

WAVELET TRANSFORM AND BINARY COALESCENCE DETECTION

JEAN-MICHEL INNOCENT

*CPT, CNRS-Luminy, Case 907
13288 Marseille Cedex 09, France
E-mail: jeanminn@indy2.imt-mrs.fr*

BRUNO TORRÉSANI

*CPT, CNRS-Luminy, Case 907
13288 Marseille Cedex 09, France
E-mail: torresan@marcptnx1.univ-mrs.fr*

Dedicated to the memory of B. Escudié

Abstract. We give a short account of some time-frequency methods which are relevant in the context of gravity waves detection. We focus on the case of wavelet analysis which we believe is particularly appropriate. We show how wavelet transforms can lead to efficient algorithms for detection and parameter estimation of binary coalescence signals. In addition, we give in an appendix some of the ingredients needed for the construction of discrete wavelet decompositions and corresponding fast algorithms.

1. Introduction and notations

1.1. Generalities. It has been recognized for a long time that a wide class of signals are efficiently described by means of *Time-Frequency representations*, i.e. representations in which time (or position) and frequency variables appear simultaneously. The prototype of such transforms is the so-called Gabor transform:

$$f(x) \in L^2(\mathbb{R}) \leftrightarrow G_f(b, \omega) \in L^2(\mathbb{R}^2) ,$$

where the function of the two variables b (time) and ω (frequency) is defined as

$$G_f(b, \omega) = \int_{\mathbb{R}} f(x) e^{-i\omega(x-b)} \overline{g(x-b)} dx . \quad (1)$$

1991 *Mathematics Subject Classification*: 83C35, 94A12, 94A13.

The paper is in final form and no version of it will be published elsewhere.

Here, $g(x)$ is a window, generally chosen in such a way that $g(x)$ (resp. $\hat{g}(\xi)$) is well localized near the origin of times $x = 0$ (resp. the origin of frequencies $\xi = 0$). Under these assumptions, one may think of the coefficient $G_f(b, \omega)$ as describing the “content of the signal $f(x)$ near time $x = b$ and frequency $\xi = \omega$ ”. Of course, the localization with respect to time and frequency variables simultaneously has to be understood in a “fuzzy sense”, because of Heisenberg’s uncertainty principle. By improving precision in time, we lose precision in frequency.

It is a standard result that the set of coefficients $G_f(b, \omega)$ characterize the signal $f(x)$, in the sense that $f(x)$ may be “reconstructed” from its Gabor transform as

$$f(x) = \frac{1}{2\pi\|g\|^2} \int_{\mathbb{R}^2} G_f(b, \omega) e^{i\omega(x-b)} g(x-b) db d\omega . \quad (2)$$

The inversion formula (2) has to be understood in the weak $L^2(\mathbb{R})$ sense, i.e. in the sense of “energy conservation”:

$$\frac{1}{2\pi\|g\|^2} \int_{\mathbb{R}^2} |G_f(b, \omega)|^2 db d\omega = \int_{\mathbb{R}} |f(x)|^2 dx . \quad (3)$$

As we said, the Gabor representation is one among many other time-frequency representations. Several examples may be found in monographs such as [14, 22, 40, 49] or papers [4, 7, 25]. Throughout this paper we shall concentrate on the wavelet transform which seems to be especially well adapted to binary coalescence signals and more particularly on continuous wavelet transform. The paper is organized as follows. The rest of the current section is devoted to some generalities and notations. In Section 2 we recall the basic definitions and properties of continuous wavelet transform. We describe in Section 3 some elementary facts on the wavelet analysis of stationary stochastic processes. Section 4 is devoted to a description of wavelet-based methods for detecting amplitude and frequency modulated signals in noisy environment, and we address the problem of detection of binary coalescence signals in Section 5. Section 6 is devoted to conclusions. Finally, we give in the Appendix some aspects of discrete wavelet transforms and their numerical implementation.

1.2. Fourier analysis. Let us start with some notions of Fourier analysis. We shall work in the framework of the space of complex valued square-integrable functions, denoted by $L^2(\mathbb{R})$, equipped with a natural inner product which turns it into a Hilbert space. We shall use the following convention for the inner product. For any two functions $f(x)$ and $g(x)$ in $L^2(\mathbb{R})$, we denote:

$$\langle f, g \rangle = \int f(x) \overline{g(x)} dx. \quad (4)$$

Our convention for the Fourier transform is the following: for $f(x) \in L^1(\mathbb{R})$, its Fourier transform $\hat{f}(\xi)$ is defined as:

$$\hat{f}(\xi) = \int f(x) e^{-i\xi x} dx . \quad (5)$$

In fact one shows that:

$$\int |f(x)|^2 dx = \frac{1}{2\pi} \int |\hat{f}(\xi)|^2 d\xi \quad (6)$$

whenever $f(x)$ is a smooth function decaying rapidly at infinity. Relation (6) (the so-called *Plancherel formula*) expresses the fact that the Fourier transform can be extended to the whole space $L^2(\mathbb{R})$ as an isometry and more precisely as a Hilbert space unitary equivalence between $L^2(\mathbb{R}, dx)$ and $L^2(\mathbb{R}, d\xi/2\pi)$. The inverse transform is given by:

$$\check{f}(x) = \frac{1}{2\pi} \int f(\xi) e^{i\xi x} d\xi. \tag{7}$$

1.3. *Hilbert transform, analytic signal.* In addition to $L^2(\mathbb{R})$, we shall often make use of the *complex Hardy space* sometimes called the *space of analytic signals*:

$$H^2(\mathbb{R}) = \left\{ f(x) \in L^2(\mathbb{R}); \hat{f}(\xi) = 0 \ \forall \xi \leq 0 \right\}. \tag{8}$$

$H^2(\mathbb{R})$ is intimately related to the *Hilbert transform* H , defined by:

$$H \cdot f(x) = \frac{1}{\pi} P.V. \int f(x-y) \frac{dy}{y}, \tag{9}$$

(where *P.V.* denotes principal value) and conveniently expressed in the Fourier domain as

$$\widehat{H \cdot f}(\xi) = -i \operatorname{sgn}(\xi) \hat{f}(\xi). \tag{10}$$

Notice that it transforms sine waves into cosine waves, and vice versa.

Given a real valued function $f(x)$, the associated *analytic signal* is defined as (up to a factor 2) its orthogonal projection $Z_f(x)$ onto $H^2(\mathbb{R})$. It is given by the formula:

$$Z_f(x) = [Id + iH]f(x), \tag{11}$$

where Id denotes the identity operator. Equivalently, its Fourier transform is given by:

$$\widehat{Z_f}(\xi) = 2\theta(\xi) \hat{f}(\xi), \tag{12}$$

where $\theta(\xi)$ denotes the Heaviside step function which is equal to 1 when $\xi \geq 0$ and to 0 otherwise. The analytic signal representation has been proven to be useful in many applications. In particular, the notion of time-dependent frequency, or *instantaneous frequency*, makes sense as the derivative of the instantaneous phase of the analytic signal:

$$\nu(x) = \frac{1}{2\pi} \frac{d \arg Z_f(x)}{dx} \tag{13}$$

1.4. *Stationary processes and their spectral representation.* We give here the basic properties of stationary processes, without going into sophisticated mathematical details. Our goal is rather to provide the reader with the main expressions which are needed in order to follow the discussion below, at least with formal calculations. The interested reader may want to consult [33] for more details.

Let us consider a real-valued stationary stochastic process $n(x) = n_\omega(x)$ of mean zero (for convenience, we suppress the explicit dependence on the random parameter ω throughout this paper). Then the autocovariance function $C(\tau) = \mathbb{E}\{n(x+\tau)\overline{n(x)}\}$ is non-negative definite and by Bochner's theorem, there exists a non-decreasing function $F(\xi)$ such that

$$C(\tau) = \mathbb{E}\{n(x+\tau)\overline{n(x)}\} = \frac{1}{2\pi} \int e^{i\xi\tau} F(d\xi). \tag{14}$$

For the sake of simplicity, we shall stick to the case where the measure $F(d\xi)$ is absolutely continuous with respect to the Lebesgue measure $d\xi$, so that we may write $F(d\xi) = \mathcal{E}(\xi)d\xi$, and we write

$$C(\tau) = \frac{1}{2\pi} \int e^{i\xi\tau} \mathcal{E}(\xi) d\xi . \quad (15)$$

Here, $\mathcal{E}(\xi)$ is the *spectral density* of the process.

The Cramér representation states that $n(x)$ may be obtained through linear filtering of white noise. More precisely, the Cramér representation of the stochastic process $n(x)$ is given by

$$n(x) = \frac{1}{2\pi} \int \sqrt{\mathcal{E}(\xi)} e^{i\xi x} dW_\xi , \quad (16)$$

where dW_ξ is a (real) white noise measure, such that

$$\mathbb{E}\{dW_\xi\} = 0 \quad \forall \xi , \quad (17)$$

and

$$\mathbb{E}\{dW_\xi dW_\zeta\} = 2\pi\delta(\xi - \zeta)d\xi . \quad (18)$$

We will not go into further details on this point.

1.5. Spectral estimation. It is a standard problem in signal analysis to estimate the spectral density from a unique (discrete) realization of finite length. Spectral estimation is a technical subject, often involving subtle choices.

To start with let us assume that we are given a discrete stationary time series, consisting of a finite number of samples $X_i = f(x_i)$, $i = 0, \dots, N-1$ of a continuous time function $f(x)$. Then the spectral density $\mathcal{E}(k)$ is usually estimated from the *sample periodogram*

$$\hat{\mathcal{E}}(\ell) = \left| \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{-ik\xi_\ell} \right|^2 \quad \ell = 0, 1, \dots, N-1 \quad (19)$$

where $\xi_\ell = 2\pi\ell/N$. However, it may be shown that such an estimator is 1) biased (but asymptotically unbiased) and 2) inconsistent. The situation is usually improved by considering *tapered* sample periodograms of the form

$$\hat{\mathcal{E}}(\ell) = \left| \frac{1}{N} \sum_{k=0}^{N-1} w_k X_k e^{-ik\xi_\ell} \right|^2 \quad \ell = 0, 1, \dots, N-1 \quad (20)$$

with a well chosen weighting sequence w_k . Several choices are possible. Ours consists in the sequence

$$w_k = \begin{cases} |\sin(k\pi/L)|^n & \text{if } 0 \leq k \leq L \\ 1 & \text{if } L \leq k \leq N-L \\ |\sin((N-k)\pi/L)|^n & \text{if } N-L \leq k \leq N-1 \end{cases} \quad (21)$$

for suitably chosen integers $L < N/2$ and n .

As an example we display in Figure 1 the spectral density of the simulated VIRGO detector noise, estimated from a sample size of 2^{15} .

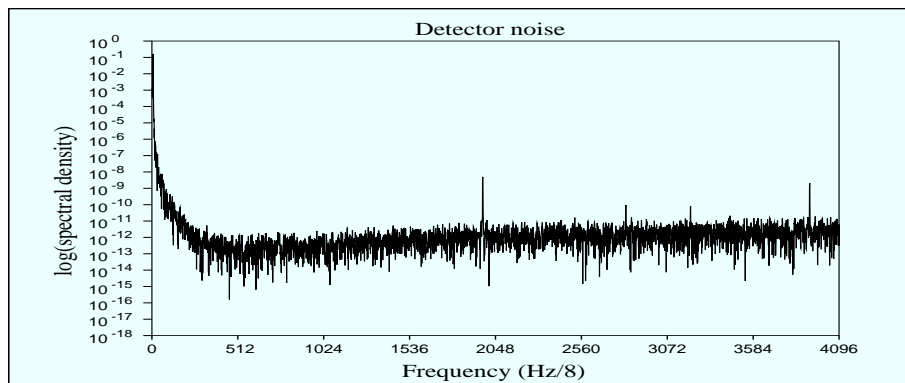


Figure 1: Power spectrum of the VIRGO detector noise.

