Characterization of Acoustic Signals Through Continuous Linear Time-Frequency Representations

PHILIPPE GUILLEMAIN AND RICHARD KRONLAND–MARTINET

Invited Paper

One important field in the framework of computer music concerns the modelization of sounds. In order to design digital models mirroring as closely as possible a real sound and permitting in addition intimate transformations by altering the synthesis parameters, we look for a signal model based on additive synthesis whose parameters are estimated by the analysis of real sounds. This model is relevant from both the physical and perceptual points of view, especially when the sound to be analyzed comes from a musical instrument. We will present some techniques, mostly unpublished, based on time-frequency representations which make possible the estimation of relevant parameters such as frequency and amplitude modulation laws corresponding to each spectral component of the sound. The techniques described will extend the results presented in [3]. These methods will then be transposed to broadband signals, allowing the characterization of transients.

I. Introduction

It has long been thought that audiophonic signals, and in particular the sounds produced by musical instruments, were wholly characterized by their spectral energy density or, in other words, by the knowledge of the energy ratio between each of its components.

The appearance of analogical sound treatment systems has rapidly shown that this information, although pertinent, was not in itself capable of characterizing a natural sound. To demonstrate this, one simply has to listen to a sound backward (i.e., by reversing the magnetic tape). The sound created in this way has the same energy spectrum as the original sound, only its phase spectrum has changed. However, the sound result is very different. In the same way, one can easily show that modifications on the "shape" of the Fourier transform of the components can drastically change the corresponding sound.

If both the phase and the amplitude of the Fourier spectrum play an important role, we could ask: what essential psychoacoustic information do they contain? The manipulation of natural sounds in the context of music known as "musique concrète" has brought a fair number of answers to this question [18]. Thus Schaeffer [19] gave a group of examples aiming to prove that the dynamic aspect of sound plays a role as important, if not more so, as the simple spreading of energy over the spectrum. In this way, the simple fact of modifying the global energy envelope of a sound allows one to give a drawn-out sound the "color" of a percussive sound. From a mathematical point of view, this dynamic imformation is contained in both the phase and the modulus of the Fourier representation, but is too hidden to be easily recovered.

It is only since the 1960's, with the appearance of digital synthesis, that we have been able to get a better comprehension of the perceptive criteria. Risset has shown, in a study of the trumpet sound [17], the importance of the temporal evolution linked to each component of the spectrum. The analysis techniques that he used were based on the Fourier analysis with synchronic windows (over a number of samples of a signal corresponding to the fundamental period). This analysis resolutely time-frequency has thus permitted the characterization of the laws of temporal evolution of each component and later the building of efficient synthesis algorithms.

From this point, most techniques for analysis and synthesis of audiophonic signals have concentrated on both temporal and frequential data. Many applications have been developed using either parametric methods taking into account an *a priori* knowledge of the signal (linear prediction in synthesis and analysis of speech signals) or nonparametric (phase vocoder).

In 1946, Gabor proposed an analysis-synthesis technique by "information grains" made up of elementary signals localized both in the time and the frequency domains. This type of decomposition, although aiming to optimize the transport of the information, gave rise to many audiophonic

Manuscript received June 1, 1995; revised December 15, 1995.
The authors are with the CNRS-Laboratoire de Mécanique et d'Acoustique, 13402 Marseille Cedex 20, France.
Publisher Item Identifier S 0018-9219(96)02351-1.

applications impossible to achieve using classical signal treatment methods. The wavelet transform can be considered as a "cousin" of the Gabor transform, since it also generally decomposes an arbitrary signal into elementary contributions localized in time and in frequency. The essential difference resides in the construction of the elementary functions and moreover in the fact that it is possible to select the reference function.

When we began to work in the wavelet field, we first addressed the problem of characterization of sounds through the behavior of the representations [13]. This problem is simpler if one requires the picture obtained to be invariant under time shifts and also the ability to choose the time and/or the frequency accuracy of the analysis. Moreover, in order to obtain physical information such as the energy distribution, we have shown that the choice of a progressive wavelet is suitable. These requirements led us to use continuous representations instead of "orthonormal" wavelet basis. We then investigated the potential of modifying the original sound by altering the pictures between the analysis and synthesis processes [14], [91]. Once again, psychoacoustic considerations led us to act on the phase and the modulus of the transforms. In that way, we gave the musicians a set of tools aiming to "sculpt" the sound and provide exotic effects such as harmonization, scattering, time stretching or transposition. More than mathematical, these nonlinear transformations were based on our expertise and some failures remainded unexplained. In order to proceed more precisely, we finally addressed the problem of the modelization of the sound by a combination of elementary components corresponding to physical aspects of the sound source, i.e., eigenfrequencies of the structure, or perturbations due to the playing (vibrato) [10]. The model we are interested in can be written as

$$s(t) = \sum_{k=1}^{N} A_k(t) \cos\left(\int_0^t \omega_k(u)\, du\right)$$

and the problem of simulating and transforming a sound consists in estimating the amplitude $(A_k(t))$ and the frequency $(\omega_k(t))$ modulations laws.

The aim of this paper is to present recent works in the field of the modelization of sounds using time-frequency related methods such as Gabor and wavelet representations. Even though the "pictures" obtained by these techniques are of great interest, the constraints due to the reproducing kernel do not allow one to estimate properly the parameters corresponding to a synthesis model and more sophisticated techniques have been derived. We will describe how some considerations on the signal allow one to construct good estimates of the modulation laws for different kind of signals (quasi-periodic or transient). For that purpose, we will mainly use the property of linearity of the transforms and most of the techniques have been based on the Gabor transform which is more appropriated to a time-frequency description of the sound. Nevertheless, the techniques described can often be supported by both wavelet and Gabor transforms.

We shall first recall our notations and conventions and then focus on the study of the behavior of the representations of the so-called "asymptotic" signals, which represent one single component of our synthesis model. After recalling the approximations obtained through stationary phase arguments, we shall introduce a new approximation which gives rise to a more intuitive description of the transforms, as well as new algorithms which enable a more accurate estimation of the modulation laws. We will then transpose these asymptotic approximations in the Fourier domain in order to describe the behavior of the representations of transients. Similar to the time-asymptotic case, we shall derive algorithms allowing the estimation of the group delay and the spectral density. These two approaches, corresponding to locally linear frequency or group delay modulation laws (and consequently well adapted to the Gabor transform), will then be extended to hyperbollically frequency modulated signals through the study of the wavelet transform of homogenous signals. We shall finally point out the advantage of the linearity of the representations, by building evolutive filters through linear combinations of restrictions of the transforms. These filters, aimed at the separation of components, will enable an accurate estimation of the modulation laws needed for the synthesis model.

II. Definitions and Notations

A. Fourier Transform

The Fourier transform of a function $f(t)$ denoted $\hat{f}(\omega)$ will be given by

$$\hat{f}(\omega) = \int f(t) e^{-i\omega t}\, dt$$
$$f(t) = \frac{1}{2\pi} \int \hat{f}(\omega) e^{i\omega t}\, d\omega.$$

B. Gabor Transform

The continuous Gabor transform of a signal $s(t)$ will be denoted by [5]

$$L_g(\tau, \alpha) = \int s(t) \overline{W}(t-\tau) e^{-i\alpha(t-\tau)}\, dt = \frac{1}{2\pi} \int \hat{s}(\omega) \overline{\hat{W}}(\omega - \alpha) e^{i\omega\tau}\, d\omega.$$

The parameter τ denotes the time translation. The parameter α denotes the frequency translation; it has the dimensions of frequency.

$W(t)$ is the localization window (the bar denotes the complex conjugate); it is a square integrable function localized and smooth, in both the time and the frequency domains. It is convenient to assume that $\hat{W}(\omega)$ is localized around $\omega = 0$, with an unique maximum for this value, and with compact support, at least from a numerical point of view. For instance, a gaussian is a good candidate.

From signal processing point of view, continuous Gabor transform for a fixed α is equivalent to linear bandpass

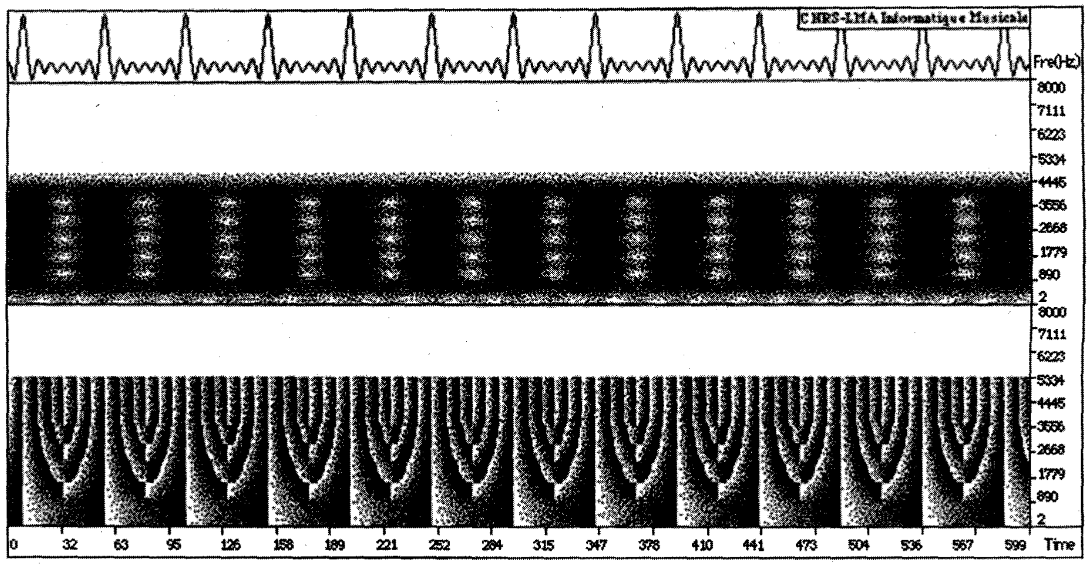


Fig. 1. Signal, modulus, and phase of the Gabor transform of a sum of six harmonic components.

filtering, centered around frequency α. In the frequency domain, since the filters are all obtained through the translation of the function $\hat{W}(\omega)$ by α, the decomposition has constant bandwidth: $\Delta\omega = Cst$.

When dealing with a real-valued signal, one can restrict the variations of α to R^+, since the values of $L_g(\tau, \alpha)$ for negative values of α can be deduced from $L_g(\tau, \alpha)$ for positive values of α.

It is convenient to represent the complex-valued 2-D functions obtained this way with the help of two pictures corresponding to the modulus and to the phase of the transform. For that purpose we use a density of black points proportional to the values of the modulus and of the principal values of the phase.

Figs. 1 and 2 show such representations. They correspond to the Gabor transform of the same signal (a sum of six harmonic components) analyzed with two different windows $W(t)$. In Fig. 1, the window is well localized with respect to the time, leading to a bad separation of the components in the frequency domain, but showing impulses in time due to the fact that the signal can also be considered as a filtered Dirac comb. In Fig. 2, the window $W(t)$ is well-localized in frequency, allowing the resolution of each component. In both figures, the phase behaves similarly, showing the periodicity of each component. This property has been used to estimate precisely the frequency of the components [10].

C. Wavelet Transform

The wavelet transform is defined as follows [1]

$$T(b, a) = \frac{1}{a}\int \overline{g}\left(\frac{t-b}{a}\right) s(t)\, dt$$
$$= \frac{1}{2\pi}\int \hat{s}(\omega)\overline{\hat{g}}(a\omega)e^{i\omega b}\, d\omega.$$

The parameter b denotes the time translation. The parameter a denotes the dilation or the scale, which is dimensionless.

For our purpose, we shall only consider wavelets $g(t)$ of the form: $g(t) = W(t)\exp(i\omega_o t)$. We shall assume that $W(t)$ is real and symmetric (its Fourier transform is real), and that its Fourier transform vanishes at $\omega = \omega_o$ (admissible wavelet), the other requirements remaining the same as for the Gabor transform.

The restriction of the continuous wavelet transform for a fixed scale a is equivalent to bandpass linear filtering operation centered around frequency $\frac{\omega_o}{a}$. In the time domain, since the filters are all obtained through the dilation of the function $W(t)e^{i\omega_o t}$ by a, the decomposition has constant relative bandwidth: $\frac{\Delta\omega}{\omega} = Cst$.

When one works with real-valued signals, one can restrict the variations of the parameter a to R^{+*}, since all the information contained in the signal is available by looking at the positive frequencies. In this case, it is convenient to assume that the wavelet is progressive:

$$\hat{W}(\omega - \omega_o) = 0 \qquad \omega \leq 0.$$

Under this assumption, the components of negative frequency do not interfere with the positive ones, and the modulus and the phase of the wavelet transform can be linked to physical parameters [Gros]. From a practical point of view, if $W(t)$ is a gaussian, the progressivity condition is effectively verified when ω_o is greater than six. In the Gabor case, this condition cannot be satisfied for any frequency translation α.

Fig. 3 shows the wavelet transform of a sum of six harmonic components. In contrast with Figs. 1 and 2, one can see that the wavelet transform privileges the frequency accuracy at low frequency (large a) and the time accuracy at high frequency (small a).

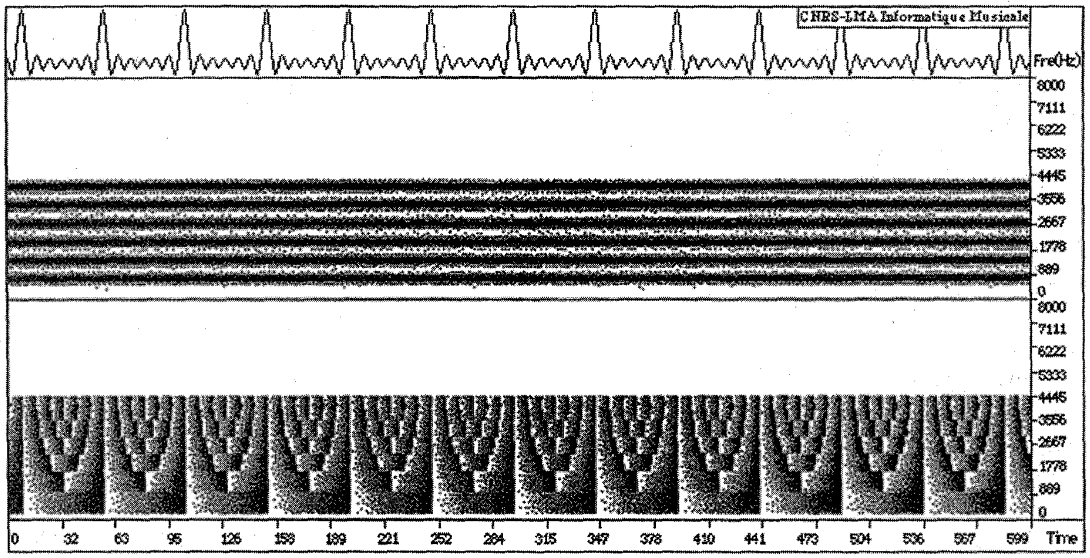


Fig. 2. Signal, modulus, and phase of the Gabor transform of a sum of six harmonic components.

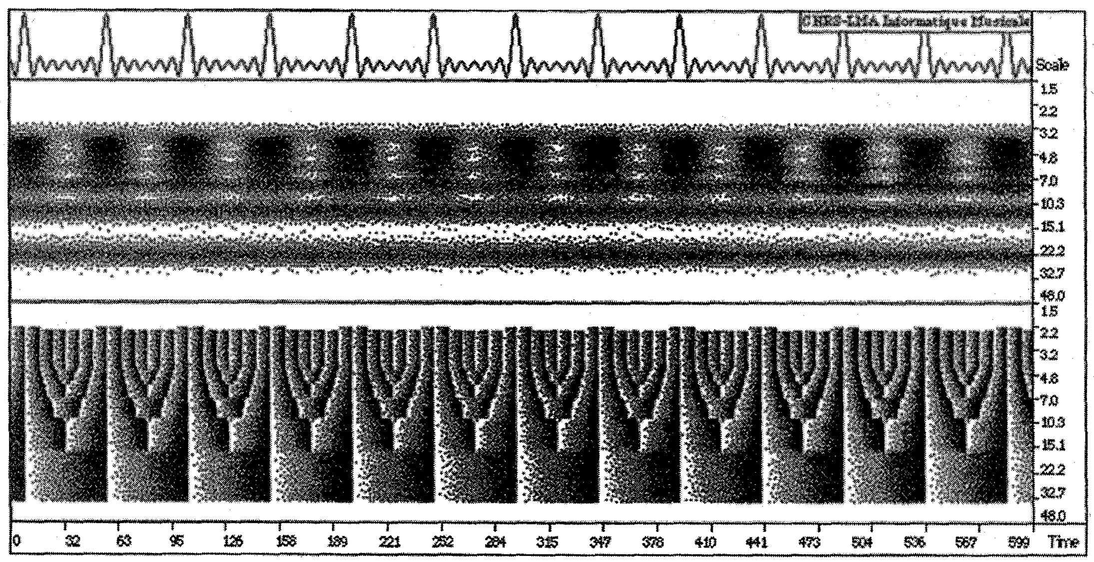


Fig. 3. Signal, modulus, and phase of the wavelet transform of a sum of six harmonic components.

