a vector space.
Examples

1. $L^1_{loc}(\Omega)$: set of mesurable functions on Ω, integrable on any compact set of Ω (for instance, $\frac{1}{\sqrt{|x|}} \in L^1_{loc}(\mathbb{R})$).
   Let $f \in L^1_{loc}(\Omega)$. For $\varphi \in \mathcal{D}(\Omega)$ one puts: $\langle T_f, \varphi \rangle = \int_\Omega f(x)\varphi(x)\, dx$
   $T_f \in \mathcal{D}'(\Omega)$:
   - T_f is well defined since
     $|f(x)\varphi(x)| \leq \|\varphi\|_\infty \mathbb{1}_{\mathrm{Supp}(\varphi)}(x)|f(x)| \in\ L^1(\Omega)$
   - T_f is linear (linearity of the integral).
   - T_f is continuous on $\mathcal{D}(\Omega)$:
     Let $\varphi_n \in \mathcal{D}(\Omega)$ be such that $\varphi_n \to 0$ dans $\mathcal{D}(\Omega)$.
     $|\langle T_f, \varphi_n \rangle| = \left| \int_\Omega f(x)\varphi_n(x)dx \right| \leq \int |f(x)|\, |\varphi_n(x)|\, dx \leq \|\varphi_n\|_\infty \int_K |f(x)|\, dx$

2. Let $a \in \mathbb{R}$. For all $\varphi \in \mathcal{D}(\mathbb{R})$ one puts: $\langle \delta_a, \varphi \rangle = \varphi(a)$

   $\delta_a \in \mathcal{D}'(\mathbb{R})$:

   - Linearity :
     $\langle \delta_a, \varphi_1 + \lambda\varphi_2 \rangle = (\varphi_1 + \lambda\varphi_2)(a) = \varphi_1(a) + \lambda\varphi_2(a) = \langle \delta_a, \varphi_1 \rangle + \lambda \langle \delta_a, \varphi_2 \rangle$
   - Continuity :
     If $\varphi_n \longrightarrow 0$ in $\mathcal{D}(\mathbb{R})$: $|\langle \delta_a, \varphi_n \rangle| = |\varphi_n(a)| \leq \|\varphi_n\|_\infty \longrightarrow 0$

   When $a = 0$, we put $\delta = \delta_0$.

Proposition 1

The application from $L^1_{loc}(\Omega)$ on $\mathcal{D}'(\Omega)$ which maps f to T_f is linear and injective.

Proof

- $\forall\, f_1, f_2 \in L^1_{loc}(\Omega)$, $\forall\, \lambda \in \mathbb{C}$, $T_{f_1+\lambda f_2} = T_{f_1} + \lambda T_{f_2}$
  Indeed, let $\varphi \in \mathcal{D}(\Omega)$,
  $$\langle T_{f_1+\lambda f_2}, \varphi \rangle = \int_\Omega (f_1 + \lambda f_2)\, \varphi = \int_\Omega f_1\varphi + \lambda \int_\Omega f_2\varphi = \langle T_{f_1}, \varphi \rangle + \lambda \langle T_{f_2}, \varphi \rangle$$
- If $\forall \varphi \in \mathcal{D}(\Omega), \langle T_f, \varphi \rangle = \displaystyle\int_\Omega f(x)\varphi(x)dx = 0$, alors $f = 0$. Indeed, since $\mathcal{D}(\Omega)$ is dense in $L^2(\Omega)$, letting φ_n a sequence coverging to f in L^2, then $\int_\Omega f(x)\varphi_n(x) = 0$ tends to $\int_{\mathbb{R}} |f(x)|^2 = 0$ and so f is null almost everywhere.

Remark 3 *The application defined by proposition 1 is not surjective, but enables to identify $L^1_{loc}(\Omega)$ to a subspace $\mathcal{D}'(\Omega)$ called* regular distributions.

2.1.3 Convergence in the distribution space

Definition 4 *A sequence of distribution $T_n \in \mathcal{D}'(\Omega)$ converges to the distribution $T \in \mathcal{D}'(\Omega)$ if for all $\varphi \in \mathcal{D}(\Omega)$, $\langle T_n, \varphi \rangle \longrightarrow \langle T, \varphi \rangle$.*

$\displaystyle\sum_{n\geq 0} T_n$ is said to converge and sums to T if the sequence $S_p = \sum_{n=0}^{p} T_n$ converges to T.

Examples 2

1. *Let $f_n(x) = \cos(nx)$, $f_n \in L^1_{loc}(\mathbb{R})$. $T_{f_n} \in \mathcal{D}'(\mathbb{R})$.*
   $\lim\limits_{n\to+\infty} T_{f_n} = 0$ (because $\forall\, \varphi \in \mathcal{D}(\mathbb{R})$, $\langle T_{f_n}, \varphi \rangle = \int \cos(nx)\varphi(x)dx \xrightarrow[n\to\infty]{} 0$).
2. *Let $n \in \mathbb{N}$, $\delta_n \longrightarrow 0$ in $\mathcal{D}'(\mathbb{R})$.*
   Let $\varphi \in \mathcal{D}(\mathbb{R})$, $\langle \delta_n, \varphi \rangle = \varphi(n) = 0$ for a large enough n (since φ is compactly).

Theorem 2

Let $T_n \in \mathcal{D}'(\Omega)$, $n \in \mathbb{N}$.
If for all $\varphi \in \mathcal{D}(\Omega)$, $\langle T_n, \varphi \rangle$ has a limit in $\mathbb{C}$, then T_n has a limit in $\mathcal{D}'(\Omega)$.

Proof

- $\varphi \in \mathcal{D}(\Omega)$ implies $\lim_{n\to+\infty} \langle T_n, \varphi \rangle$ exists.
  $\varphi_1, \varphi_2 \in \mathcal{D}(\Omega)$ et $\lambda \in \mathbb{C}$.
  $\lim\limits_{n\to+\infty} \langle T_n, \varphi_1 + \lambda\varphi_2 \rangle = \lim\limits_{n\to+\infty} \langle T_n, \varphi_1 \rangle + \lambda \langle T_n, \varphi_2 \rangle = \lim\limits_{n\to+\infty} \langle T_n, \varphi_1 \rangle + \lambda \lim\limits_{n\to+\infty} \langle T_n, \varphi_2 \rangle$, hence the linearity.
- The continuity of $\lim T_n$ is a consequence of Banach-Steinhaus theorem (admitted)

Example 3 $\forall\ n,\ T_n = \sum\limits_{p=0}^{n} \delta_p \in \mathcal{D}'(\mathbb{R})$. *Indeed, let* $\varphi \in \mathcal{D}(\mathbb{R})$, **Exercise** $n_o \in \mathbb{N}\mathrm{Supp}\,(\varphi) \subset [-n_0, n_0]$. $\langle T_n, \varphi \rangle = \sum\limits_{p=0}^{n_0} \varphi(p) \xrightarrow[n\to\infty]{} \sum\limits_{p=0}^{n_0} \varphi(p)$, *so there exists* $T \in \mathcal{D}(\mathbb{R})$ *such that* $T_n \longrightarrow T$ *dans* $\mathcal{D}'(\mathbb{R})$ *:* $T = \sum\limits_{p\geq 0} \delta_p$.

2.1.4 Derivation in the distribution space

Let $f \in C^1(\Omega)$ (so $\in L^1_{loc}(\Omega)$). For $\varphi \in \mathcal{D}(\Omega)$:

$$\int_\Omega f'(x)\varphi(x)dx = -\int_\Omega f(x)\varphi'(x)dx \Leftrightarrow \langle T_{f'}, \varphi \rangle = -\langle T_f, \varphi' \rangle$$

Extending this to more general distributions, we get:

Definition 5 *Let* $T \in \mathcal{D}'(\Omega)$, *one defines* T' *as :* $\langle T', \varphi \rangle \stackrel{def}{=} -\langle T, \varphi' \rangle$

Proposition 2

T is indefinitely differentiable, and one has: $\forall\, \varphi \in \mathcal{D}(\Omega)$, $\langle T^{(\alpha)}, \varphi \rangle = (-1)^\alpha \langle T, \varphi^{(\alpha)} \rangle$

Proof (du 1.)

- Let φ and $\psi \in \mathcal{D}(\Omega)$, $\lambda \in \mathbb{C}$.
  $$\begin{aligned} \langle T', \varphi + \lambda\psi \rangle &= -\langle T, (\varphi + \lambda\psi)' \rangle = -\langle T, \varphi' + \lambda\psi' \rangle \\ &= -\langle T, \varphi' \rangle - \lambda \langle T, \psi' \rangle = \langle T', \varphi \rangle + \lambda \langle T', \psi \rangle \end{aligned}$$
- Let $\varphi_n \longrightarrow 0$ in $\mathcal{D}(\Omega)$, then $\varphi_n' \longrightarrow 0$ in $\mathcal{D}(\Omega)$. One has: $\langle T', \varphi_n \rangle = -\langle T, \varphi_n' \rangle \longrightarrow 0$

Proposition 3

The derivation is a continuous operation on $\mathcal{D}'(\Omega)$.
If $T_n, T \in \mathcal{D}'(\Omega)$ and $T_n \longrightarrow T$ in $\mathcal{D}'(\Omega)$, then $\forall\, \alpha \in \mathbb{N}$, $T_n^{(\alpha)} \longrightarrow T^{(\alpha)}$ in $\mathcal{D}'(\Omega)$.

Proof
Let $\varphi \in \mathcal{D}(\Omega)$, $\langle T_n^{(\alpha)}, \varphi\rangle = (-1)^\alpha \langle T_n, \varphi^{(\alpha)}\rangle \longrightarrow (-1)^\alpha \langle T, \varphi^\alpha\rangle = \langle T^\alpha, \varphi\rangle$

Example 4 *If $f \in C^1(\Omega)$: $(T_f)' = T_{f'}$, the derivative is still a regular distribution.*

Examples 5 *Let us define T_Y, with Y the Heaviside function defined by:* $Y(x) = \begin{cases} 1 & si\ x > 0 \\ 0 & si\ x < 0 \end{cases}$

$Y \in L^1_{loc}(\mathbb{R})$, so Y is a distribution $T_Y \in \mathcal{D}'(\mathbb{R})$. Then, let $\varphi \in \mathcal{D}'(\mathbb{R})$, we get
$\langle T'_Y, \varphi\rangle = -\langle T_Y, \varphi'\rangle = -\displaystyle\int_0^{+\infty} \varphi'(x)dx = \varphi(0) = \langle \delta_0, \varphi\rangle$, *meaning that $T'_Y = \delta_0$.*

2.2 Fourier Transform of Distributions

The Fourier transform of distributions is going to be defined on a subset of $\mathcal{D}'(\mathbb{R})$, called *tempered distributions*. These are defined as continuous linear forms on the Schwartz space which we first introduce.

2.3 The Schwartz class

Definition 6 $\mathcal{S}(\mathbb{R})$, *called the* Schwartz class *is the set of functions $\phi : \mathbb{R} \to \mathbb{C}$ such that:*

- $\phi \in C^\infty(\mathbb{R})$
- $\forall \alpha \in \mathbb{N},\ n \in \mathbb{N}, \exists C$ *such that* $|\phi^{(\alpha)}(x)| \leq \frac{C}{(1+\|x\|)^n}$ *(fast decay)*

In other words, the functions in $\mathcal{S}(\mathbb{R})$ are $C^\infty(\mathbb{R})$ functions having all their derivatives with fast decay.
Example : $\phi(x) = e^{-x^2}$

Theorem 3

$\mathcal{S}(\mathbb{R})$ has the following properties:

1. $\forall \phi \in \mathcal{S}(\mathbb{R}), \quad \forall P \in \mathbb{C}[X] \quad P\phi \in \mathcal{S}(\mathbb{R})$
2. $\forall \phi \in \mathcal{S}(\mathbb{R}), \quad \phi' \in \mathcal{S}(\mathbb{R})$
3. $1 \leq p \leq \infty \quad \mathcal{S}(\mathbb{R}) \subset L^p(\mathbb{R})$

Theorem 4

$\mathcal{F}$ (i.e. the Fourier transform, $\mathcal{F}(\phi) = \hat{\phi}$) is a linear bijection from $\mathcal{S}(\mathbb{R})$ onto itself with inverse $\overline{\mathcal{F}}$.

Proof Let $f \in \mathcal{S}(\mathbb{R})$, since for all k, $x^k f(x)$ is in $L^1(\mathbb{R})$, $\hat{f}$ belongs to $C^\infty(\mathbb{R})$. Let $n \in \mathbb{N}$ and $p \in \mathbb{N}$, one has:

$$\begin{aligned}\xi^n \hat{f}^{(p)}(\xi) &= \xi^n \mathcal{F}((-2i\pi x)^p f(x)) \text{ by differentiation of an integral depending on a parameter} \\ &= \frac{1}{(2i\pi)^n}\mathcal{F}(((-2i\pi x)^p f(x))^{(n)}) \text{ using properties of the Fourier transform of derivatives.}\end{aligned}$$

Using the stability properties of $\mathcal{S}(\mathbb{R})$ by multiplication with a polynom and by derivation, we get that the function of which we compute the Fourier transform is in $\mathcal{S}(\mathbb{R})$, so that its Fourier transform is bounded, which proves that $\hat{f}$ belongs to $\mathcal{S}(\mathbb{R})$.
Furthermore, as f and $\hat{f}$ are in $L^1(\mathbb{R})$ and, as f is continuous, one gets $f(x) = \overline{\mathcal{F}}(\hat{f})(x)$.