2. Time-Frequency transforms and wavelet analysis. The starting point of time-frequency analysis was the fact that a wide class of signals may be represented more adequately by using simultaneously time and the frequency variables. Let us quote for example musical signals, which are interpreted by human ear in terms of time (i.e. time of emission, and duration) and frequency (the height of the sound), or chirp signals such as gravitational waves generated by coalescing binaries. Throughout this paper, we shall in particular be concerned with model signals of the form

$$f(x) = \sum_k A_k(x) \cos \phi_k(x) , \quad (22)$$

where the functions $A_k(x)$ (termed local amplitudes) are assumed to be slowly varying compared with the oscillations corresponding to $\phi_k(x)$ (local phases).

Of course, the notion of “time-frequency content” of a signal cannot make sense in an infinitely precise way. For example, Heisenberg’s inequality prevents us from localizing perfectly functions simultaneously in time and frequency. As a result, some arbitrariness is necessarily introduced into time-frequency representations. Let us simply stress some of the main consequences.

- First, time-frequency representations are not unique: there are many different ways of describing the “time-frequency content” of a signal.
- Second, for a given time-frequency representation, it is impossible to achieve perfect time-frequency localization, because of the Heisenberg uncertainty principle. This means that we shall always have to look for a compromise between time localization and frequency localization.

There exists now a large class of time-frequency representations enjoying different properties. It is not the purpose of this paper to review them (for this we refer to [22] for example), and we shall stick to the particular case of the continuous wavelet transform.

2.1. The continuous wavelet transform. Let us start by introducing the continuous wavelet transform (CWT for short). Let $\psi(x) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ be a fixed function, called the *analyzing wavelet*, or *mother wavelet*. The corresponding family of wavelets is the

family of shifted and scaled copies of $\psi(x)$ defined by:

$$\psi_{(b,a)}(x) = \frac{1}{a} \psi \left(\frac{x-b}{a} \right). \quad (23)$$

Given an analyzing wavelet $\psi(x)$, the associated continuous wavelet transform is defined as follows

DEFINITION 2.1. Let $\psi(x) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ be an analyzing wavelet. The *continuous wavelet transform* of $f(x) \in L^2(\mathbb{R})$ is defined by the integral transform

$$T_f(b, a) = \langle f, \psi_{(b,a)} \rangle = \frac{1}{a} \int f(x) \overline{\psi \left(\frac{x-b}{a} \right)} dx \quad (24)$$

The (real or complex) number $T_f(b, a)$ carries information concerning the signal $f(x)$ at scale a around the point b .

REMARK 2.1. If the wavelet $\psi(x)$ is *progressive*, i.e. if $\psi(x) \in H^2(\mathbb{R})$, then the CWT of a signal $f(x)$ reads

$$T_f(b, a) = \langle f, \psi_{(b,a)} \rangle = \frac{1}{2} \langle Z_f, \psi_{(b,a)} \rangle.$$

A crucial property of the continuous wavelet transform is that, under a mild condition on the analyzing wavelet (see equation (25) below), the transform is invertible on its range (see e.g. [25] for a proof):

THEOREM 2.1. Let $\psi(x) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, and let

$$c_\psi = \int_0^\infty |\hat{\psi}(a\xi)|^2 \frac{da}{a}. \quad (25)$$

If c_ψ is finite, nonzero and independent of $\xi \in \mathbb{R}$ (resp. finite and nonzero), then every $f(x) \in L^2(\mathbb{R})$ (resp. $f(x) \in H^2(\mathbb{R})$) may be decomposed as

$$f(x) = \frac{1}{c_\psi} \int_{-\infty}^\infty \int_0^\infty T_f(b, a) \psi_{(b,a)}(x) \frac{da}{a} db, \quad (26)$$

with strong convergence in $L^2(\mathbb{R})$.

Assuming that c_ψ is independent of ξ actually amounts to assuming independence with respect to $\text{sgn}(\xi)$. Obviously, such an assumption is not needed in the $H^2(\mathbb{R})$ case. The finiteness of c_ψ implies the vanishing of the integral of the wavelet $\psi(x)$:

$$\int \psi(x) dx = \hat{\psi}(0) = 0.$$

Condition (25) is often referred to as the *admissibility condition*. An admissible wavelet is then essentially a band pass filter (we shall come back to this comment later on). Such a condition may be enforced by assuming vanishing moments for the wavelet: for example

$$\int x^m \psi(x) dx = 0, \quad \forall m = 0, 1, \dots, M-1.$$

In the Fourier domain, the vanishing moments essentially control the behaviour of the Fourier transform of the wavelet at the origin. Such a property turns out to be essential for the analysis of singularities and transients in signals.

As a consequence, we have the following partial isometry between $L^2(\mathbb{R})$ and the target space of the transform, namely $\mathcal{H} = L^2(\mathbb{R} \times \mathbb{R}_+^*, a^{-1}dad b)$:

$$\|f\|^2 = \frac{1}{c_\psi} \int_{-\infty}^{\infty} \int_0^{\infty} |T_f(b, a)|^2 \frac{da}{a} db \tag{27}$$

for all $f(x) \in L^2(\mathbb{R})$. This allows for the interpretation of the squared-modulus of the wavelet transform (suitably normalized) as a time-frequency or more precisely a time-scale energy density.

2.2. Redundancy and reproducing kernels. For a given admissible wavelet $\psi(x)$ fulfilling the admissibility condition, the image of $L^2(\mathbb{R})$ by the wavelet transform is a closed subspace \mathcal{H}_ψ of $L^2(\mathbb{R} \times \mathbb{R}_+^*, a^{-1}dad b)$. This space is called the *reproducing kernel Hilbert space*. It is the space of solutions $F(b, a)$ of the integral equation

$$F(b', a') = P_\psi F(b', a') = \int_{-\infty}^{\infty} \int_0^{\infty} \mathcal{K}_\psi(b', a'; b, a) F(b, a) \frac{da}{a} db, \tag{28}$$

where the reproducing kernel \mathcal{K}_ψ is given by:

$$\mathcal{K}_\psi(b', a'; b, a) = \frac{1}{c_\psi} \langle \psi_{(b,a)}, \psi_{(b',a')} \rangle. \tag{29}$$

This fact is readily proved by taking the inner product of both sides of equation (26) with the wavelet $\psi_{(b',a')}(x)$. The corresponding integral operator P_ψ is easily shown to be an orthogonal projection on the \mathcal{H}_ψ space (i.e. $P_\psi^* = P_\psi^2 = P_\psi$).

Remark 2.2. Equation (28) expresses the redundancy of the CWT. As before, a consequence of this redundancy is the existence of many different inversion formulas for the CWT, or otherwise stated the possibility of using in the inversion formula (26) a reconstruction wavelet different from the analyzing wavelet $\psi(x)$: if the function $\chi(x) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is such that the number

$$c_{\psi\chi} = \int_0^{\infty} \overline{\hat{\psi}(a\xi)} \hat{\chi}(a\xi) \frac{da}{a} \tag{30}$$

is finite, nonzero and independent of ξ , then equation (26) may be replaced with:

$$f(x) = \frac{1}{c_{\psi\chi}} \int_{-\infty}^{\infty} \int_0^{\infty} T_f(b, a) \chi_{(b,a)}(x) \frac{da}{a} db, \tag{31}$$

where the wavelet coefficients $T_f(b, a)$ are still defined by (24).

Remark 2.3. There exists a simpler version of continuous wavelet analysis, known under the name of continuous *Littlewood-Paley decompositions*. Given a wavelet $\psi(x) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, and assuming that the number

$$k_\psi = \int_0^{\infty} \overline{\hat{\psi}(a\xi)} \frac{da}{a} \tag{32}$$

is well-defined, finite, nonzero and independent of ξ , we have the following simple inversion formula (known as *Morlet's inversion formula*) for the corresponding continuous wavelet transform

$$f(x) = \frac{1}{k_\psi} \int_0^{\infty} T_f(x, a) \frac{da}{a}, \tag{33}$$

the proof of which is elementary.

2.3. *Translation and scaling covariance.* The wavelet transform enjoys built-in covariance properties. For example, the CWT of a shifted copy of the signal $f(x)$ equals the corresponding time-shifted copy of the CWT of $f(x)$. A similar property holds with dilations. More generally we have the following

LEMMA 2.1. *Let $f(x) \in L^2(\mathbb{R})$, and set*

$$\tilde{f}(x) = f\left(\frac{x - x_0}{\lambda}\right).$$

Then

$$T_{\tilde{f}}(b, a) = T_f\left(\frac{b - x_0}{\lambda}, \frac{a}{\lambda}\right). \quad (34)$$

REMARK 2.4. Lemma 2.1 may be given an instructive geometric interpretation, which we sketch here. The space of scale and translation variables may be endowed with a (Lie) group structure, with product given by $(b, a) \cdot (b', a') = (b + ab', aa')$ and inverse $(b, a)^{-1} = (-b/a, 1/a)$. This group is termed the *affine group* and denoted by G_{aff} . The natural action of G_{aff} on $L^2(\mathbb{R})$ given by

$$\pi(b, a)f(x) = \frac{1}{\sqrt{a}}f\left(\frac{x - b}{a}\right) \quad (35)$$

is actually an *unitary representation* of G_{aff} , in the sense that $\pi(b, a)$ is a unitary operator for all $(b, a) \in G_{aff}$, and that $\pi(b, a)\pi(b', a') = \pi((b, a) \cdot (b', a'))$, and the connection to the wavelet transform is as follows: if $f(x) \in L^2(\mathbb{R})$

$$T_f(b, a) = \frac{1}{\sqrt{a}}\langle f, \pi(b, a)\psi \rangle, \quad (b, a) \in G_{aff}. \quad (36)$$

Back to Lemma 2.1, we have that $\tilde{f}(x) = \sqrt{\lambda}\pi(x_0, \lambda)f(x)$, and $T_{\tilde{f}}(b, a) = \sqrt{\frac{a}{\lambda}}\langle f, \pi((x_0, \lambda)^{-1} \cdot (b, a))\psi \rangle = \sqrt{\frac{a}{\lambda}}\langle f, \pi\left(\frac{b - x_0}{\lambda}, \frac{a}{\lambda}\right)\psi \rangle$, which yields the lemma. The invariance properties of the wavelet transform then have a deeper geometric interpretation in terms of the action of the affine group. We shall come back to that point in Section 5.2.

Such properties have found a lot of applications, for example for the study of fractal and multifractal functions and measures. We shall see below their implications for the particular case of binary coalescence detection.