D. Time and Frequency Asymptotic Signal

1) Time-Asymptotic Signal: A signal $s(t) = A_s(t)e^{i\Phi_s(t)}$ is called asymptotic with respect to the window $W(t)$ under the following assumptions [16]

$$\left|\frac{d\Phi_s(t)}{dt}\right| \gg \left|\frac{1}{A_s(t)}\,\frac{dA_s(t)}{dt}\right|$$
$$\left|\frac{1}{A_s(t)}\,\frac{dA_s(t)}{dt}\right| \ll \left|\frac{1}{W(t)}\,\frac{dW(t)}{dt}\right|.$$

These two conditions imply that $s(t)$ oscillates fast with respect to its amplitude variations, which are supposed to be locally constant on the support of $W(t)$.

2) Frequency-Asymptotic Signal: If we transpose all the requirements made on the signal from the time domain to the frequency domain, the signal $s(t)$ is called a frequency-asymptotic signal with respect to the window $W(t)$ if its Fourier transform $\hat{s}(\omega) = A_s(\omega)e^{i\Phi_s(\omega)}$ satisfies the following assumptions:

$$\left|\frac{d\Phi_s(\omega)}{d\omega}\right| \gg \left|\frac{1}{A_s(\omega)}\,\frac{dA_s(\omega)}{d\omega}\right|$$
$$\left|\frac{1}{A_s(\omega)}\,\frac{dA_s(\omega)}{d\omega}\right| \ll \left|\frac{1}{\hat{W}(\omega)}\,\frac{d\hat{W}(\omega)}{d\omega}\right|.$$

These conditions imply that the modulus of $\hat{s}(\omega)$ varies slowly on the support of $\hat{W}(\omega)$. As a consequence, the time support of the signal is small with respect to the time support of the analyzing function.

3) Instantaneous Frequency: The instantaneous frequency of an analytic signal $s(t) = A_s(t)e^{i\Phi_s(t)}$ is generally defined as its phase derivative [16], but it also corresponds to the "frequency centroid" of the signal:

$$\omega_i(t) = \frac{d\Phi_s(t)}{dt} = \mathrm{Re}\left\{ \frac{\int \omega \cdot \hat{s}(\omega)e^{i\omega t}\, d\omega}{\int \hat{s}(\omega)e^{i\omega t}\, d\omega} \right\}.$$

4) Group Delay: In the same way, the group delay of a signal $s(t)$ is generally defined through the phase derivative of its Fourier transform, but also corresponds to the "time centroid" of the signal

$$\tau_g(\omega) = -\frac{d\Phi_s(\omega)}{d\omega} = \mathrm{Re}\left\{ \frac{\int t \cdot s(t)e^{-i\omega t}\, dt}{\int s(t)e^{-i\omega t}\, dt} \right\}$$

where $\Phi_s(\omega)$ is the phase of the Fourier transform of $s(t)$.

5) Local Frequency: We shall call "local frequency" associated to the wavelet and Gabor transforms, the derivative of the phase of the transforms with respect to the time translation parameter.

Applying this definition to the transforms yields

$$\frac{\partial\Phi(\tau,\alpha)}{\partial\tau} = \mathrm{Re}\left\{ \frac{\int \omega \cdot \hat{s}(\omega)\overline{\hat{W}}(\omega-\alpha)e^{i\omega\tau}\, d\omega}{\int \hat{s}(\omega)\overline{\hat{W}}(\omega-\alpha)e^{i\omega\tau}\, d\omega} \right\}$$

$$\frac{\partial\Phi(b,a)}{\partial b} = \mathrm{Re}\left\{ \frac{\int \omega \cdot \hat{s}(\omega)\overline{\hat{g}}(a\omega)e^{i\omega b}\, d\omega}{\int \hat{s}(\omega)\overline{\hat{g}}(a\omega)e^{i\omega b}\, d\omega} \right\}$$

and the local frequency, in both the wavelet and the Gabor case, appears as the "centroid" in the frequency domain of the product of the signal by the analyzing function.

6) Local Delay: In the same way, we shall define the "local delay" of the Gabor transform as the phase derivative of the transform with respect to the frequency translation parameter

$$\frac{\partial\Phi(\tau,\alpha)}{\partial\alpha} = \tau - \mathrm{Re}\left\{ \frac{\int t \cdot s(t)\overline{W}(t-\tau)e^{-i\alpha(t-\tau)}\, dt}{\int s(t)\overline{W}(t-\tau)e^{-i\alpha(t-\tau)}\, dt} \right\}.$$

The local delay is related to the "centroid" in the time domain of the product of the signal and the analyzing function. In the wavelet case, an equivalent definition cannot be obtained, since the phase derivative of the transform with respect to the scale is dimensionless. Nevertheless, we shall see below that in a specific context, a similar definition can be proposed.

III. Representations of Mono-Component Asymptotic Signals

In this section, we shall study evolutive signals whose time-frequency representations displays a localization of the energy along a unique trajectory. We shall show that, depending on the assumptions made on the signal, this trajectory is related either to the instantaneous frequency (time-asymptotic signals) or to the group delay (frequency-asymptotic signals). Several algorithms, allowing a continuous estimation of this trajectory will be proposed.

A. Time-Asymptotic Signals

Let us first consider a time-asymptotic signal as defined previously. This assumption makes sense in the framework of musical sounds, since it corresponds to a signal, the frequency of which is locally (with respect to the time support of the analyzing window) linear, and the amplitude of which is locally constant. These assumptions are not too restrictive from a perceptive point of view, if one remembers that the duration of the window generally corresponds to a few milliseconds of sound. In this case, it has been shown [3], [4] that the stationary phase method leads to a correct approximation of the transforms of such a signal as well as continuous algorithms for the estimation of the modulation laws. The interested reader will find more details on this approach in the Appendix.

Even though this approach gave new tools to estimate parameters of a synthesis model, some problems still remained:

1) These approximations fail in the case of a monochromatic signal.
2) Contrary to what we would expect, the behavior of the transforms does not generalize the monochromatic case. Indeed, it is not possible to relate the restriction of the transforms for a fixed time to the Fourier transform of the analyzing function.
3) In the Gabor case, two different criteria should give rise to the same ridge, but in practice, the curves extracted are generally different.

In order to give answers to these problems, in the Gabor case, we shall propose a different approach allowing a more accurate approximation of the transform as well as a more intuitive description of its behavior by relaxing the hypothesis made on the amplitude modulation.

1) Local Chirp-Based Approximation: Roughly speaking, the stationary phase approximation consists of expanding the phase of the signal around a point in time that cancels the term involving the first derivative of the phase of the expression to be integrated. In a certain sense, this approximation "forgets" that the modulus of the analysis function localizes a portion of signal that also contributes to the integral. Here, we shall consider an approximation which takes into account this contribution. The hypothesis on the signal still remains the same, but in order to be more general, we rather use a window $W(t)$ defined by $W(t) = \exp(\frac{-t^2}{2\sigma^2})\exp(i\beta(\frac{t^2}{2}))$.

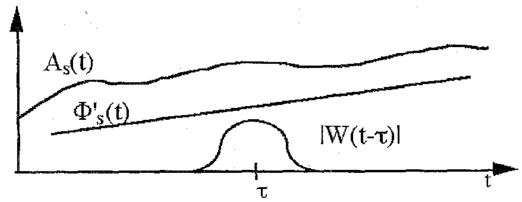


Fig. 4. Time-asymptotic hypothesis made on the signal and the window.

As one will see soon, this "chirped" window will be very useful. For the calculation, one has first to assume that around $t = \tau$, the phase of the signal can be well approximated by its second-order limited expansion around $t = \tau$

$$\Phi_s(t) \simeq \Phi_s(\tau) + (t-\tau)\Phi'_s(\tau) + \tfrac{1}{2}(t-\tau)^2\Phi''_s(\tau).$$

In the same way, we consider an expansion of the amplitude modulation law around $t = \tau$

$$A_s(t) = \sum_{k=0}^{k=\infty} \frac{(t-\tau)^k}{k!} A_s^{(k)}(\tau).$$

Notice that, contrary to the stationary phase approximation, we do not assume that the amplitude modulation law is locally constant. This will make the calculus more relevant for musical sounds, especially when they rapidly vary (percussive sounds).

These assumptions on the signal and the window are illustrated in Fig. 4, where both the amplitude and the frequency modulation laws of the signal are represented.

Under these hypotheses, the Gabor transform is a simple gaussian integral and is expressed by using [6] as

$$L_g(\tau,\alpha) \simeq \exp(i\Phi_s(\tau))\sigma\sqrt{\frac{2\pi}{1-i\sigma^2(\Phi''_s(\tau)-\beta)}} \cdot \sum_{k=0}^{k=\infty} \frac{(i)^k}{k!} A_s^{(k)}(\tau)\frac{\partial^k}{\partial\alpha^k} \cdot \exp\left(-\frac{\sigma^2}{2}\,\frac{(\Phi'_s(\tau)-\alpha)^2}{1-i\sigma^2(\Phi''_s(\tau)-\beta)}\right).$$

Since the signal is asymptotic, we can assume that the term of order $k = 0$ in the above serie is predominant. Consequently, the transform is mainly governed by the term

$$A_s(\tau)\exp\left(-\frac{\sigma^2}{2}\,\frac{(\Phi'_s(\tau)-\alpha)^2}{1-i\sigma^2(\Phi''_s(\tau)-\beta)}\right) = A_s(\tau)\hat{W}(\Phi'_s(\tau)-\alpha)^{\frac{1}{1-i\sigma^2(\Phi''s(\tau)-\beta)}}.$$

The behavior of the transform can then be interpreted as a superposition of Fourier transforms of $W(t)$ located along the curve $\Phi'_s(\tau) = \alpha$, called ridge, that describes the frequency modulation law of the signal in the $\{\tau, \alpha\}$ plane (see Fig. 5).

In order to extract the frequency modulation law of the signal, we have to estimate the curve $\Phi'_s(\tau) = \alpha$ with the help of the coefficients of the transform.

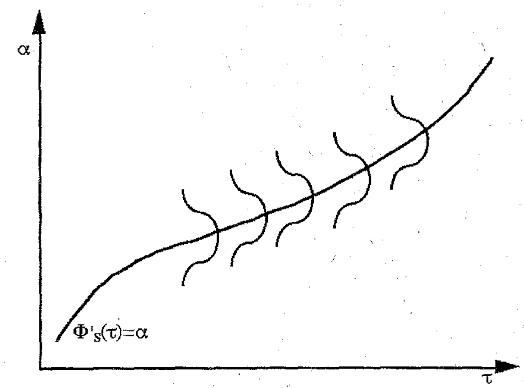


Fig. 5. Behavior of the Gabor transform of a time-asymptotic signal as a superposition of $\hat{W}$.

Since the expression of the transform is not easy to handle, we shall restrict the discussion to four particular cases leading to explicit results. The first one is the case of a constant frequency, the second one is the case of a constant amplitude, the third one is the case of a linear amplitude approximation, and the fourth one is the case of any polynomial amplitude, with an adapted window satisfying $\beta = \Phi''_s(\tau)$.

a) Constant frequency and spectral line: Let us first discuss the spectral line case which corresponds to a signal with constant frequency $(\Phi''_s(\tau) = 0)$. By assuming that $\beta = 0, \hat{W}(\omega)$ is symmetric about zero, and by denoting $\omega_s = \Phi'_s(\tau)$, the Gabor transform becomes

$$L_g(\tau,\alpha) = \exp(i\omega_s\tau)\left[A_s(\tau)\hat{W}(\omega_s-\alpha) + \sum_{k=1}^{\infty}\frac{(-i)^k}{k!}\hat{W}^{(k)}(\omega_s-\alpha)A_s^{(k)}(\tau)\right].$$

This expression needs some comments, especially when the analyzing frequency is $\alpha_s = \omega_s$.

At this frequency, the previous expression reads

$$L_g(\tau,\alpha_s) = \exp(i\omega_s\tau) \cdot \left[A(\tau)\hat{W}(0) + \sum_{k=1}^{\infty}\frac{(-i)^k}{k!}\hat{W}^{(k)}(0)A^{(k)}(\tau)\right] = K \cdot s(\tau) + \exp(i\omega_s\tau)\cdot \text{remainder}.$$

Under the assumptions made on W, one can say that at the frequency $\alpha_s = \omega_s$:

1) The first derivative of $\hat{W}(\omega)$ at $\omega = 0$ is zero, and consequently, the remainder is of order two.
2) Since $\hat{W}(\omega)$ is symmetric about zero, all its odd order derivatives at $\omega = 0$ are zero. The remainder is purely real and as a consequence, the local frequency of the Gabor transform directly provides ω_s, consequently, $\Phi'_s(\tau) = \alpha$ is verified at $\alpha = \omega_s$.
3) From the two previous statements, the modulus of the Gabor transform gives the amplitude modulation law of the signal, up to a second order error term. The

more $\hat{W}(\omega)$ and $A(t)$ are smooth, the more this error is small.

b) Linear frequency and constant amplitude: This approximation corresponds to the one presented in the appendix devoted to the stationary phase approximation. In this case, the transform is given by

$$L_g(\tau,\alpha) = \exp(i\Phi_s(\tau))\sigma\sqrt{\frac{2\pi}{1-i\sigma^2(\Phi_s''(\tau)-\beta)}}A_s(\tau) \cdot \exp\left(-\frac{\sigma^2}{2}\,\frac{(\Phi_s'(\tau)-\alpha)^2}{1-i\sigma^2(\Phi_s''(\tau)-\beta)}\right).$$

It is easy to check that the set of points $\Phi_s'(\tau) = \alpha$ satisfies both

$$\frac{\partial\Phi(\tau,\alpha)}{\partial\tau} = \alpha \quad \text{and} \quad \frac{\partial\Phi(\tau,\alpha)}{\partial\alpha} = 0.$$

This result is obviously the same as the one obtained by the use of the stationary phase approximation. Though these two criterias are equivalent, we shall only consider the first one as significant, since it involves quantities the dimensions of which are frequency, which is the quantity to estimate.

c) Linear frequency and amplitude modulation laws: If both the amplitude and the frequency modulation laws are linear, the transform is given by

$$L_g(\tau,\alpha) = \exp(i\Phi_s(\tau))\sigma\sqrt{\frac{2\pi}{1-i\sigma^2(\Phi_s''(\tau)-\beta)}} \cdot \exp\left(-\frac{\sigma^2}{2}\,\frac{(\Phi_s'(\tau)-\alpha)^2}{1-i\sigma^2(\Phi_s''(\tau)-\beta)}\right) \cdot \left(A_s(\tau) + i\sigma^2 A_s'(\tau)\frac{(\Phi_s'(\tau)-\alpha)}{1-i\sigma^2(\Phi_s''(\tau)-\beta)}\right).$$

After some calculation, we find that the set of points $\Phi_s'(\tau) = \alpha$ satisfies

$$\frac{\partial\Phi(\tau,\alpha)}{\partial\tau} + \Phi_s''(\tau)\frac{\partial\Phi(\tau,\alpha)}{\partial\alpha} = \alpha.$$

The ridge is here defined by a more complicated equation, taking into account the derivative of the phase of the transform with respect to both the time and the frequency. We shall see below that this equation can be generalized to signals the amplitude of which is arbitrary, up to the condition of a matching window.

d) Polynomial amplitude and matched window: A more general case is obtained if one uses a polynomial development of the amplitude modulation law. Moreover, if the second derivative of the phase of the analyzing function fits the second derivative of the phase of the signal ($\Phi_s''(\tau) = \beta$), the transform is given by

$$L_g(\tau,\alpha) = \exp(i\Phi_s(\tau))\sigma\sqrt{2\pi}\sum_{k=0}^{\infty}\frac{(i)^k}{k!}A_s^{(k)}(\tau) \cdot \frac{\partial^k}{\partial\alpha^k}\exp\left(-\frac{\sigma^2}{2}(\Phi_s'(\tau)-\alpha)^2\right).$$

One can then show that: The set of points $\Phi_s'(\tau) = \alpha$ and $\Phi_s''(\tau) = \beta$ satisfies

$$\left(\frac{\partial}{\partial\tau} + \Phi_s''(\tau)\frac{\partial}{\partial\alpha}\right)\Phi(\tau,\alpha) = \Phi_s'(\tau) = \alpha$$
$$\left(\frac{\partial}{\partial\tau} + \Phi_s''(\tau)\frac{\partial}{\partial\alpha}\right)^2\Phi(\tau,\alpha) = \Phi_s''(\tau) = \beta.$$

This new result (the "crossed criteria") shows that either the "horizontal" derivative or the "vertical" derivative of the phase introduce a bias in the extraction of the frequency modulation law of the signal if its amplitude is not constant. To avoid this bias, one must use a particular combination of these two derivatives, with a matched window.

This combination actually acts as the derivation of the phase of the transform in the direction of the slope of the frequency modulation law of the signal, given by $\Phi_s''(\tau)$.

2) Algorithmic Aspects: In order to use these criteria in practice, we shall now address the problem of the algorithmic estimation of the modulation laws from the Gabor transforms coefficients. The aim is to compute the transform only along the ridge rather than the whole transform. Then how can we achieve this since the ridge is unknown?

We shall first present an algorithm aiming to solve the equation of the ridge under the approximation of a constant amplitude linear chirp. Then we shall consider the linear chirp with polynomial amplitude approximation (the "crossed criterion").

1) Constant amplitude linear chirp approximation.

The problem to consider is: For a given parameter τ_o, find α such that $\frac{\partial\Phi(\tau_o,\alpha)}{\partial\tau} = \alpha$.

Several algorithmic methods can be found to solve this equation but, since it is a fixed-point equation under the form $F(\alpha) = \alpha$, with $F(\alpha) = \frac{\partial\Phi(\tau_o,\alpha)}{\partial\tau}$ it seems natural to use the classical fixed-point algorithm, which reads:

Let $\alpha_o(\tau_o)$ an "arbitrary" initial value.
$\alpha_{i+1}(\tau_o) = \frac{\partial\Phi(\tau_o,\alpha_i(\tau_o))}{\partial\tau}$.
The convergence criteria is $|\alpha_{i+1} - \alpha_i| < \epsilon, \epsilon$ fixed arbitrary small.
One proceeds to time $\tau_o + d\tau$ by the relation: $\alpha_o(\tau_o + d\tau) = \alpha_\infty(\tau_o)$.

