Reconstruction de phase pour la transformée en ondelettes

Irène Waldspurger
CEREMADE, Université Paris Dauphine
(en collaboration avec Stéphane Mallat)

19 mai 2017
Séminaire du LJLL
Audio processing

Raw representation of a signal:

\[ f : \mathbb{R} \to \mathbb{R}. \]

A priori difficult to analyze.

We use the \textbf{wavelet transform}

\[ Wf : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{C}. \]

But we keep only the \textbf{modulus}:

\[ |W|f : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}^+. \]
The scalogram has two desirable properties [Risset and Wessel, 1999; Balan et al., 2006]

**Stability to unaudible transformations**

Two signals that are identical to the human ear have very similar scalograms.

**Discriminability**

Two signals that sound different have different scalograms.

Theoretical analysis of these properties?
We consider the **inverse problem**:

To what extent is it possible to reconstruct a signal from its scalogram?

This is a **phase retrieval problem**.

Inverting the wavelet transform is easy, but we have only the modulus. We first need to recover the phase.
Motivations

- Understand the theoretical properties of the scalogram.
- Some audio processing tasks require to reconstruct a signal from the modified scalogram of another signal. [Virtanen, 2007]
- It adds an original item to the family of phase retrieval problems whose properties we understand.
Main questions

Theoretical aspects

- Uniqueness:
  Is the signal uniquely determined from its scalogram?

- Stability to noise:
  If the scalogram is only approximately known, what information does it give on the signal?

Algorithmical aspects

- Design an algorithm that is both accurate and fast.
Summary

1. Definitions; presentation of phase retrieval problems
2. Reconstruction from the scalogram: theoretical aspects
   ▶ Uniqueness
   ▶ No « strong » stability, but a « local » stability
3. Reconstruction from the scalogram: algorithms
Definition of the wavelet transform

**Wavelet**: $\psi \in L^1 \cap L^2(\mathbb{R})$ such that $\int_{\mathbb{R}} \psi(t) dt = 0$

\[
\forall j \in \mathbb{Z} \quad \psi_j(t) = 2^{-j} \psi(2^{-j} t) \\
\iff \quad \hat{\psi}_j(\omega) = \hat{\psi}(2^j \omega)
\]
Definition of the wavelet transform

**Wavelet**: $\psi \in L^1 \cap L^2(\mathbb{R})$ such that $\int_{\mathbb{R}} \psi(t) dt = 0$

$\forall j \in \mathbb{Z} \quad \psi_j(t) = 2^{-j} \psi(2^{-j}t)$

$\iff \hat{\psi}_j(\omega) = \hat{\psi}(2^j \omega)$

**Wavelet transform**: $W : f \in L^2(\mathbb{R}) \rightarrow \{ f \ast \psi_j \}_{j \in \mathbb{Z}}$
Phase retrieval problems

Example

The $f \ast \psi_j$ are complex-valued. Only real parts are displayed.
Phase retrieval problems

Example

\[ |W| f = \{|f \ast \psi_j|\} \]

The \( f \) are complex-valued. Only real parts are displayed.

Scalogram = modulus of the wavelet transform
Reconstruct $f$ from $|W|f = \{|f \ast \psi_j|\}_{j \in \mathbb{Z}}$?

This is an example of a **phase retrieval problem**.
Reconstruct \( f \) from \( |W|f = \{|f * \psi_j|\}_{j \in \mathbb{Z}} \)?

This is an example of a \textbf{phase retrieval problem}.

A generic phase retrieval problem:

Reconstruct \( x \in V \) from \( \{|L_i(x)|\}_{i \in I} \)?

where:

\( \rightarrow \) \( V \) is a known complex vector space;

\( \rightarrow \) \( \{L_i\}_{i \in I} \) is a fixed set of linear forms on \( V \).

Reconstruction \textbf{up to a global phase}:

\( x \sim ux \) if \( |u| = 1 \).
Main questions

Reconstruct $x \in V$ from $\{ |L_i(x)| \}_{i \in I}$?