2.4. *The case of (complex) progressive wavelets.* If $\psi(x)$ is a progressive wavelet (see Remark 2.1 for a definition), Eq. (26) holds for functions $f(x) \in H^2(\mathbb{R})$. Progressive wavelets are also well suited for $L^2(\mathbb{R})$ real signals. Indeed, if $f(x)$ is a real valued function, then its Fourier transform possesses Hermitian symmetry (i.e. $\hat{f}(-\xi) = \overline{\hat{f}(\xi)}$) and is completely characterized by its projection on $H^2(\mathbb{R})$. Then the $H^2(\mathbb{R})$ version of wavelet analysis may be used as well, and is particularly convenient as we shall see.

The wavelet transform of real signals with respect to progressive wavelets is a complex-valued function, and is also progressive with respect to the variable b . As such, it may be uniquely written in the form

$$T_f(b, a) = |T_f(b, a)|e^{i\Omega(b, a)}, \quad (37)$$

where $\Omega(b, a) = \arg T_f(b, a)$. Let us denote by $\omega(b, a)$ the local frequency of $T_f(b, a)$, i.e.

$$\omega(b, a) = \partial_b \Omega(b, a) . \tag{38}$$

Then it is easy to see that

$$\omega(b, a) = \frac{1}{a} \frac{\tilde{R}_f(b, a)I_f(b, a) - \tilde{I}_f(b, a)R_f(b, a)}{|T_f(b, a)|^2} , \tag{39}$$

where we have set

$$\tilde{T}_f(b, a) = \frac{1}{a} \int f(x) \psi' \left(\frac{x-b}{a} \right) dx , \tag{40}$$

and R_f and I_f (resp. \tilde{R}_f and \tilde{I}_f) are the real and imaginary parts of T_f (resp. \tilde{T}_f). Note that $\tilde{T}_f(b, a)$ is a wavelet transform of $f(x)$ as well, the wavelet being the derivative of $\psi(x)$.

3. Wavelet transform of stationary processes We now turn to the description of the CWT of stochastic processes. Let us consider first a stochastic process, and denote by C its covariance operator. Then if $\psi(x) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ we have

$$\mathbb{E}\{T_n(b, a)\overline{T_n(b', a')}\} = \langle C\psi_{(b', a')}, \psi_{(b, a)} \rangle \quad \forall (b, a), (b', a') . \tag{41}$$

In the case of stationary time series the covariance operator is a convolution operator, with the spectral density $\mathcal{E}(\xi)$ as multiplier. Let $n(x)$ be such a time series, and consider its Cramér representation given in Eq. (16). Then, its CWT takes the form of a stochastic integral

$$T_n(b, a) = \frac{1}{2\pi} \int e^{i\xi b} \sqrt{\mathcal{E}(\xi)} \hat{\psi}(a\xi) dW_\xi , \tag{42}$$

and we have the following

PROPOSITION 3.1. *Let $n(x)$ be a stationary Gaussian stochastic process, with spectral density denoted by $\mathcal{E}(\xi)$, and let $T_n(b, a)$ denote its CWT, with respect to the progressive wavelet $\psi(x) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Then*

1. $T_n(b, a)$ is a Gaussian process.
2. For fixed scale a , $T_n(b, a)$ is a zero mean stationary process, with power density

$$\mathcal{E}_a(\xi) = \mathcal{E}(\xi) |\hat{\psi}(a\xi)|^2 \tag{43}$$

3. In particular, one has

$$\mathbb{E}\{|T_n(b, a)|^2\} = \frac{1}{2\pi} \int \mathcal{E}(\xi) |\hat{\psi}(a\xi)|^2 d\xi . \tag{44}$$

4. Assume now that $\psi(x) \in H^2(\mathbb{R})$. Then for fixed a and b , the real and imaginary parts of $T_n(b, a)$ are independent Gaussian random variables.

Let us consider as an example the case of a white noise process $n(x)$. In this case, Eq. (41) becomes

$$\mathbb{E}\{T_n(b, a)\overline{T_n(b', a')}\} = \langle \psi_{(b, a)}, \psi_{(b', a')} \rangle = c_\psi \mathcal{K}_\psi(b, a; b', a') \quad \forall (b, a), (b', a') .$$

and we have in particular

$$\mathbb{E}\{|T_n(b, a)|^2\} = \|\psi\|^2/a .$$

Let us now consider the “signal + noise” case, i.e.

$$f(x) = f_0(x) + n(x) ,$$

where $f_0(x)$ is a deterministic signal, and $n(x)$ is a weakly stationary process. Then clearly

$$T_f(b, a) = T_{f_0}(b, a) + T_n(b, a) .$$

In addition, we may write

$$\mathcal{M}(b, a) = |T_f(b, a)|^2 = |T_{f_0}(b, a)|^2 + N(b, a) , \quad (45)$$

where

$$N(b, a) = 2\Re \left(T_{f_0}(b, a) \overline{T_n(b, a)} \right) + |T_n(b, a)|^2 . \quad (46)$$

Thus by Proposition 3.1, we have

$$\mathbb{E}\{N(b, a)\} = \frac{1}{2\pi} \int \mathcal{E}(\xi) |\hat{\psi}(a\xi)|^2 d\xi . \quad (47)$$

4. Ridge detection methods for time-varying frequencies. We now address the problem of characterizing time and amplitude modulated signals from the behavior of a given time-frequency representation. Such a problem has been addressed by several authors in various contexts. We just give here a few methods that seem to us well suited to the gravitational waves detection problem, and we focus on the wavelet transform case.

4.1. Generalities. Let us consider as a toy model a signal of the form

$$f(x) = A(x) \cos \phi(x) \quad (48)$$

and assume that the amplitude $A(x)$ is slowly varying compared to the oscillations. Let $\psi(x) \in H^2(\mathbb{R})$ be a progressive wavelet, and assume that $|\hat{\psi}(\xi)|$ has a (unique) maximum at $\xi = \omega_0$. Then we have

$$T_f(b, a) \approx \frac{1}{2} A(b) e^{i\phi(b)} \overline{\hat{\psi}(a\phi'(b))} , \quad (49)$$

which indicates that the wavelet transform is essentially localized near a curve, called the *ridge* of the wavelet transform, of equation

$$a = a_r(b) = \frac{\omega_0}{\phi'(b)} . \quad (50)$$

As an illustration of this fact, we display in Figure 2 the modulus square of the wavelet transform of a (Newtonian) gravitational wave signal generated (or at least expected to be generated) by a coalescing binary system (the model for such signals is given in Eq. (70) below). More precisely, we computed a (complex, progressive) wavelet transform with Morlet’s wavelet

$$\psi(x) = e^{-x^2/2} e^{i\omega_0 x}$$

with $\omega_0 = 2\pi$, and scales of the form $2a_0^n$, with $a_0 = 2^{1/8}$ and $n = 1, 2, \dots, 40$. With such a choice, it may be verified that the scale variable is equivalent to a period (at least when the sampling frequency is set to 1). The wavelet transform modulus square is represented with gray levels: in our convention, the gray level at point (b, a) is directly proportional to the value $|T_f(b, a)|^2$. We can clearly see the localization properties of the

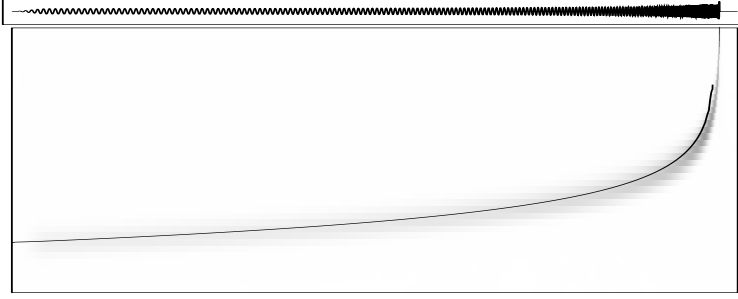


Figure 2: Modulus square of the wavelet transform of a binary coalescence signal.

wavelet transform. The transform is localized in a neighborhood of a ridge (shown in the figure). The algorithm used to estimate the ridge is described in Section 4.2 below.

It is possible to derive a more accurate formula for the wavelet transform using stationary phase approximations. For this, let us suppose that the wavelet $\psi(x)$ is progressive, and may be written in its canonical form as

$$\psi(x) = A_\psi(x)e^{i\phi_\psi(x)} . \quad (51)$$

Then we may write

$$T_f(b, a) \approx \sqrt{\frac{\pi}{2}} \frac{e^{i\frac{\pi}{4}\text{sgn}(\Phi'_{(b,a)}(x_0))}}{\sqrt{a^2|\Phi''_{(b,a)}(x_0)|}} \overline{\psi\left(\frac{x_0 - b}{a}\right)} Z_f(x_0) , \quad (52)$$

where

$$\Phi_{(b,a)}(x) = \phi(x) - \phi_\psi\left(\frac{x - b}{a}\right) , \quad (53)$$

and $x_0 = x_0(b, a)$ is a stationary point of the integrand, i.e. a time such that

$$\Phi'_{(b,a)}(x_0) = 0 . \quad (54)$$

In addition, it is assumed that for any (b, a) under consideration, there exists only one such point, and that $\Phi''_{(b,a)}(x_0) \neq 0$. We refer to [19] for a more detailed analysis.

4.2. Local analysis of the wavelet transform. The basic formulae (49) and (50) have two immediate and important consequences. First, if the ridge equation $a = a_r(b)$ is known, then equation (50) yields the local frequency of the signal:

$$\nu(x) = \frac{1}{2\pi}\phi'(x) = \frac{1}{2\pi} \frac{\omega_0}{a_r(x)} \quad (55)$$

Second, the local amplitude of the signal is obtained by putting $a = a_r(b)$ into equation (49) :

$$A(x) = 2 \frac{|T_f(x, a_r(x))|}{|\hat{\psi}(\omega_0)|} \quad (56)$$

This shows the significance of the ridge of the wavelet transform. More precise estimates of the wavelet coefficients, such as the stationary phase approximation described above

lead to efficient methods using the phase of the wavelet transform. The general framework is given in [19]. Let us show how it works with the Morlet wavelet,

$$\psi(x) = e^{-x^2/2} e^{i\omega_0 x} , \quad (57)$$

with Fourier transform

$$\hat{\psi}(\xi) = \sqrt{2\pi} e^{-(\xi-\omega_0)^2/2} . \quad (58)$$

For ω_0 large enough (say $\omega_0 > 5$) ψ is (at least numerically) admissible and progressive. The wavelets coefficients of the signal $f(x)$

$$T_f(b, a) = \frac{1}{2a} \int A(x) e^{-\frac{1}{2} \left(\frac{x-b}{a}\right)^2} e^{i[\phi(x) - \omega_0 \frac{x-b}{a}]} dx \quad (59)$$

are approximately equal to the leading term in the stationary phase expansion of this oscillatory integral [19]

$$T_f(b, a) \approx |T_0(b, a)| e^{i\Omega_0(b, a)} \quad (60)$$

where

$$|T_0(b, a)| = \sqrt{\frac{\pi}{2}} \frac{e^{-\frac{1}{2} \frac{(x_0-b)^2 \phi''(x_0)^2}{a^2 \phi''(x_0)^2 + a^{-2}}}}{[1 + a^4 \phi''(x_0)^2]^{1/4}} \quad (61)$$

and

$$\Omega_0(b, a) = \phi(x_0) - \omega_0 \frac{x_0 - b}{a} + \frac{1}{2} \frac{(x_0 - b)^2 \phi''(x_0)}{1 + a^4 \phi''(x_0)^2} + \frac{1}{2} \arctan [a^2 \phi''(x_0)] , \quad (62)$$

(see[19]). In these formulae $x_0 = x_0(b, a)$ is the stationary point given by

$$\phi'(x_0) = \frac{\omega_0}{a} \quad (63)$$

We assume that $\phi'' > 0$ hence for each (b, a) there is a unique first order stationary point. Now the equation

$$x_0(b, a) = b \quad (64)$$

appears as another version of the ridge equation . It is easy to see that in the present case of the Morlet wavelet x_0 depends only on a and therefore for fixed a

$$\frac{\partial}{\partial b} \Omega_0(b, a) = \frac{\omega_0}{a} - (x_0 - b) \frac{\phi''(x_0)}{1 + a^4 \phi''(x_0)^2} \quad (65)$$

and on the ridge

$$\frac{\partial}{\partial b} \Omega_0(b, a) = \frac{\omega_0}{a} . \quad (66)$$

This suggests to look for the ridge by solving for a the implicit equation

$$\frac{\partial}{\partial b} \Omega(b, a) = \frac{\omega_0}{a} , \quad (67)$$

where $\Omega(b, a)$ is the phase of the wavelet coefficient $T_f(b, a)$. This can be done in practice by various fixed point methods , for example direct iterations or a Newton method.