The topology of $\mathcal{S}(\mathbb{R})$ is not defined by a norm but by a numerable family of norms:

$$\forall \phi \in \mathcal{S}(\mathbb{R}), \quad \mathcal{N}_p(\phi) = \max_{0 \le \alpha, \beta \le p} \sup_{x \in \mathbb{R}} |x^\alpha \phi^{(\beta)}((x)|, \quad p \in \mathbb{N}$$

One says that a sequence ϕ_n converges to ϕ in $\mathcal{S}(\mathbb{R})$ if:

$$\mathcal{N}_p(\phi_n - \phi) \to 0 \text{ for all } p \ge 0, \text{ when } n \to \infty.$$

One can alternatively define the convergence in $\mathcal{S}(\mathbb{R})$ as follows:

$$\forall \alpha, \beta \in \mathbb{N}\ x^\alpha \phi_n^{(\beta)}(x) \to x^\alpha \phi^{(\beta)}(x) \text{ uniformly on } \mathbb{R}$$

2.4 The space of tempered distributions $\mathcal{S}'(\mathbb{R})$

We need to define the Fourier transform in a more general framework than that of the functions so that the Fourier transform of sampled signals makes sense. The set of tempered distributions $\mathcal{S}'(\mathbb{R})$ is defined by:

$$\left\{ \begin{array}{cccc} T: & \mathcal{S}(\mathbb{R}) & \to & \mathbb{C} \quad \textit{linear, continuous} \\ & \varphi & \mapsto & \langle T, \varphi \rangle \end{array} \right\}$$

Here, the continuity has to be understood in the following sense:

$$\exists m, C_m \text{ tel que } \forall \phi \in \mathcal{S}(\mathbb{R}), \quad |\langle T, \phi \rangle| \le C_m \mathcal{N}_m(\phi)$$

or, using sequences:

$$\phi_n \to \phi \text{ in } \mathcal{S}(\mathbb{R}) \Rightarrow \langle T, \phi_n \rangle \to \langle T, \phi \rangle \text{ in } \mathbb{C}.$$

Remark 4 *$\mathcal{D}(\mathbb{R}) \subset \mathcal{S}(\mathbb{R})$ and if T is continuous for the topology on $\mathcal{S}(\mathbb{R})$, it is also continuous for the topology on $\mathcal{D}(\mathbb{R})$, so $\mathcal{S}'(\mathbb{R}) \subset \mathcal{D}'(\mathbb{R})$. Furthermore, one can show that $\mathcal{D}(\mathbb{R})$ is dense in $\mathcal{S}(\mathbb{R})$ for the topology of $\mathcal{S}(\mathbb{R})$. A consequence is that to prove the continuity in $\mathcal{S}'(\mathbb{R})$, one can restrain to functions in $\mathcal{D}(\mathbb{R})$, i.e.:*

$$\exists m, C_m \text{ tel que } \forall \phi \in \mathcal{D}(\mathbb{R}), \quad |\langle T, \phi \rangle| \leq C_m \mathcal{N}_m(\phi)$$

or, using sequences: $\phi_n \in \mathcal{D}(\mathbb{R}),\ \phi_n \to \phi$ in $\mathcal{S}(\mathbb{R}) \Rightarrow \langle T, \phi_n \rangle \to \langle T, \phi \rangle$ in $\mathbb{C}$

Examples 6

- $f \in L^1(\mathbb{R}),\ L^2(\mathbb{R})$ *ou* $L^\infty(\mathbb{R}) \Rightarrow T_f \in \mathcal{S}'(\mathbb{R})$
- *Dirac* $\delta_a \in \mathcal{S}'(\mathbb{R})$
- *Dirac comb* $\sum_{n\in\mathbb{Z}} \delta_n \in \mathcal{S}'(\mathbb{R})$
- $\mathcal{E}'(\mathbb{R}) \subset \mathcal{S}'(\mathbb{R})$
- *Functions of slow increase are in $\mathcal{S}'(\mathbb{R})$. A function is of slow increase if:*

$$\exists c > 0 \quad \exists N \in \mathbb{N}, \quad \forall x \in \mathbb{R}\ |f(x)| \leq c(1 + |x|)^N$$

- *If the sequence $(y_n)_{n\in\mathbb{Z}}$ is of slow increase (i.e. $\exists m \in \mathbb{N},\ c \in \mathbb{R}$ such that $|y_n| \leq C(1 + |n|)^m,\ n \in \mathbb{Z}$), the distribution*

$$T = \sum_{n\in\mathbb{Z}} y_n \delta_{na}$$

is tempered.

2.5 Fourier Transform in $\mathcal{S}'(\mathbb{R})$

Definition 7 *Let $T \in \mathcal{S}'(\mathbb{R})$, one defines its Fourier transform as follows*

$$\hat{T} : \begin{pmatrix} \varphi & \mapsto & \langle \hat{T}, \varphi \rangle = \langle T, \hat{\varphi} \rangle \\ \mathcal{S}(\mathbb{R}) & \to & \mathbb{C} \end{pmatrix}$$

One has: $\hat{T} \in \mathcal{S}'(\mathbb{R})$

Remark 5

- $\varphi \in \mathcal{S}(\mathbb{R}) \Rightarrow \hat{\varphi} \in \mathcal{S}(\mathbb{R})$
- *So $\hat{T}$ is well defined, linear by linearity of T, continuous using the continuity of T (for that we use the fact that $\mathcal{S}(\mathbb{R})$ is stable through Fourier transform).*

Proposition 4

$T_n \in \mathcal{S}'(\mathbb{R})$ converges to T in $\mathcal{S}'(\mathbb{R})$, if $\forall \varphi \in \mathcal{S}(\mathbb{R}),\ \langle T_n, \varphi \rangle \to \langle T, \varphi \rangle$.

Proposition 5

The Fourier transform is a continuous application from $\mathcal{S}'(\mathbb{R})$ onto itself.

Proof Let T_n tending to T in $\mathcal{S}'(\mathbb{R})$ then for all φ in $\mathcal{S}(\mathbb{R})$, one has:

$$\langle \hat{T}_n, \varphi \rangle = \langle T_n, \hat{\varphi} \rangle \to \langle T, \hat{\varphi} \rangle = \langle \hat{T}, \varphi \rangle$$

Theorem 5

$$\begin{array}{rccc} \mathcal{F}: & T & \mapsto & \hat{T} \\ & \mathcal{S}'(\mathbb{R}) & \to & \mathcal{S}'(\mathbb{R}) \end{array}$$

is invertible, and its inverse is:

$$\begin{array}{rccc} \bar{\mathcal{F}}: & T & \mapsto & \bar{\mathcal{F}}(T) \\ & \mathcal{S}'(\mathbb{R}) & \to & \mathcal{S}'(\mathbb{R}) \end{array}$$

with $\langle \bar{\mathcal{F}}(T), \varphi \rangle = \langle T, \bar{\mathcal{F}}(\varphi) \rangle$

Proof Let $T \in \mathcal{S}'(\mathbb{R})$, $\forall \varphi \in \mathcal{S}(\mathbb{R}), \langle \bar{\mathcal{F}}\mathcal{F}(T), \varphi \rangle = \langle \mathcal{F}(T), \bar{\mathcal{F}}(\varphi) \rangle = \langle T, \mathcal{F}\bar{\mathcal{F}}(\varphi) \rangle = \langle T, \varphi \rangle$, because $\mathcal{F}\bar{\mathcal{F}} = Id$ in $\mathcal{S}(\mathbb{R})$.

Examples 7

(i) Let $f \in L^1(\mathbb{R})$ or $L^2(\mathbb{R})$, then $T_f \in \mathcal{S}'(\mathbb{R})$.

$$\langle \widehat{T_f}, \varphi \rangle = \langle T_f, \hat{\varphi} \rangle = \int_{\mathbb{R}} f(y)\hat{\varphi}(y)dy = \int_{\mathbb{R}} \hat{f}(y)\varphi(y)dy = \langle T_{\hat{f}}, \varphi \rangle$$

so $\widehat{T_f} = T_{\hat{f}}$.

Conclusion: if the Fourier transform exists in the functional sense, and is denoted by $\hat{f}$, its Fourier transform in the sense of distributions will be $T_{\hat{f}}$.

(ii) $\forall \varphi \in \mathcal{S}(\mathbb{R}), \langle \hat{\delta}, \varphi \rangle = \langle \delta, \hat{\varphi} \rangle = \hat{\varphi}(0) = \int_{-\infty}^{+\infty} \varphi(x)e^{-2i\pi 0x}dx = \int_{-\infty}^{+\infty} \varphi(x)dx = \langle T_1, \varphi \rangle$, so $\hat{\delta} = T_1$.

(iii) $\forall \varphi \in \mathcal{S}(\mathbb{R}), \langle \hat{\delta_a}, \varphi \rangle = \langle \delta_a, \hat{\varphi} \rangle = \hat{\varphi}(a) = \int_{-\infty}^{+\infty} \varphi(x)e^{-2i\pi ax}dx = \langle e^{-2i\pi ax}, \varphi \rangle$, so $\hat{\delta_a} = T_{e^{-2i\pi ax}}$.

(iv) $\forall \varphi \in \mathcal{S}(\mathbb{R}), \langle \widehat{T_{e^{2i\pi k_0 x}}}, \varphi \rangle = \langle T_{e^{2i\pi k_0 x}}, \hat{\varphi} \rangle = \int_{-\infty}^{+\infty} e^{2i\pi k_0 y}\hat{\varphi}(y)dy = \varphi(k_0) = \langle \delta_{k_0}, \varphi \rangle$, so $\widehat{T_{e^{2i\pi k_0 x}}} = \delta_{k_0}$ (in particular : $\widehat{T_1} = \delta_0$).

(v) Let T be a strictly positive real and $(y_n)_{n\in\mathbb{Z}}$ a sequence of slow increase, then

$$\mathcal{F}(\sum_{n\in\mathbb{Z}} y_n \delta_{nT}) = \sum_{n\in\mathbb{Z}} y_n e^{-2i\pi nTx},$$

which is a consequence of the continuity of the Fourier transform on $\mathcal{S}'(\mathbb{R})$:

$$\widehat{\sum_{n\in\mathbb{Z}} y_n \delta_{nT}} = \sum_{n\in\mathbb{Z}} y_n \hat{\delta}_{nT} = \sum_{n\in\mathbb{Z}} y_n e^{-2i\pi nTx}.$$

This last example is very important in signal processing, in which community one defines (often without proof of existence) the so-called discrete time Fourier transform (DTFT), as follows

Definition 8 *The discrete-time Fourier transform (DTFT) of a sequence (x_n) is defined by:*

$$X(e^{2i\pi\omega}) = \sum_{n\in\mathbb{Z}} x_n e^{-2i\pi n\omega} \tag{2.1}$$

which exists in $\mathcal{S}'(\mathbb{R})$ as soon as (x_n) is of low increase.

In particular, when (x_n) is in $l_1(\mathbb{Z})$, one has normal convergence, and X is continuous. When (x_n) is in $l_2(\mathbb{Z})$, $X(e^{2i\pi\omega})$ can be viewed as a Fourier series of a 1-periodic function and the convergence takes place in $L^2([-1/2, 1/2[)$, so that we can write:

$$x_n = \int_{-1/2}^{1/2} X(e^{2i\pi\omega}) e^{2i\pi n\omega} d\omega,\ n \in \mathbb{Z} \tag{2.2}$$

Another consequence is the Parseval equality when (x_n) is in $l_2(\mathbb{Z})$:

$$\int_{-1/2}^{1/2} |X(e^{2i\pi\omega})|^2 d\omega = \sum_{n\in\mathbb{Z}} |x_n|^2.$$

2.5.1 Distributions with compact support

Definition 9 *Let $T \in \mathcal{D}'(\Omega)$ and $\omega \subset \Omega$ an open set. T is <u>null</u> on ω if for any $\varphi \in \mathcal{D}(\omega)$, $\langle T, \varphi\rangle = 0$.*

Example 8 *Let $a \in \mathbb{R}$, δ_a is null on $\mathbb{R} \setminus \{a\}$. If $\omega \subset \mathbb{R}$ is an open set and if $a \notin \omega$, then δ_a is null on ω.*

Definition 10 *Let $T \in \mathcal{D}'(\Omega)$, the <u>support</u> of T, denoted* Supp (T)*, is the complement set (in Ω) of the largest open set ω on which T is null.*

One can show that ω exists.

Examples 9

- Supp $(\delta_a) = \{a\}$
- *For all $a \in \mathbb{R}$ and all $\alpha \in \mathbb{N}$, one has* Supp $(\partial^\alpha \delta_a) = \{a\}$
- *If $f \in C^0(\Omega)$ and* Supp (f) *is a compact set, then $T_f \in \mathcal{E}'(\Omega)$.*

Proposition 6

Let $T \in \mathcal{E}'(\mathbb{R})$, then $\widehat{T}$ belongs to $C^\infty(\mathbb{R})$ and is with slow increase.

2.5.2 Convolution $\mathcal{E}'(\mathbb{R}) * \mathcal{D}'(\mathbb{R})$

Let $u \in C_0(\mathbb{R})$ and $v \in L^1_{loc}(\mathbb{R})$ then for all $\varphi \in \mathcal{D}(\mathbb{R})$, since $(x,y) \mapsto u(y)v(x-y)\varphi(x) \in L^1(\mathbb{R}\times\mathbb{R})$, using Fubini theorem one may write:

$$\begin{aligned}\int_{\mathbb{R}} u*v(x)\varphi(x)dx &= \int_{\mathbb{R}}\left(\int_{\mathbb{R}} u(y)v(x-y)dy\right)\varphi(x)dx = \int_{\mathbb{R}} u(y)\left(\int_{\mathbb{R}} v(x)\varphi(x+y)dx\right)dy \\ &= \int_{\mathbb{R}} v(x)\left(\int_{\mathbb{R}} u(y)\varphi(x+y)dy\right)dx.\end{aligned}$$

which can be rewrite using distributions notations as:

$$\langle T_{u*v},\varphi\rangle = \langle T_u, \langle T_v, \varphi(.+y)\rangle\rangle = \langle T_v, \langle T_u, \varphi(.+y)\rangle\rangle.$$

One can then generalize this remark through the following definition.