In order to insure the convergence of this algorithm to $\alpha_\infty(\tau_o)$ satisfying: $\alpha_\infty(\tau_o) = F(\alpha_\infty(\tau_o))$, one has to check that the function $F(\alpha)$ is a contraction on a frequency subset $]\alpha_-, \alpha_+[$ containing both the initial value α_o and the final value α_∞. In other words: For any two consecutive iterations, i and $i+1$: $|F(\alpha_{i+1}) - F(\alpha_i)| < |\alpha_{i+1} - \alpha_i|$.

Thanks to the regularity of the Gabor transform, the function F is differenciable, and the previous condition becomes $|\frac{\partial F(\alpha)}{\partial\alpha}| < 1$.

Under the constant amplitude approximation, the derivative of $F(\alpha)$ with respect to α is given by

$$\frac{\partial F(\alpha)}{\partial\alpha} = \frac{\sigma^4\Phi_s''(\tau_o)(\Phi_s''(\tau_o)-\beta)}{1+\sigma^4(\Phi_s''(\tau_o)-\beta)^2}.$$

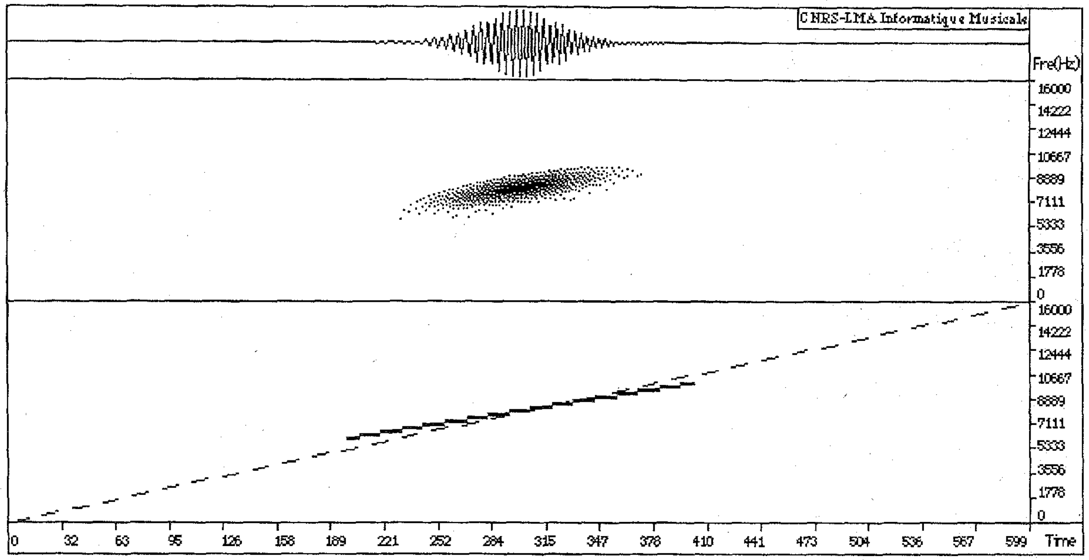


Fig. 6. Signal, modulus, and ridge of the Gabor transform of a gaussian linear chirp.

For the "classical" Gabor analysis ($\beta = 0$), as well as for the "matched window" case ($\beta = \Phi_s''(\tau)$), the absolute value of this expression is always smaller than one. Since $\frac{\partial F(\alpha)}{\partial \alpha}$ is independant of α, F is a contraction on R^+.

Such an algorithm can be easily transposed to the wavelet case.

2) Linear chirp with polynomial amplitude approximation.

The problem reads: For a given parameter τ_o, find α and β such that

$$\left(\frac{\partial}{\partial \tau} + \beta \frac{\partial}{\partial \alpha}\right) \Phi(\tau_o, \alpha) = \alpha$$
$$\left(\frac{\partial}{\partial \tau} + \beta \frac{\partial}{\partial \alpha}\right)^2 \Phi(\tau_o, \alpha) = \beta.$$

Then the natural algorithm is once more a fixed point one, but with the two variables α and β.

$$\alpha_{i+1} = \frac{\partial \Phi(\tau_o, \alpha_i)}{\partial \tau} + \beta_i \frac{\partial \Phi(\tau_o, \alpha_i)}{\partial \alpha}$$
$$\beta_{i+1} = \frac{\partial}{\partial \tau}\left(\frac{\partial \Phi(\tau_o, \alpha_i)}{\partial \tau} + \beta_i \frac{\partial \Phi(\tau_o, \alpha_i)}{\partial \alpha}\right) + \beta_i \frac{\partial}{\partial \alpha}\left(\frac{\partial \Phi(\tau_o, \alpha_i)}{\partial \tau} + \beta_i \frac{\partial \Phi(\tau_o, \alpha_i)}{\partial \alpha}\right)$$

— $\forall \epsilon$ fixed arbitrary small, the convergence criteria are the following:

$$|\alpha_{i+1} - \alpha_i| < \epsilon$$
$$|\beta_{i+1} - \beta_i| < \epsilon.$$

— One proceeds to time $\tau + d\tau$ by the relations

$$\alpha_o(\tau + d\tau) = \alpha_\infty(\tau) + \beta_\infty(\tau)\, d\tau$$
$$\beta_o(\tau + d\tau) = \beta_\infty(\tau).$$

This expression of β_{i+1} is numerically easy to compute, since it can be viewed as the difference of two α_{i+1} values computed at two different frequencies (the difference of which is β_i) and two consecutive samples. The two α_{i+1} values are obtained by the derivation of the phase of the transform between two consecutive samples and different frequencies (the difference of which is also β_i). Using the same arguments as in (1), by assuming $\beta = \Phi_s''(\tau_o)$, one can proove the convergence of the algorithm giving α. The convergence of the algorithm which gives β has been only checked numerically.

3) Example: The following example demonstrates the use of the crossed criterion with a matched window for the extraction of the amplitude and frequency modulation laws associated to a chirped gaussian wave packet. In order to show the improvement obtained, it is compared to the stationary phase criterion.

The analysis has been performed over 600 samples; the sampling rate is 32 kHz. The theoretical frequency modulation law of the signal is linear, going from 0 Hz to 16 kHz over the duration of the analysis.

Fig. 6 represents the signal, the modulus and the ridge of the Gabor transform computed by the criterion $\frac{\partial \Phi(\tau, \alpha)}{\partial \tau} = \alpha$. The theoretical frequency modulation law is also represented in dotted line. One can see that in this case, the estimation is biased.

Fig. 7 represents the algorithmic estimation of the ridge zoomed between the samples 200 and 400. Here, we used the criterion $\frac{\partial \Phi(\tau, \alpha)}{\partial \tau} + \Phi_s''(\tau) \frac{(\partial \Phi(\tau, \alpha)}{\partial \alpha)} = \alpha$, and the automatic matching of the window to the slope of the frequency modulation law. In this case, the estimation is exact.

B. Frequency-Asymptotic Signals

In the framework of time-asymptotic signals, we have obtained approximations of the Gabor transform of frequency modulated signal, which were concentrated in the frequency domain. Now we shall study the case of signals

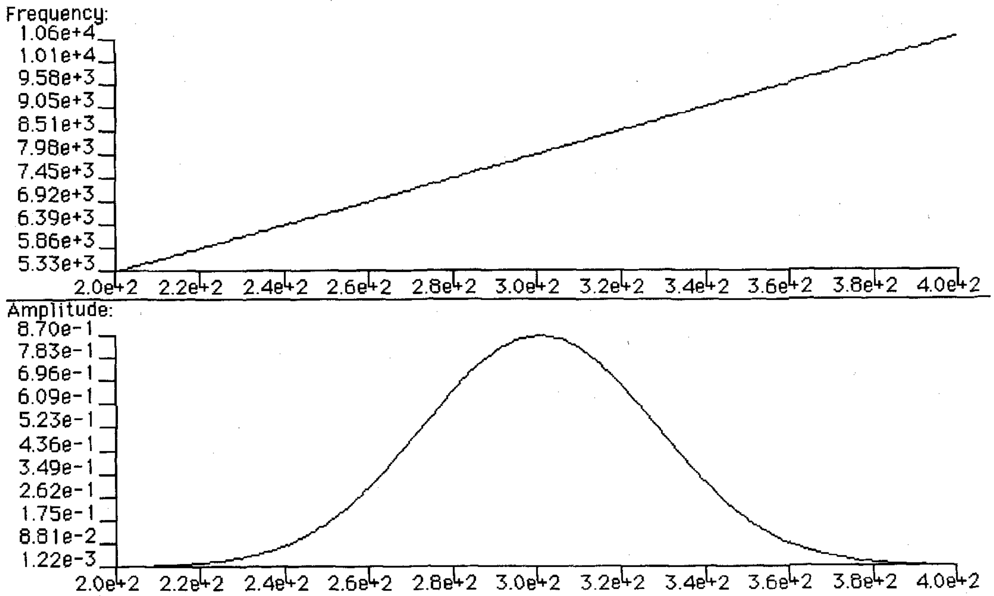


Fig. 7. Frequency and amplitude modulation laws estimated by the use of the crossed criterion.

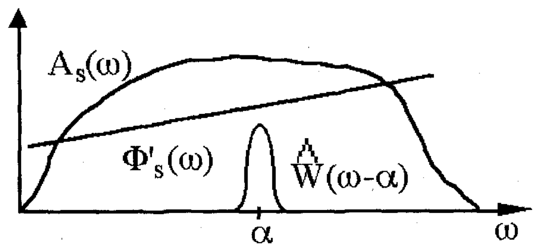


Fig. 8. Frequency-asymptotic hypothesis made on the signal and the window.

localized in the time domain, being then asymptotic in the frequency domain. We shall assume the modulus of the Fourier transform to be only locally linear, but as it has been done in the time-asymptotic case, we could have assumed it to be locally polynomial. Actually, even though this case does not correspond to the situations usually encountered in music, the importance of transients in acoustics (i.e., sonars and nondestructive evaluation) motivated the presentation of these results.

1) Local Chirp-Based Approximation: Similar to the time-asymptotic case, we shall present an approximation of the Gabor transform of a frequency-asymptotic signal. The assumptions made on the signal let us assume that the modulus and the phase of the Fourier transform of the signal are smooth enough around the frequency $\omega = \alpha$ to write

$$A_s(\omega) = A_s(\alpha) + (\omega - \alpha)A'_s(\alpha)$$
$$\Phi_s(\omega) = \Phi_s(\alpha) + (\omega - \alpha)\Phi'_s(\alpha) + \tfrac{1}{2}(\omega - \alpha)^2\Phi''_s(\alpha).$$

We require the analyzing function to be a gaussian: $W(t) = \exp(\frac{-t^2}{2\sigma^2})$, the Fourier transform of which is $\hat{W}(\omega) = \sigma\sqrt{2\pi}\exp(\frac{-\sigma^2\omega^2}{2})$. These assumptions on the signal and the window are shown in Fig. 8.

Under these requirements, the Gabor transform reduces to a simple gaussian integral, and is expressed by [6]

$$L_g(\tau,\alpha) \simeq \sigma \exp\left(i(\Phi_s(\alpha) + \alpha\tau)\right)\sqrt{\frac{1}{\sigma^2 - i\Phi''_s(\alpha)}} \cdot \exp\left(-\frac{1}{2}\,\frac{(\tau + \Phi'_s(\alpha))^2}{\sigma^2 - i\Phi''_s(\alpha)}\right) \cdot \left(A_s(\alpha) + iA'_s(\alpha)\frac{\tau + \Phi'_s(\alpha)}{\sigma^2 - i\Phi''_s(\alpha)}\right).$$

Since the signal is asymptotic, we can assume that the transform is mainly governed by the term

$$A_s(\alpha)\exp\left(-\frac{1}{2}\,\frac{(\tau + \Phi'_s(\alpha))^2}{\sigma^2 - i\Phi''_s(\alpha)}\right) = A_s(\alpha)W(\tau + \Phi'_s(\alpha))^{\frac{\sigma^2}{\sigma^2 - i\Phi''_s(\alpha)}}.$$

The behavior of the transform can then be interpreted as a superposition of windows $W(t)$ located along the curve $\Phi'_s(\alpha) = -\tau$ which describes the group delay of the signal in the $\{\tau, \alpha\}$ plane. (See Fig. 9.)

As it has been done in the case of a time-asymptotic signal, we shall present criteria involving only the transform which permit the estimation of the set of points verifying $\Phi'_s(\alpha) = -\tau$ that carry the group delay information contained in the signal. We shall consider two different cases, that parallel the time-asymtpotic cases: the locally constant and the locally linear spectral density.

a) Constant spectral density: In that case, the transform is given by

$$L_g(\tau,\alpha) = A_s(\alpha)\sigma\exp(i(\Phi_s(\alpha) + \alpha\tau)) \cdot \sqrt{\frac{1}{\sigma^2 - i\Phi''_s(\alpha)}}\exp\left(-\frac{1}{2}\,\frac{(\tau + \Phi'_s(\alpha))^2}{\sigma^2 - i\Phi''_s(\alpha)}\right)$$

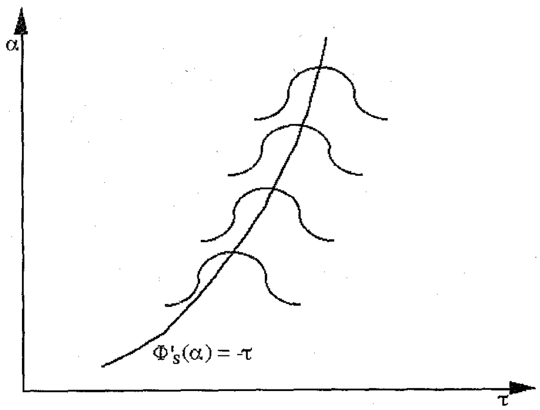


Fig. 9. Behavior of the Gabor transform of a frequency-asymptotic signal as a superposition of $\hat{W}$.

and it is easy to check that the set of points $\Phi_s'(\alpha) = -\tau$ verifies

$$\frac{\partial\Phi(\tau,\alpha)}{\partial\alpha} = 0 \quad \text{and} \quad \frac{\partial\Phi(\tau,\alpha)}{\partial\tau} = \alpha.$$

This result is the same as the one obtained in the case of a constant amplitude time-asymptotic signal. This is not surprising, since the approximations made describe the same signal. Actually, the Fourier transform of a linear chirp with constant amplitude is a linear chirp with a constant spectral density and a group delay being the reciprocal of the instantaneous frequency. Though these two criterias are equivalent, we shall only consider the first one as significant, since it involves quantities the dimensions of which are time, which is the quantity to estimate.

b) Linear spectral density: In this case, one can show that the set of points $\Phi_s'(\alpha) = -\tau$ satisfies

$$\frac{\partial\Phi(\tau,\alpha)}{\partial\tau} - \frac{1}{\Phi_s''(\alpha)}\frac{\partial\Phi(\tau,\alpha)}{\partial\alpha} = \alpha.$$

As for the linear amplitude time-asymptotic signal, we find that the estimation of the group delay law requires a combination of "horizontal" and "vertical" derivatives of the phase of the transform. This crossed criterion seems different from the one obtained in the framework of time asymptotic signals: $\frac{\partial\Phi(\tau,\alpha)}{\partial\tau} + \Phi_s''(\tau)(\frac{\partial\Phi(\tau,\alpha)}{\partial\alpha}) = \alpha$, but they can be related. Indeed, if one considers the signal $s(t) = \exp(i(\omega_1 t + \beta(\frac{t^2}{2})))$, the second derivative of the phase is β, and the second derivative of the phase of its Fourier transform is $-(\frac{1}{\beta})$. In this case, the two combined criteria become identical and lead to the same trajectory in the half-plane of the Gabor transform. Nevertheless, if the amplitude modulation law or if the spectral density are not constant, the curves obtained are different. This is consistent with the fact that under this condition, the group delay and the instantaneous frequency are not reciprocal functions.

2) Algorithmic Estimation of the Group Delay: Similar to the time-asymptotic case, we shall focus on the algorithmic estimation of the group delay. For that purpose, we shall limit the discussion to the constant spectral density approximation. In this case, the problem reads:

For a given frequency parameter α_o, find τ such that $\frac{\partial\Phi(\tau,\alpha_o)}{\partial\alpha} = 0$.

Contrary to the time-asymptotic case, this is not a fixed-point equation, but it is actually easy to obtain such an equation. Let us consider the Gabor transform of a delta function located at time t_o, which is the simplest function satisfying our assumptions, since its spectral density and group delay are constant

$$L_g(\tau,\alpha) = \overline{W}(t_o - \tau)e^{i\alpha(\tau - t_o)}.$$

Then for any frequency parameter α_o, the time t_o is given by

$$t_o = \tau - \frac{\partial\Phi(\tau,\alpha_o)}{\partial\alpha}.$$

Consequently, $\frac{\partial\Phi(\tau,\alpha_o)}{\partial\alpha=0}$ gain to be replaced by $\tau = \tau - \frac{\partial\Phi(\tau,\alpha_o)}{\partial\alpha}$, which is a fixed-point equation of the type $F(\tau) = \tau$, with $F(\tau) = \tau - \frac{\partial\Phi(\tau,\alpha_o)}{\partial\alpha}$.