For the sake of comparison, we show in Figure 3 the wavelet transform of a section of the signal analyzed in Figure 2. Notice in particular the behavior of the phase of the wavelet transform in the right hand image. At a given point on the ridge, the wavelet transform has a tendency to oscillate at the same frequency as the wavelet itself, as predicted by (67). This is the property which is exploited by the ridge extraction algorithm.

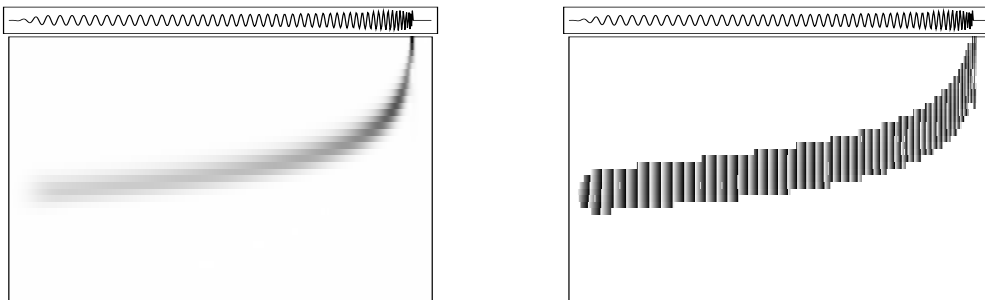


Figure 3: Wavelet transform of a smaller part of the binary coalescence signal; left: modulus square; right: phase.

Remark 4.1. It is worth noticing that very little a priori information about the signal is used. This is a main interest of this method. This way of extracting the ridge gives good results in the case of a signal of the form

$$f(x) = A(x) \cos \phi(x) ,$$

provided that the stationary phase assumptions are fulfilled and the additional noise is not too strong. When the input signal-to-noise ratio is low (in practice say - 5 db) the algorithm is unstable and fails to provide a reliable extraction of the ridge. This is basically due to its local nature.

4.3. *Penalization approaches.* In very noisy situations, the local approaches described above may no longer be suitable, and it may be necessary to turn to global methods. Such methods were introduced in [11, 12] and are developed in great details in [13]. They are based on a different setting of the problem, which makes use explicitly of the *a priori* assumptions made on the signal. Let us give here an example of such methods.

We start with a function of two variables $M(b, \omega)$ (where we consider for convenience the variable $\omega = \log(a)$), supposed to be localized near a ridge φ_0 . $M(b, \omega)$ may be for example a modulus square of wavelet transform, or some modified version of it. We shall be more specific later on. The starting point is the assumption that the ridge is a one-dimensional object, i.e. a curve, and then has to be modeled in that form. Let us for simplicity consider the case of parametric curves. A ridge is then modeled as a mapping

$$\varphi : s \in [0, 1] \rightarrow \varphi(s) \in \mathbb{R}^2 , \quad (68)$$

where the first component of $\varphi(s)$ is a time component $b(s)$ and the second one is for convenience taken to be the logarithm of the scale $\log a(s)$. Given a wavelet transform, the problem is to find the *optimal ridge*, in a sense to be specified. Following [11, 13] we state the problem as a minimization problem, for a conveniently chosen penalty function $\Gamma(\varphi)$. A natural candidate for such a penalty function is the following

$$\Gamma(\varphi) = - \int |T_f(\varphi(s))|^2 ds + \int |\lambda \cdot \varphi'(s)|^2 ds \quad (69)$$

Such a function is the sum of two terms. Let us consider them independently. The first one involves only the “concentration” of $M(b, \omega)$. However it cannot be utilized alone. Indeed, minimizing only the first term would produce a curve trying to occupy densely the domain of (b, ω) . Some rigidity constraints have to be imposed on the curve. This is the purpose of the second term.

The penalty functional $\Gamma(\varphi)$ is not quadratic, and then cannot be minimized explicitly. The minimization has to be done numerically. In noise-free situations, the task is easy and may be achieved using one of the standard minimization techniques (see for example [44, 45]). However, in noisy situations, special care has to be paid to the problem of local minima of $\Gamma(\varphi)$. Indeed, the number of such minima turns out to increase with the noise level, and standard minimization techniques fail in such situations.

A convenient alternative is provided by stochastic relaxation algorithms, for example simulated annealing. A method for solving numerically the minimization problem described above is described in great details in [13] (see also [11]).

5. The gravitational wave detection problem. Let us now turn to the problem of gravity waves detection, and the corresponding parameter estimation problem. We shall see that wavelet techniques are well adapted to such problems.

5.1. The model signal. We shall focus on the case of the binary star coalescence signal, since it is one for which time-frequency analysis may be expected to perform well. In the Newtonian approximation, the model is the following:

$$f(x) = A\theta(x_0 - x)(x_0 - x)^\alpha \cos\left(\Phi - \frac{2\pi}{\beta + 1}F(x_0 - x)^{\beta+1}\right), \quad (70)$$

(we recall that $\theta(x)$ is the Heaviside step function) with

$$\alpha = -\frac{1}{4}, \quad \beta = -\frac{3}{8}. \quad (71)$$

Here, A (resp. F) is generally interpreted as the numerical values of the signal’s amplitude (resp. frequency) one second before coalescence, and Φ as a global phase term.

In more general restricted post-Newtonian approximations, the signal is of the form

$$f(x) = A\nu(x)^{2/3} \cos\phi(x), \quad (72)$$

where $\nu(x) = \phi'(x)/2\pi$ is the local frequency, and the time-dependent phase $\phi(x)$ is the sum of several post-Newtonian terms

$$\phi(x) = \phi_0(x) + \phi_1(x) + \phi_{1.5}(x) + \phi_2(x) + \dots, \quad (73)$$

where $\phi_0(x)$ is the dominant Newtonian part. The correspondence between local frequency and time is given by

$$x = 1 + \tau_0 \left(1 - \left(\frac{\nu(x)}{F}\right)^{-8/3}\right) + \tau_1 \left(1 - \left(\frac{\nu(x)}{F}\right)^{-2}\right) + \tau_{1.5} \left(1 - \left(\frac{\nu(x)}{F}\right)^{-5/3}\right) + \dots, \quad (74)$$

which provides an expression for the *group delay*, i.e. the time of appearance of a given

frequency:

$$x = \tau(\nu) = 1 + \tau_0 \left(1 - \left(\frac{\nu}{F}\right)^{-8/3}\right) + \tau_1 \left(1 - \left(\frac{\nu}{F}\right)^{-2}\right) + \dots$$

Remark 5.1. In both Newtonian or post Newtonian case the signal is a *locally monochromatic* one in the sense that time and frequency are in one-to-one correspondence. This property is crucial for the applicability of the wavelet analysis.

For the sake of simplicity, we shall restrict our investigations to the case of the Newtonian approximation, i.e. we shall set $\tau_1, \tau_{1.5}, \tau_2 \dots$ to zero. However, most of the techniques we are about to discuss may be extended without difficulties to post-Newtonian situations.

To proceed, we need to analyze more carefully the Fourier transform and the analytic signal of the model signals.

Using formal arguments, the Fourier transform of the signal in (70) may be evaluated via the stationary phase approximation, which yields

$$\hat{f}(\xi) \approx \frac{A}{\sqrt{2\pi\beta F}} \left(\frac{\xi}{2\pi F}\right)^{(2\alpha-\beta+1)/(2\beta)} \exp\left\{i\left(\Phi + \frac{\beta\xi}{\beta+1} \left(\frac{\xi}{2\pi F}\right)^{1/\beta} - \xi x_0 + \frac{\pi}{4}\right)\right\} \tag{75}$$

Remark 5.2. Notice that the function given in (70) is neither integrable nor square integrable. Therefore, the meaning of the Fourier transform in (75) is very problematic. However, this problem may be circumvented as follows. Let F be a band-pass filter, such that $\hat{F}(0) = 0$, and consider the filtered signal $F * f(x)$. The effect of such a filtering is to force to zero both very high and very low frequencies. However, since we are dealing with a locally monochromatic signal, forcing to zero very high and very low frequency amounts to enforcing the decay of $F * f(x)$ at $x \rightarrow -\infty$ and forcing $F * f(x)$ to zero as $x \rightarrow x_0$. Then, if one works with a filtered signal instead of the original one, the Fourier transform can be defined, and approximated as in (75). Again, let us stress that such an approximation is acceptable only in a limited frequency range, excluding both very low and very high frequencies, i.e. excluding values of the time variable close to or very far away from the coalescence time.

Remark 5.3. In practice, the experimental signal has to be “prewhitened” before sampling, in order to reduce quantization noise. This means that the signal to be processed by detection algorithms is of the form

$$C^{-1/2} f(x) = \frac{1}{2\pi} \int e^{i\xi x} \frac{\hat{f}(\xi)}{\sqrt{\mathcal{E}(\xi)}} d\xi$$

where $\mathcal{E}(\xi)$ is the spectral density of the detector noise, or at least an approximation of it. Note that the transfer function $1/\sqrt{\mathcal{E}(\xi)}$ of the convolution operator $C^{-1/2}$ has precisely the properties required for the filter F .

At first sight, it is tempting to state the analytic signal of a real signal like $A(x) \cos(\phi(x))$ equals $A(x) \exp\{i\phi(x)\}$. However, this is not true in general (for example, the Hilbert transform of $\sin(1/x)$ is $1 - \cos(1/x)$ and not $-\cos(1/x)$ as one would naively expect). However we have the following

LEMMA 5.1. *Let λ be a (large) positive number, and let $f(x) = A(x) \cos(\lambda\phi(x)) \in L^1(\mathbb{R})$, where $A(x)$ and $\phi(x)$ are twice and four times differentiable functions respectively. Then*

$$Z_f(x) = A(x)e^{i\lambda\phi(x)} + O(\lambda^{-1}) \quad (76)$$

as $\lambda \rightarrow \infty$.

Such a result is not of direct application, since the limit $\lambda \rightarrow \infty$ is not suited for practical situations. However, a “weak interpretation” of it would be the following. Suppose that 1) the variations of the amplitude are much slower than the variations coming from the oscillations, and 2) the variations of the frequency $\phi'(x)$ are small enough. Then the analytic signal of $f(x) = A(x) \cos \phi(x)$ is approximately equal to $A(x) \exp\{i\phi(x)\}$.