Definition 11 *Let $S \in \mathcal{E}'(\mathbb{R})$ and $T \in \mathcal{D}'(\mathbb{R})$.*

- *There exists a distribution, called convolution of S with T which we write $S*T$ and such that for all $\varphi \in \mathcal{D}(\mathbb{R})$, one has:*

$$\langle S*T,\varphi\rangle = \langle S_t, \langle T_x, \varphi(x+t)\rangle\rangle = \langle T_u, \langle S_x, \varphi(x+u)\rangle\rangle$$

- *The application $(S,T) \to S*T$ from $\mathcal{E}'(\mathbb{R})\times\mathcal{D}'(\mathbb{R})$ onto $\mathcal{D}'(\mathbb{R})$ is continuous with respect to each variable.*
- *If $T \in \mathcal{S}'(\mathbb{R})$ then $S*T \in \mathcal{S}'(\mathbb{R})$.*

Examples: Let $T \in \mathcal{D}'(\mathbb{R})$, $\delta_a * T = T * \delta_a = \tau_a T$.
$\delta^{(k)} * T = T * \delta^{(k)} = T^{(k)}$

Proposition 7

Let $S \in \mathcal{E}'(\mathbb{R})$ and $T \in \mathcal{S}'(\mathbb{R})$, one has

$$\widehat{S*T} = \hat{S}\hat{T}$$

We here make the link with applications in signal processing. Let us consider the convolution of the discrete sequences (h_n) and (x_n) by:

$$y_n = \sum_{k\in\mathbb{Z}} x_{n-k}h_k \tag{2.3}$$

Il the support of (h_k) is finite, one can associate to this sequence the following compactly supported distribution $h = \sum_{k\in\mathbb{Z}} h_k\delta_k$ and if (x_n) is of slow increase, it is associated with the tempered distribution $x = \sum_{k\in\mathbb{Z}} x_k\delta_k$. Applying the definition of the convolution of two distributions, one obtains the following tempered distribution $y = \sum_{k\in\mathbb{Z}} y_n\delta_n$, of which we can take the Fourier transform.

Usual framework in signal processing: If (h_n) is compactly supported, and if (x_n) belongs to $l^1(\mathbb{Z})$ or $l^2(\mathbb{Z})$, then (y_n) belongs to the same space and one has:

Proposition 8

$$Y(e^{2i\pi\omega}) = X(e^{2i\pi\omega})H(e^{2i\pi\omega}),$$

where Y, X and H are the $DTFT$s of sequences y, x and $h$respectively.

2.6 Exercises

Exercise 1 Are the following applications T, defined for $\varphi \in \mathcal{D}(\mathbb{R})$, distributions?

1. $\langle T, \varphi \rangle = \int_0^1 \varphi(x)dx$

2. $\langle T, \varphi \rangle = \int_0^1 |\varphi(x)|dx$

3. $\langle T, \varphi \rangle = \sum_{n=0}^{+\infty} \varphi(n)$

4. $\langle T, \varphi \rangle = \sum_{n=1}^{+\infty} \varphi(1/n)$

Exercise 2 Let $\varphi \in \mathcal{D}(\mathbb{R})$.

1. Show that there exist a constant $C(\varphi)$ such that:

$$\forall n \in \mathbb{Z}, \left| \int_{-\infty}^{+\infty} e^{inx}\varphi(x)dx \right| \leq \frac{C(\varphi)}{1+n^2}$$

2. Let $(a_n)_{n\in\mathbb{Z}}$ be a bounded sequence. Show that the series with general term:

$$a_n \int_{-\infty}^{+\infty} e^{inx}\varphi(x)dx$$

converges, and that the application which maps φ to the sum of this series is a distribution.

3. Show that when the sequence $n^2 a_n$ is bounded, the distribution is indeed a function.

Exercise 3 Let us consider the regular distributions e^{inx}.

1. Show that for all $n \neq 0$:

$$\forall \varphi \in \mathcal{D}(]-\pi, \pi[), \quad \left|\langle e^{inx}, \varphi \rangle\right| \leq \frac{C(\varphi)}{n^2}$$

where $C(\varphi)$ is a constant which does not depend on φ.

2. Let us define:

$$u_N(x) = \sum_{n=-N}^{N} e^{inx}$$

Show that the sequence of distributions T_N associated with functions u_N converges in the distributional sense on $]-\pi, \pi[$. Let T be its limit.

3. Show that:

$$u_N(x) = \frac{\sin(N+\frac{1}{2})x}{\sin(\frac{x}{2})}$$

4. Show that if $\varphi \in \mathcal{D}(]-\pi,\pi[)$ is such that $\varphi(0)=0$, then $\frac{\varphi(x)}{\sin(\frac{x}{2})}$ belongs to $C_c^\infty(]-\pi,\pi[)$ (we recall that $\frac{\sin x}{x}$ is a smoot function).

5. Show that if $\varphi \in \mathcal{D}(]-\pi,\pi[)$ is such that $\varphi(0)=0$, then $\langle T,\varphi\rangle = 0$.

6. Deduce that there exists a constant C such that:

$$T = \sum_{n=-\infty}^{+\infty} e^{inx} = C\delta.$$

We will admit that $C = 2\pi$.

Exercise 4 After having shown the following distributions are tempered distributions, compute the Fourier transform of the following distributions:

1. 1

2. x^n

3. $\delta^{(n)}$

4. $e^{2i\pi\nu_0 x}$

Exercise 5 Compute, using the definition of the Fourier transforms of distributions the following integral:

$$\int_{\mathbb{R}} e^{-\pi x^2}\cos(2\pi x)dx$$

Exercise 6 Show that if f belongs to $L^1(\mathbb{R})$ or $L^2(\mathbb{R})$ that $\widehat{T_f} = T_{\hat{f}}$.

Exercise 7 *Fourier transform of $vp(1/x)$*

1. Show that $vp(1/x)$ is a tempered distribution

2. We recall that $xvp(1/x) = 1$, deduce the Fourier transform of $vp(1/x)$

Exercise 8 *Fourier transform of the Heavyside function*
Remarking that $U(x) = \frac{1}{2}(sign(x)+1)$, compute its Fourier transform.

Chapter 3

Linear Time-Frequency Analysis

Here we focus on linear time-frequency techniques, that is we are going to define linear transforms that map a function to its time-frequency representation. The focus is put on the short-time Fourier transform both in the continuous and discrete setting. In the latter case, the emphasis will be put on the reaction with the Fourier transform of distributions.

3.1 Linear Time-Frequency analysis: the continuous time framework

Time-Frequency analysis is related to the definition of Short-Time Fourier Transform (STFT), the definition of which we now recall in different contexts.

3.1.1 Continuous Time Short Time Fourier Transform

Definition 1 *The STFT of a given signal $f \in L^1(\mathbb{R}) \bigcap L^2(\mathbb{R})$ and g a real window also in $L^2(\mathbb{R})$ is given by:*

$$V_f^g(t,\omega) = \int_{\mathbb{R}} f(u)g(u-t)e^{-i2\pi\omega(u-t)}du. \tag{3.1}$$

Remark 1 *The existence of the STFT is a direct consequence of Cauchy-Schwartz theorem.*

This transform is invertible under some assumptions:

Proposition 1

Assume $\int_{\mathbb{R}} g = 1$, and that $\hat{f}$ is in $L^1(\mathbb{R})$, the following reconstruction formula holds :

$$f(t) = \int\int_{\mathbb{R}^2} V_f^g(u,\omega)e^{i2\pi\omega(t-u)}dud\omega.$$

Proof Because of the hypothesis made on g, one has: $f(t) = \int_{\mathbb{R}} f(t)g(t-u)du$ Now, the Fourier transform of f reads:

$$\begin{aligned}\hat{f}(\omega) &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(t)g(t-u)e^{-2i\pi\omega(t-u)}dte^{-2i\pi\omega u}du \\ &= \int_{\mathbb{R}} V_f^g(u,\omega)e^{-2i\pi\omega u}du.\end{aligned}$$

So, since $\hat{f}$ is in $L^1(\mathbb{R})$, it is invertible and we can write: $f(t) = \int_{\mathbb{R}^2} V_f^g(t,\omega)e^{2i\pi\omega(t-u)}dud\omega$.

Proposition 2

Assume $\|g\|_2 = 1$, the following reconstruction formula holds (in $L^2(\mathbb{R})$) :

$$f(t) = \int\int_{\mathbb{R}^2} V_f^g(u,\omega)g(t-u)e^{i2\pi\omega(t-u)}dud\omega.$$

Proof The proof uses the fact that $\{g(t-u)e^{i2\pi\omega(t-u)}\}_u$ is a frame of $L^2(\mathbb{R})$, but it will not be detailed here.

Proposition 3

If one assumes that f is analytic (namely $\hat{f}(\omega) = 0$ if $\omega < 0$), g is continuous, iand both f and g are in $L^1(\mathbb{R})\bigcap L^2(\mathbb{R})$, one may also write:

$$f(t) = \frac{1}{g(0)}\int_0^\infty V_f^g(t,\omega)d\omega.$$

Proof

$$\begin{aligned}\int_0^\infty V_f^g(t,\omega)d\omega &= \int_0^\infty\int_{\mathbb{R}} f(u)g(u-t)e^{-i2\pi\omega(u-t)} = \int_0^\infty\int_{\mathbb{R}} \hat{f}(\omega)\hat{g}(\omega-\nu)^* e^{i2\pi\omega t}d\omega d\nu \\ &= \int_0^\infty \hat{f}(\omega)e^{i2\pi\omega t}d\omega\int_{\mathbb{R}}\hat{g}(\nu-\omega)d\nu = f(t)g(0),\end{aligned}$$

so one has the following reconstruction formula: $f(t) = \frac{1}{g(0)}\int_0^\infty V_f^g(t,\omega)d\omega$.

3.1.2 Discrete-Time Short-Time Fourier Transform

Since this part of the course is more signal processing oriented, we replace the notation x_n for a sequence by $x[n]$. For a sequence $(f[n])_{n\in\mathbb{Z}}$ in $l_1(\mathbb{Z})$, and a discrete real window g also in $l_1(\mathbb{Z})$, the STFT is defined for each ω by:

$$V_{f,d}^g(m,\omega) = \sum_{n\in\mathbb{Z}} f[n]g[n-m]e^{-i2\pi\omega(n-m)}. \tag{3.2}$$

The STFT can be viewed as the Fourier transform of $\sum_{n\in\mathbb{Z}} f[n]g[n-m]\delta_n$ times the phase shift term $e^{2i\pi m\omega}$. For STFT, we have the following reconstruction formula:

Proposition 4

Assume $g(0) \neq 0$, then:

$$f[m] = \frac{1}{g(0)} \int_0^1 V_{f,d}^g(m,\omega) d\omega.$$

Proof Since $V_{f,d}^g(m,\omega)$ is 1-periodic with respect to ω, using Fourier series theory we get:

$$f[n]g[n-m] = \int_0^1 V_{f,d}^g(m,\omega) e^{i2\pi\omega(n-m)} d\omega,$$

and then considering $n = m$ and $g(0) \neq 0$, we obtain:

$$f[m] = \frac{1}{g(0)} \int_0^1 V_{f,d}^g(m,\omega) d\omega. \tag{3.3}$$

Proposition 5

Note that with the hypothesis put on g, $(V_{f,d}^g(m,\omega))_{m\in\mathbb{Z}}$ is also in $l_1(\mathbb{Z})$, and further assuming $\|g\|_2 = 1$, we get :

$$f[n] = \int_0^1 \sum_{m\in\mathbb{Z}} V_{f,d}^g(m,\omega) g[n-m] e^{2i\pi\omega(n-m)} d\omega.$$

Proof Indeed, we have:

$$\begin{aligned}\int_0^1 \sum_{m\in\mathbb{Z}} V_{f,d}^g(m,\omega) g[n-m] e^{2i\pi\omega(n-m)} d\omega &= \sum_{m,k\in\mathbb{Z}} f[k]g[k-m]g[n-m] \int_0^1 e^{i2\pi\omega(n-k)} d\omega \\ &= \sum_{m\in\mathbb{Z}} f[n]g[n-m]^2 = f[n] \sum_{m\in\mathbb{Z}} g[m]^2 = f[n]\end{aligned}$$

Proposition 6

Alternatively, if one considers a filter $g \in l_1(\mathbb{Z})$ such that $\sum_m g[m] = 1$, the reconstruction of f is as follows:

$$f[n] = \int_0^1 \sum_{m\in\mathbb{Z}} V_{f,d}^g(m,\omega) e^{i2\pi\omega(n-m)} d\omega.$$

Proof The proof is similar to the previous one and is thus left as an exercice.