*Fixed-point algorithm:

This last equation can be solved by the process:

Let $\tau_o(\alpha_o)$ be an "arbitrary" initial value.
$\tau_{i+1}(\alpha_o) = \tau_i(\alpha_o) - \frac{\partial\Phi(\tau_i(\alpha_o),\alpha_0)}{\partial\alpha}$.
The convergence criteria is $|\tau_{i+1} - \tau_i| < \epsilon$, ϵ fixed arbitrary small.
One proceeds to frequency $\alpha_o + d\alpha$ by the relation: $\tau_o(\alpha_o + d\alpha) = \tau_\infty(\alpha_o)$.

Proof of the Convergence: In order to insure that this algorithm converges toward $\tau_\infty(\alpha_o)$ verifying: $\tau_\infty(\alpha_o) = F(\tau_\infty(\alpha_o))$, one has to check that the function $F(\tau)$ is a contraction on a time interval $]\tau_-, \tau_+[$ that contains both the initial value τ_o and the final value τ_∞, that is, for any two consecutive iterations, i and $i+1$: $|F(\tau_{i+1}) - F(\tau_i)| < |\tau_{i+1} - \tau_i|$. Thanks to the regularity of the Gabor transform, the function F is differenciable, and the previous condition then reads $|\frac{\partial F(\tau)}{\partial\tau}| < 1$.

Under the constant spectral density approximation, the derivative of $F(\tau)$ with respect to τ is given by

$$\frac{\partial F(\tau)}{\partial\tau} = \frac{\Phi_s''(\alpha_o)^2}{\sigma^4 + \Phi_s''(\alpha_o)^2}.$$

The absolute value of this expression is obviously always smaller than one. Since $\frac{\partial F(\tau)}{\partial\tau}$ is independent of τ, F is then a contraction on R.

From a strictly mathematical point of view, though this algorithm provides the set of points satisfying $\frac{\partial\Phi(\tau,\alpha_o)}{\partial\alpha} = 0$ with any desired accuracy, one has to remember that the trajectory $\tau(\alpha)$ is representative of the group delay of the signal if it has a constant spectral density and a linear group delay.

3) Example: Figs. 10 and 11 correpond to the Gabor analysis of a signal, the group delay of which is parabolic, and the spectral density of which is gaussian. This is a typical example in which the group delay and the instantaneous frequency are not reciprocal functions. Fig. 10 represents the signal, the modulus, and the ridge of the

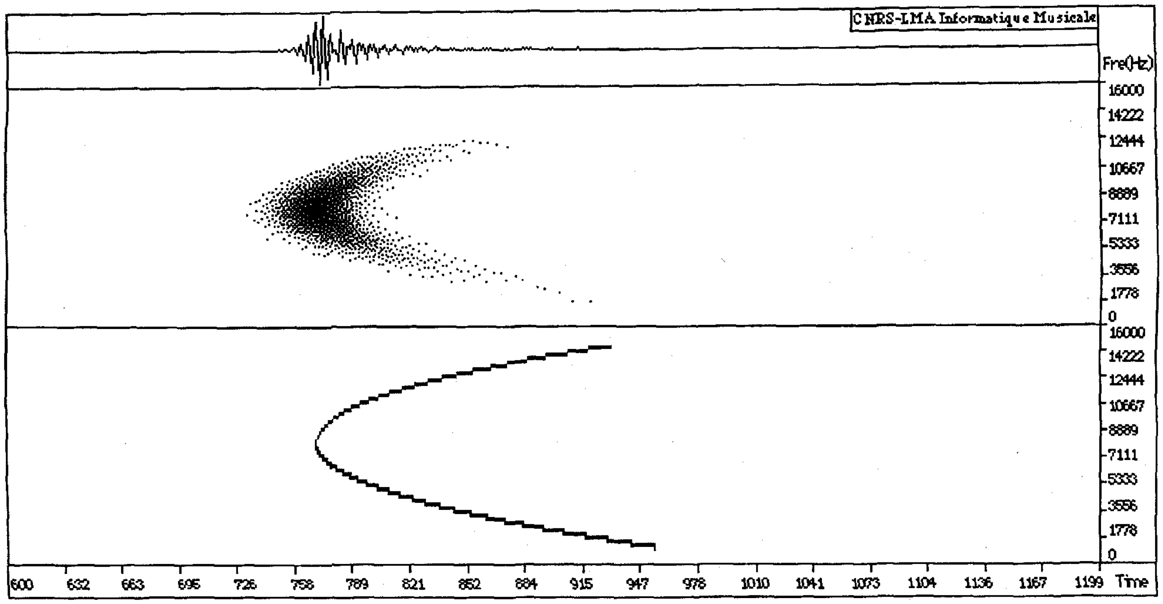


Fig. 10. Signal, modulus, and ridge of the Gabor transform of a signal with an hyperbolic group delay.

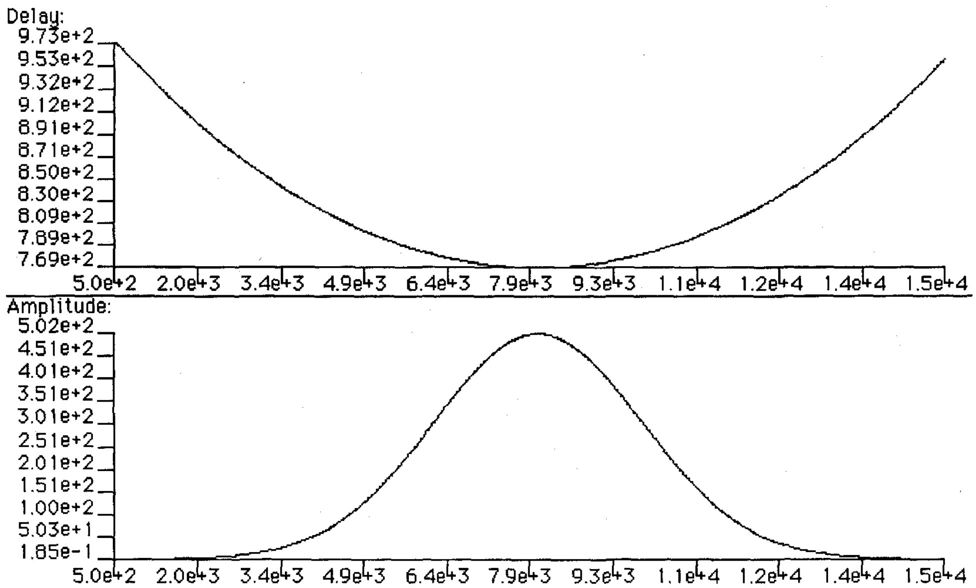


Fig. 11. Group delay and spectral density estimated by the use of the algorithm.

Gabor transform. Fig. 11 shows the estimated group delay and spectral density obtained by the algorithm.

C. Wavelet Transform of Complex Homogenous Signals

As it has been pointed out in the two previous sections that the Gabor transform is well adapted to signals whose instantaneous frequency or group delay are linear. The purpose of this section is to describe the behavior of the wavelet transform of signals being eigen functions of the dilation operator, namely the homogenous functions [9]. These functions can modelize local discontinuity of derivatives appearing in a signal, as well as hyperbolic frequency modulation laws. Actually, let us consider the signal

$$s(t) = 0 \qquad t \leq 0$$
$$s(t) = t^{\nu} \qquad t > 0$$

the νth derivative of which is discontinuous.

This signal is homogenous of order ν, since it obviously satisfies

$$\forall \lambda \in R^{+*}, \quad s(\lambda t) = \lambda^{\nu} s(t).$$

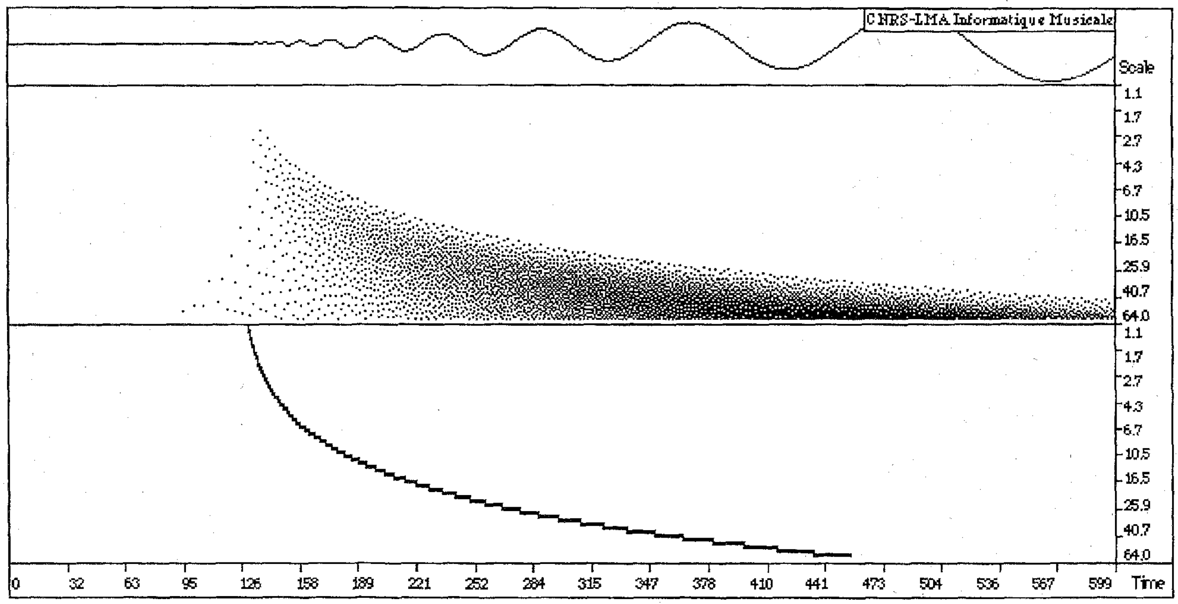


Fig. 12. Signal, modulus, and ridge of the wavelet transform of an homogenous signal.

Moreover, the Fourier transform of any homogenous function of order ν is any homogenous function of order $-\nu - 1$

$$\forall \lambda \in R^{+*}, \quad \hat{s}(\lambda\omega) = \lambda^{-\nu-1}\hat{s}(\omega).$$

As a consequence, the signal of the Fourier transform

$$\hat{s}(\omega) = K\omega^{-\nu-1}$$

is an homogenous function of ω, of order $-\nu-1$. Actually, it corresponds to the Fourier transform of $s(t)$ as it has been defined, with $K = i^{-\nu-1}\Gamma(1+\nu)$.

As we pointed out, these functions modelize hyperbolically frequency modulated signals. Indeed, splitting ν into its real and imaginary parts, respectively, ν_r and ν_i yields

$$s(t) = t^{\nu_r} e^{i\nu_i ln(t)}$$

the instantaneous frequency of which is $\frac{\nu_i}{t}$. Applying the same calculus in the Fourier domain yields a group delay which is $\frac{\nu_i}{\omega}$, being consequently the reciprocal function of the instantaneous frequency.

Let us assume that the wavelet $g(t)$ is progressive and has at least $\nu+1$ vanishing moment. The wavelet transform of an homogenous function located at $t = t_o$ is then given by

$$T(b,a) = Ka^{\nu} i^{-\nu-1}\overline{g}^{(-\nu-1)}\left(\frac{t_o - b}{a}\right)$$

where $\overline{g}^{(-\nu-1)}$ denotes the fractional derivative of order $-\nu-1$ of $\overline{g}$, defined by $f^{(\alpha)}(x) = \frac{1}{2\pi}\int (i\omega)^{\alpha}\hat{f}(\omega)e^{i\omega x}\, d\omega$.

Let us describe the behavior of the wavelet transform of such functions [7]. Along the lines in the half-plane of the wavelet transform defined by $\frac{t_o - b}{a} = \text{Const}$, the transform behaves like a^{ν}. This means that the logarithm of the modulus of the transform as a function of the logarithm of the scale is a linear function of slope ν_r, while the phase as a function of the logarithm of the scale is a linear function of slope ν_i. As previously, we shall focus on the estimation of the instantaneous frequency by the use of the phase of the wavelet transform. By derivation of the wavelet transform, one can show that the phase satisfies the relation

$$\frac{\partial\Phi(b,a)}{\partial a} = \frac{\nu_i}{a} + \frac{(t_o - b)}{a}\frac{\partial\Phi(b,a)}{\partial b}.$$

As we pointed out, the frequency modulation law of the signal is $\Phi_s'(t) = \frac{\nu_i}{t - t_o}$. Then, in the half-plane of the wavelet transform, the trajectory which represents this frequency modulation (ridge) is $\Phi_s'(b) = \frac{\nu_i}{b - t_o} = \frac{\omega_o}{a}$, where ω_o represents a reference frequency.

Since $\frac{\partial\Phi(b,a)}{\partial b}$ has the dimensions of frequency, let us search for the set of points satisfying: $\frac{\partial\Phi(b,a)}{\partial b} = \Phi_s'(b) = \frac{\nu_i}{b - t_o}$.

The relations between the partial derivatives of the phase show that the set of points such that $\frac{\nu_i}{b - t_o} = \frac{\omega_o}{a}$ corresponds to the points (b,a) of the transform where the three following equivalent relations hold:

$$\frac{\partial\Phi(b,a)}{\partial a} = 0$$
$$\frac{\partial\Phi(b,a)}{\partial b} = \frac{\omega_o}{a}$$
$$\frac{\partial\Phi(b,a)}{\partial b} + \frac{\omega_o}{\nu_i}\frac{\partial\Phi(b,a)}{\partial a} = \frac{\omega_o}{a}.$$

The first two relations parallel the results obtained in the framework of time or frequency asymptotic signal with a locally constant amplitude or a locally constant spectral density. It is also interesting to compare the third relation to the crossed criterion exhibited in the framework of frequency asymptotic signals. For that purpose, let us introduce a function Ψ such that $\Psi(b, \frac{\omega_o}{a}) = \Phi(b,a)$. Then

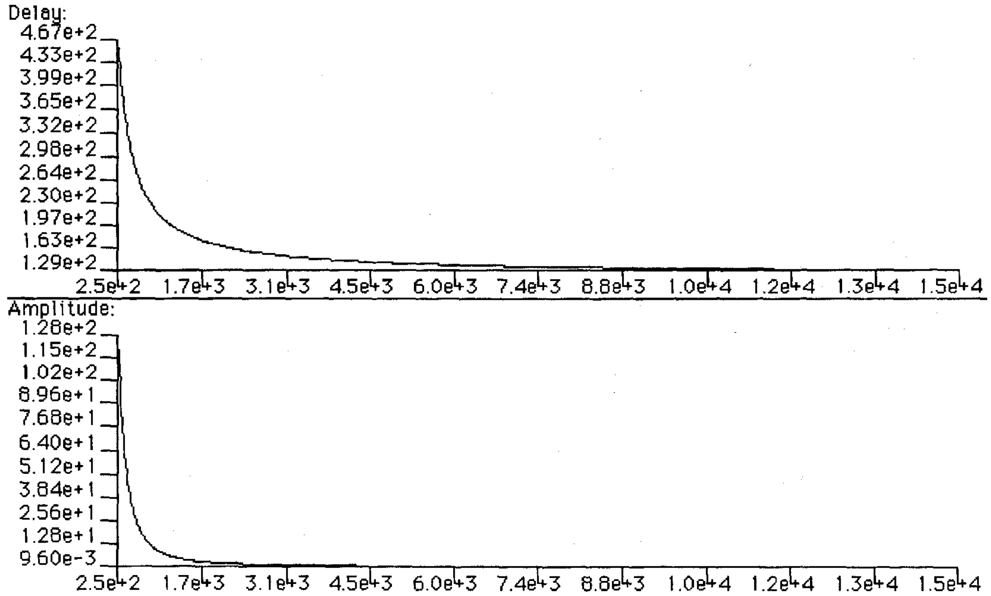


Fig. 13. Group delay and spectral density estimated by the use of the algorithm.

the third equation can be rewritten as

$$\frac{\partial\Psi\left(b,\frac{\omega_o}{a}\right)}{\partial b} - \frac{\left(\frac{\omega_o}{a}\right)^2}{\nu_i}\,\frac{\partial\Psi\left(b,\frac{\omega_o}{a}\right)}{\partial\left(\frac{\omega_o}{a}\right)} = \frac{\omega_o}{a}.$$

By remembering that the signal has hyperbolic group delay, the second derivative of the phase of its Fourier transform is $\Phi_s''(\omega) = \frac{\nu_i}{\omega^2}$ and the third equation can be written as

$$\frac{\partial\Psi\left(b,\frac{\omega_o}{a}\right)}{\partial b} - \frac{1}{\Phi_s''\left(\frac{\omega_o}{a}\right)}\,\frac{\partial\Psi\left(b,\frac{\omega_o}{a}\right)}{\partial\left(\frac{\omega_o}{a}\right)} = \frac{\omega_o}{a}.$$

This equation corresponds exactly to the crossed criterion previously obtained in the Gabor case, the frequency α being replaced by the frequency $\frac{\omega_o}{a}$.

We shall now present an algorithm which enables the estimation of the group delay law, by solving the equation: $\frac{\partial\Phi(b,a)}{\partial a} = 0$.