Let us come back to the binary coalescence signal. A naive calculation would erroneously suggest that the analytic signal of (70) is of the form

$$Z_f(x) = A\theta(x_0 - x)(x_0 - x)^\alpha \exp \left\{ i \left[\Phi - \frac{2\pi}{\beta + 1} F(x_0 - x)^{\beta+1} \right] \right\}, \quad (77)$$

and that the instantaneous frequency then reads

$$\nu(x) = F(x - x_0)^\beta \quad (78)$$

Such a conclusion is only approximately true, for the same reasons as before, in particular because of a possible correction affecting the low frequencies. However, within a limited range of frequencies, such an approximation is definitely sensible.

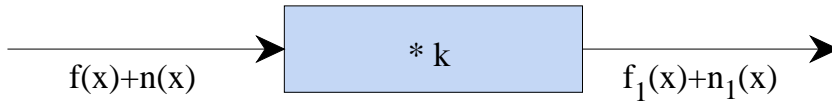


Figure 4: Filtering scheme.

A standard approach for detecting such signals amounts to use the *matched filter* technique, which we outline here (see [5, 18, 34, 35, 36, 47]) for a detailed analysis in the context of binary coalescence detection). The basic situation is the following: a linear filter of impulse response $k(x)$ (convolution by $k(x)$) is applied to an input signal which is the sum of a (deterministic) signal of interest $f(x)$ and some noise modeled as a second order stationary process $n(x)$ with known spectral density $\mathcal{E}(\xi)$. See Figure 4 for the scheme.

$$f_1(x) = (f * k)(x) = \int k(x - u)f(u)du \quad (79)$$

$$n_1(x) = (n * k)(x) = \int k(x - u)n(u)du \quad (80)$$

Then n_1 is also a second order stochastic process with spectral density $\mathcal{E}_1(x)$. The filtering formulas in the Fourier domain read:

$$\hat{f}_1(\xi) = \hat{k}(\xi)\hat{f}(\xi) \quad (81)$$

$$\mathcal{E}_1(\xi) = |\hat{k}(\xi)|^2\mathcal{E}(\xi) \quad (82)$$

It is the purpose of matched filtering to maximize, at a given time x_0 the *output signal to noise ratio*

$$\text{output SNR} = \frac{|f_1(x_0)|}{\sigma_1} \tag{83}$$

where σ_1 is the standard deviation of the output noise $n_1(x)$. For this we assume that $\{\xi \text{ s.t. } \hat{f}(\xi) \neq 0 \text{ and } \mathcal{E}(\xi) = 0\}$ is a zero-measure set.

Then it follows directly from the Cauchy-Schwarz inequality that in the class of all linear filters the optimal one is the so-called *matched filter* given by its transfer function:

$$\hat{k}_{opt}(\xi) = C e^{-i\xi x_0} \frac{\overline{\hat{f}(\xi)}}{\mathcal{E}(\xi)}, \tag{84}$$

with a non-zero arbitrary constant C . The maximum output SNR is then given by

$$\max(\text{output SNR})^2 = \frac{1}{2\pi} \int \frac{|\hat{f}(\xi)|^2}{\mathcal{E}(\xi)} d\xi. \tag{85}$$

A detection is declared if the output of the matched filter is beyond a given threshold, based upon the statistics of the output noise. In realistic situations such as the GW detection, the signal to be detected has several unknown parameters, then the problem turns into an estimation/detection one (see e.g. [5, 9, 35] for a review). The practical way currently under study by many authors is the construction of a family of filters (called templates) which will form a net in the space of all filters. Of course one looks for a good compromise between accuracy and calculation cost (which is directly related to the number and the sizes of the templates). In the following we shall explore some wavelet-based alternatives.

5.2. Why wavelets? Let us now explain why wavelet analysis seems to be a good candidate for the detection problem described above. First, wavelets have a certain number of intrinsic nice properties, such as the existence of fast algorithms (described in Appendix below). But in addition, a closer look at the signal model in (77) shows that wavelet analysis turns out to be naturally adapted to the problem. Let us neglect for the time being the phase factor Φ . We see that Eq. (77) is (up to a normalization factor) a shifted and scaled copy of a reference signal

$$f_{\alpha,\beta}(x) = \theta(-x)x^\alpha \exp\{ix^{\beta+1}\}. \tag{86}$$

(In the terminology of [30] $f_{\alpha,\beta}(x)$ is called a *chirp of type* $(\alpha, -\beta - 1)$; note however that the numerical value of β is out of the range of the analysis of [30]; gravity waves are not oscillatory enough to qualify as true trigonometric chirps.) Therefore, as a consequence of the covariance properties of the wavelet transform, the transforms of the corresponding signals are obtained from the transform of (86) by the action of the affine group. More precisely, we have that

$$Z_f(x) = A F^{-\alpha/(\beta+1)} e^{i\Phi} f_{\alpha,\beta}\left(F^{1/(\beta+1)}(x_0 - x)\right) \tag{87}$$

and then by Lemma 2.1

$$T_f(b, a) = AF^{-\alpha/(\beta+1)} e^{i\Phi} T_{f_{\alpha, \beta}} \left(F^{1/(\beta+1)}(x_0 - b), F^{1/(\beta+1)}a \right), \quad (88)$$

and the detection problem may be reformulated as a detection problem in the time-scale plane: find a translation parameter x_0 and a scale parameter $a_0 = F^{1/(\beta+1)}$ such that $T_f(b, a) = T_{f_{\alpha, \beta}}((-x_0, a_0)^{-1} \cdot (-b, a))$ where \cdot is the product in the affine group given above.

This actually opens the problem of developing a detection theory for functions defined on the affine group, where translation is replaced with the group action. To our knowledge, such a theory has not been developed so far.

Another reason for which it is natural to use wavelet-based techniques is the fact that the expected signal is mainly characterized by a time-varying frequency. It turns out that time-frequency representations of such signals have particular localization properties. We describe below the behavior of the wavelet transform of such signals.

5.3. The wavelet transform. As we saw in the previous sections, the wavelet transform has a tendency to concentrate in the neighborhood of a ridge. In the case of signals of the form (70), such ridges take the form

$$a_r(b) = \frac{\omega_0}{2\pi F} (x_0 - b)^{3/8}, \quad (89)$$

or equivalently, if we consider the reciprocal function

$$b_r(a) = x_0 - \left(\frac{2\pi a F}{\omega_0} \right)^{8/3}. \quad (90)$$

The problem amounts to that of finding an efficient and robust algorithm for detecting the ridge and estimating its parameters. We now describe the application to this problem of the methods mentioned to above.

5.4. Looking for ridges. Let us consider numerical examples of the continuous wavelet transform for the simulated binary coalescence signals. In Figure 2 we analysed the noise-free signal given by Eq.(70) (with $A = 1$ and $F = 50$). The localization of the wavelet transform near the ridge (shown in the figure) appears clearly. In that example, the ridge was estimated by the method described in Section 4.2. In Figure 5 we consider the same signal but with an additive stationary Gaussian noise with spectral density modelling the

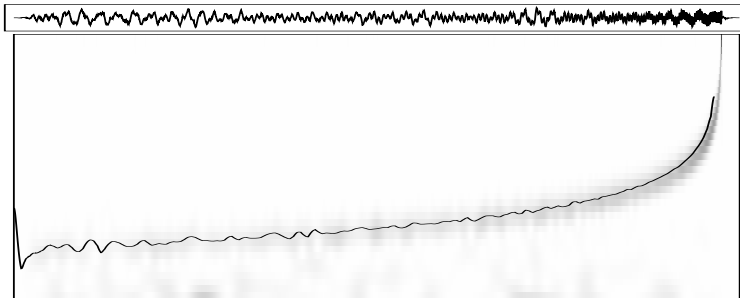


Figure 5: Noisy binary coalescence signal, buried in shot and thermal noises, with $\kappa = .1$.

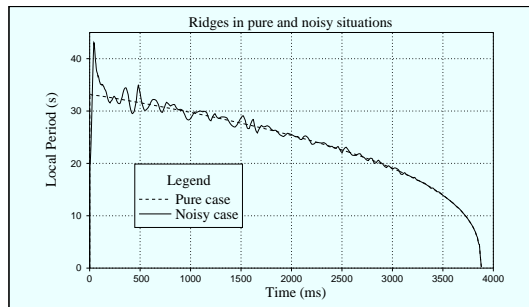


Figure 6: Comparison of ridges in noise-free and noisy situations.

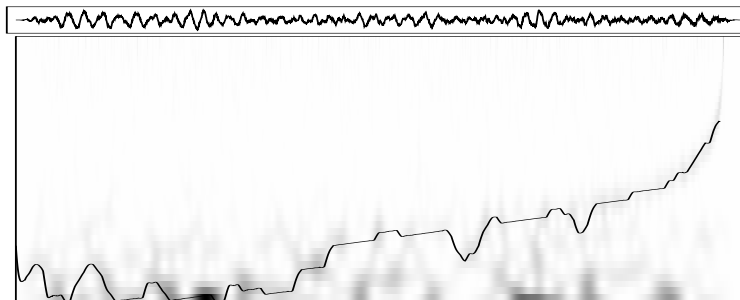


Figure 7: Noisy binary coalescence signal, buried in shot and thermal noises, with $\kappa = .5$.

thermal and shot noises in the interferometric detectors of the form

$$\mathcal{E}(\xi) = \kappa (S_s + S_t \xi^{-4}) ,$$

where $S_s = 1$, $S_t = 10^8$, and $\kappa = .1$. (In our simulations the seismic noise has been replaced with a cutoff at a frequency of $10Hz$.)

The same localization properties of the wavelet transform may be observed (notice that in our graphical conventions, the gray levels are automatically adjusted to the range of the transform; thus even though the contribution of the frequency modulated signal seems weaker in Figure 5 than in Figure 2, this is just a graphical effect), but the localization is now somewhat blurred by the presence of the noise. As a consequence, the ridge (again shown in the figure) is not extracted as accurately as before. The two ridges (in noise-free and noisy cases) are compared in Figure 6.

Finally, we analysed the same signal but now with $\kappa = .5$ which means higher noise. To improve the accuracy of the ridge extraction we applied the penalization algorithm described in Section 4.3. The result is shown Figure 7. As may be seen from the wavelet transform, the signal is still visible at small scales, i.e. at high frequencies, but is perturbed by noise at larger scales. The ridge is correctly estimated in the small scales domain, i.e. when its “energy” is at least comparable with that of the noise. On the contrary, in the large scales, the estimated ridge is attracted by the low frequency part of the noise (in this case the thermal noise). This clearly shows that the non-parametric methods such as the ones we have described here are limited to a certain range of signal to noise ratio.

5.5. Parametric methods. Up to now, we have only described a series of methods which *do not* take into account the explicit model we have for the signal. Clearly, the performances of the algorithms may be greatly enhanced by taking such information into account. We describe here a number of possible approaches based upon wavelet transform. For the sake of simplicity, we stick to the case of the continuous wavelet transform as described above. We refer to [28] for a precise description of related methods.

Time-frequency template matching. This approach is a time-frequency version of Wiener's filter. The main idea here is to replace the classical analysis, which matches phases, with a frequency matching. Similar ideas have been developed in a different form in [21, 50].