3.1.3 Short-Time Fourier Transform for finite length signal and filter

Now we assume the signal is of length L and that the filter g is supported on $\{-M, \cdots, M\}$ such that:

$$2^{\log_2 \lfloor 2M+1 \rfloor + 1} = N \leq L, \tag{3.4}$$

then we have the following reconstruction formula:

Proposition 7

Assume $g[0] \neq 0$, then we may write:

$$f[m] = \frac{1}{g[0]N} \sum_{k=0}^{N-1} V_{f,d}^{g}(m, \frac{k}{N}). \tag{3.5}$$

Proof Indeed,

$$\begin{aligned} V_{f,d}^{g}(m, \frac{k}{N}) &= \sum_{n \in \mathbb{Z}} f[n] g[n-m] e^{-i2\pi \frac{k(n-m)}{N}} = \sum_{n=-M}^{M} f[m+n] g[n] e^{-i2\pi \frac{kn}{N}}, \\ &= \sum_{n=0}^{2M} f[m+n-M] g[n-M] e^{-i2\pi \frac{k(n-M)}{N}}. \end{aligned}$$

Since g is null on $\{M+1, \cdots, N-1-M\}$, the STFT can be rewritten as:

$$V_{f,d}^{g}(m, \frac{k}{N}) e^{-i2\pi \frac{kM}{N}} = \sum_{n=0}^{N-1} f[m+n-M] g[n-M] e^{-i2\pi \frac{kn}{N}}.$$

Using the properties of the discrete Fourier transform, one obtains, for any $n \in \{0, \cdots, N-1\}$:

$$f[m+n-M] g[n-M] = \frac{1}{N} \sum_{k=0}^{N-1} V_{f,d}^{g}(m, \frac{k}{N}) e^{i2\pi \frac{k(n-M)}{N}}. \tag{3.6}$$

Finally, taking $n = M$ and assuming $g[0] \neq 0$:

$$f[m] = \frac{1}{g[0]N} \sum_{k=0}^{N-1} V_{f,d}^{g}(m, \frac{k}{N}).$$

Remark 2 *To reconstruct $f[m]$, one only needs the knowledge of $\left(V_{f,d}^{g}(m, \frac{k}{N})\right)_k$, while $\left(V_{f,d}^{g}(m, \frac{k}{N})\right)_k$ is non zero for $m \in \{-M, \cdots, L-1+M\}$, but the transform outside the support of f is not used in the reconstruction.*

Proposition 8

Now, if g is with unit energy, one has:

$$f[n] = \sum_{m=n-M}^{m=n+M} \frac{1}{N} \sum_{k=0}^{N-1} V_{f,d}^{g}(m, \frac{k}{N}) g[n-m] e^{i2\pi \frac{k(n-m)}{N}}.$$

Proof Indeed, we may write:

$$\begin{aligned}\sum_{m=n-M}^{m=n+M} \frac{1}{N} \sum_{k=0}^{N-1} V_{f,d}^{g}(m, \frac{k}{N}) g[n-m] e^{2i\pi \frac{k(n-m)}{N}} &= \sum_{m=n-M}^{m=n+M} \sum_{p=0}^{m=n+M} f[p] g[p-m] g[n-m] \frac{1}{N} \sum_{k=0}^{N-1} e^{i2\pi \frac{k(n-p)}{N}} \\ &= \sum_{m=n-M}^{m=n+M} f[n] g[n-m]^2 = f[n] \sum_{m=-M}^{m=M} g[m]^2 = f[n].\end{aligned}$$

This time, one needs the knowledge of $\left(V_{f,d}^{g}(m, \frac{k}{N})\right)_k$ for $m \in \{-M, \cdots, L-1+M\}$, while one would like to be able to reconstruct f using only $\left(V_{f,d}^{g}(m, \frac{k}{N})\right)_k$ for $m \in \{0, \cdots, L-1\}$.
To circumvent this difficulty, one can assume f is L-periodic instead of finite: the STFT is no longer in $l_1(\mathbb{Z})$ but is also L-periodic (in the sum defining the STFT p varies from $m-M$ to $m+M$). In this case, we may write:

Proposition 9

Assuming g is with unit energy, one has the following reconstruction formula assuming f is L-periodic:

$$f[n] = \sum_{m=n-M}^{n+M} \sum_{k=0}^{N-1} V_{f,d}^{g}(m \bmod L, \frac{k}{N}) g[n-m] \frac{e^{i2\pi \frac{k(n-m)}{N}}}{N}.$$

Note that the hypothesis that f is periodic could be avoided easily as well as on the unit energy for the filter. Indeed,

Proposition 10

assuming f is null outside its boundary, we have:

$$f[n] = \frac{\sum_{m=\max(n-M,0)}^{m=n+M} \frac{1}{N} \sum_{k=0}^{N-1} V_{f,d}^{g}(m, \frac{k}{N}) g[n-m] e^{2i\pi \frac{k(n-m)}{N}}}{\sum_{m=\max(n-M,0)}^{m=n+M} g[m]^2}$$

Proof We may actually write

$$\sum_{m=\max(n-M,0)}^{m=n+M} \frac{1}{N} \sum_{k=0}^{N-1} V_{f,d}^{g}(m, \frac{k}{N}) g[n-m] e^{2i\pi \frac{k(n-m)}{N}}$$

$$= \sum_{m=\max(n-M,0)}^{m=n+M} \sum_{p=0}^{p=m+M} f[p] g[p-m] g[n-m] \frac{1}{N} \sum_{k=0}^{N-1} e^{i2\pi \frac{k(n-p)}{N}}$$

$$= \sum_{m=\max(n-M,0)}^{m=n+M} f[n] g[n-m]^2 = f[n] \sum_{m=\max(n-M,0)}^{m=n+M} g[m]^2 = f[n].$$

Similarly to what was done previously in the continuous time case, if we further assume that f is L-periodic, and that $\sum_{m=-M}^{M} g(m) = 1$, one has the following reconstruction formula:

$$f[p] = \sum_{m=p-M}^{p+M} \frac{1}{N} \sum_{k=0}^{N-1} V_{f,d}^{g}(m \bmod L, \frac{k}{N}) e^{i2\pi \frac{k(p-m)}{N}}.$$

Again, not assuming any periodicity hypothesis we also have.

$$f[p] = \frac{\sum_{m=max(p-M,0)}^{p+M} \frac{1}{N} \sum_{k=0}^{N-1} V_{f,d}^{g}(m \bmod L, \frac{k}{N}) e^{i2\pi \frac{k(p-m)}{N}}}{\sum_{m=max(p-M,0)}^{p+M} g[p]}.$$

To conclude on this part we have shown different reconstruction procedures for associated with the STFT and considering different hypothesis on the filter. We are going to use these developments in the study of reassignment technique in the following chapter.

3.2 Applications in Matllab

One has to answer the following questions. The Matlab function mentioned in the following section are listed at the end of the chapter.

3.2.1 Lab session

1. In the files tfrstft.m and itfrstft.m we have implemented the STFT and its inverse respectively using some of the formulae introduced above. Explain to which cases correspond "cas" equal 1, 2 or 3 in theses procedures.

2. Write a program that decomposes the following signal and then reconstruct it with each of the above mentioned STFT, once the STFT is computed plot its modulus using the command *imagesc*, use a number of frequency bins equal to 512.

```
N  =  4096;
t    = (0:N-1)/N;
```

```
a   = 2;
s1  = a.*exp(2*pi*1i*(1000*t+60*cos(3*pi*t)));
s2  = a.*exp(2*pi*1i*(400*t+30*cos(3*pi*t)));
s   = s1+s2;
s   = s(:);
```

Check that the reconstruction is correct, in each of the studied case.

3. We add some Gaussian complex white noise to the original signal, using the sigmerge.m procedure, given an input SNR, through the following formula:

```
noise    =  randn(N,1)+1i*randn(N,1);
[Snoise]      =  sigmerge(s,noise,SNR);
```

Show that, if *noise* is a white noise with variance σ^2:

$$Var\left(\Re\{V^g_{noise,d}(m,\frac{k}{N})\}\right) = Var\left(\Im\{V^g_{noise,d}(m,\frac{k}{N})\}\right) = \sigma^2\|g\|^2_{2,m}$$

with $\|g\|^2_{2,m} = \sum_{n=\max(m-M,0)}^{n=m+M} g[n]^2$, if f is assumed to be zero outside its boundaries and

$$Var\left(\Re\{V^g_{noise,d}(m,\frac{k}{N})\}\right) = Var\left(\Im\{V^g_{noise,d}(m,\frac{k}{N})\}\right) = \sigma^2\|g\|^2_2$$

in the case f periodic.

4. Assuming the real and the imaginary parts of $V^g_{noise,d}(m,\frac{k}{N})$ are independent, show that $\frac{|V^g_{noise,d}(m,\frac{k}{N})|^2}{\sigma^2\|g\|^2_{2,m}}$ is χ_2 distributed with two degrees of freedom (when f is assumed to be null outside its boundaries, otherwise if one considers a periodic hypothesis then $\frac{|V^g_{noise,d}(m,\frac{k}{N})|^2}{\sigma^2\|g\|^2_2}$ has to be considered instead). Then, threshold the STFT using the following procedure (called hard-thresholding):

$$\overline{V}^g_{Snoise}(m,\frac{k}{N}) = \begin{cases} V^g_{Snoise}(m,\frac{k}{N}), \text{if } |V^g_{Snoise}(m,\frac{k}{N})| \geq 3\sigma\|g\|_2 \\ 0 \text{ otherwise.} \end{cases} \tag{3.7}$$

Explain why the choice of threshold is a good one. Propose a reconstruction s_r for signal s from the denoised STFT. Assuming σ is known, Write a program to plot the output SNR with respect to the input SNR. The output SNR is computed as follows:

$$SNR_{output} = 20\log_{10}\left(\frac{\|s\|_2}{\|s-s_r\|_2}\right).$$

Give a illustration of the denoised STFT and of the denoised signal (real and imaginary parts).

5. Based on the hypothesis the signal part is sparse in the representation of the STFT, the absolute median deviation is a good choice to estimate $\gamma = \sigma\|g\|_2$, i.e.

$$Y2 = real(V^g_{Snoise}(m,k/N));$$

$$\hat{\gamma} = median(abs(Y2(:)))/0.6745;$$

Having, computed this estimate, use it in the thresholding algorithm. Show that it makes little difference in terms of output SNR, i.e. knowing the true variance of the noise or only an estimation.