Let us first assume that the wavelet $g(t)$ is of the form: $g(t) = A_g(t)\exp(i\phi_g(t))$. As it has been done in the case of a frequency-asymptotic signal, in order to exhibit a fixed point equation, we shall consider the wavelet transform of a delta function located at $t = t_o$

$$T(b,a) = \frac{1}{a}A_g\left(\frac{t_o-b}{a}\right)e^{-i\phi_g\left(\frac{t_o-b}{a}\right)}.$$

Then the point t_o is given by the relation

$$t_o = b + \frac{a^2}{\phi_g'\left(\frac{t_o-b}{a}\right)}\,\frac{\partial\Phi(b,a)}{\partial a}.$$

This implies that the natural fixed-point algorithm is given by

$$b_{n+1} = b_n + \frac{a^2}{\phi_g'(0)}\,\frac{\partial\Phi(b_n,a)}{\partial a}.$$

Using the same arguments as in the Gabor case, it is easy to show that, at least for a delta function, and a wavelet with a fixed frequency, this algorithm converges. For more general cases, the convergence has been checked numerically with a fixed frequency wavelet.

1) Example: Figs. 12 and 13 correspond to the wavelet transform of an homogenous signal of complex degree. Fig. 12 represents the signal, the modulus, and the ridge of the wavelet transform. Fig. 13 shows the exact estimation of the group delay and the spectral density obtained by the algorithm.

IV. Representations of Multicomponents Time-Asymptotic Signals

Generally, the transforms of signals with several elementary components do not behave like a superposition of the transforms corresponding to each component. This is due to the fact that the interesting quantities, such as the modulus (related to an energy density) and the phase (related to the oscillating features of the signal) are nonlinear functions. The purpose of this section is to present an algorithm using the linearity of the Gabor transform and the approximations presented in the previous section, which enable a true separation of the components. We shall first consider the case of components whose frequencies are constant and generalize the approach to frequency-modulated components.

A. Fixed-Frequency Components: Spectral Lines

The additive synthesis model, used in the context of the resynthesis of a real musical sound, requires a very accurate estimation of the amplitude modulation laws of the components. Our own experience has proved that the perceptual difference between a perfect resynthesis, and a good resynthesis relies on whether the microvariations and abrupt changes are taken into account or not.

It may seem simple to estimate these quick variations, but in fact it is not possible to find accurately the amplitude modulation law associated with each component just by considering the horizontal restrictions of the transform, since a time-frequency analysis cannot be precise in both the time and the frequency domains. If one wishes to be accurate in time, one can be tempted to choose a short-time analyzing window. In this case the Fourier transform of such a window is spread out and several components will be merged, leading to strong oscillations in the modulus of the transform. Nevertheless, in order to avoid such a problem, it is possible to choose a long-time analyzing window. In this case, on the time-frequency representation, the components are well separated but the amplitude modulation law associated with each one is strongly smoothed (see Figs. 1–3 in Section II). In fact, merging of components is better than the smoothing of the amplitudes and the aim of the algorithm we shall present is the estimation of these laws, whatever the window one uses.

1) Algorithm: As we shall see, this algorithm requires knowledge of the number of components, and their frequencies. The estimation of the frequency of the spectral lines, by the use of averaged versions of the local frequency of the transforms, have been described in [10].

In the case of spectral lines with constant amplitudes, the signal is written as

$$s(t) = \sum_{k=1}^{N} A_k \exp(i\omega_k t).$$

The Gabor transform of $s(t)$ is given by

$$\begin{aligned} L_g(\tau, \alpha) &= \sum_{k=1}^{N} \frac{A_k}{2}(\hat{W}(\omega_k - \alpha)\exp(i\omega_k\tau) \\ &\quad + \hat{W}(-\omega_k - \alpha)\exp(-i\omega_k\tau)) \\ &= \sum_{k=1}^{N} \frac{A_k}{2}(\hat{W}(\omega_k - \alpha) + \hat{W}(-\omega_k - \alpha)) \\ &\quad \cdot \cos(\omega_k\tau) + i\sum_{k=1}^{N}\frac{A_k}{2} \\ &\quad \cdot (\hat{W}(\omega_k - \alpha) - \hat{W}(-\omega_k - \alpha))\sin(\omega_k\tau). \end{aligned}$$

Let us consider the N restrictions of $L_g(\tau, \alpha)$ for $\alpha_p = \omega_p, 1 \le p \le N$.

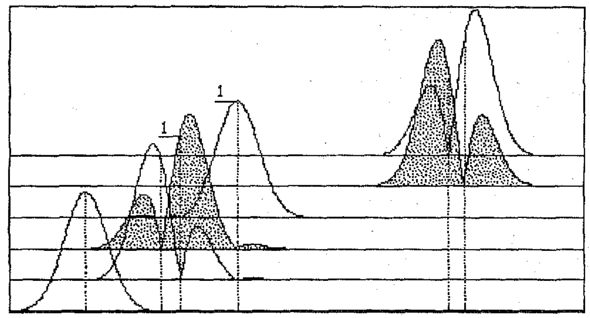


Fig. 14. Filters for a constant amplitude estimation.

Then $\forall p$ one can write

$$\begin{aligned} &\sum_{k=1}^{N} \frac{A_k}{2}(\hat{W}(\omega_k - \alpha_p) + \hat{W}(-\omega_k - \alpha_p))\cos(\omega_k\tau) \\ &\quad = \mathrm{Re}\{\mathrm{L}_g(\tau, \alpha_p)\} \\ &\sum_{k=1}^{N} \frac{A_k}{2}(\hat{W}(\omega_k - \alpha_p) - \hat{W}(-\omega_k - \alpha_p))\sin(\omega_k\tau) \\ &\quad = \mathrm{Im}\{\mathrm{L}_g(\tau, \alpha_p)\}. \end{aligned}$$

These formulas correspond to two linear systems of N equations with N unknowns whose form is

$$\sum_{k=1}^{N} W_{pk}X_k(\tau) = L_p(\tau) \quad \forall p, \qquad 1 \le p \le N.$$

The elements of the first matrix are given by ${}^1W_{pk} = \frac{1}{2}(\hat{W}(\omega_k - \alpha_p) + \hat{W}(-\omega_k - \alpha_p))$.

The elements of the second matrix are given by ${}^2W_{pk} = \frac{1}{2}(\hat{W}(\omega_k - \alpha_p) - \hat{W}(-\omega_k - \alpha_p))$.

Then at any time τ, the solution vector $\boldsymbol{X}(\tau)$ is given by $\boldsymbol{X}(\tau) = {}^1W^{-1}\ \mathrm{Re}\{\boldsymbol{L_g}(\tau)\}\ + i\ {}^2W^{-1}\ \mathrm{Im}\{\boldsymbol{L_g}(\tau)\}$, and the modulus of $X_k(\tau)$ is A_k.

Question: beyond the calculus, what have we really done?

Consider a given vector $\boldsymbol{\Gamma}$ of $\{\alpha_p, 1 \le p \le N\}$ parameters. With this vector and a given function $W(t)$, one has built two matricial operators: ${}^1\boldsymbol{\Omega} = {}^1W^{-1}$ and ${}^2\boldsymbol{\Omega} = {}^2W^{-1}$, acting on the restrictions of the transform at any time τ for the parameters α_p. Since these operators and the Gabor transform are linear, their action on the restrictions of the transform can be directly transposed on the Gabor functions themselves, and one can write

$$\begin{aligned} \boldsymbol{X}(\tau) &= {}^1\boldsymbol{\Omega}\ \mathrm{Re}\{\boldsymbol{L_g}(\tau)\} + i\ {}^2\boldsymbol{\Omega}\ \mathrm{Im}\{\boldsymbol{L_g}(\tau)\} \\ &= \tfrac{1}{2}({}^1\boldsymbol{\Omega}(\boldsymbol{L_g}(\tau) + \overline{\boldsymbol{L}}_{\boldsymbol{g}}(\tau)) + i\ {}^2\boldsymbol{\Omega}(\boldsymbol{L_g}(\tau) - \overline{\boldsymbol{L}}_{\boldsymbol{g}}(\tau))) \\ &= \tfrac{1}{2}\int ({}^1\boldsymbol{\Omega}(\hat{W}(\omega - \boldsymbol{\Gamma}) + \hat{W}(\omega + \boldsymbol{\Gamma})) \\ &\quad + {}^2\boldsymbol{\Omega}(\hat{W}(\omega - \boldsymbol{\Gamma}) - \hat{W}(\omega + \boldsymbol{\Gamma})))\hat{s}(\omega)e^{i\omega\tau}\, d\omega. \end{aligned}$$

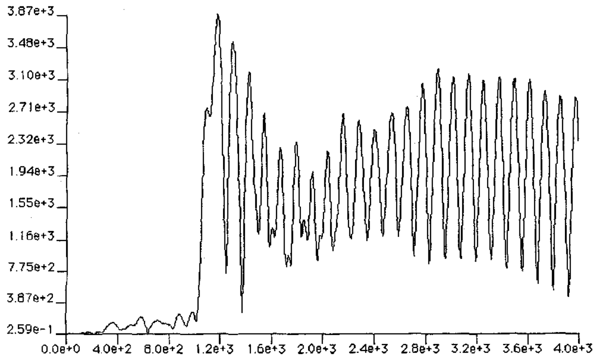


Fig. 15. Modulus of the restriction of the Gabor transform for $\alpha = 5\omega_{\text{fond}}$ computed with a window spread out in frequency. Many harmonics are merged and the amplitude shows significant oscillations.

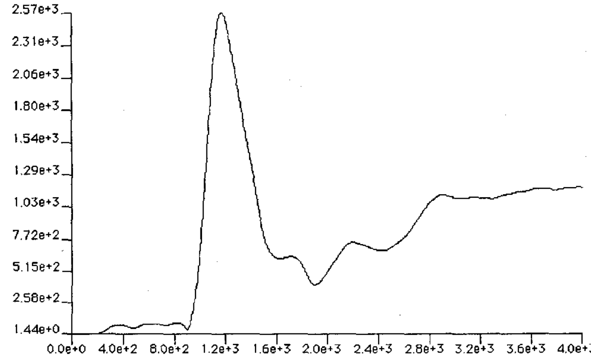


Fig. 16. Modulus of the restriction of the Gabor transform for $\alpha = 5\omega_{\text{fond}}$ computed with a window narrow in frequency, the localization of which has been matched to resolve only one component (the time window has to be three times longer than in Fig. 15). In this case, the amplitude is heavily smoothed.

Then, finally

$$\begin{aligned} X_k(\tau) &= \tfrac{1}{2} \int \Bigg(\sum_{p=1}^{N} {}^1W_{kp}^{-1} (\hat{W}(\omega - \alpha_p) + \hat{W}(\omega + \alpha_p)) \\ &\quad + {}^2W_{kp}^{-1} (\hat{W}(\omega - \alpha_p) - \hat{W}(\omega + \alpha_p)) \Bigg) \\ &\quad \cdot \hat{s}(\omega) e^{i\omega\tau} \, d\omega \\ &= \int \hat{F}_k(\omega) \hat{s}(\omega) e^{i\omega\tau} \, d\omega. \end{aligned}$$

The vector $\hat{F}(\omega)$ is a filter bank composed by N elements. By construction, each element $\hat{F}_k(\omega)$ has the following properties: $\hat{F}_k(\alpha_p) = \delta_{pk}, \hat{F}_k(-\alpha_p) = 0, 1 \leq p \leq N, 1 \leq k \leq N$. This means the following:

1) An elementary filter $F_k(t)$ is orthogonal to any function $\cos(\omega_p t + \phi)$ with $\omega_p = \alpha_p$ and p_k. Moreover, its time support equals the time support of the function $W(t)$.
2) For any k_o, the convolution between $F_{k_o}(t)$ and $s(t) = \Sigma_{k=1}^{N} A_k \cos(\omega_k t + \phi_k)$ is $A_{k_o} \exp(i(\omega_{k_o} t + \phi_{k_o}))$.

The Fourier transform of such a family of filters is represented in Fig. 14. The horizontal axis is the frequency. The input signal is a sum of six sine waves. The dotted vertical lines indicate the position of each frequency in the spectrum. The six filters are displayed vertically stacked, two of them being shaded gray. The Fourier transform of a filter equals one for its corresponding spectral line, and equals zero for the five other frequencies.

The generalization of this algorithm when the amplitudes of the components vary with respect to the time is obvious if, in the Gabor transform expression, one can replace A_k by $A_k(\tau)$. Such an approximation is generally a reasonable one, but in some cases it is not sufficient and one has to use a more accurate algorithm. In order to make easier the reading of this paper, this algorithm is described in the Appendix.

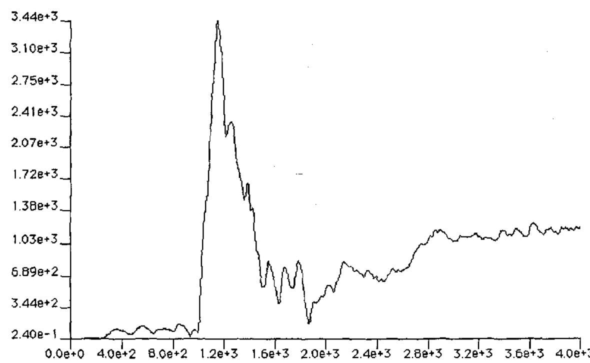


Fig. 17. Result obtained by the algorithm described in the Appendix, by considering locally constant amplitudes. The time support of the window used is the same as in Fig. 15, but the amplitude shows fewer oscillations.

In the same way, it leads to the construction of a filter bank, where each filter enables the extraction of an amplitude modulation law being locally a polynomion of a given order J on the time support of $W(t - \tau)$.

2) Example: In order to illustrate the method, we have represented in Figs. 15–18 the amplitude modulation laws of the 125 first milliseconds of the fifth harmonic ($5\omega_{\text{fond}}$) of a sitar sound, extracted with different techniques. The sampling rate is 32 kHz. The horizontal axis is the time, and displays 4000 samples. The window $W(t)$ used for the computation is a gaussian.

B. Frequency-Modulated Components

Up to now, we have presented several algorithms which allow the estimation of pertinent parameters when the signal is a single asymptotic signal or composed of a sum of spectral lines. In the first case, one of the conditions for the estimation of a ridge corresponding to the frequency modulation law was the asymptotism of the signal with respect to the analyzing function (generally, a sum of asymptotic signals is not asymptotic). In the second case, the algorithmic development became possible through a

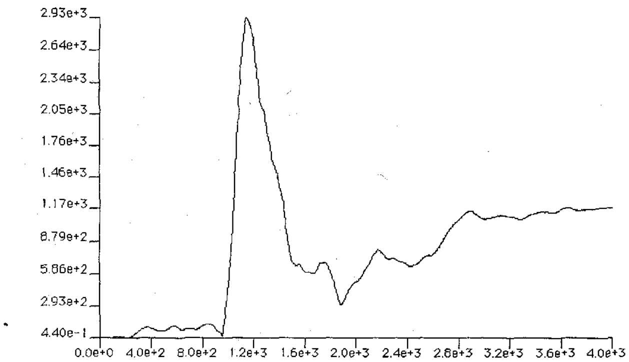


Fig. 18. Result obtained by the algorithm described, by considering locally linear amplitudes. The time support of the window used is the same as in Fig. 15, but the amplitude does not oscillate and is not smoothed.

strong hypothesis on the signal: it is a sum of elementary contributions, and one can associate a supposenly horizontal ridge to each component.

The purpose of this part is the description of a technique combining the algorithms previously presented, allowing the estimation of the frequency modulation law of each component when the signal is a sum of asymptotic signals. We will focus on two different cases: noncrossing ridges and crossing ridges.

For that purpose, we shall consider a signal $s(t)$ composed by N complex-valued frequency modulated components (the requirement of complexity is just for clarity in the following):

$$s(t) = \sum_{k=1}^{N} A_k \exp(i\Phi_k(t)) = \sum_{k=1}^{N} s_k(t).$$

1) Noncrossing Ridges: We shall first assume that the intersection between the theoretical ridges is empty. The aim is to associate a ridge $\alpha_k(\tau)$ to each contribution.

We shall suppose that the ridge estimation algorithm could be applied if the input signal was a single component $s_k(t)$. This means that at any time τ, one can define an interval I_k associated with the transform, inside which the iterative scheme converges. What happens now if one has several components? Two situations have to be considered.

The first one is the simplest: in the half-plane of the transform, one can associate a closed interval I_k to each elementary contribution and the intersection between two different intervals can be set empty, just by choosing a window sufficiently narrow in frequency. In this case, the transform completely separates each contribution, leading to N disjointed ridges, as long as such a long-time window does not merge the components in the time domain.

The second situation to be considered is the overlapping of the intervals, corresponding to arbitrary close theoretical ridges. In this case, even if it is sometimes possible to associate a ridge to each contribution by defining subintervals inside which the algorithm converges, these estimated

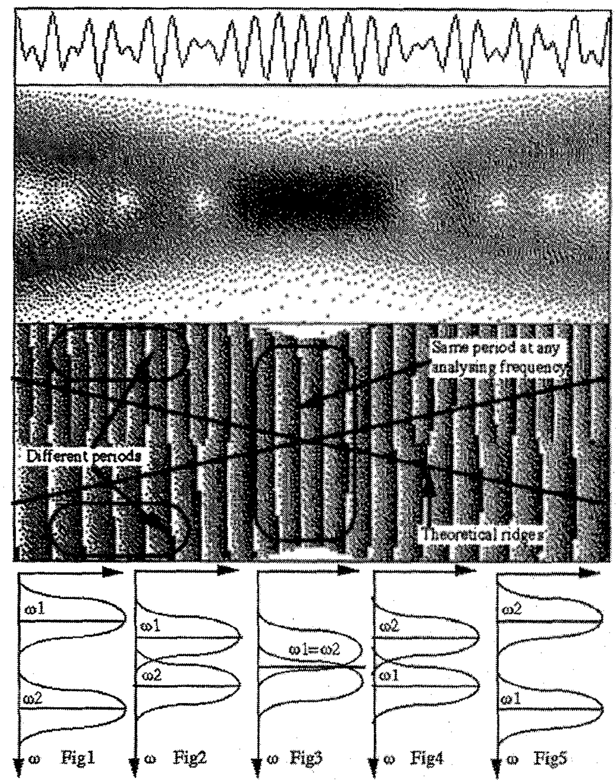


Fig. 19. Signal, modulus, and phase of the Gabor transform of two crossing linear chirps.

ridges will be false, due to the interferences between the contributions. Let us focus on this situation.