Let us start with the additive noise model

$$f(x) = Af_F(x_0 - x) + n(x) .$$

We have

$$\mathcal{M}(b, a) = |T_{f_F}(x_0 - b, a)|^2 + N(b, a)$$

where $\mathcal{M}(b, a)$ is defined in (45). $N(b, a)$ is a noise term, whose statistics has been described above.

Fix a date τ , and consider a domain $\Omega_\tau = [\tau - T, \tau] \times [a_{min}, a_{max}]$ in the time-scale domain. For simplicity, we shall denote by $L^2(\Omega_\tau)$ the space $L^2(\Omega_\tau, dadb/a)$. Let us now consider the following *time-scale templates*

$$\Gamma_{(\tau, \gamma)}(b, a) = \frac{\tilde{\Gamma}_{(\tau, \gamma)}(b, a)}{\|\Gamma_{(\tau, \gamma)}\|_{L^2(\Omega_\tau)}} \quad (91)$$

where

$$\tilde{\Gamma}_{(\tau, \gamma)}(b, a) = |T_{f_\gamma}(\tau - b, a)|^2 . \quad (92)$$

Then, minimizing $\|M - A\Gamma_{(\tau, \gamma)}\|_{L^2(\Omega_\tau)}$ with respect to A and γ yields the following maximization problems:

$$\hat{F}(\tau) = \arg \max_{\gamma} \langle M, \Gamma_{(\tau, \gamma)} \rangle_{L^2(\Omega_\tau)} \quad (93)$$

$$\hat{A}(\tau) = \max_{\gamma} \langle M, \Gamma_{(\tau, \gamma)} \rangle_{L^2(\Omega_\tau)} \quad (94)$$

This leads to the following *meta-algorithm* for detection

- FOR $\tau = \tau_{min}$ TO $\tau = \tau_{max}$ DO
 1. Solve the maximization problem in (93) with respect to γ .
 2. Store the values $\hat{F}(\tau)$ and $\hat{A}^2(\tau)$.
- Scan the local maxima $\hat{A}(\tau_m)$ of the function $\hat{A}(\tau)$.
- IF $\hat{A}(\tau_m) \geq \text{THRESHOLD}$: mark τ_m as a possible date for an event.

Of course, the choice of the **THRESHOLD** depends on several parameters, and in particular it relies on some a priori knowledge on the behavior of the algorithm when only noise is present. Such a knowledge may easily be obtained through Monte-Carlo simulations.

Line integral methods. As an alternative, let us simply discuss the line integral approach, described in more details in [28]. The idea is there to exploit the expected energy concentration of the wavelet transform by considering restrictions of it to specific curves, in the same spirit as the algorithms described above and in [13]. Let us consider for the sake of simplicity the Newtonian signal, and let $\mathcal{M}(b, a)$ be as in (45).

Let (τ, γ) be a candidate for the pair (x_0, F) , and let us consider the corresponding ridge, expressed in terms of group delay

$$b_{(\tau, \gamma)}(a) = \tau - \left(\frac{2\pi a \gamma}{\omega_0} \right)^{8/3}. \quad (95)$$

Finally, consider the following *line integral*

$$\mathcal{L}_f(\tau, \gamma) = \int \mathcal{M}(b_{(\tau, \gamma)}(a), a) \frac{da}{a}. \quad (96)$$

Let us now consider the case of a gravitational wave generated by a binary system collapsing at time x_0 , with chirp parameter F . Using the stationary phase approximation derived above, we can see that in the non-noisy situation, we have that

$$\mathcal{L}_f(x_0, F) \approx \frac{2\pi}{\phi'_\psi(0)} |\psi(0)|^2 \|f\|^2 = \|f\|^2$$

In addition,

$$\mathcal{L}_f(\tau, \gamma) \leq \mathcal{L}_f(x_0, F).$$

This suggests to use the line integral $\mathcal{L}_f(\tau, \gamma)$ instead of the quantity $A(\tau)$ in the previous algorithm. More precisely, the algorithm is based on the following scheme: for any fixed τ , solve

$$\max_{\gamma} \mathcal{L}_f(\tau, \gamma) \quad (97)$$

and among the local extrema of the latter quantity, keep those which are above a certain threshold.

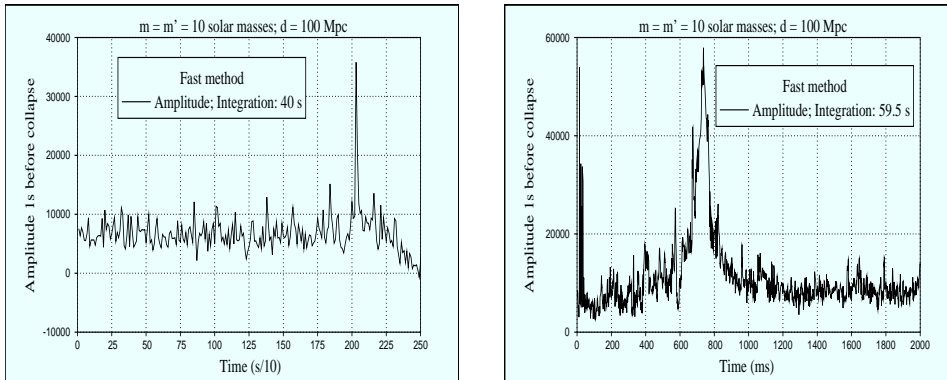


Figure 8: Line integral method for a pair of stars of 10 solar masses, at a distance of 100 Mpc; left plot: 20 seconds simulation, with a time step of 100 ms; right plot: 2 seconds simulation, with a time step of 1 ms.

The advantage of such an approach is that it can be made extremely fast, since any computation of $\mathcal{L}_f(\tau, \gamma)$ requires the evaluation of a single integral instead of a double one. In our implementation, the maximization is performed using an adapted version of Brent's method, and the usual wavelet transform is replaced with a predenoised one. We refer to [28] for more details on this method.

As an illustration, we show in Figure 8 the result of the method for the case of a binary system made of 2 stars of 10 solar masses, at a distance of 100 Mpc. The signal was simulated with the SIESTA software, and includes a Newtonian approximation for the signal and the VIRGO detector noise (provided by the VIRGO collaboration), whose spectral density is given in Figure 1.

Remark 5.4. Several variations around these two schemes are possible. For instance, it is shown in [28] that it is convenient to replace the wavelet transform of the signal with the so-called *prewhitening wavelet transform*, in which the spectral density of the noise (which in that case has to be known in advance, from a model or from previous experiments) is taken into account. It may be shown that in such a case, the output of the algorithm is equal to that of the matched filter. Other variations were given in [21].

6. Conclusions. In this paper we have given a brief description of continuous wavelet transform, focusing on some particular aspects which we believe relevant for gravitational waves detection.

More precisely, we have described a set of methods for analyzing and detecting amplitude and frequency modulated signals embedded in noise. Some of the signals which are expected at the gravitational waves detectors, namely gravity waves generated by coalescing binaries, fall into this class, and the techniques we described in this paper apply to these.

The first type of methods we have described are non-parametric. They amount to searching for the expected signal as a set of salient points or a curve in the time-scale plane. We gave two different formulations of this approach, based on local [19] or integral [13] techniques. They may be used if signal-to-noise ratio is low enough both for detection [29] and for parameter estimation [37].

The second methods are parametric methods. They also amount to searching for curves in the time-scale plane, but the curves are now given a specific functional form, based on Newtonian or post-Newtonian approximations [28]. They represent interesting alternatives to matched filter techniques, and may easily be implemented on line.

In addition in the Appendix we describe the main aspects of discrete wavelet decompositions, from a subband coding perspective. Let us stress that subband coding and quadrature mirror filters were originally introduced in order to reduce quantization noise in speech. Quantization noise seems to be a relevant issue for data acquisition at interferometric detectors, since the detector noise has a wide dynamical bandwidth (which will require a prewhitening prior to quantization). For this reason, discrete wavelets should be considered as a serious candidate from the data acquisition point of view as well.

J.M. Innocent is also ESM2, IMT-Technopole de Chateau-Gombert, 13451 Marseille, Cedex 20, FRANCE , and the VIRGO group, Orsay.

Acknowledgements. We are very indebted to A. Królak, for inviting us to Banach Center mini-semester “Mathematical Aspects of Theories of Gravitation” where these results were presented. We thank J.K. Blackburn, S. Dhurandhar, R. Flaminio, A. Królak, A. Rüdiger and M. Tinto for stimulating discussions. Thanks are also due to R. Carmona, A. Grossmann, M. Holschneider, W.L. Hwang, Y. Meyer, Ph. Tchamitchian and J.Y. Vinet for discussions on related topics, and to F. Cavalier and the VIRGO collaboration for providing us with simulated samples.

Dedication. This paper was written a few weeks after the untimely departure of Bernard Escudié. Bernard was a great source of inspiration, and a large part of the material presented in this paper is based on his ideas.

We dedicate this paper to his memory.

7. APPENDIX: Discrete wavelet decompositions from a sub-band coding perspective and fast algorithms. In this appendix, we describe the main aspects of discrete wavelet decompositions, and their connections to fast algorithms. Starting from the sampling theorem, we describe the quadrature mirror filters technique and the corresponding sub-band coding algorithms. We then describe the construction of wavelet bases and show how they fit into the sub-band coding schemes. Finally we turn to the algorithms for non-orthogonal wavelet decompositions.

Our starting point will be the sampling theorem, which asserts that any band-limited $L^2(\mathbb{R})$ function whose Fourier transform’s support is included in the interval $[-\pi\nu, \pi\nu]$ may be sampled without information loss with a sampling frequency $\nu_s \geq \nu/2$.

7.1. Perfect reconstruction quadrature mirror filters. Let us start with a discrete sequence $\{f_n, n \in \mathbb{Z}\}$, assumed for the sake of simplicity to consist of samples $f_n = f(n)$ of a band limited function $f(x)$ with unit sampling frequency. As a consequence of Poisson’s summation formula, the Fourier transform of the sequence is a 2π -periodic function, hereafter denoted by $F(\xi) = \sum_k \hat{f}(\xi + 2\pi k)$. Let us now consider the 2π -periodic functions $H(\xi)$ and $G(\xi)$ defined by $H(\xi) = \sum_k \chi_{[-\pi/2, \pi/2]}(\xi + 2\pi k)$ and $G(\xi) = H(\xi + \pi)$. Let

$$h_k = \frac{\sqrt{2}}{2\pi} \int_{-\pi}^{\pi} H(\xi) e^{-ik\xi} d\xi$$

$$g_k = \frac{\sqrt{2}}{2\pi} \int_{-\pi}^{\pi} G(\xi) e^{-ik\xi} d\xi$$

denote the (appropriately normalized) Fourier coefficients of $H(\xi)$ and $G(\xi)$ respectively, and introduce the sequences $\tilde{s}_n = \sum_k h_{n-k} f_k$ and $\tilde{t}_n = \sum_k g_{n-k} f_k$, with Fourier transforms $F(\xi)H(\xi)$ and $F(\xi)G(\xi)$ respectively. Clearly, the bandwidth of the sequences $\{\tilde{s}_k\}$ and $\{\tilde{t}_k\}$ is half that of the sequence $\{f_k\}$, so that these two sequences may be subsampled by a factor two without loss of information. Since in addition we have $H(\xi) + G(\xi) = 1 \forall \xi$, we deduce that the sequence $\{f_k\}$ is completely characterized by the two sequences

$$s_n = \sum_k h_{2n-k} f_k \tag{98}$$

$$t_n = \sum_k g_{2n-k} f_k \tag{99}$$

The sub-band coding technique (which was introduced in a signal processing context in order to reduce quantization noise) is an extension of this simple calculation. The goal is to replace the perfect filters with smoother ones, in order to reduce the number of operations in Eqs (98) and (99). By doing so, one introduces aliasing, which may be cancelled by an appropriate choice of the filters.