3.2.2 Files used in Lab session

Tools functions

```
function sig=sigmerge(x1,x2,ratio);

%SIGMERGE Add two signals with given energy ratio in dB.
% SIG=SIGMERGE(X1,X2,RATIO) adds two signals so that a given
% energy ratio expressed in deciBels is satisfied.
%
% X1, X2 : input signals.
% RATIO  : Energy ratio in deciBels (default : 0 dB).
% X      : output signal.
% X= X1+H*X2, such that 10*log(Energy(X1)/Energy(H*X2))=RATIO

if (nargin<2)
 error('At least two parameters are required');
elseif nargin==2,
 ratio=0;
end;

[x1row,x1col] = size(x1);
[x2row,x2col] = size(x2);

if (x1col~=1)|(x2col~=1),
 error('X1 and X2 must have only one column');
elseif (x1row~=x2row),
 error('X1 and X2 must have the same number of rows');
elseif (length(ratio)~=1),
 error('RATIO must be a scalar');
elseif (ratio==inf),
 sig = x1;
else
 Ex1=mean(abs(x1).^2);
 Ex2=mean(abs(x2).^2);
 h=sqrt(Ex1/(Ex2*10^(ratio/10)));
 sig=x1+h*x2;
end;

function res = snr(s,n)

%SNR computes the SNR. s : signal and n: noise

rms = @(x) sqrt(mean(abs(x).^2));
res = 20 * log10(rms(s)/rms(n));

end
```

```
function y = amgauss(N,t0,T);

%AMGAUSS Generate gaussian amplitude modulation.
% Y=AMGAUSS(N,T0,T) generates a gaussian amplitude modulation
% centered on a time T0, and with a spread proportional to T.
% This modulation is scaled such that Y(T0)=1
% and Y(T0+T/2) and Y(T0-T/2) are approximately equal to 0.5 .
%
% N  : number of points.
% T0 : time center (default : N/2).
% T  : time spreading (default : 2*sqrt(N)).
% Y  : signal.

if (nargin == 0),
 error ( 'The number of parameters must be at least 1.' );
elseif (nargin == 1),
 t0=N/2; T=2*sqrt(N);
elseif (nargin ==2),
 T=2*sqrt(N);
end;

if (N<=0),
 error('N must be greater or equal to 1.');
else
 tmt0=(1:N)'-t0;
 y = exp(-(tmt0/T).^2 * pi);
end;

function h=tftb_window(N,name,param,param2);

%tftb_window Window generation.
% H=tftb_window(N,NAME,PARAM,PARAM2)
% yields a window of length N with a given shape.
%
% N      : length of the window
% NAME   : name of the window shape (default : Hamming)
% PARAM  : optional parameter
% PARAM2 : second optional parameters
%
% Possible names are :
% 'Hamming', 'Hanning', 'Nuttall',  'Papoulis', 'Harris',
% 'Rect',    'Triang',  'Bartlett', 'BartHann', 'Blackman'
% 'Gauss',   'Parzen',  'Kaiser',   'Dolph',    'Hanna'.
% 'Nutbess', 'spline',  'Flattop'
%
% For the gaussian window, an optionnal parameter K
```

```
% sets the value at both extremities. The default value is 0.005
%
% For the Kaiser-Bessel window, an optionnal parameter
% sets the scale. The default value is 3*pi.
%
% For the Spline windows, h=tftb_window(N,'spline',nfreq,p)
% yields a spline weighting function of order p and frequency
% bandwidth proportional to nfreq.

if (nargin==0), error ( 'at least 1 parameter is required' ); end;
if (N<=0), error('N should be strictly positive.'); end;
if (nargin==1), name= 'Hamming'; end ;
name=upper(name);
if strcmp(name,'RECTANG') | strcmp(name,'RECT'),
 h=ones(N,1);
elseif strcmp(name,'HAMMING'),
 h=0.54 - 0.46*cos(2.0*pi*(1:N)'/(N+1));
elseif strcmp(name,'HANNING') | strcmp(name,'HANN'),
 h=0.50 - 0.50*cos(2.0*pi*(1:N)'/(N+1));
elseif strcmp(name,'KAISER'),
 if (nargin==3), beta=param; else beta=3.0*pi; end;
 ind=(-(N-1)/2:(N-1)/2)' *2/N; beta=3.0*pi;
 h=besselj(0,j*beta*sqrt(1.0-ind.^2))/real(besselj(0,j*beta));
elseif strcmp(name,'NUTTALL'),
 ind=(-(N-1)/2:(N-1)/2)' *2.0*pi/N;
 h=+0.3635819 ...
   +0.4891775*cos(    ind) ...
   +0.1363995*cos(2.0*ind) ...
   +0.0106411*cos(3.0*ind) ;
elseif strcmp(name,'BLACKMAN'),
 ind=(-(N-1)/2:(N-1)/2)' *2.0*pi/N;
 h= +0.42 + 0.50*cos(ind) + 0.08*cos(2.0*ind) ;
elseif strcmp(name,'HARRIS'),
 ind=(1:N)' *2.0*pi/(N+1);
 h=+0.35875 ...
   -0.48829 *cos(    ind) ...
   +0.14128 *cos(2.0*ind) ...
   -0.01168 *cos(3.0*ind);
elseif strcmp(name,'BARTLETT') | strcmp(name,'TRIANG'),
 h=2.0*min((1:N),(N:-1:1))'/(N+1);
elseif strcmp(name,'BARTHANN'),
 h=  0.38 * (1.0-cos(2.0*pi*(1:N)/(N+1))') ...
   + 0.48 * min((1:N),(N:-1:1))'/(N+1);
elseif strcmp(name,'PAPOULIS'),
 ind=(1:N)'*pi/(N+1); h=sin(ind);
elseif strcmp(name,'GAUSS'),
 if (nargin==3), K=param; else K=0.005; end;
```

```
 h= exp(log(K) * linspace(-1,1,N)'.^2 );
elseif strcmp(name,'PARZEN'),
 ind=abs(-(N-1)/2:(N-1)/2)'*2/N; temp=2*(1.0-ind).^3;
 h= min(temp-(1-2.0*ind).^3,temp);
elseif strcmp(name,'HANNA'),
 if (nargin==3), L=param; else L=1; end;
 ind=(0:N-1)';h=sin((2*ind+1)*pi/(2*N)).^(2*L);
elseif strcmp(name,'DOLPH') | strcmp(name,'DOLF'),
 if (rem(N,2)==0), oddN=1; N=2*N+1; else oddN=0; end;
 if (nargin==3), A=10^(param/20); else A=1e-3; end;
 K=N-1; Z0=cosh(acosh(1.0/A)/K); x0=acos(1/Z0)/pi; x=(0:K)/N;
 indices1=find((x<x0)|(x>1-x0));
 indices2=find((x>=x0)&(x<=1-x0));
 h(indices1)= cosh(K*acosh(Z0*cos(pi*x(indices1))));
 h(indices2)= cos(K*acos(Z0*cos(pi*x(indices2))));
 h=fftshift(real(ifft(A*real(h))));h=h'/h(K/2+1);
 if oddN, h=h(2:2:K); end;
elseif strcmp(name,'NUTBESS'),
 if (nargin==3), beta=param; nu=0.5;
 elseif (nargin==4), beta=param; nu=param2;
 else beta=3*pi; nu=0.5;
 end;
 ind=(-(N-1)/2:(N-1)/2)' *2/N;
 h=sqrt(1-ind.^2).^nu .* ...
   real(besselj(nu,j*beta*sqrt(1.0-ind.^2)))/real(besselj(nu,j*beta));
elseif strcmp(name,'SPLINE'),
 if (nargin < 3),
  error('Three or four parameters required for spline windows');
 elseif (nargin==3),
  nfreq=param; p=pi*N*nfreq/10.0;
 else nfreq=param; p=param2;
 end;
  ind=(-(N-1)/2:(N-1)/2)';
  h=sinc((0.5*nfreq/p)*ind) .^ p;
elseif strcmp(name,'FLATTOP'),
 ind=(-(N-1)/2:(N-1)/2)' *2.0*pi/(N-1);
 h=+0.2810639 ...
   +0.5208972*cos(    ind) ...
   +0.1980399*cos(2.0*ind) ;
else error('unknown window name');
end;
```

Time-frequency functions

```
function [tfr] = tfrstft(x,N,cas,g,Lg)

 %x        : signal
```

```
 %N        : number of frequency bins
 %cas      : if 1, ...
 %           if 2, ...
 %           if 3, ...
 %g        : the filter h used
 %Lg       : the filter is of length 2Lh+1
 %tfr      : short time Fourier transform

 [xrow,xcol] = size(x);

 t = 1:xrow; %the time instant, we consider the time instant shitfed by a factor shift.

 tfr= zeros (N,length(t)) ;
 if (cas == 1)
  %case without periodizing
  trans  = zeros(1,length(t));

  for icol=1:length(t),
   tau = -min([Lg,t(icol)-1]):min([Lg,xrow-t(icol)]);
   tfr(1:length(tau),icol) = x(t(icol)+tau,1).*g(Lg+1+tau);
   trans(icol)  = tau(1);
  end
  tfr=fft(tfr,N);
  A = exp(-2/N*pi*1i*(0:N-1)'*trans);
  tfr = tfr.*A;
 end

 if (cas == 2)||(cas == 3)

  %cases with periodization
  tau = -Lg:Lg;
  for icol = 1:length(t),
   if (t(icol) > Lg) && (t(icol) <= xrow-Lg)
    tfr(1:length(tau),icol) = x(t(icol)+tau,1).*g(Lg+1+tau);
   else
    tfr(1:length(tau),icol) = x(1+rem((t(icol)-1)+tau+xrow,xrow),1).*g(Lg+1+tau);
   end
  end
   tfr = fft(tfr,N);
   trans = Lg*ones(1,length(t));
   A = exp(2/N*pi*1i*(0:N-1)'*trans);
   tfr = tfr.*A;
 end
end

function [x] = itfrstft(tfr,cas,g)
```

```
%tfr  : STFT of signal x
%cas  : if 1, ...
%       if 2, ...
%       if 3, ...
%h    : filter
%x    : restored signal

[N,xrow] = size(tfr);

if (cas == 1)
 %case without periodizing
 x = zeros(xrow,1);
 Lg = (length(g)-1)/2;
 for icol=1:xrow,
  x(icol) = 1/g(Lg+1)*mean(tfr(:,icol));
 end
end

if (cas == 2)
 %cases with periodization
 Lg = (length(g)-1)/2;
 x = zeros(xrow,1);
 for i = 1:xrow
  ind  = i-Lg:i+Lg;
  if (i > Lg)&&(i <= xrow - Lg)
   x(i) = mean(tfr(:,ind).*exp(2*1i*pi*(0:N-1)'*(i-ind)/N)*...
          g(Lg+1+i-ind))/norm(g(Lg+1+i-ind))^2;
  else
   x(i) = mean((tfr(:,1+rem((ind-1)+xrow,xrow)).*exp(2*1i*pi*(0:N-1)'*(i-ind)/N))...
               *g(Lg+1+i-ind))/norm(g(Lg+1+i-ind))^2;
  end
 end
end

if (cas == 3)
 %case with periodization

 Lg = (length(g)-1)/2;
 x = zeros(xrow,1);

 for i = 1:xrow
  ind = i-Lg:i+Lg;
  if (i > Lg)&&(i <= xrow-Lg)
   x(i) = mean((tfr(:,ind).*exp(2*1i*pi*(0:N-1)'*(i-ind)/N))*...
          ones(length(Lg+1+i-ind),1))/sum(g(Lg+1+i-ind));
  else
   x(i) = mean((tfr(:,1+rem((ind-1)+xrow,xrow)).*exp(2*1i*pi*(0:N-1)'*(i-ind)/N))*...
```

```
           ones(length(Lg+1+i-ind),1))/sum(g(Lg+1+i-ind));
   end
  end
 end
end

function test_three_cases_uncomplete(cas,window)

 %cas    : 1...,
 %         2...,
 %         3...,
 %window : choice for window g here either Gaussian or Hamming

 N = 4096;
 t = (0:N-1)/N;
 a  = 2;
 s1 = a.*exp(2*pi*1i*(1000*t+60*cos(3*pi*t)));
 s2 = a.*exp(2*pi*1i*(400*t+30*cos(3*pi*t)));
 s  = s1+s2;
 s = s(:);

 Nfft = 512; %number of frequency bins

 %we build the filter g
 if strcmp(window,'hamming')
  glength=floor(161);
  glength=glength+1-rem(glength,2);%the length of the filter has to be odd
  g = tftb_window(glength,window);
 else
  prec = 10^(-3);
  sigma_w = 0.15;
  L =  Nfft*sigma_w;
  Lg = floor(L*sqrt(-log(prec)/pi))+1;
  g = amgauss(2*Lg+1,Lg+1,L); %explain what is being done here
 end

 [grow,gcol]=size(g);
 Lg=(grow-1)/2;

 %compute STFT
 %plot its modulus
 %invert it

end
```

```
function test_three_cases_noise_uncomplete(cas,window)

 %cas    : 1...,
 %         2...,
 %         3...,
 %window : choice for window g here either Gaussian or Hamming

 N = 4096;
 t = (0:N-1)/N;
 a  = 2;
 s1 = a.*exp(2*pi*1i*(1000*t+60*cos(3*pi*t)));
 s2 = a.*exp(2*pi*1i*(400*t+30*cos(3*pi*t)));
 s  = s1+s2;
 s = s(:);

 Nfft = 512; %number of frequency bins

 %we build the filter g
 if strcmp(window,'hamming')
  glength=floor(161);
  glength=glength+1-rem(glength,2);%the length of the filter has to be odd
  g = tftb_window(glength,window);
 else
  prec = 10^(-3);
  sigma_w = 0.15;
  L =  Nfft*sigma_w;
  Lg = floor(L*sqrt(-log(prec)/pi))+1;
  g = amgauss(2*Lg+1,Lg+1,L); %explain what is being done here
 end

 [grow,gcol]=size(g);
 Lg=(grow-1)/2;

 %add noise to s
 %compute STFT
 %plot its modulus
 %denoise STFT
 %plot its modulus
 %invert it

end
```

Chapter 4

Reassignment Techniques and Synchrosqueezing

4.1 Introduction

Over the last 30 years, numerous methods have been proposed to extend Fourier analysis to non-stationary signals, resulting in a body of work that is referred to (at-large) as "time-frequency" (TF) methods [1, 2, 3]. Broadly speaking, generalizing Fourier analysis to take into account possible variations in the frequency content of a signal can be understood in two complementary ways. The first attempts to make time-dependent the *Fourier transform*, while the second focuses on the associated *spectral density*. The main difference is that the first approach is *linear*, with a complex-valued frequency description that involves magnitude and phase contributions, whereas the second is *quadratic* and leads to real-valued transforms in most cases.

Linear TF methods include short-time Fourier transforms (STFTs) and wavelet transforms (WTs), while most quadratic methods can be seen as variations of the celebrated Wigner-Ville distribution (WVD), with squared STFTs (spectrograms) and WTs (scalograms) as special cases. Perhaps the key point is that none of these approaches allow for the definition of one and only one transform. This follows in some sense from the uncertainty relation that links time and frequency, with the consequence that the result of any transform depends not only on intrinsic characteristics of the analyzed signal, but also on the specific properties of the chosen transform; i.e. the transform should be viewed as a measurement device. In the case of linear methods, this entanglement between the measured quantity and the measuring device takes on a special importance when, e.g., the signal under study is almost as elementary (in terms of Heisenberg-Gabor uncertainty) as the window or wavelet used for its analysis: in such a situation, one could think of the signal analyzing the window as much as the window analyzing the signal! Something similar occurs for AM-FM signals: while the idealized picture of such signals would correspond to perfectly localized trajectories associated with the instantaneous frequencies in the TF plane, values of linear transforms are spread over a ribbon whose geometry depends jointly on the signal and the window (see Figures 4.1 and 4.2 for illustrations).

In order to overcome this difficulty, in the late 70's Kodera, Gendrin and de Villedary pioneered an approach aimed at "modifying" the "moving window method" (i.e., the STFT) [4, 5]. Their analysis pointed out that the spreading of the STFT magnitude (the quantity that is usually displayed in graphical representations) can be compensated by taking into account the phase information that is usually discarded. This offered a dramatic improvement in terms of readability but because no

inversion formula exists this approach did not receive much attention.
Subsequently, in the 80's came the development of Wigner-type distributions, that could be tailored to guarantee perfect localization of signals with specific FM laws (linear for the Wigner distribution and, e.g., hyperbolic for some of its generalizations), though at the expense of new difficulties, e.g. cross-terms that hampered readability in the multicomponent case. Nevertheless, this new way of interpreting squared linear transforms permitted a revisit of Kodera's approach and an extension of its applicability beyond the STFT. Moreover, Auger and Flandrin (who coined the term "reassignment") showed in the early 90's that the explicit use of the STFT phase can be efficiently replaced by a combination of STFTs with suitable windows [6]. This was the starting point of its use in a variety of new domains, such as audio [7], physics [8] or ecology [9].
In parallel and independently, Maes and Daubechies developed another phase-based technique that they termed "synchrosqueezing" [10]. Its purpose was very similar to that of reassignment (indeed it is a special case), with the additional advantage of allowing for reconstruction.