- **Uniqueness**

  $$(\forall i, |L_i(x)| = |L_i(y)|) \implies (x = y) \quad ?$$

- **Stability**

  $$(\forall i, |L_i(x)| \approx |L_i(y)|) \implies (x \approx y) \quad ?$$

- **Algorithm** ?
Random linear forms

When the $L_i$ are randomly chosen:
- uniqueness and stability;
- polynomial algorithm.

[Balan et al., 2006; Candès et al., 2013]

Deterministic linear forms

« Built on purpose » linear forms
[Balan et al., 2009; Bodmann and Hammen, 2014]

« Physical » linear forms
Often more difficult to analyze; problems tend to be ill-posed.
The wavelet transform is one of the few cases that can be precisely analyzed, at least for a specific choice of wavelets.
The wavelet transform is one of the few cases that can be precisely analyzed, at least for a specific choice of wavelets.

**Overview**

- **Uniqueness**
- No "strong" stability
- Instabilities can be described
- Reconstruction algorithm

} for Cauchy wavelets

} for relatively general wavelets
Cauchy wavelets

\[ \hat{\psi}(\omega) = \omega^p e^{-\omega} 1_{\omega \geq 0} \]

\((p > 0 : \text{arbitrary parameter})\)

These wavelets are \textit{analytical}: their Fourier transform is zero on \(\mathbb{R}^-\).

\[ \Rightarrow \{ |f \ast \psi_j| \}_{j \in \mathbb{Z}} \text{ does not depend upon } \hat{f}(\omega) \text{ for } \omega \leq 0. \]
Theorem (Uniqueness)

Let $f, g \in L^2(\mathbb{R})$ be such that $\hat{f}(\omega) = \hat{g}(\omega) = 0$ if $\omega \leq 0$. We assume that, for any $j \in \mathbb{Z}$:

$$|f \ast \psi_j| = |g \ast \psi_j|.$$

Then there exists $\phi \in \mathbb{R}$ such that:

$$f = ge^{i\phi}$$
Theorem (Uniqueness)

Let $f, g \in L^2(\mathbb{R})$ be such that $\hat{f}(\omega) = \hat{g}(\omega) = 0$ if $\omega \leq 0$. We assume that, for any $j \in \mathbb{Z}$:

$$|f \ast \psi_j| = |g \ast \psi_j|.$$

Then there exists $\phi \in \mathbb{R}$ such that:

$$f = ge^{i\phi}$$

Remark: A similar result holds for real-valued functions (instead of functions with no negative frequencies).
Idea of proof in a simpler case

Simplification
- \( \hat{f}(\omega) \) and \( \hat{g}(\omega) \) converge to non-zero limits when \( \omega \to 0^+ \)

Key property of Cauchy wavelets

\[
\forall j \in \mathbb{Z}, \forall x \in \mathbb{R} \quad f \ast \psi_j(x) = F(x + i2^j)
\]

where \( F \) is the holomorphic extension of \( f^{(p)} \) to the complex upper half plane.
Knowledge of $|f \ast \psi_j|$ for all $j$

$\iff$ Knowledge of $|F|$ on a set of horizontal lines of $\mathbb{C}$

$\mathbb{R} + i2^2$

$\mathbb{R} + i2^1$

$\mathbb{R} + i2^0$

$\cdots$

$\mathbb{R}$

**Reformulation of the problem**: can we determine the holomorphic $F$ from its modulus on these horizontal lines?
Lemma

From $|F(x + i2^j)|$, $\forall x \in \mathbb{R}$, we can compute:

$$\forall x \in \mathbb{R}, \quad F \left(x + i \frac{2.2^{j+1}}{3} \right) \overline{F \left(x + i \frac{2.2^j}{3} \right)}$$