Let us then consider a pair of 2π periodic functions

$$\begin{aligned} H(\xi) &= \frac{1}{\sqrt{2}} \sum_k h_k e^{ik\xi} \\ G(\xi) &= \frac{1}{\sqrt{2}} \sum_k g_k e^{ik\xi}, \end{aligned}$$

and consider the sequences introduced in (98) and (99). Consider also the *reconstructed sequence*

$$f_k^r = \sum_n (\bar{h}_{2n-k} s_n + \bar{g}_{2n-k} t_n). \quad (100)$$

Imposing the perfect reconstruction, i.e. $f_n^r = f_n$ imposes constraints on the filters $H(\xi)$ and $G(\xi)$. The classical solution to these constraints yields the so-called *Quadrature Mirror Filters* (QMF for short), for which

$$G(\xi) = e^{i\xi} \overline{H(\xi + \pi)}, \quad (101)$$

and

$$|H(\xi)|^2 + |G(\xi)|^2 = 1. \quad (102)$$

The *sub-band coding* is based on a recursive implementation of the perfect reconstruction “convolution-subsampling” scheme described above. More precisely, we start again with a sequence $\{f_n\}$ and we set $s_0^n = f_n$. Then, defining

$$s_n^j = \sum_k h_{2n-k} s_k^{j+1} \quad (103)$$

$$t_n^j = \sum_k g_{2n-k} s_k^{j+1} \quad (104)$$

we know how to reobtain the sequence $\{s_k^{j+1}\}$ from the sequences $\{s_n^j\}$ and $\{t_n^j\}$:

$$s_k^{j-1} = \sum_n (\bar{h}_{2n-k} s_n^j + \bar{g}_{2n-k} t_n^j). \quad (105)$$

A sub-band coding of the sequence $\{f_n = s_0^n\}$ amounts to representing it by the coefficients $\{s_n^{-J}, t_n^{-J}, t_n^{-J+1}, \dots, t_n^{-1}\}$ instead of the original coefficients $\{f_n\}$. In the case of a finite sequence of length say $N = 2^L$, we then have (because of the subsampling)

$$2^{L-1} + 2^{L-2} + \dots + 2^{L-J} + 2^{L-J} = N$$

coefficients, i.e. the same number exactly.

Let us stress that the same filters are used throughout all the stages of the algorithm, i.e. for all values of j . It is easy to see that to complete a decomposition at all scales of a finite sequence of length N , the computational cost goes as $O(MN)$, where M is the length of the sequences $\{h_k\}$ and $\{g_k\}$. It is then an extremely efficient algorithm.

Remark 7.1. The original motivation for the introduction of sub-band coding was the need of reducing quantization noise. Such a problem appears as soon as the dynamical range of an (analog) signal is large. Then the dynamical range of the signal is generally much smaller within each of the sub-bands, making the quantization task easier. Since this seems to be the case with gravitational waves detector signals (where the spectral density of the noise varies over several orders of magnitude), we believe that sub-band coding could be an appropriate strategy.

We shall now see the close connection of sub-band coding with wavelets.

7.2. *Multiresolution analysis and its connections to sub-band coding.* The construction of orthonormal bases of wavelets relies on the concept of *multiresolution analysis*, which we discuss here for the sake of completeness (see [17] for a self contained and pedagogical introduction to the subject).

DEFINITION 7.1. A *multiresolution analysis* (MRA) of $L^2(\mathbb{R})$ is a collection of nested closed subspaces $V_j \subset L^2(\mathbb{R})$

$$\dots \subset V_{-1} \subset V_0 \subset V_1 \subset \dots \tag{106}$$

such that the following properties hold

1. $\overline{\cup V_j} = L^2(\mathbb{R})$ and $\cap V_j = \{0\}$.
2. If $f(x) \in V_0$, then $f(x-k) \in V_0$ for all $k \in \mathbb{Z}$; $f(x) \in V_j$ if and only if $f(x/2) \in V_{j-1}$.
3. There exists a function $\phi(x) \in V_0$ such that the collection of the integer translates $\phi(x-k), k \in \mathbb{Z}$ is an orthonormal basis of V_0 .

Many examples of MRA have been constructed. We refer to [14, 17, 40] for reviews. We shall not go into mathematical details here and we shall rather focus on the implications of the above definition.

It follows directly from the inclusion of the V_j spaces that $\phi(x)$ may be expressed as a linear combination of the functions $\phi(2x-k)$ (which form a basis of V_1). This yields the so-called *two-scale difference equation* (or *refinement equation*)

$$\phi(x) = \sqrt{2} \sum_k h_k \phi(2x+k) . \tag{107}$$

The coefficients h_k are the Fourier coefficients of a 2π -periodic function, denoted by

$$H(\xi) = \frac{1}{\sqrt{2}} \sum h_k e^{ik\xi} . \tag{108}$$

Denote by W_j the orthogonal complement of V_j in V_{j+1} . Then one may prove that there exists a function $\psi(x) \in W_0$ such that the collection $\{\psi(x-k), k \in \mathbb{Z}\}$ is an orthonormal basis of W_0 . More precisely, let m be an arbitrary integral number, and set

$$G(\xi) = e^{i(2m+1)\xi} \overline{H(\xi + \pi)} = \frac{1}{\sqrt{2}} \sum g_k e^{ik\xi} . \tag{109}$$

The coefficients g_k are related to the coefficients h_k by

$$g_k = -(-1)^k \overline{h_{2m+1-k}} \tag{110}$$

The function $\psi(x)$, called the wavelet associated with the MRA, is defined by

$$\psi(x) = \sqrt{2} \sum_k g_k \phi(2x + k) . \quad (111)$$

Remarkably enough, the functions $H(\xi)$ and $G(\xi)$ are quadrature mirror filters, i.e. they satisfy equations (101) and (102).

Let us introduce the following notation for the shifted and scaled wavelets and scaling functions

$$\begin{cases} \psi_{jk}(x) &= 2^{j/2} \psi(2^j x - k) , \\ \phi_{jk}(x) &= 2^{j/2} \phi(2^j x - k) , \end{cases} \quad (112)$$

and associate with any function $f(x) \in L^2(\mathbb{R})$ the following family of coefficients

$$\begin{cases} t_k^j &= \langle f, \psi_{jk} \rangle , \\ s_k^j &= \langle f, \phi_{jk} \rangle . \end{cases} \quad (113)$$

Then the results outlined above may be summarized as follows

THEOREM 7.1. *Let $\{V_j, j \in \mathbb{Z}\}$ be a MRA, with scaling function $\phi(x)$ and wavelet $\psi(x)$. Then the family $\{\psi_{jk}, j, k \in \mathbb{Z}\}$ is an orthonormal basis of $L^2(\mathbb{R})$. More precisely, any $f(x) \in L^2(\mathbb{R})$ may be decomposed as*

$$f(x) = \sum_{j,k} t_k^j \psi_{jk}(x) = \sum_k s_k^{j_0} \psi_{j_0 k}(x) + \sum_{j \geq j_0, k} t_k^j \psi_{jk}(x) . \quad (114)$$

7.3. Fast algorithms for orthonormal wavelet decompositions. Let us suppose now that we are given a multiresolution analysis with scaling function $\phi(x)$ and wavelet $\psi(x)$, and denote by $H(\xi)$ and $G(\xi)$ the associated QMFs as before. Then it is a direct consequence of the refinement equations that we have

PROPOSITION 7.1. *The coefficients t_k^j and s_k^j are related by*

$$\begin{cases} t_k^j &= \sum_{\ell} \overline{g_{\ell}} s_{2k-\ell}^{j-1} , \\ s_k^j &= \sum_{\ell} \overline{h_{\ell}} s_{2k-\ell}^{j-1} . \end{cases} \quad (115)$$

In other words, we are exactly in a sub-band coding situation. This result is remarkable in many respects. Let us just mention that it provides for free a fast algorithm for orthonormal wavelet decompositions (let us stress that even though the wavelets get larger and larger as the scale grows, the computational cost itself does not depend on the scale). The connection between wavelet bases and subband coding was established first by S. Mallat [39], and clarified later on by A. Cohen and W. Lawton. For more details, we refer to [17, 51].

7.4. Fast algorithms for non-orthonormal wavelet decompositions. Let us now consider a slightly different situation, closer to the continuous wavelet transform described in the main body of this paper.

To start with, we consider scales which are still restricted to be powers of 2, but we now allow the values of the shift variable to belong to a given lattice, independent of the

scale, say \mathbb{Z} . This leads to consider the following set of coefficients

$$\begin{cases} T^j(k) &= 2^j \int f(x)\psi(2^j(x-k)), \\ S^j(k) &= 2^j \int f(x)\phi(2^j(x-k)). \end{cases} \quad (116)$$

Again, it follows directly from the refinement equations that such coefficients may be computed as follows

PROPOSITION 7.2. *The coefficients $T^j(k)$ and $S^j(k)$ are related by*

$$\begin{cases} T^j(k) &= \sum_{\ell} \overline{g_{\ell}} S^{j-1}(k - 2^{j-1}\ell), \\ S^j(k) &= \sum_{\ell} \overline{h_{\ell}} S^{j-1}(k - 2^{j-1}\ell). \end{cases} \quad (117)$$

The corresponding algorithm is as efficient as the previous one. Indeed, assuming that we have at hand $N = 2^L$ discrete values $S^0(k)$ to start with, it is easy to see that the number of operations required to compute the wavelet coefficients $T^j(k)$ for $k = 1, \dots, N$ and $j = -1, \dots, -L$ goes as $O(MN \log(N))$ (to compute $N \log(N)$ coefficients).

To deal with scales which are not restricted to powers of two, the situation is somewhat more complicated, and one has to turn to approximate algorithms (if one wants to stick to sub-band coding techniques; extremely efficient alternatives relying on FFT-based implementations are also available). We refer to [2, 42] and references therein for a discussion of such approaches.

7.5. *Some classical examples*

Haar's wavelets. Let us consider the following pair of filters:

$$h_0 = h_{-1} = g_{-1} = -g_0 = \frac{1}{\sqrt{2}}, \quad h_k = g_k = 0 \quad \forall k \neq 0, -1$$

Thus,

$$H(\xi) = \frac{1}{2} (1 + e^{-i\xi}) .$$

The corresponding pyramid algorithm reads

$$\begin{cases} s_k^j &= \frac{1}{\sqrt{2}} (s_{2k}^{j+1} + s_{2k+1}^{j+1}), \\ t_k^j &= \frac{1}{\sqrt{2}} (s_{2k+1}^{j+1} - s_{2k}^{j+1}). \end{cases}$$

It may be expressed simply in terms of sums and differences.

It is easy to check that such a choice leads to

$$\phi(x) = \chi_{[0,1]}(x) ,$$

and

$$\psi(x) = \chi_{[\frac{1}{2},1]}(x) - \chi_{[0,\frac{1}{2}]}(x) .$$

The corresponding wavelet basis is known to as the *Haar basis*. It is made of compactly supported functions, thus achieving optimal localization in the time domain. However, Haar wavelets are poorly localized in the frequency domain (since both $\hat{\varphi}(\xi)$ and $\hat{\psi}(\xi)$ decay as $1/\xi$ at infinity).