The basic idea is the following:

As in the case of the spectral lines, it is possible to define, at any time τ_o, an operator $\boldsymbol{\Omega}(\tau_o)$ acting now on a vector of scalar products $\boldsymbol{L_g}(\tau_o)$ (instead of horizontal restrictions of the transform) between the signal and Gabor functions at the different frequencies $\alpha_k(\tau_o)$. The way to build such an operator, by considering either a locally constant frequency approximation or a locally linear frequency approximation, is detailed in the Appendix.

As in the spectral line case, this algorithm can be interpreted as a way to build an evolutive filter bank. The results presented in the appendix are valid in the case of components with constant amplitude, but they can be easily generalized to the case of any locally polynomial amplitude modulation laws, as it has been done in the spectral line case.

2) Crossing Ridges: We shall first investigate the case of a locally constant frequency approximation, then switch to the linear one.

a) Constant frequency approximation—Cooking method: Let us imagine that two theoretical ridges are crossing. Two problems then arise. First, without taking into account some continuity conditions on the amplitude or on the frequency of each component, it is impossible to say after the crossing point which ridge is the prolongation of a ridge before the crossing point, that is to say, one cannot make the

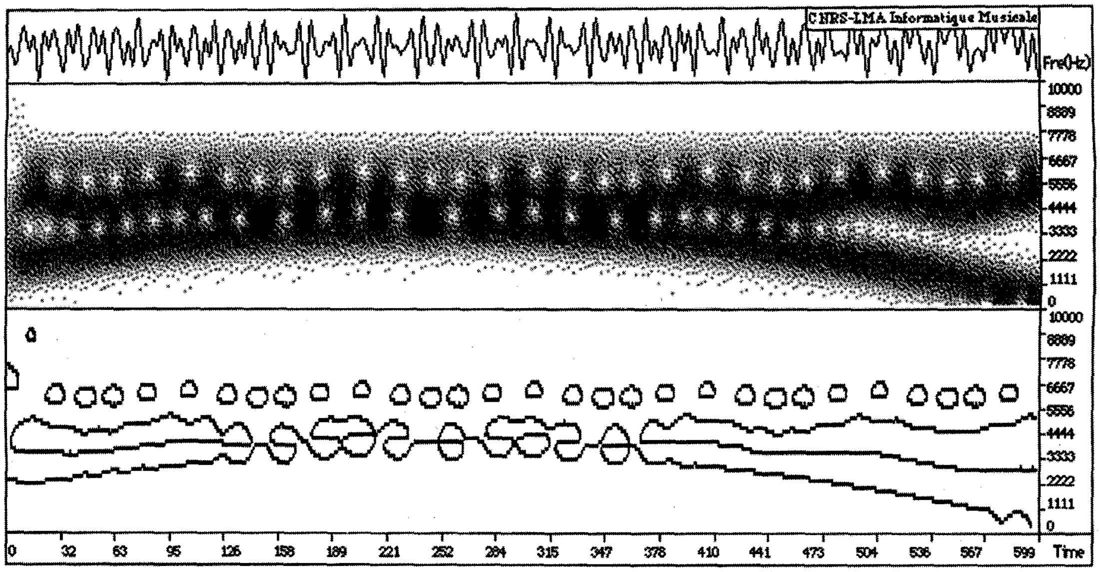


Fig. 20. Modulus and ridge associated to the Gabor transform of a signal composed by a sine wave with amplitude 0.5, a sinusoidal frequency modulated component and a parabolic chirp, both with amplitude one. The ridge only shows bubbles.

difference between two ridges effectively crossing and two ridges getting closer before the crossing point and moving away afterwards. The second problem is more algorithmic. At the crossing point, one can no longer consider that there are two ridges since they are identical. Consequently, the operator Ω is not defined since the corresponding matrix W is not invertible. Nevertheless, it is possible to get around the second problem by considering the phase behavior of the transform. For that purpose, let us decide that the ridges must not cross, even if theoretically they are crossing. It means that in the neighborhood of the crossing point, one ridge is "above" the other one. Let us remember that near the ridge, the local frequency of the transform is very close to its value on the ridge. Moreover, the operator Ω acting on Gabor windows builds a special filter bank. Each filter equals one for a given frequency and zero for the other ones. This filter is as smooth as the functions used for its construction. When the two ridges are drawing nearer, the matrix W starts to be ill-conditionned. The idea, justified by the two previous remarks is the following. One can perturb the linear system to keep an invertible matrix when the distance between the estimated ridges or the determinant of the matrix W becomes smaller than a given constant D. In this case, at any step of the iteration, one will add a small quantity $d\alpha$ to the estimated frequency parameter corresponding to the "upper" ridge, while one will substract $d\alpha$ to the estimated frequency parameter corresponding to the "lower" ridge. One can hope that such a trick will prevent the ridges to cross, while it keeps the matrix invertible and allow the ridges to become very close since the local frequency corresponding to the perturbed frequency is very close to the not perturbed one. Obviously, the problem is the appropriate choice of D and $d\alpha$. It has been solved by numerical tests.

Fig. 19 displays the transform of a signal composed by two crossing linear chirps:

Fig1 and Fig5: The two components are naturally separated by the analysis, and the estimation is done without any problem.
Fig2 and Fig4: The components are getting nearer. Beats appear on the modulus. Nevertheless, the phase clearly shows two different periodicities. In this case, the algorithm enables the estimation of the components without any modifications.
Fig3: Around the crossing point, the signal is locally monochromatic. The phase has the same periodicity at any analyzing frequency. Any linear filtering can then give the correct frequency. One can then freeze the central frequencies of the functions involved in the filter bank, in order to keep the inversibility of the matrix W.

b) Linear frequency approximation: In this case, the problem is different. One assumes continuity conditions by taking into account the local slope of each frequency modulation law in the definition of the matrix W. Then, this matrix is mathematically invertible near and even at the crossing point as far as the two slopes are different. This implies that the two frequency modulation laws must not be tangent. This is the mathematical point of view, but actually, from the numerical one, since the corrections introduced are in fact very small, the effective crossing of two components can be obtained only by the use of a window $W(t)$, the time support of which is large enough. From one hand, this condition is natural, since in this case, the window "sees" a large portion of only one component, and the corrections introduced by taking the slope into account become significant. From the other hand, the choice of such a window can lead to numerical

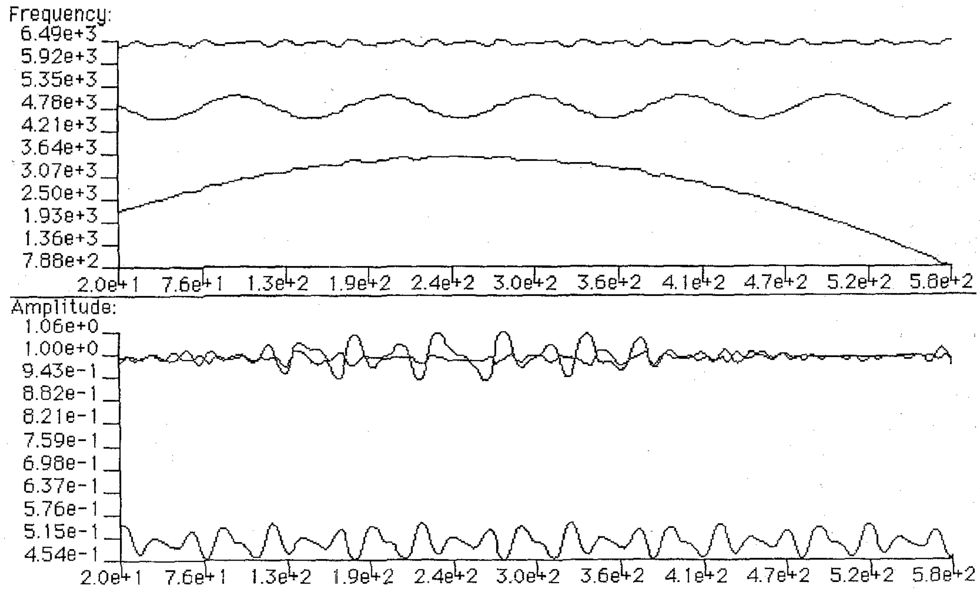


Fig. 21. Frequency and amplitude modulation laws of each component obtained with the algorithm under the constant frequency approximation.

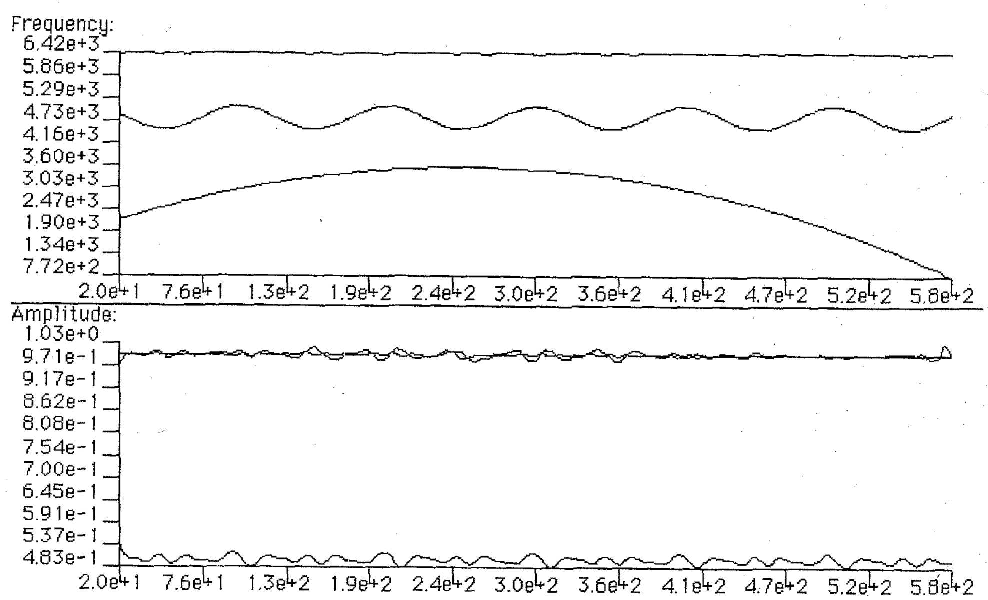


Fig. 22. Frequency and amplitude modulation laws of each component obtained with the algorithm under the linear frequency approximation. The error of the estimation is smaller than in Fig. 21.

problems, especially aliasing ones. When one uses a short-time window, the problem seems to be located in the iterative scheme that solves for the slopes, since one has checked, in the upper example of two crossing linear chirps, that once the slope parameters are fixed to their exact value, the iterative scheme converges toward the theoretical values of α, and the chirps really cross.

3) Examples: The following pictures illustrate the techniques described above. For Figs. 20 and 23, the graphical conventions are:

1) The sampling rate is 32 kHz.
2) The horizontal axis is the time, expressed in samples.
3) The vertical axis is the frequency, expressed in hertz.

Figs. 20 and 23 are divided into three parts:

1) The upper part is the signal.
2) The middle part is the modulus of the transform,

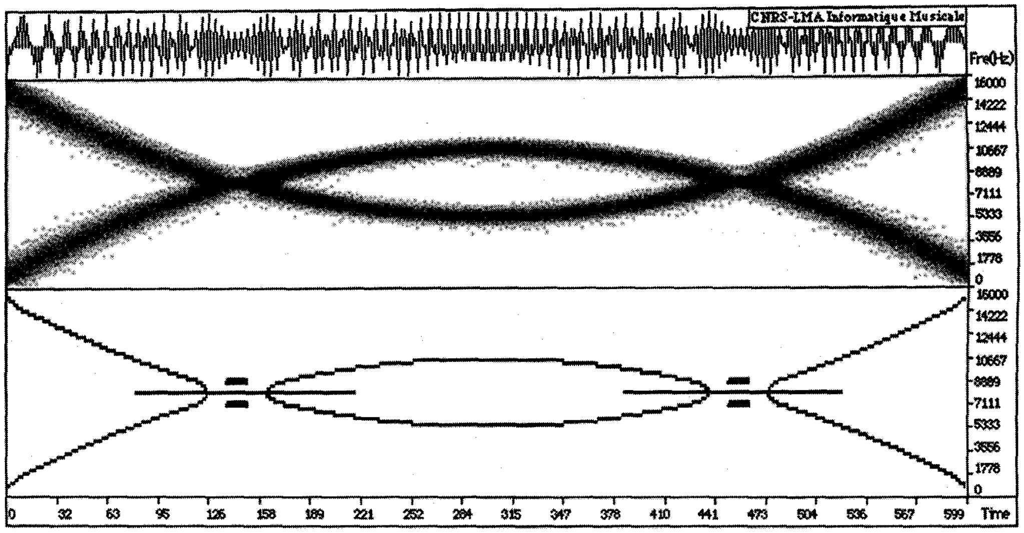


Fig. 23. Modulus and ridge associated to the Gabor transform of two crossing parabolic chirps, both with amplitude one. Near the crossing points the ridges are false.

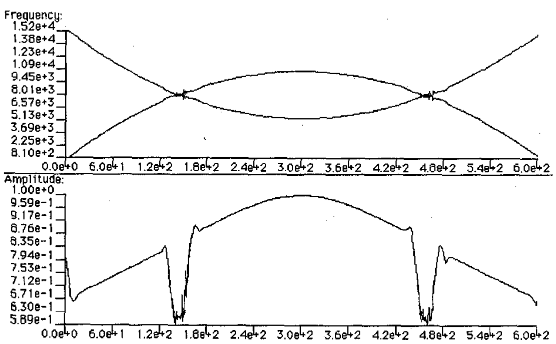


Fig. 24. Frequency and amplitude modulation laws of each component obtained with the algorithm under the constant frequency approximation. Near the two crossing points, the frequencies involved in the definition of the matrix W have been modified to keep this matrix invertible and to avoid the crossing.

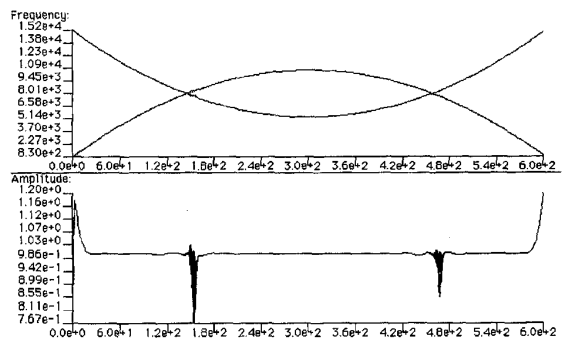


Fig. 25. Frequency and amplitude modulation laws of each component obtained with the algorithm under the linear frequency approximation. Near the two crossing points, the matrix W stays invertible, by taking into account the different slopes of each component, and using an adaptive window. The two components cross, and the error is smaller than in Fig. 24.

shown in a linear scale by density of black dots. It is maximum (minimum) when it is black (white).

3) The lower part is the ridge corresponding to the set of points verifying $\frac{\partial\Phi(\tau,\alpha)}{\partial\tau} = \alpha$.

Figs. 21 and 24 correspond to the extraction of the modulation laws by the use of the iterative algorithm, with a constant frequency hypothesis. It is divided into two parts:

1) The upper part represents the frequency modulation law of all the components obtained through the algorithmic estimation.
2) The lower part represents the amplitude modulation law of all the components obtained through the algorithmic estimation.

Figs. 22 and 25 correspond to the extraction of the modulation laws by the use of the iterative algorithm, with a linear frequency hypothesis. It is divided into two parts:

1) The upper part represents the frequency modulation law of all the components obtained through the algorithmic estimation.
2) The lower part represents the amplitude modulation law of all the components obtained through the algorithmic estimation.

Example 1: Signal with three components (Figs. 20–22).
Example 2: Crossing parabolic chirps (Figs. 23–25).
Example 3: Saxophone sound.

This last example demonstrates the use of the evolutive filter bank, in order to improve the spectral line estimation on a saxophone sound composed of contributions modulated both in amplitude and in frequency. The sampling rate is 32 kHz, and the duration of the analysis is 265 ms (8500 samples). For the construction of the filters, we have assumed that the frequency was locally constant and that the amplitude was locally linear.

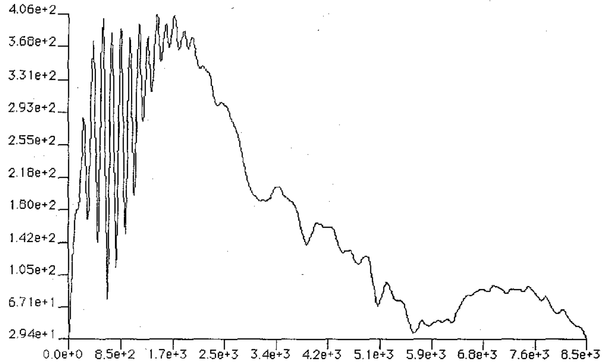


Fig. 26. Amplitude modulation law extracted by the use of the spectral line estimation algorithm.