Proof

$x \to |F(x + i2^j)|^2$ and $x \to F \left(x + i \frac{2.2^{j+1}}{3} \right) \overline{F \left(x + i \frac{2.2^j}{3} \right)}$ are the restriction of the same holomorphic function to two different horizontal lines.
Reconstruction algorithm

- From $|F(x + i2^j)|$, compute $F(x + i\frac{2.2^{j+1}}{3}) F(x + i\frac{2.2^j}{3})$
- For all $j \in \mathbb{Z}$, $k \in \mathbb{N}$, compute:
  
  \[
  \frac{F(x + i\frac{2.2^j}{3})}{F(x + i\frac{2.2^{j+2k}}{3})}
  \]

- Let $k$ go to $\infty$ : $F(x + i\frac{2.2^{j+2k}}{3})$ goes to a constant.
- From $\left\{ F(x + i\frac{2.2^j}{3}) \right\}_{x \in \mathbb{R}, j \in \mathbb{Z}}$, reconstruct $F$. 

Strong stability

\[ \left( \| W \| f - W \| g \|_2 \leq \epsilon \right) \Rightarrow \left( \| f - g \|_2 \leq C \epsilon \right) \]

This does not hold: the reconstruction is not uniformly continuous.
Counter-example

\( \hat{f} \)

\( \hat{g} \)

\((f \ast \psi_j)\)

\((g \ast \psi_j)\)
Theoretical aspects

Counter-example

\[ \hat{f}, \hat{g} \]

\[ |f \ast \psi_j|, |g \ast \psi_j| \]
More general construction

- any $f \in L^2(\mathbb{R})$
- phases $(\phi_j(t))_{j \in \mathbb{Z}}$, varying slowly in both $j$ and $t$

There exists $g \in L^2(\mathbb{R})$ such that:

$$\forall j \in \mathbb{Z} \quad f \ast \psi_j \approx e^{i\phi_j}(g \ast \psi_j)$$

$$\Rightarrow \forall j \in \mathbb{Z} \quad |f \ast \psi_j| \approx |g \ast \psi_j|$$

but we can have $f \not\approx g$. 
For Cauchy wavelets, if $\forall j, |f \ast \psi_j| \approx |g \ast \psi_j|$, there exist slow-varying phases $(j, t) \rightarrow \phi_j(t)$ such that:

$$\forall j \in \mathbb{Z}, \quad f \ast \psi_j(t) \approx e^{i \phi_j} (g \ast \psi_j(t))$$

except in the neighborhood of points $(j, t)$ where $f \ast \psi_j(t) \approx 0$.

Local stability

$$\{f \ast \psi_j\}_{j \in \mathbb{Z}} \approx \{g \ast \psi_j\}_{j \in \mathbb{Z}}$$

up to a global phase in the neighborhood of each $(j, t)$. 
Idea of proof
Study the stability of the reconstruction algorithm introduced for proving uniqueness.
Algorithmical aspects

Reconstruct $f \in L^2(\mathbb{R})$ from $\{|f \ast \psi_j|\}_{j \in \mathbb{Z}}$?

Wavelets are now generic, not Cauchy.

**Algorithms for generic phase retrieval problems**

- **Iterative methods**
  [Gerchberg and Saxton, 1972; Fienup, 1982; Candès et al., 2015]
  → relatively fast, but suffer from local minima

- **Convexification methods**
  [Candès et al., 2013; Waldspurger et al., 2015]
  → more precise, but too slow for large problems
For the wavelet transform, none yields satisfying results.

→ Specialized algorithm, that uses the structure of the wavelet transform to improve over existing algorithms.
First element: multiscale method

Reconstruct $f$ from $\{|f \ast \psi_j|\}_{j \in \mathbb{Z}}$.

- We reconstruct $f$ on the band $[0; 2^{-J}]$, for large $J$.
- Once $f$ is reconstructed on the band $[0; 2^{-j}]$, we reconstruct the band $[2^{-j}; 2^{1-j}]$.