4.2 Notation and Multicomponent Signal Definition

We recall that the STFT of f is defined by:

$$V_f^g(t,\omega) = \int_{\mathbb{R}} f(\tau) g(\tau - t) e^{-i2\pi\omega(\tau - t)}\, d\tau \tag{4.1}$$

The spectrogram $S_f^g(t,\omega)$ is then usually defined as $|V_f^g(t,\omega)|^2$. One of the most popular cases is when g is the Gaussian window $\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}}$. Multicomponent signals f to be considered hereafter in either reassignment or synchrosqueezing techniques in the STFT framework are defined by:

$$f(t) = \sum_{k=1}^{K} f_k(t), \text{ with } f_k(t) = a_k(t) e^{i2\pi\phi_k(t)}, \tag{4.2}$$

for some finite K, where $a_k(t) > 0$ is a continuously differentiable function, ϕ_k is a two times continuously differentiable function satisfying $\phi_k'(t) > 0$ and $\phi_{k+1}'(t) > \phi_k'(t)$ for all t. In the following, f_k will be referred to as an AM-FM component or a mode of f. In that context, ideal TF (ITF) representations can be defined as:

$$ITF_q(t,\omega) \quad = \quad \sum_{k=1}^{N} a_k(t)^q \delta(\omega - \phi_k'(t)), \tag{4.3}$$

where q is a positive integer depending on the chosen TF distribution (TFD), the STFT (resp. spectrogram) being associated with $q = 1$ (resp. $q = 2$). In that context, $\phi_k'(t)$ is called the instantaneous frequency (IF) of the kth mode at time t.

4.3 Uncertainty Principle for Multicomponent Signals

The most significant issue in TF signal analysis is the uncertainty principle which stipulates that one cannot localize a signal with an arbitrary precision both in time and frequency. TF representations often include parameters to allow for the balance between frequency resolution and time localization. In the case of the STFT (or spectrogram), this can be achieved by varying the size of the analysis

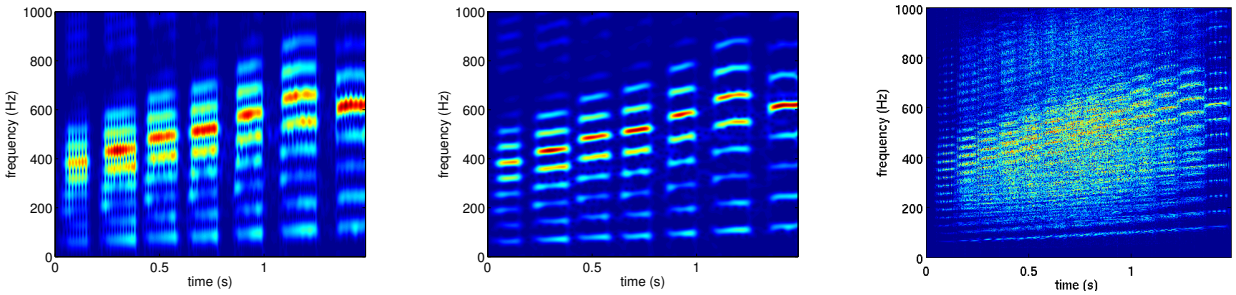


Figure 4.1: A and B: Illustration of the Heisenberg uncertainty principle: spectrogram of a saxophone sound computed with two different Gaussian windows (A: $\sigma = 9$ ms, B: $\sigma = 15$ ms); C: Wigner-Ville distribution of the saxophone sound.

window. As an illustration, we display the spectrogram of a saxophone sound (i.e. a succession of several notes) on Figure 4.1 for different sizes of the window g: a small window localizes the transients well (beginning of each note), while a large one gives precise frequency information. There have been many attempts to optimize this trade-off among which the Wigner-Ville distribution, other quadratic representations from the Cohen's class [1, 2], or multi-linear distributions (see [11] for instance). However, improvements in terms of TF resolution brought about by these techniques usually rely on strong assumptions, so that each method is suited only for a specific class of signals.

4.4 Time-Frequency Representation Enhancement with Reassignment

Reassignment techniques offer an alternative approach. They aim to sharpen the TF representation while keeping the temporal localization and are particularly well adapted to multicomponent signals. Starting with the definition (4.1) of the STFT, the spectrogram can be written as [1]:

$$S_f^g(t,\omega) = \iint_{\mathbb{R}^2} W_g(\tau - t, \nu - \omega)\, W_f(\tau,\nu)\, d\tau\, d\nu, \tag{4.4}$$

where $W_f(t,\omega)$ is the Wigner-Ville distribution (WVD), defined for any f in $L^2(\mathbb{R})$ by

$$W_f(t,\omega) := \int_{\mathbb{R}} f(t+\tau/2)\, f^*(t-\tau/2)\, e^{-i2\pi\omega\tau}\, d\tau. \tag{4.5}$$

The spectrogram is thus the 2D smoothing of the WVD of the analyzed signal by the WVD of the analyzing window. This alternative formulation allows for a simple understanding of the main features of a spectrogram when compared to a WVD. On the one hand, a WVD is known to sharply localize individual linear chirps in the TF plane but the 2D smoothing involved in the spectrogram computation results in a *smearing* of their energy distribution. On the other hand, the quadratic nature of the WVD is known to create oscillatory interference between individual components ([3, chapter 3], see Figure 4.1 C for an illustration) which can be removed by the 2D smoothing used in the spectrogram computation. The analysis is therefore faced with a trade-off between TF localization and the interference level.

From (4.4), it follows that the value of the spectrogram at (t,ω) is the sum of all WVD signal contributions within the TF domain over which the WVD of the window is essentially nonzero. The principle of the *reassignment method* (RM) illustrated here on the spectrogram is to compensate for the TF shifts induced by the 2D smoothing defining the spectrogram. To do so, a meaningful TF location to which to assign the local energy given by the spectrogram is first determined. It corresponds to the *centroid* of the distribution (4.4), whose coordinates are defined by

$$\hat{\omega}_f(t,\omega) := \frac{1}{S_f^g(t,\omega)} \iint_{\mathbb{R}^2} \nu \, W_g(\tau - t, \nu - \omega) \, W_f(\tau,\nu) \, d\tau \, d\nu$$
$$\hat{t}_f(t,\omega) := \frac{1}{S_f^g(t,\omega)} \iint_{\mathbb{R}^2} \tau \, W_g(\tau - t, \nu - \omega) \, W_f(\tau,\nu) \, d\tau \, d\nu.$$

Both quantities, which define *locally* an instantaneous frequency and a group delay, enable perfect localization of linear chirps, i.e. $\phi'(\hat{t}_f(t,\omega)) = \hat{\omega}_f(t,\omega)$. RM then consists in moving the value of the spectrogram from the point of computation to this centroid [6]:

$$\hat{S}_f^g(t,\omega) = \iint_{\mathbb{R}^2} S_f^g(\tau,\nu) \, \delta\left(\omega - \hat{\omega}_f(\tau,\nu)\right) \delta\left(t - \hat{t}_f(\tau,\nu)\right) \, d\tau \, d\nu, \tag{4.6}$$

where δ stands for the Dirac distribution. Due to the above mentioned property of $(\hat{t}_f(t,\omega), \hat{\omega}_f(t,\omega))$, RM perfectly localizes linear chirps while removing most of the interference. However, in practice, the centroid $(\hat{t}_f, \hat{\omega}_f)$ is not evaluated as above. Remarking that the above operators can be written as:

$$\hat{t}_f(t,\omega) := t - \frac{1}{2\pi} \partial_\omega \arg V_f^g(t,\omega),$$
$$\hat{\omega}_f(t,\omega) := \frac{1}{2\pi} \partial_t \arg V_f^g(t,\omega),$$

a more efficient [6] procedure computes them according to

$$\hat{\omega}_f(t,\omega) = \omega - \frac{1}{2\pi} \Im \left\{ \frac{V_f^{g'}(t,\omega)}{V_f^g(t,\omega)} \right\}$$
$$\hat{t}_f(t,\omega) = t + \Re \left\{ \frac{V_f^{tg}(t,\omega)}{V_f^g(t,\omega)} \right\},$$

where tg stands for the function $tg(t)$ and $\Re\{Z\}$ (resp. $\Im\{Z\}$) is the real (resp. imaginary) part of the complex number Z. Compared to the standard spectrogram, its reassigned version can thus be computed with a moderate increase in the computational cost, since three STFTs are evaluated (and combined) instead of one. An illustration of RM is given in the sub-figure called "RM" of Figure 4.2.

4.5 Multicomponent Signal Reconstruction with Synchrosqueezing

While RM provides a direct and powerful representation of a multicomponent signal in the TF plane, no mode reconstruction technique using the reassigned transform is straightforward. In contrast,

the SynchroSqueezing Transform (SST), introduced by Maes and Daubechies in [10], enhances the TFD given by the STFT in a manner similar to RM with the spectrogram, but still enables mode retrieval. This property is of great importance since the understanding of a multicomponent signal as defined in (4.2) is tightly related to the analysis of its constituent modes. Furthermore, in contrast to EMD, mode reconstruction using SST is carried out in a convenient mathematical framework. Indeed, a recent result [16] shows that the SST is a good approximation to the ideal TF representation of the signal f (with $q = 1$ see (4.3)) and enables mode reconstruction when f is made of weakly modulated modes. We now present the SST in the STFT framework with the emphasis on the differences with RM providing insights into some theoretical results.

4.5.1 SST in a Nutshell

In contrast with RM which enhances the TFD given by the spectrogram, the SST operates directly on the STFT. The construction of the SST is closely related to the synthesis formula

$$f(t) = \frac{1}{g(0)} \int_{\mathbb{R}} V_f^g(t,\omega)\, d\omega, \tag{4.7}$$

which sums the STFT in frequency for each time t. However, when f is made of separate components f_k in the TF plane, i.e. verifying (4.2) and conditions (4.11) and (4.12) (see below), the main part of the STFT of f_k is localized in the vicinity of the *ridge* $(t, \phi_k'(t))$, which can be seen as the TF trajectory associated with component k. The interesting angular frequencies (AF) ω used to reconstruct f_k are then selected as those such that $\hat{\omega}_f(t,\omega)$ is a good approximation of $\phi_k'(t)$ [17]. Based on this idea, the *synchrosqueezing operator* reassigns the STFT as follows:

$$T_f^g(t,\omega) = \frac{1}{g(0)} \int_{\mathbb{R}} V_f^g(t,\nu)\delta(\omega - \hat{\omega}_f(t,\nu))\, d\nu. \tag{4.8}$$

Note that this definition is similar to that used by RM, except that time reassignment is not considered and V_f^g is used instead of the spectrogram (an illustration of what the operator T_f^g does is given in the sub-figure called "SST" of Figure 4.2).

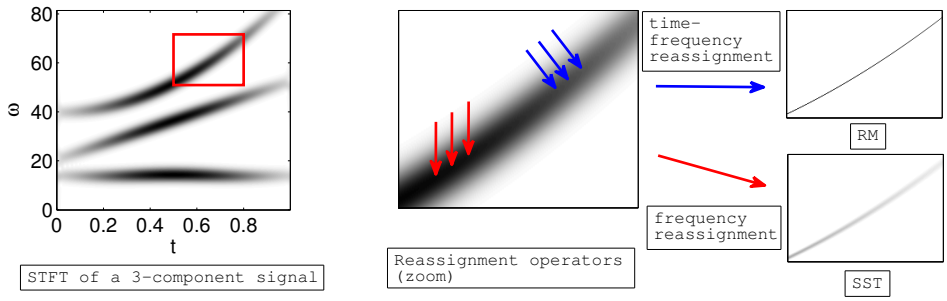


Figure 4.2: center: blue (resp. red) arrows symbolizing how the reassignment is performed with RM (resp. SST) on a small patch (delimited by the red segments) extracted from the 3-component signal STFT modulus depicted on the left; right: reassignment carried out with RM (resp. SST) for the signal STFT depicted in the central subfigure.

Having computed the synchrosqueezing operator T_f^g, the kth mode is then reconstructed by integrating T_f^g in the vicinity of the corresponding ridge:

$$f_k(t) \approx \int_{\{\omega,|\omega-\phi_k'(t)|<d\}} T_f^g(t,\omega)\ d\omega, \tag{4.9}$$

for some small parameter d. Since $\phi_k'(t)$ is unknown and needs to be estimated in practice, this local averaging of T_f^g in frequency (t being fixed) is compulsory to retrieve f_k.
Given T_f^g and assuming the number K of components is known, a ridge extraction technique can be used to estimate the ϕ_k's before proceeding with mode retrieval. For this, there exists a variety of methods [18][19], but those based on ridge estimation are particularly well adapted to the SST case. Briefly, the principle of the latter techniques is to minimize the following energy functional initially proposed by Carmona et *al.* [20]:

$$E_f(\varphi) = \sum_{k=1}^{K} -\int_{\mathbb{R}} |T_f^g(t,\varphi_k(t))|^2\,dt + \int_{\mathbb{R}} \lambda\varphi_k'(t)^2 + \beta\varphi_k''(t)^2\,dt. \tag{4.10}$$

Doing so, one finds smooth curves φ_k along which the magnitude of T_f^g is maximal, λ and β enabling a trade-off between smoothness of the curve and energy maximization. Although this general variational formulation looks very appealing, it is hard to implement. Heuristic algorithms such as for instance simulated annealing [20] or the crazy climbers algorithm [21] which are particularly appropriate in a noisy context have to be used. Another simple yet efficient approach related to the resolution of (4.10) was developed in [22] and consists of a local determination of the ridges starting from different initializations and then in an averaging of the results obtained. All these three methods behave very similarly on the examples studied, and the illustrative examples that follow will use the last method as ridge estimator. A summary of how the SST performs mode retrieval is given in Figure 4.2. To conclude, it is worth noting here alternative approaches that assume a polynomial phase for the modes [23], whereas the SST does not require such an assumption.