Spline wavelets. Let $V_0 = \{f \in C^{r-1}, f(x) = \text{polynomial of degree } r \text{ on } [k, k+1]\}$, and define V_j by scaling of V_0 . Let $\chi(x) = \chi_{[0,1]} * \chi_{[0,1]} * \dots * \chi_{[0,1]}(x)$ ($r+1$ times) and set

$$\hat{\phi}(\xi) = \frac{\hat{\chi}(\xi)}{\sum_k |\hat{\chi}(\xi + 2\pi k)|^2}.$$

It may be checked that this yields a multiresolution analysis with scaling function $\phi(x)$, from which the wavelet $\psi(x)$ may be computed easily. The corresponding wavelets are called spline wavelets, and have been described in great details in [14] (with several generalizations). Neither $\phi(x)$ nor $\psi(x)$ are compactly supported, but they have exponential decay. In addition, $\hat{\phi}(\xi)$ decays as ξ^{-r} at infinity. The wavelet has the same localization and regularity properties as the scaling function. In addition, $\psi(x)$ has $r+1$ vanishing moments.

Daubechies' wavelets. Another classical strategy consists in looking for compactly supported quadrature mirror filters which would generate orthonormal wavelet bases. This approach was developed by I. Daubechies [17], who proposed to look for filters of the form

$$H(\xi) = \left(\frac{1 + e^{-i\xi}}{2} \right)^r \mathcal{F}(\xi),$$

and search for trigonometric polynomials $\mathcal{F}(\xi)$ such that the resulting wavelet $\psi(x)$ is in $L^2(\mathbb{R})$ and yields an orthonormal basis of $L^2(\mathbb{R})$. This leads to compactly supported wavelets, whose frequency localization is described by $|\hat{\psi}(\xi)| \sim |\xi|^{-\alpha r}$ as $|\xi| \rightarrow \infty$. Tables for the corresponding filter coefficients g_k and h_k are given in [17], as well as precise estimates for the coefficient α .

The two constructions outlined above have found a lot of generalizations in the literature. We have no room here to give a precise account of these, and refer the reader to [14, 17, 40, 51, 53] for example.

References

- [1] A. Abramovici et al. (1992): LIGO: the Laser Interferometer Gravitational-Wave Observatory, *Science*, **256**, p. 325-333.
- [2] P. Abry and A. Aldroubi (1995): Designing Multiresolution Analysis Type Wavelets and their Fast Algorithms, *Int. J. of Fourier Anal. and Appl.*, **2**, 135-159.
- [3] A. Antoniadis & G. Oppenheim Eds (1994): proceedings of the conference *Wavelets and Statistics*, Villard de Lans, France, Lecture Notes in Statistics.
- [4] F. Auger and P. Flandrin (1993): Improving the Readability of Time-Frequency and Time-Scale Representations by the Reassignment Method, Technical Report 93.05, Laboratoire d'automatique, Ecole Centrale de Nantes.
- [5] R. Balasubramanian, B.S. Sathyaprakash and S.V. Dhurhandar (1995): Gravitational Waves from Coalescing Binaries: Detection Strategies and Monte Carlo Estimation of Parameters, *Phys. Rev.* **D 53**, vol. 6 pp. 3033-3055.
- [6] L. Blanchet, T. Damour and B.R. Iyer (1995): *Phys. Rev.* **D 51**, p. 5360.
- [7] B. Boashash (1992): Estimating and Interpreting the Instantaneous Frequency of a Signal,

- Part I: Fundamentals. *Proc. IEEE* **80**, pp. 520–538. Part II: Applications and Algorithms. *Proc. IEEE* **80**, pp. 540–568.
- [8] C. Bradaschia et al. (1990): *Nucl. Instrum. Methods Phys. Res. A* 518.
 - [9] D.R. Brillinger (1981): *Time Series; Data Analysis and Theory*, Holden Day Inc.
 - [10] R. Carmona (1993): Wavelet Identification of Transients in Noisy Signals. in *Mathematical Imaging: Wavelet Applications in Signal and Image Processing* pp. 392-400.
 - [11] R. Carmona, W.L. Hwang, B. Torr sani (1995): Characterization of signals by the ridges of their wavelet transform, preprint, submitted to *IEEE Trans. Signal Processing*.
 - [12] R. Carmona, W.L. Hwang, B. Torr sani (1995): Multiridge Detection and Time-Frequency Reconstruction, preprint, submitted to *IEEE Trans. Signal Processing*.
 - [13] R. Carmona, W.L. Hwang, B. Torr sani, *Practical Time-Frequency Analysis*, monograph, to appear.
 - [14] C.K. Chui (1992): *An Introduction to Wavelets*, Academic Press.
 - [15] J.M. Combes, A. Grossmann, Ph. Tchamitchian Eds. (1989): *Wavelets, Time-Frequency Methods and Phase Space*, Springer Verlag.
 - [16] E. Copson (1965): *Asymptotic expansions*, Cambridge University Press.
 - [17] I. Daubechies (1992): *Ten Lectures on Wavelets*, CBMS-NFS Regional Series in Applied Mathematics 61.
 - [18] M.H.A. Davis (1989): A review of the statistical theory of signal detection in *Gravitational Wave Data Analysis*, B.F.Schutz Ed., Kluwer Academic Publishers.
 - [19] N. Delprat, B. Escudi , P. Guillemain, R. Kronland-Martinet, Ph. Tchamitchian, B. Torr sani (1992): Asymptotic wavelet and Gabor analysis: extraction of instantaneous frequencies. *IEEE Trans. Inf. Th.* **38**, special issue on Wavelet and Multiresolution Analysis 644-664.
 - [20] R.B. Dingle (1973): *Asymptotic expansions, their derivation and interpretation*, Academic Press.
 - [21] R. Flaminio, L. Massonnet, B. Mours, S. Tissot, D. Verkindt and M. Yvert (1994): Fast Trigger Algorithms for Binary Coalescences, *Astroparticle Physics* 2, pp. 235-248.
 - [22] P. Flandrin (1993): *Temps-Fr quence*, Trait  des Nouvelles Technologies, s rie Traitement du Signal, Hermes.
 - [23] D. Gabor (1946): Theory of communication, *J. Inst. Elec. Eng.* **903**, 429.
 - [24] A. Grossmann, R. Kronland-Martinet, J. Morlet (1989): Reading and understanding the continuous wavelet transform, in [15].
 - [25] A. Grossmann, J. Morlet (1984): Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM J. of Math. An.* 15 ,723.
 - [26] P. Hall, W. Qian and D.M. Titterington (1992): Ridge Finding from Noisy Data. *J. Comput. and Graph. Statist.* **1**, 197-211.
 - [27] J.M. Innocent (1994): Remarks about the Detection of Coalescent Binaries, VIRGO technical report.
 - [28] J.M. Innocent and B. Torr sani (1996): A Multiresolution Strategy for Detecting Gravitational Waves Generated by Binary Collapses, Preprint.
 - [29] J.M. Innocent, J.Y. Vinet (1992): Time-Frequency Analysis of Gravitational Signals from Coalescing Binaries, VIRGO technical Report.
 - [30] S. Jaffard and Y. Meyer (1995): Pointwise Behavior of Functions, preprint.
 - [31] M. Kass, A. Witkin and D.Terzopoulos (1988): Snakes: Active Contour Models, *Int. J. of Computer Vision*, 321-331.

- [32] K. Kodera, R. Gendrin, C. de Villedary (1978): Analysis of time-varying signals with small BT values, *IEEE Trans. ASSP* **26**, 64.
- [33] L.H. Koopmans (1995): *The spectral Analysis of Time Series*. Probability and Mathematical Statistics 22, Academic Press.
- [34] A. Królak (1989): Coalescing binaries to post-Newtonian order, in *Gravitational Wave Data Analysis*, B.F. Schutz Ed., Kluwer Academic Publishers.
- [35] A. Królak, K.D. Kokkotas, G. Schäfer (1995): On estimation of the post-Newtonian parameters in the gravitational-wave emission of a coalescing binary, *Phys.Rev.D***52**, 2089.
- [36] A. Królak and B.F. Schutz (1987): Coalescing binaries : probe of the Universe, *General Relativity and Gravitation*, **19**, 1163-1171.
- [37] A. Królak, P. Trzaskoma (1996): Application of the Wavelet Analysis to Estimation of Parameter of the Gravitational-Wave Signal from a Coalescing Binary, *Classical and Quantum Gravity* **13**, pp. 813-830.
- [38] P.J.M. van Laarhoven and E.H.L. Aarts (1987): *Simulated Annealing: Theory and Applications*, Reidel Pub. Co.
- [39] S. Mallat (1989): Multiresolution Approximation and Wavelets, *Trans. AMS* **615**, 69-88.
- [40] Y. Meyer (1989): *Ondelettes et opérateurs* (1989), Hermann.
- [41] Y. Meyer Ed. (1989): *Wavelets and applications*, Proceedings of the second wavelet conference, Marseille (1989), Masson.
- [42] M.A. Muschietti and B. Torrési (1995): Pyramidal Algorithms for Littlewood-Paley Decompositions. *SIAM J. Math. Anal.* **26** 925-943.
- [43] B. Picinbono, W. Martin (1983): Représentation des signaux par amplitude et phase instantanées, *Annales des Télécommunications* **38** (1983), 179-190.
- [44] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery (1992): *Numerical Recipes: the Art of Scientific Computing*, Second Edition, Cambridge University Press.
- [45] J. Stoer and R. Bulirsch (1991): *Introduction to Numerical Analysis*, Texts in Applied Mathematics 12, Springer Verlag.
- [46] Ph. Tchamitchian, B. Torrési (1991): Ridge and Skeleton extraction from wavelet transform, in *Wavelets and their Applications*, M.B. Ruskai & Al Eds, Jones&Bartlett, Boston.
- [47] K.S. Thorne (1987): Gravitational Radiation, in *300 Years of Gravitation*, Hawking & Israel Eds, Cambridge University Press.
- [48] K.S. Thorne (1996): Gravitational Waves from Compact Objects, proceedings of the IAU Symposium 165, J. van Paradijs, E. van der Heuvel and E. Kuulkers Eds, Kluwer.
- [49] B. Torrési (1995): *Analyse Continue par Ondelettes*, Coll. Savoirs Actuels, InterEditions/Éditions du CNRS (in French, to be translated into English).
- [50] D. Verkindt (1993): *Etude d'algorithmes rapides de recherche d'un signal d'onde gravitationnelle provenant de coalescence d'étoiles binaires*, PhD Thesis, Université de Savoie (in French).
- [51] M. Vetterli and J. Kovacevic (1995): *Wavelets and Sub Band Coding*, Prentice Hall Signal Processing Series.
- [52] J. Ville (1948): Théorie et Applications, de la Notion de Signal Analytique, *Cables et Transmissions* **2** 61-74.
- [53] M.V. Wickerhauser (1994): *Adapted Wavelet Analysis, from Theory to Software*, A.K. Peters Publ.
- [54] E.P. Wigner (1932): On the Quantum Corrections for the Thermodynamic Equilibrium, *Phys. Rev.* **40**, 749-759.