It works better than reconstructing all bands at once.
Second element : reformulation

Knowing $|f \star \psi_j|^2 \iff$ knowing $(f \star \psi_j^{(1)})(f \star \psi_j^{(2)})$,
where $(\psi_j^{(1)})_{j \in \mathbb{Z}}$ and $(\psi_j^{(2)})_{j \in \mathbb{Z}}$ are new families of wavelets.

$\psi_j^{(1)}$ has a lower characteristic frequency than $\psi_j$.
$\psi_j^{(2)}$ has a higher characteristic frequency than $\psi_j$. 
Propagation of the phase information

- Once the $[0; 2^{-(j+1)}]$ frequency band has been reconstructed, we have an estimate of $f \ast \psi_j^{(1)}$.

- From $(f \ast \psi_j^{(1)})(f \ast \psi_j^{(2)})$, we get an estimate of $f \ast \psi_j^{(2)}$.

- We get an estimate of $f$ on the frequency band $[0; 2^{-j}]$.

Additional advantage of the reformulation

It seems to reduce the number of local minima, when we directly apply a local optimization algorithm to it.
Numerical results

- Precise: the scalogram of the reconstructed signal is in general almost equal to the correct one.
- Complexity linear in the size of the signal, up to logarithmic factors.

Failures are more frequent when wavelet transforms are sparse.

Audio example
Morlet wavelets, 1% of noise
Numerical results

- Precise: the scalogram of the reconstructed signal is in general almost equal to the correct one.
- Complexity linear in the size of the signal, up to logarithmic factors.

Failures are more frequent when wavelet transforms are sparse.

Audio example
Morlet wavelets, 1% of noise
Original signal
Numerical results

- Precise: the scalogram of the reconstructed signal is in general almost equal to the correct one.
- Complexity linear in the size of the signal, up to logarithmic factors.

Failures are more frequent when wavelet transforms are sparse.

Audio example
Morlet wavelets, 1% of noise
Original signal
Difference between true and reconstructed modulus:

- Gerchberg-Saxton: 4.7%
- our algorithm: 0.6%
Numerical results

- Precise: the scalogram of the reconstructed signal is in general almost equal to the correct one.
- Complexity linear in the size of the signal, up to logarithmic factors.

Failures are more frequent when wavelet transforms are sparse.

Audio example
Morlet wavelets, 1% of noise
Original signal
Difference between true and reconstructed modulus:

- Gerchberg-Saxton: 4.7%
- our algorithm: 0.6%
Numerical results

- Precise: the scalogram of the reconstructed signal is in general almost equal to the correct one.
- Complexity linear in the size of the signal, up to logarithmic factors.

Failures are more frequent when wavelet transforms are sparse.

**Audio example**

Morlet wavelets, 1% of noise

Original signal

Difference between true and reconstructed modulus:

- Gerchberg-Saxton: 4.7%
- our algorithm: 0.6%
The reconstruction is not **strongly stable**:

\[
\frac{\| |W| f - |W| f_{rec} \|_2}{\| |W| f \|_2} = 0.6\% \quad \text{but} \quad \frac{\| f - f_{rec} \|_2}{\| f \|_2} = 86\%
\]

but we do not hear the difference because all the perceptual content of the audio signal is encoded in the modulus.

We can empirically confirm our **local stability** result.
The phase difference varies slower than the modulus, and much slower in the zones where the wavelet transform has no very small values.
Question 1

Can this study be extended to a more sophisticated representation, the scattering transform?

\[ f \]

\[ |f \ast \psi_{J-2}| \quad |f \ast \psi_{J-1}| \quad |f \ast \psi_J| \]

\[ \vdots \]

\[ ||f \ast \psi_J \ast \psi_{J-1}|| \quad ||f \ast \psi_J \ast \psi_J|| \]

\[ \vdots \]

[Mallat, 2012]
Question 2

For “random” phase retrieval problems, there exist fast, simple, and provably correct algorithms.

For non-random problems, algorithms must usually be specifically designed, and have no convergence guarantees.

Can we understand the gap between the two categories?