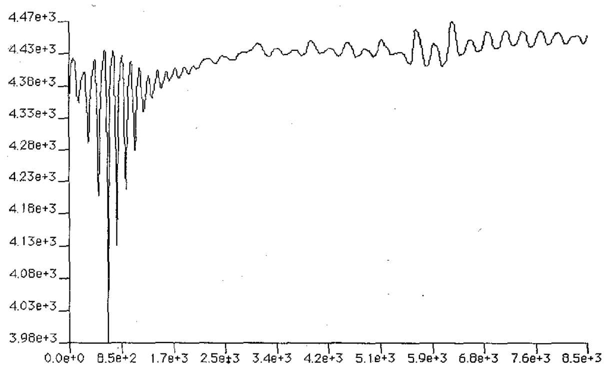


Fig. 27. Frequency modulation law extracted by the use of the spectral line estimation algorithm.

Figs. 26 and 27 display, respectively, the amplitude (on a linear scale) and the frequency modulation laws of the 20th harmonic extracted by the use of the spectral line estimation algorithm. Since the frequency of the components changes on the duration of the analysis, the fixed filters does not match the frequencies of the signal on all its duration and the estimation performed is biased during the beginning, yielding oscillations in the modulation laws due to interferences with the neighbor components.

Figs. 28 and 29 display respectively the amplitude (on a linear scale) and the frequency modulation laws of the 20th harmonic extracted by the use of the evolutive spectral line estimation algorithm. This time, the frequencies of the filter follow the frequencies of the signal and the estimation is correct.

V. Conclusion

The modelization of audiophonic signals requires a decomposition of the sound into elementary components composed of "pure tones" modulated in both amplitude and frequency. Depending on the signal, we have proposed techniques based on time-frequency representations (Gabor and wavelet transforms) allowing such a decomposition. These techniques take into account the fact that continuous transforms are redundant and that a particular set of coefficients, called the ridge, contains the most relevant

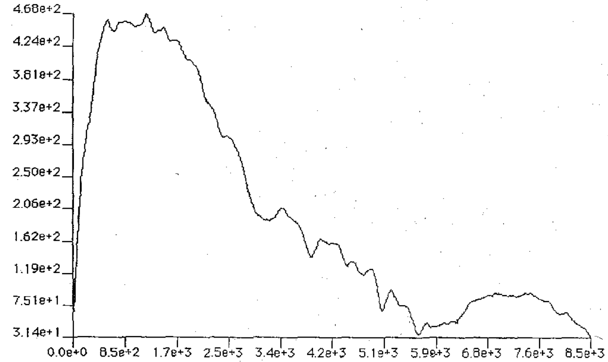


Fig. 28. Amplitude modulation law extracted by the use of the evolutive spectral line estimation algorithm.

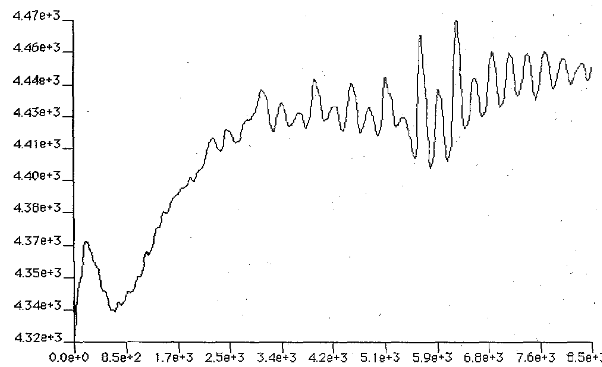


Fig. 29. Frequency modulation law extracted by the use of the evolutive spectral line estimation algorithm.

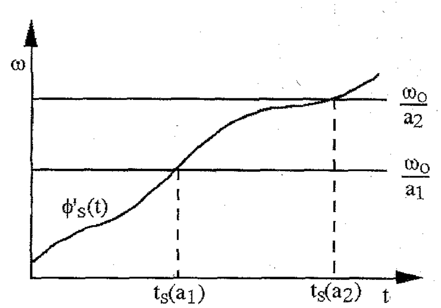


Fig. 30. Interpretation of the stationary point.

information. In the case of signals satisfying asymptotic requirements, we have shown that the ridge of the transform can be estimated by only considering "horizontal" or "vertical" derivatives of the phase of the transforms. Nevertheless, explicit calculus made under different signal approximations have led us to consider a more general definition of the ridge, given by the "crossed criterion," which necessitates a consideration of the phase derivative of the transform with respect to both parameters of the representations. In the same way, by exchanging time and frequency, we have shown that a similar criterion leads to the estimation of the group delay and the spectral density of frequency broadband signals. These techniques have been

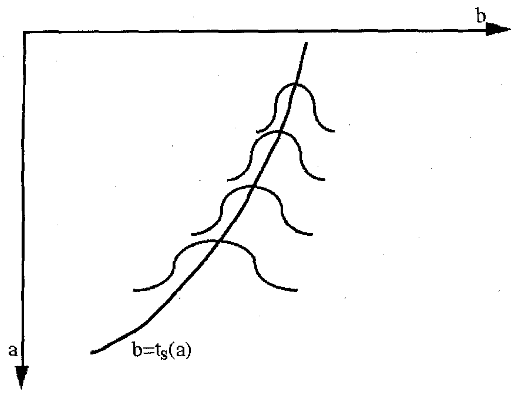


Fig. 31. Behavior of the wavelet transform.

extended to the case of multi component signals and applied to the analysis of real sounds. It is difficult to describe the sounds obtained after resynthesis through a modelization with words, but during presentations, people agree that the synthesized sound is generally indistinguishable from the original one. Another aspect, even more difficult to describe, is the possibility given by such a modelization concerning manipulations of the sounds, permitting exotic and intimate transformations.

APPENDIX

A. Stationary Phase-Based Approximations

If the signal satisfies asymptotic properties, its transform is given by a strongly oscillating integral which can be approximated by the stationary phase method [2]–[4]. We shall show that the stationary points, corresponding to the strongest contribution of the integral, can be linked to the frequency modulation law of the signal and can be retreived from the transform itself. Even though this approach does not permit to conclude on the limits of the algorithm derived this way, it gives general information on the behavior of the transforms.

Let us call $\Phi(b,a)$ (respectively, $\Phi(\tau,\alpha)$) the phase of the wavelet (respectively, Gabor) transform, $M(b,a)$ (respectively, $M(\tau,\alpha)$) its modulus, and let us assume that the function $W(t)$ is a gaussian: $W(t)=\exp(\frac{-t^2}{2\sigma^2})$.

1) The Wavelet Case: The wavelet transform is expressed by

$$T(b,a)=\frac{1}{a}\int A_s(t)W\left(\frac{t-b}{a}\right)e^{i(\Phi_s(t)-\omega_o(\frac{t-b}{a}))}\,dt.$$

Under the hypothesis of asymptotism, the phase stationarity criterion at a point $t_s(a)$ leading to an approximation of this integral is given by $\frac{d\Phi_s(t)}{dt}=\frac{\omega_o}{a}$ in $t=t_s(a)$.

Let us first discuss the physical interpretation of the point $t_s(a)$. One first has to assume that this point is unique for a given value of a. For that purpose, consider a time-frequency plan $\{t,\omega\}$, and draw the curve $\frac{d\Phi_s(t)}{dt}$ (see Fig. 30). For any scale parameter a, one can also draw the curve $\frac{\omega_o}{a}$. The stationary point t_s is then located at the intersection of the two curves. For each curve $\frac{\omega_o}{a}$, there exists a unique stationary point $t_s(a)$, being a function of a. Consequently, this set of point carries all the frequency information contained in the signal, as soon as the frequency modulation law is monotonous.

Moreover, if the wavelet transform is considered as a set of bandpass filters tuned to the frequency $\frac{\omega_o}{a}$, a point $t_s(a_o)$ represents the time at which to expect a maximum response of the filter indexed by a_o, since this one is then tuned to the instantaneous frequency of the signal.

Thus $t_s(a)$ is directly related to the reciprocal function of the instantaneous frequency of the signal, and knowledge of it leads to the frequency modulation law of the signal.

It is interesting to look at the behavior of the wavelet transform under such an approximation. By denoting $\Phi_s''=\frac{d^2\Phi_s(t_s(a))}{dt^2}$, then the transform is approximated by

$$\begin{aligned}\Phi(b,a)\simeq{}&\Phi_s(t_s(a))-\omega_o\left(\frac{t_s(a)-b}{a}\right)\\&+\frac{1}{2}\mathrm{Arctan}(a^2\sigma^2\Phi_s'')+\frac{1}{2}(b-t_s(a))^2\\&\cdot\frac{\Phi_s''}{1+a^4\sigma^4\Phi_s''^2}\\M(b,a)\simeq{}&A_s(t_s(a))\frac{\sigma\sqrt{2\pi}}{(1+a^4\sigma^4\Phi_s''^2)^{1/4}}\\&\cdot\exp\left(-\frac{1}{2}\,\frac{(b-t_s(a))^2a^2\sigma^2\Phi_s''}{1+a^4\sigma^4\Phi_s''^2}\right).\end{aligned}$$

The transform can be interpreted as a superposition of wavelets localized around the points $b=t_s(a)$ (see Fig. 31).

This set of points in the $\{b,a\}$ half-plane have a intuitive meaning, since through the definition of the stationary points, they correspond to the points (b,a) where $\frac{d\Phi_s(b)}{db}=\frac{\omega_o}{a}$.

2) The Gabor Case: The Gabor transform reads

$$L_g(\tau,\alpha)=\int A_s(t)W(t-\tau)e^{i(\Phi_s(t)-\alpha(t-\tau))}\,dt.$$

The phase stationarity criterion at a point $t_s(\alpha)$ leading to an approximation of this integral is given by $\frac{d\Phi_s(t)}{dt}=\alpha$ in $t=t_s(\alpha)$. Its interpretation remains the same as in the wavelet case, $t_s(\alpha)$ being exactly the reciprocal function of the instantaneous frequency of the signal.

By denoting $\Phi_s''=\frac{d^2\Phi_s(t_s(\alpha))}{dt^2}$, then

$$\begin{aligned}\Phi(\tau,\alpha)\simeq{}&\Phi_s(t_s(\alpha))-\alpha(t_s(\alpha)-\tau)\\&+\frac{1}{2}\mathrm{Arctan}(\sigma^2\Phi_s'')+\frac{1}{2}(\tau-t_s(\alpha))^2\\&\cdot\frac{\Phi_s''}{1+\sigma^4\Phi_s''^2}\\M(\tau,\alpha)\simeq{}&A_s(t_s(\alpha))\frac{\sigma\sqrt{2\pi}}{(1+\sigma^4\Phi_s''^2)^{\frac{1}{4}}}\\&\cdot\exp\left(-\frac{1}{2}\,\frac{(\tau-t_s(\alpha))^2\sigma^2\Phi_s''}{1+\sigma^4\Phi_s''^2}\right).\end{aligned}$$

The behavior of the transform is the same as in the wavelet case, the set of important points being defined by either $\tau = t_s(\alpha)$ or $\frac{d\Phi_s(\tau)}{d\tau} = \alpha$.

The crucial point is now the estimation of the curves $b = t_s(a)$ and $\tau = t_s(\alpha)$ from the coefficients of the transforms.

By derivations of the phase of the wavelet transform, it is easy to prove that the set of points $b = t_s(a)$ satisfies

$$\frac{\partial \Phi(b,a)}{\partial b} = \frac{\omega_o}{a} = \frac{d\Phi_s(b)}{db}.$$

In the same way, in the Gabor case, the set of points $\tau = t_s(\alpha)$ satisfies

$$\text{either} \quad \frac{\partial \Phi(\tau,\alpha)}{\partial \tau} = \alpha = \frac{d\Phi_s(\tau)}{d\tau}$$
$$\text{or} \quad \frac{\partial \Phi(\tau,\alpha)}{\partial \alpha} = 0.$$

In both cases, we shall call the set of points satisfying these equations a "ridge."

In the Gabor case, the ridge defined from the phase derivative corresponds to the local maximum of the modulus of the transform with respect to α. In the wavelet case, because of the term: $\frac{\sigma\sqrt{2\pi}}{(1+a^4\sigma^4\Phi''^2{}_s)^{\frac{1}{4}}}$, being a function of a, the maximum of the modulus of the transform with respect to a does not correspond to the ridge defined from the phase, and this maximum cannot be used to estimate the exact frequency modulation law. On the ridge, the difference between the signal and the restriction of the transforms is smallest. Consequently, the modulus of these restrictions along the ridge gives a good estimate of the amplitude modulation law of the signal while their phase derivative with respect to τ or b gives a good estimate of its frequency modulation law [3].

B. Specral Lines Estimation

In this section, we describe an accurate spectral lines estimation algorithm. The interested reader can find complements in [12]. We shall assume that on the time support of $W(t-\tau)$, localized around $t = \tau$, the amplitude modulation law of each component can be locally described by a polynomial of order J. One can then use the expression of the Gabor transform of a single spectral line up to order J.

In this case, the gabor transform of the signal is exactly given by

$$\begin{aligned} L_g(\tau,\alpha) = \frac{1}{2} \sum_{k=1,j=0}^{k=N,j=J} \frac{(-i)^j}{j!} A_k^{(j)}(\tau) \\ \cdot (\hat{W}^{(j)}(\omega_k - \alpha)\exp(i\omega_k\tau) \\ + \hat{W}^{(j)}(-\omega_k - \alpha)\exp(-i\omega_k\tau)). \end{aligned}$$

We consider the restriction of $L_g(\tau,\alpha)$ for $\alpha_p = \omega_p, 1 \le p \le N$.

Let us denote $c_j = \cos(j\frac{\pi}{2})$ and $s_j = \sin(j\frac{\pi}{2})$ and

$$\begin{aligned} {}^1W_{pk}^{(j)} &= \tfrac{1}{2}(\hat{W}^{(j)}(\omega_k - \alpha_p) + \hat{W}^{(j)}(-\omega_k - \alpha_p)) \\ {}^2W_{pk}^{(j)} &= \tfrac{1}{2}(\hat{W}^{(j)}(\omega_k - \alpha_p) - \hat{W}^{(j)}(-\omega_k - \alpha_p)). \end{aligned}$$

For any p, one can write the two linear systems

$$\begin{aligned} &\sum_{k=1,j=0}^{k=N,j=J} \frac{1}{j!}(A_k^{(j)}(\tau)\cos(\omega_k\tau)c_j\, {}^1W_{pk}^{(j)} \\ &\quad + A_k^{(j)}(\tau)\sin(\omega_k\tau)s_j\, {}^2W_{pk}^{(j)}) = \mathrm{Re}\{\mathrm{L}_g(\tau,\alpha_p)\} \\ &\sum_{k=1,j=0}^{k=N,j=J} \frac{1}{j!}(A_k^{(j)}(\tau)\sin(\omega_k\tau)c_j\, {}^2W_{pk}^{(j)} \\ &\quad - A_k^{(j)}(\tau)\cos(\omega_k\tau)s_j\, {}^1W_{pk}^{(j)}) = \mathrm{Im}\{\mathrm{L}_g(\tau,\alpha_p)\}. \end{aligned}$$

These two equations are not coupled, since $\forall j, c_j s_j = 0$. Moreover, $\forall j, c_{2j+1} = 0$ and $s_{2j} = 0$. Consequently, each system is composed of $(J+1)\cdot N$ unknowns that are, respectively:

$$\begin{aligned} &A_k^{(2j)}(\tau)\cos(\omega_k\tau), A_k^{(2j+1)}(\tau)\sin(\omega_k\tau) \\ &A_k^{(2j)}(\tau)\sin(\omega_k\tau), A_k^{(2j+1)}(\tau)\cos(\omega_k\tau). \end{aligned}$$

At this time, one only has N equations, but the $J \cdot N$ missing equations can be obtained by considering the derivatives with respect to α upto order J of the restrictions of the transform at the frequencies α_p. The two systems then become

$\forall q, 1 \le q \le J; \forall p, 1 \le p \le N$

$$\begin{aligned} &\sum_{k=1,j=0}^{k=N,j=J} \frac{(-1)^q}{j!}(A_k^{(j)}(\tau)\cos(\omega_k\tau)c_j\, {}^1W_{pk}^{(j+q)} \\ &\quad + A_k^{(j)}(\tau)\sin(\omega_k\tau)s_j\, {}^2W_{pk}^{(j+q)}) \\ &\quad = \mathrm{Re}\left\{\frac{\partial^q L_g(\tau,\alpha_p)}{\partial\alpha^q}\right\} \\ &\sum_{k=1,j=0}^{k=N,j=J} \frac{(-1)^q}{j!}(A_k^{(j)}(\tau)\sin(\omega_k\tau)c_j\, {}^2W_{pk}^{(j+q)} \\ &\quad - A_k^{(j)}(\tau)\cos(\omega_k\tau)s_j\, {}^1W_{pk}^{(j+q)}) \\ &\quad = \mathrm{Im}\left\{\frac{\partial^q L_g(\tau,\alpha_p)}{\partial\alpha^q}\right\}. \end{aligned}$$

Since all these equations are linear, and since the derivation of the restrictions of the Gabor transform with respect to α can be transposed directly on the Fourier transform of the window, the filter bank interpretation is still valid. This time, each filter is composed by linear combinations of the functions $\hat{W}(\omega - \alpha_k)$, and also their derivatives with respect to ω until order J.