4.5.2 Mathematical Foundations of the SST

The discussion above on mode retrieval is reinforced by some theoretical results mostly derived in the wavelet framework in [16] and then adapted to the STFT context in [17]. Our goal here is not to delve too deeply into the derivation of these results, but to focus on what type of signals they are valid for. They are obtained under the following assumptions on the signal f, assuming the window g is Gaussian:

A1 the f_ks have weak frequency modulation, implying the existence of a small ε such that for each t, one has:

$$\sigma^2|\phi_k''(t)| \leq \varepsilon \text{ and } |a_k'(t)| \leq \varepsilon\phi_k'(t), \tag{4.11}$$

where σ is the size of the Gaussian window.

A2 The modes are well *separated* in frequency which corresponds, assuming the frequency bandwidth of g (in rad/s) is $[-\Delta,\Delta]$ ($\Delta = \frac{\sqrt{2\log(2)}}{\sigma}$ since g is the Gaussian window), to an inequality of type

$$|\phi_k'(t) - \phi_l'(t)| \geq 2\Delta \tag{4.12}$$

for each t and $k \neq l$.

Note that because the modes f_k are such that $\phi'_k(t) < \phi'_{k+1}(t)$ for all t, (4.12) can always be satisfied by choosing a a Gaussian window with the appropriate size. A mode f_k satisfying A1 will be called in the sequel a $\mathcal{A}_\varepsilon$ signal. Under these hypotheses, one has that:

1. For any (t,ω) in $\bigcup_k \{(t,\omega);\ |\omega - \phi'_k(t)| \leq \Delta\} = \bigcup_k B_k$, there exists a k such that $|\hat{\omega}_f(t,\omega) - \phi'_k(t)| \leq C\varepsilon$, where C is some constant.

2. The reconstruction error associated with the retrieval of f_k by summing the coefficients around $T_f^g(t,\phi'_k(t))$ following (4.9) tends to zero as ε goes to zero.

This clearly establishes, on one hand, the relation between the amplitude of ε and the quality of the modes retrieval and, on the other hand, the role played by the window's size in the separation of the components.

4.5.3 Denoising Multicomponent Signals using SST

Here we illustrate how the SST provides us with a naive denoising procedure that can outperform a state-of-the-art method based on wavelets. To do so, we consider the signal whose STFT is shown on the left sub-figure of Figure 4.2 to which we add white Gaussian noise with varying standard deviation, leading to different SNR (SNR in). Then we denoise this signal with the Block-Thresholding (BT) method developed in [24], a TF technique designed for audio recordings, and also with the SST, by simply selecting 3 ridges as explained in section 4.5.1 (with $\lambda = 0$ and $\beta = 0.02$; these parameters, though not optimal, lead to good results in practice). The results displayed on Figure 4.3 in terms of the SNR before and after denoising (SNR in and out respectively, the SNR out being computed as $\frac{\|f\|^2}{\|f-\tilde{f}\|^2}$ where $\tilde{f}$ is the denoised signal) show a much better denoising performance of the SST over BT on this particular example, provided σ is appropriately chosen. As soon as the ridges are detected, the SST thus enables signal denoising in a straightforward manner. However, when the signal contains highly modulated modes, as in the studied case, the use of too large a window should be proscribed, as is reflected by the first inequality of (4.11). In this regard, we will try later to go beyond this limitation by taking into account the modulation in the synchrosqueezing operator. It is also of note that, as the modes are not perfectly monochromatic, their SST is not perfectly reassigned onto the ridges. In this regard, the parameter d introduced in the reconstruction formula (4.9) enables compensation for the lack of accuracy in the estimation of $\phi'_k(t)$ by means of $\hat{\omega}_f(t,\omega)$. Indeed, it has to be chosen all the larger that the modulation is important ($d = 8$ rad/s being satisfactory in the studied case). Finally, we remark that by considering modified versions

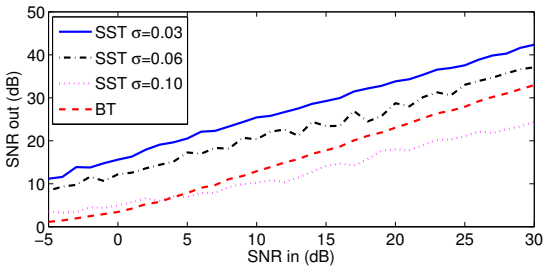


Figure 4.3: Denoising performance with the SST and BT techniques for the signal whose STFT is depicted on the left of Figure 4.2 and for various window's size σ, d being fixed to 8 rad/s.

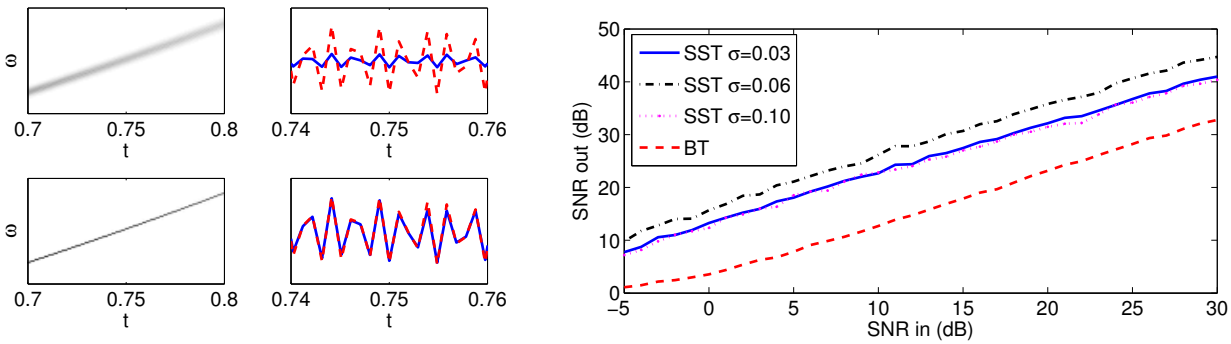
of the spectrogram one could obtain better ridge estimations than with the SST in such a noisy context [14]. However, doing so, the reconstruction properties inherent to the SST would be lost.

4.5.4 Adapting the SST to a Stronger Modulation

Our goal is here to propose a new development to better take mode modulation into in the SST. For a mode $f(t) = a\, e^{i2\pi\phi(t)}$ such that $\phi''(t) = 0$ for all t, one exactly has $\phi'(t) = \hat{\omega}_f(t,\omega)$, so that when ϕ'' is small compared to ϕ' in the vicinity of t, the approximation of $\phi'(t)$ by $\hat{\omega}_f(t,\omega)$ is fully justified. However, when it is not the case, e.g. when the signal f is a linear chirp $f(t) = ae^{i\alpha t^2} = ae^{i\phi(t)}$, and when a Gaussian analysis window is used, then:

$$\phi'(t) = \hat{\omega}_f^c(t,\omega) := \hat{\omega}_f(t,\omega) + \phi''(t)^2\sigma^4(\hat{\omega}_f(t,\omega) - \omega). \tag{4.13}$$

This expression *a posteriori* explains why the denoising performance of the method based on the SST introduced in section 4.5.3 deteriorates when σ is chosen too large. Indeed, when dealing with strongly modulated components the factor $\phi_k''(t)^2\sigma^4$ is no longer negligible, making the approximation of $\phi'(t)$ by $\hat{\omega}_f(t,\omega)$ very inaccurate.



A B

Figure 4.4: A: top left: $|T_f^g|$ of the signal of the central sub-figure of Figure 4.2; top right : reconstructed mode (solid line) and the ground truth f_3 (dashed line), with $d = 1$ rad/s. Second row: same figures using $\tilde{T}_f^g$ instead of T_f^g. B: denoising results using the algorithm presented in section 4.5.3 but using $\tilde{T}_f^g$ (the parameter d being still fixed to 8 rad/s)

For a linear chirp, one has $\hat{t}_f(t,\omega) = t + \frac{\sigma^4\phi''(t)}{1+\phi''(t)^2\sigma^4}(\omega - \phi'(t))$, therefore using (4.13) one obtains $\phi''(t) = -\frac{1}{\sigma^4}\frac{\hat{t}_f(t,\omega)-t}{\hat{\omega}_f(t,\omega)-\omega}$, and finally the following closed form for $\omega_f^c(t,\omega)$: $\hat{\omega}_f(t,\omega) + \frac{1}{\sigma^4}\frac{(\hat{t}_f(t,\omega)-t)^2}{\hat{\omega}_f(t,\omega)-\omega}$. Defining a new synchrosqueezing operator $\tilde{T}_f^g$ by replacing $\hat{\omega}_f(t,\omega)$ by $\hat{\omega}_f^c(t,\omega)$ leads to a sharper representation. An illustration of this is given in Figure 4.4 A. The reconstruction of the high frequency mode of the left sub-figure of Figure 4.2, carried out using only the coefficients on the ridge associated with that mode, shows that the signal energy is much more concentrated around the ridge when $\tilde{T}_f^g$ is used instead of T_f^g (SNR after reconstruction equals 40 dB (resp. 10 dB) in the first (resp. second) case).

To consider the ridge estimation based on $\hat{\omega}_f^c(t,\omega)$ rather than $\hat{\omega}_f(t,\omega)$ not only improves the quality of the reconstruction of strongly modulated modes but also impacts the denoising performance of the SST-based technique. Indeed, the denoising algorithm obtained by selecting the coefficients around the ridges obtained from T_f^g was found to be sensitive to the choice for the window's size. On the

contrary, because of relation (4.13), $\hat{\omega}_f^c(t,\omega)$ leads to a stable approximation of $\phi'(t)$ in the case of a strong modulation when σ varies. This has the following consequence: the denoising results obtained with the SST-based technique are improved with this new estimator when σ increases (comparison between Figure 4.3 and 4.4 B).

4.5.5 Computation of the new IF estimate

Based on the above developments we would like to define a new IF estimate in a proper mathematical framework. Let us define $\tilde{\omega}_f(t,\eta) = \frac{\partial_t V_f^g(t,\eta)}{2i\pi V_f^g(t,\eta)}$ and then introduce:

$$\tilde{t}_f(t,\eta) \quad = \quad t - \frac{\partial_\eta V_f^g(t,\eta)}{2i\pi V_f^g(t,\eta)}. \tag{4.14}$$

Similarly to $\widehat{\omega}_f(t,\eta) = \mathcal{R}(\tilde{\omega}_f(t,\eta))$, we define $\widehat{t}_f(t,\eta) = \mathcal{R}(\tilde{t}_f(t,\eta))$. First, we define an estimate of the frequency modulation as follows.

Definition 1 *Let $f \in L^2(\mathbb{R})$ and consider when $V_f^g(t,\eta) \neq 0$ and $\partial_t \left(\frac{\partial_\eta V_f^g(t,\eta)}{V_f^g(t,\eta)} \right) \neq 2i\pi$ the quantity*

$$\tilde{q}_f(t,\eta) \quad = \quad \frac{\partial_t \tilde{\omega}_f(t,\eta)}{\partial_t \tilde{t}_f(t,\eta)} = \frac{\partial_t \left(\frac{\partial_t V_f^g(t,\eta)}{V_f^g(t,\eta)} \right)}{2i\pi - \partial_t \left(\frac{\partial_\eta V_f^g(t,\eta)}{V_f^g(t,\eta)} \right)}. \tag{4.15}$$

An estimate of the frequency modulation is then defined by

$$\hat{q}_f(t,\eta) = \mathcal{R}\left(\tilde{q}_f(t,\eta)\right). \tag{4.16}$$

Definition 2 *Let $f \in L^2(\mathbb{R})$, we define the second order IF complex estimate of f as:*

$$\tilde{\omega}_f^{(2)}(t,\eta) \quad = \quad \begin{cases} \tilde{\omega}_f(t,\eta) + \tilde{q}_f(t,\eta)(t - \tilde{t}_f(t,\eta)) & \text{if } \partial_t \tilde{t}_f(t,\eta) \neq 0 \\ \tilde{\omega}_f(t,\eta) & \text{otherwise,} \end{cases} \tag{4.17}$$

and then its real part $\hat{\omega}_f^{(2)}(t,\eta) = \mathcal{R}(\tilde{\omega}_f^{(2)}(t,\eta))$.

Remark 1 *It can be proved that $\hat{q}_f(t,\eta) = \phi''(t)$ and $\hat{\omega}_f^{(2)}(t,\eta) = \phi'(t)$, when f is a Gaussian modulated linear chirp, i.e. a chirp where both ϕ and $\log(A)$ are quadratic.*

Proposition 1

Let $f \in L^2(\mathbb{R})$, then the IF estimate $\widehat{\omega}_f^{(2)}(t,\eta)$ can be expressed by means of five different STFTs, since we have:

$$\begin{aligned}
\tilde{\omega}_f(t,\eta) &= \eta - \frac{1}{2i\pi}\frac{V_f^{g'}(\eta,t)}{V_f^{g}(\eta,t)},\\
\tilde{q}_f(t,\eta) &= \frac{1}{2i\pi}\frac{V_f^{g''}(t,\eta)V_f^{g}(t,\eta) - (V_f^{g'}(t,\eta))^2}{V_f^{tg}(t,\eta)V_f^{g'}(t,\eta) - V_f^{tg'}(t,\eta)V_f^{g}(t,\eta)},\\
t - \tilde{t}_f(\eta,t) &= -\frac{V_f^{tg}(t,\eta)}{V_f^{g}(t,\eta)}.
\end{aligned}$$

provided g, g', g'', tg, tg' are in $L^1(\mathbb{R})$.

Proof The expressions for $\tilde{\omega}_f(t,\eta)$ and $t-\tilde{t}_f(t,\eta)$ are straightforward. Indeed, since g is in $C^2(\mathbb{R})$, the STFT of f belongs to $C^\infty(\mathbb{R})$, and we have:

$$\begin{aligned}
\partial_\eta V_f^g(t,\eta) &= -2i\pi V_f^{tg}(t,\eta)\\
\partial_t V_f^g(t,\eta) &= 2i\pi\eta V_f^g(t,\eta) - V_f^{g'}(t,\eta).
\end{aligned} \tag{4.18}$$

Based on these equalities, the expression for $\tilde{q}_f(t,\eta)$ is easy to obtain.

The second order *synchrosqueezing operator* then reassigns the STFT as follows:

$$T_{f,2}^g(t,\omega) = \frac{1}{g(0)}\int_{\mathbb{R}} V_f^g(t,\nu)\delta(\omega - \hat{\omega}_f^{(2)}(t,\nu))\,d\nu. \tag{4.19}$$

and then the modes are reconstructed through:

$$f_k(t) \approx \int_{\{\omega,|\omega-\phi_k'(t)|<d\}} T_{f,2}^g(t,\omega)\ d\omega, \tag{4.20}$$

4.6 Applications of SST to Physiological Signals

Particular applications of SST are related to the study of ECG signals [26] of which we give an illustration hereafter. During anesthesia, the anesthetic agents exert differential effects on the neural activity of different regions of the brain. While the cortical activity is commonly recorded by electroencephalography (EEG), this technique is not adapted to assist in the control of the essential components of anesthesia including motor suppression, analgesia and autonomic activity, which are largely governed by the sub-cortical regions, such as the autonomic nervous system (ANS). It is well known that the ANS regulates the vital physiological functions and controls emergency responses [27, chapter 12]. A non-invasive and common technique is to assess the ANS activity using ECG recordings (see Figure 4.5 A) by measuring the variability of the time intervals between sequential heart beats, called the heart rate variability (HRV). This measurement analyzed using SST may be able to provide information of use in anesthesia. In clinical practice, the heart beat rate is assessed

by computing the number of beats (or R peaks) in a minute. However, it is likely that information is hidden in the measurement, which can be uncovered by analyzing the R-to-R peak signal by means of the RRI signal, defined as the cubic-spline interpolant of $\left(t_i, \frac{1}{t_{i+1}-t_i}\right)$, the R peaks being located in $\{t_i\}_{i=1}^N$. The RRI signal of an anesthetized patient changes drastically at the time of waking as illustrated by Figure 4.5 B around the time 800 s.

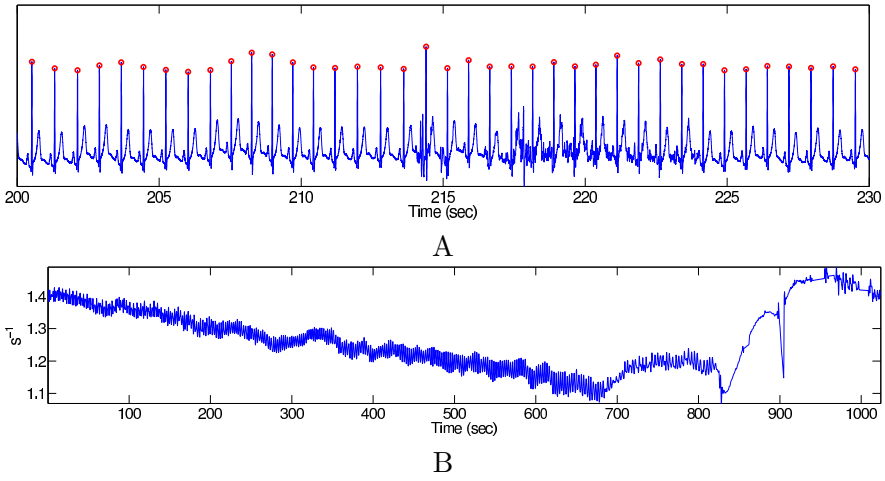


Figure 4.5: A: small portion of an ECG signal (the blue curve), the red circles indicate the R peaks. B: RRI signal associated with an ECG signal of a patient waking up from anesthesia (the waking time is around 800 s)

Our concern is to show how to use SST to study RRI signals of patients waking up from anesthesia with spontaneous breathing [28], by studying the RRI signal of Figure 4.5 B. To start with, we compute a detrended RRI, called RRI_D, by subtracting to the original RRI signal its local mean $m(t)$ computed using a median filter, i.e. $RRI(t) = m(t) + RRI_D(t)$. We then apply the SST to RRI_D (computed using a Kaiser window g) to finally obtain the representation of Figure 4.6.

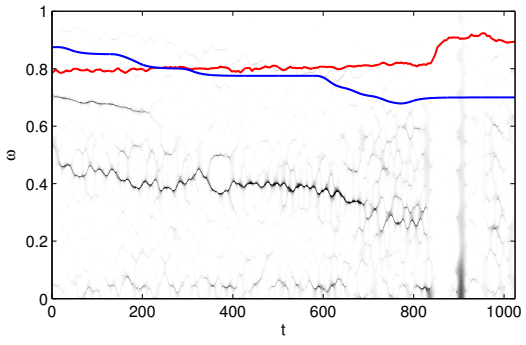


Figure 4.6: moduli of the SST of RRI_D computed using a Kaiser window. The red curve superimposed is the BIS index reflecting the anesthetic level – the lower the value, the deeper the anesthetic level. The blue curve superimposed is the anesthetic drug concentration. Notice the dramatic change of BIS and that of the reassigned spectrogram around 800 s

To analyze Figure 4.6, we consider a signal in the $\mathcal{A}_\varepsilon$ class rhythmic (see section 4.5.2), otherwise non-rhythmic. On Figure 4.6, a transition from rhythmic to non-rhythmic dynamics in the RRI signal is visible: before 800 s, there are two dominant curves corresponding to two $\mathcal{A}_\epsilon$ functions – from 0 to 200 s and from 0 to 800 s, while after 800 s, no such behavior persists. In other words, RRI_D can be written in the form $RRI_D(t) = \sum_{k=1}^{2} a_k(t)\cos(\phi_k(t))$ with a_k, ϕ_k satisfying (4.11) before 800 s (one of the component vanishing after 200 s), whereas is irrelevant after this time. Furthermore this change of behavior also corresponds to the transition from deep to light level of anesthesia as corroborated by the objective anesthetic depth index, the Bi-spectral index (BIS) (the red curve superimposed on Figure 4.6). BIS is evaluated from the simultaneously recorded electroencephalography; the higher the BIS index, the lighter the anesthetic depth. We also superimpose on Figure 4.6 the anesthetic drug concentration for comparison (blue curve).
Cortical activity is known to become non rhythmic when the anesthetic level decreases. Combined with the physiological fact that HRV is mainly controlled by the sub-cortical level, the above results show that the same phenomenon exists in RRI signals, which in turn suggests that the sub-cortical activity becomes non rhythmic when the anesthetic level decreases.

Bibliography

[1] F. Hlawatsch and G.F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Sig. Proc. Mag.*, vol. 9, no. 2, pp. 21–67, 1992.

[2] L. Cohen, *Time-frequency analysis: theory and applications*, Prentice-Hall, Inc., 1995.

[3] P. Flandrin, *Time-Frequency/Time-Scale analysis*, Academic Press, San Diego (CA), 1999.

[4] K. Kodera, C. De Villedary, and R. Gendrin, "A new method for the numerical analysis of non-stationary signals," *Phys. Earth Plan. Inter.*, vol. 12, no. 2-3, pp. 142–150, 1976.

[5] K. Kodera, R. Gendrin, and C. Villedary, "Analysis of time-varying signals with small BT values," *IEEE Trans. Acoust., Speech and Sig. Proc.*, vol. 26, no. 1, pp. 64–76, 1978.

[6] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Sig. Proc.*, vol. 43, no. 5, pp. 1068–1089, 1995.

[7] S.A. Fulop and K. Fitz, "Algorithms for computing the time corrected instantaneous frequency (reassigned) spectrogram, with applications," *Journal of the Acoustical Society of America*, vol. 119, no. 1, pp. 360–371, 2006.

[8] O. Kotte, M. Niethammer, and L.J. Jacobs, "Lamb waves characterization by differential reassignment and non-linear anisotropic diffusion," *NDT&E International*, vol. 39, pp. 96–105, 2006.

[9] B. Dugnol, C. Fernandez, G. Galiano, and J. Velasco, "Implementation of a diffusive differential reassignment method for signal enhancement: An application to wolf population counting," *Applied Mathematics and Computation*, vol. 193, pp. 374–384, 2007.

[10] I. Daubechies and S. Maes, "A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models," *Wavelets Med. Biol.*, pp. 527–546, 1996.

[11] B. Boashash and P. O'Shea, "Polynomial Wigner-Ville distributions and their relationship to time-varying higher order spectra," *IEEE Trans. Sig. Proc.*, vol. 42, no. 1, pp. 216–220, 1994.

[12] F. Auger, E. Chassande-Mottin, and P. Flandrin, "On phase-magnitude relationships in the short-time Fourier transform," *IEEE Sig. Proc. Lett.*, vol. 19, no. 5, pp. 267–270, 2012.

[13] B. Boashash and V. Sucic, "Resolution measure criteria for the objective assessment of the performance of quadratic time-frequency distributions," *IEEE Trans. Sig. Proc.*, vol. 51, no. 5, pp. 1253–1263, 2003.

[14] I. Djurovic and L.J. Stankovic, "Time-frequency representation based on the reassigned S-method," *Sig. Proc.*, vol. 77, pp. 115–120, 1999.

[15] N. E. Huang, Z. Shen, S. R. Long, Wu, M. L., H. H. Shih, Q. Zheng, Yen N. C., C. C. Tung, and H. H. Liu, "The Empirical Mode Decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. of London A: Math.l, Phys. and Eng. Sc.*, vol. 454, pp. 903–995, 1998.

[16] I. Daubechies, J. Lu, and H.T. Wu, "Synchrosqueezed wavelet transforms: an Empirical Mode Decomposition-like tool," *Appl. Comp. Harm. Anal.*, vol. 30, no. 2, pp. 243–261, 2011.

[17] J. Thakur and H-T. Wu, "Synchrosqueezing based recovery of instantaneous frequency from nonuniform samples," *SIAM J. Math. Anal.*, vol. 43, no. 5, pp. 2078–2095, 2011.

[18] I. Djurovic and L.J. Stankovic, "An algorithm for the Wigner distribution based instantaneous frequency estimation in a high noise environment," *Sig. Proc.*, vol. 84, no. 3, pp. 631–643, 2004.

[19] B. Boashash and P. O'Shea, "Use of the cross Wigner-Ville distribution for estimation of instantaneous frequency," *Sig. Proc.*, vol. 41, no. 3, pp. 1439–1445, 1993.

[20] R.A. Carmona, W.L. Hwang, and B. Torresani, "Characterization of signals by the ridges of their wavelet transforms," *IEEE Trans. Sig. Proc.*, vol. 45, no. 10, pp. 2586–2590, 1997.

[21] R.A. Carmona, W.L. Hwang, and B. Torresani, "Multiridge detection and time-frequency reconstruction," *IEEE Trans. Sig. Proc.*, vol. 47, no. 2, pp. 480–492, 1999.

[22] G. Thakur, E. Brevdo, N.S Fuckar, and H.-T. Wu, "The Synchrosqueezing algorithm for time-varying spectral analysis: robustness properties and new paleoclimate applications," *Sig. Proc.*, vol. 93, no. 5, pp. 1079–1094, 2013.

[23] S. Barbarossa, A. Scaglione, and G.B. Giannakis, "Product high-order ambiguity function for multicomponent polynomial phase signal modeling," *IEEE Trans. Sig. Proc.*, vol. 46, no. 3, pp. 691–708, 1998.

[24] G. Yu, S. Mallat, and E. Bacry, "Audio denoising by time-frequency block thresholding," *IEEE Trans. Sig. Proc.*, vol. 56, no. 5, pp. 1830–1839, 2008.

[25] F. Auger, E. Chassande-Mottin, and P. Flandrin, "Making reassignment adjustable: the Levenberg-Marquardt approach," in *Proc. IEEE-ICASSP*, Kyoto, pp. 3889–3892, 2012.

[26] H-T. Wu, "Instantaneous frequency and wave shape functions (I)," *Appl. Comp. Harm. Anal., http://dx.doi.org/10.1016/j.acha.2012.08.008*, 2012.

[27] R.D. Miller, L.I. Eriksson, L.A. Fleisher, J.P. Wiener-Kronish, and W.L. Young, *Miller's anesthesia*, Churchill Livingstone, 7th edition, 2009.

[28] Y.T. Lin, S.S. Hseu, H.W. Yien, and J. Tsao, "Analyzing autonomic activity in electrocardiography about general anesthesia by spectrogram with multitaper time-frequency reassignment," *Proc. IEEE-BMEI*, vol. 2, pp. 628–632, 2011.