Since each filter enables the extraction of an amplitude modulation law being a polynom of order J on the time support of $W(t-\tau)$, it satisfies the following conditions:

$$\begin{aligned} &\hat{F}_k(\alpha_p) = \delta_{pk} \\ &\hat{F}_k(-\alpha_p) = 0 \quad 1 \le p \le N \quad 1 \le k \le N \\ &\frac{d^j\hat{F}_k(\alpha_p)}{d\omega^j} = 0 \\ &\frac{d^j\hat{F}_k(-\alpha_p)}{d\omega^j} = 0 \quad 1 \le p \le N \quad 1 \le k \le N \quad 1 \le j \le J. \end{aligned}$$

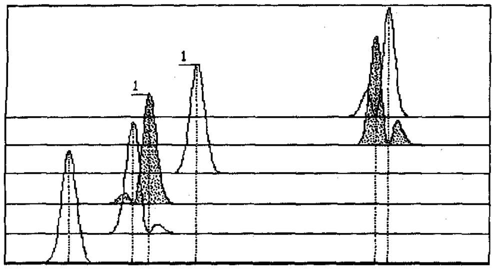


Fig. 32. Filters for a linear amplitude estimation.

The Fourier transform on a linear scale of a family of filters allowing the extraction of components with locally linear amplitudes is shown in Fig. 32. The horizontal axis is the frequency. The input signal is a sum of six sine waves. The dotted vertical lines indicate the position of each frequency in the spectrum. The six filters are displayed vertically stacked, two of them being shaded gray. The Fourier transform of a filter equals one for its corresponding spectral line and equals zero for the five other frequencies. The first derivative of the Fourier transform of each filter vanishes for the six frequencies.

C. Noncrossing Multiridges Estimation

We present here the way to build an auto-adaptive filter bank aimed at the disentangling of frequency-modulated asymptotic components.

As in the case of the spectral lines, it is possible to define, at any time τ_o, an operator $\boldsymbol{\Omega}(\tau_o)$ acting now on a vector of scalar products $\boldsymbol{Lg}(\tau_o)$ (instead of horizontal restrictions of the transform) between the signal and Gabor functions at the different frequencies $\alpha_k(\tau_o)$.

1) An Intuitive Definition of the Operator: $\boldsymbol{\Omega}(\tau_o)$: Let us suppose that at time τ_o, one knows the exact parameters $\alpha_k(\tau_o)$ defined by

$$\alpha_k(\tau_o) = \frac{d\Phi_k(\tau_o)}{d\tau}.$$

Let us call $W(\tau_o)$ the matrix of elements defined by

$$W_{pk}(\tau_o) = \int \exp(i(\phi_k(t) - \phi_k(\tau_o)))\overline{W}(t-\tau_o) \cdot \exp(-i\alpha_p(\tau_o)(t-\tau_o))\, dt.$$

Then $\boldsymbol{\Omega}(\tau_o) = W^{-1}(\tau_o)$ is obviously a good operator, since it directly uses the frequency modulation law of each component.

Indeed

$$L_g(\tau_o,\alpha) = \sum_{k=1}^{N} \int s_k(t)\overline{W}(t-\tau_o)\exp(-i\alpha(t-\tau_o))\, dt$$

$\forall p, 1 \leq p \leq N$, one can write

$$\begin{aligned} L_g(\tau_o,\alpha_p(\tau_o)) &= \sum_{k=1}^{N} \int s_k(t)\overline{W}(t-\tau_o)\exp(-i\alpha_p(\tau_o)(t-\tau_o))\, dt \\ &= \sum_{k=1}^{N} A_k \exp(i\phi_k(\tau_o))W_{pk}(\tau_o). \end{aligned}$$

Then, each component can be found just by applying $\boldsymbol{\Omega}(\tau_o)$ to $L_g(\tau_o,\alpha_p(\tau_o))$. Unfortunately, this calculation enables the extraction of independant components under the assumption that the frequency modulation laws are known. Let us try now to see if one can obtain an approximation of each frequency modulation law.

2) A Suboptimal Operator: $\boldsymbol{\Omega}'(\tau)$: Let us assume that each elementary component can be considered locally either as a sine wave, or as a linear chirp. Such an hypothesis is relevant since it just says that the frequency of each component does not move too fast on the wavelet support around τ_o, and is equivalent to replace the phase of each component either by its first or second order limited expansion around $t = \tau$. We shall describe the constant frequency approximation. The linear frequency approximation has been developed in the previous section.

Let us consider a single component $A_k \exp(i\Phi_k(t))$ and write

$$\Phi_k(t) = \Phi_k(\tau) + (t-\tau)\Phi'_k(\tau).$$

Its Gabor transform is then given by

$$\begin{aligned} L_g(\tau,\alpha) &= A_k \int \overline{W}(t-\tau)\exp(-i\alpha(t-\tau)) \\ &\quad \cdot \exp(i(\Phi_k(\tau) + (t-\tau)\Phi'_k(\tau)))\, dt \\ &= A_k \exp(i(\Phi_k(\tau) - \tau\Phi'_k(\tau))) \\ &\quad \cdot \int \overline{W}(t-\tau)\exp(-i\alpha(t-\tau)) \\ &\quad \cdot \exp(it\Phi'_k(\tau))\, dt. \end{aligned}$$

The integral term is then the Gabor transform of a monochromatic signal with frequency $\Phi'_k(\tau)$. The Gabor transform is then approximated by

$$L_g(\tau,\alpha) = A_k\overline{\hat{W}}(\Phi'_k(\tau) - \alpha)\exp(i\phi_k(\tau)).$$

Let $\alpha_k(\tau_o)$ be the fixed point of the ridge estimation algorithm, corresponding to the k^{th} component, and assume that $\alpha_k(\tau_o) = \Phi'_k(\tau_o)$ As previously, one defines a new operator $\boldsymbol{\Omega}'(\tau)$ with the help of

$$\begin{aligned} W_{pk}(\tau_o) &= \int \exp(i\alpha_k(\tau_o)t)\overline{W}(t-\tau_o) \\ &\quad \cdot \exp(-i\alpha_p(\tau_o)(t-\tau_o))\, dt \\ &= \overline{\hat{W}}(\alpha_k(\tau_o) - \alpha_p(\tau_o)). \end{aligned}$$

Obviously, in this case, one has

$$A_k \exp(i\Phi_k(\tau_o)) = \sum_{p=1}^{N} W_{kp}^{-1}(\tau_o)L_g(\tau_o,\alpha_p(\tau_o)).$$

This approximation can seem to be strong, but as we will see in the examples, the results obtained are satisfying.

Nevertheless, by assuming that $W(t)$ is a gaussian, instead of using a constant frequency approximation, one can use a linear one. If one assumes that $W(t)$ is a gaussian, one then obtain an other definition of $\boldsymbol{\Omega}'(\tau)$ through $W_{pk}(\tau_o)$, coming directly from the results of the previous section

$$W_{pk}(\tau_o) = \sigma\sqrt{\frac{2\pi}{1 - i\sigma^2(\Phi_k''(\tau_o) - \beta)}} \cdot \exp\left(-\frac{\sigma^2}{2}\frac{(\alpha_k(\tau_o) - \alpha_p(\tau_o))^2}{1 - i\sigma^2(\Phi_k''(\tau_o) - \beta)}\right).$$

Concurrently to the definition of $\boldsymbol{\Omega}(\tau)$, one has replaced, on the support of $W(t)$ located at τ_o, the frequency modulation law of each component either by a constant or a linear function corresponding to its estimated version at time τ_o if the component was alone. Obviously, the more the signal is asymptotic with respect to $W(t)$, the more $\boldsymbol{\Omega}'(\tau)$ will be close to $\boldsymbol{\Omega}(\tau)$. Nevertheless, such approximations of the frequency modulation laws are no more available when the signal is a sum of elementary contributions. Moreover, in the linear frequency approximation, the second derivative of the phase of each component is unknown.

3) A Self-Building Operator: $\boldsymbol{\Omega}'(\tau)$: automatic extraction of multiridges. For clarity, one will start by a constant frequency approximation, and then switch to the linear one.

We can achieve a self definition of $\boldsymbol{\Omega}'(\tau)$, by a generalization of the ridge estimation algorithm to multicomponent signals. Let us denote I_k the interval of convergence corresponding to the transform of a single component $s_k(t)$ and let us call $\alpha_k^o(\tau)$ initial values of scale parameters, belonging respectively to I_k.

Let us define the initial vector $\boldsymbol{X}^o(\tau) = \boldsymbol{\Omega}'^o(\tau)\boldsymbol{L}^o(\tau)$, where $\boldsymbol{L}^o(\tau)$ represents the vector of scalar products between the signal and the Gabor functions at the frequencies $\alpha_k^o(\tau)$.

In the case of a constant frequency approximation, the following algorithm builds iteratively (on index n) the operator $\boldsymbol{\Omega}'(\tau)$, the trajectories $\alpha_k(\tau)$, and the functions $A_k \exp(i\Phi_k(\tau))$ at any time τ:

$$\begin{aligned}
\boldsymbol{X}^n(\tau) &= \boldsymbol{\Omega}'^n(\tau)\boldsymbol{L}^n(\tau) \quad \text{with} \\
X_k^n(\tau) &= \sum_{p=1}^{N} W_{kp}^{-1} L_g(\tau, \alpha_k^n) \\
\alpha_k^{n+1}(\tau) &= \frac{\partial}{\partial\tau}\mathrm{Arg}\{X_k^n(\tau)\} \\
\boldsymbol{\Omega}'^{n+1}(\tau) &= W^{-1}(\tau) \quad \text{with} \\
W_{pk}(\tau) &= \hat{W}(\alpha_k^{n+1}(\tau) - \alpha_p^{n+1}(\tau)).
\end{aligned}$$

For a fixed ϵ, arbitrary small, the convergence criteria are the following:

$$\forall k, \quad 1 \le k \le N \quad |\alpha_k^{n+1}(\tau) - \alpha_k^n(\tau)| \le \epsilon.$$

One proceeds to time $\tau + d\tau$ by the relation

$$\alpha_k^o(\tau + d\tau) = \alpha_k^\infty(\tau).$$

In the case of a linear frequency approximation, one also has to estimate iteratively $\Phi_k''(\tau)$. As we have seen before, such an estimation requires the calculus of a combination of the "horizontal" and "vertical" derivatives of the phase of the transform, as well as the use of a chirped Gabor function.

The algorithm then reads

$$\begin{aligned}
\boldsymbol{X}^n(\tau) &= \boldsymbol{\Omega}'^n(\tau)\boldsymbol{L}^n(\tau) \quad \text{with} \\
X_k^n(\tau) &= \sum_{p=1}^{N} W_{kp}^{-1} L_g(\tau, \alpha_k^n) \\
\alpha_k^{n+1}(\tau) &= \frac{\partial}{\partial\tau}\mathrm{Arg}\{X_k^n(\tau)\} + \beta_k^n \frac{\partial}{\partial\alpha}\mathrm{Arg}\{X_k^n(\tau)\} \\
\beta_k^{n+1}(\tau) &= \frac{\partial}{\partial\tau}\left(\frac{\partial}{\partial\tau}\mathrm{Arg}\{X_k^n(\tau)\}\right. \\
&\qquad \left. + \beta_k^n \frac{\partial}{\partial\alpha}\mathrm{Arg}\{X_k^n(\tau)\}\right) \\
&\qquad + \beta_k^n \frac{\partial}{\partial\alpha}\left(\frac{\partial}{\partial\tau}\mathrm{Arg}\{X_k^n(\tau)\}\right. \\
&\qquad \left. + \beta_k^n \frac{\partial}{\partial\alpha}\mathrm{Arg}\{X_k^n(\tau)\}\right) \\
\boldsymbol{\Omega}'^{n+1}(\tau) &= W^{-1}(\tau) \quad \text{with}. \\
W_{pk}(\tau) &= \sigma\sqrt{\frac{2\pi}{1 - i\sigma^2(\beta_k^{n+1}(\tau) - \beta_p^{n+1}(\tau))}} \\
&\qquad \cdot \exp\left(-\frac{\sigma^2}{2}\frac{(\alpha_k^{n+1}(\tau) - \alpha_p^{n+1}(\tau))^2}{1 - i\sigma^2(\beta_k^{n+1}(\tau) - \beta_p^{n+1}(\tau))}\right).
\end{aligned}$$

For a fixed ϵ, arbitrary small, the convergence criteria are the following:

$$\begin{aligned}
\forall k, \quad 1 \le k \le N \quad &|\alpha_k^{n+1}(\tau) - \alpha_k^n(\tau)| \le \epsilon \\
&|\beta_k^{n+1}(\tau) - \beta_k^n(\tau)| \le \epsilon.
\end{aligned}$$

One proceeds to time $\tau + d\tau$ by the relations

$$\alpha_k^o(\tau + d\tau) = \alpha_k^\infty(\tau) + \beta_k^\infty(\tau)\, d\tau, \quad \beta_k^o(\tau + d\tau) = \beta_k^\infty(\tau).$$

At each iteration of the ridge estimation algorithm, we have introduced the algorithm allowing the separation of the components. In fact, at the first iteration, the initial frequency values do not have any reason to match the theoretical ones. Consequently, a component $X_k(\tau)$ will mainly contain the component k but also a few of the other ones. At each new iteration, if one estimated ridge comes closer to the theoretical one, then the "weight" associated to the other contributions will necessarily decrease. Consequently at the next iteration, the estimated ridge will be closer to the theoretical one.

References

[1] J. M. Combes, A. Grossmann, and P. Tchamitchian, Eds., *Wavelets, Time-Frequency Methods and Phase Space.* Berlin: Springer-Verlag, 1989.

[2] E. T. Copson, *Asymptotic Expansions.* New York: Cambridge, 1967.

[3] N. Delprat *et al.,* "Asymptotic wavelet and Gabor analysis: extraction of instantaneous frequencies," *IEEE Trans. Inform. Theory,* vol. 38, pp. 644–665, Mar. 1992.

[4] B. Escudié, A. Grossmann, R. Kronland–Martinet, and B. Torrésani, "Analyse par ondelettes de signaux asymptotiques," in *12th GRETSI Conf.*, June 1989.

[5] D. Gabor, "Theory of communication," *J. IEEE London*, vol. 93, 1946, pp. 429–441.

[6] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products*. New York: Academic.

[7] A. Grossmann, M. Holschneider, R. Kronland–Martinet, and J. Morlet, "Detection of abrupt changes in sound signals with the help of wavelet transforms," in *Inverse Problems*. New York: Academic, 1987, pp. 289–306.

[8] A. Grossmann, R. Kronland–Martinet, and J. Morlet, "Reading and understanding continuous wavelet transform," in *Wavelets, Time-Frequency Methods and Phase Space,* Combes, Grossmann, Tchamitchian, Eds. Berlin: Springer-Verlag, 1989.

[9] I. M. Guelfand and G. E. Chilov, "Les distributions," in *Editions Dunod,* 1962.

[10] P. Guillemain, R. Kronland-Martinet, and B. Martens, "Estimation of spectral lines with the help of the wavelet transform. Applications in NMR spectroscopy," in *Wavelets and Applications,* Meyer, Ed. Masson: Springer-Verlag 1991, pp. 38–60.

[11] P. Guillemain and R. Kronland–Martinet, "Parameters estimation through continuous wavelet transform for synthesis of audio sounds," preprint, *19th AES Convent.*, Feb. 1991, Paris, France.

[12] P. Guillemain, "Analyse et modélization de signaux sonores par des représentations temps-fréquence linéaires," Ph.D. dissertation, Université Aix-Marseille II, June 1994.

[13] R. Kronland–Martinet, J. Morlet, and A. Grossmann, "Analysis of sound patterns through wavelet transforms," in *Int. J. Patt. Recog. and Art.* vol. 1, no. 2, World Scientific, 1987, pp. 273–302.

[14] R. Kronland–Martinet, "The wavelet transform for analysis, synthesis and processing of speech and music sounds," *Computer Music J.*, vol. 12, no. 4, MIT Press, 1988, pp. 11–20.

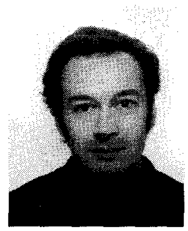
[15] R. Kronland–Martinet and A. Grossmann, "Application of time-frequency and time-scale methods (wavelet transform) to the analysis, synthesis and transformations of natural sounds" of speech and music sounds" in *Representation of Musical Signals.* Cambridge, MA: MIT Press, 1991, pp. 45–85.

[16] B. Picinbono and W. Martin, "Représentation des signaux par amplitude et phase instantanées," *Ann. des Télécommunications,* vol. 38, pp. 179–190, 1983.

[17] J. C. Risset and M. Mathews, "Analysis of musical instruments tones," *Physics Today,* vol. 22, no 2, pp. 23–30, Feb. 1969.

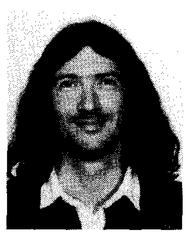
[18] J. C. Risset and D. Wessel, "Exploration of timbre by analysis and synthesis," in *The Psychology of Music.* New York: Academic, 1982, pp. 25–58.

[19] P. Schaeffer, "Traité des objets musicaux," Paris, *Éditions du Seuil,* 1966.



Philippe Guillemain received the Ph.D. degree for modelization of sounds from Université Aix-Marseille II, France, in 1994.

He is a researcher in the Centre National de la Recherche Scientique (CNRS) in Marseille, France. He primarily works on computer music research. Since 1989, he has been working in the field of modelization of sounds produced by musical instruments using time-frequency representations and the additive synthesis model.



Richard Kronland–Martinet received the Ph.D. degree from the Université Aix-Marseille II for his work on digital active control of vibrations in 1983, and the "Doctorat d'état" in 1989 for work on the applications of wavelets in acoustics.

He is a researcher in the Centre National de la Recherche Scientique (CNRS) in Marseille, France. He primarily works on computer music research. Since 1985, he has been working in the field of analysis and synthesis of real sounds using the continuous wavelet transform.