# INTERMODULATION EFFECTS ANALYSIS USING COMPLEX BANDPASS FILTERBANKS

*José R. Beltrán, Jesús Ponce de León and Eduardo Estopiñán*

Digital Audio Group, Dept. Electronic Engineering and Communications
University of Zaragoza, Zaragoza, Spain
`jrbelbla@unizar.es, jponce@unizar.es`

## ABSTRACT

The objective of this paper is to show the ability of complex bandpass filterbanks to extract the intermodulation information that appears when two audio signals interact inside the same analysis band. To perform the analysis a sinusoidal model of the signals has been assumed. Three kinds of signals have been analyzed: a sum of two cosines, a sum of two linear chirps and a sum of two exponential chirps. The complex bandpass filtering of the signals is carried out using a new algorithm based on the Complex Continuous Wavelet Transform. The developed algorithm has been validated comparing the practical results with the theoretical instantaneous amplitude and instantaneous phase of the obtained model of the signals. With the appropriate width, the complex bandpass filters show the same behaviour as our perceptual ability to discriminate interacting tones when they fall inside a critical band of the human ear.

## 1. INTRODUCTION

It is well known that human hearing is very conditioned by the *critical bands* of the ear. The existence of the critical bands influences the consonance and dissonance of sounds and the masking effect, among other effects. As pointed out in [1] the bandwidth of the critical bands lies between one-third and one-sixth of a octave. This fact leads to the well-known 24 critical bands when the whole frequency axis is covered. These critical bands also affect the perception of beats when two pure tones get closer. When two tones get close enough we are not able to resolve them in spectral terms and the perceived sound can be described by a signal that exhibits amplitude and frequency variations. The *instantaneous amplitude* and the *instantaneous frequency* of the signal can describe these variations. The *canonical pair* of such a signal models exactly this fact.

In this work to extract the canonical pair of a signal and its instantaneous amplitude and frequency the *Complex Continuous Wavelet Transform* (CCWT) is used. The CCWT can be viewed as the filtering of the input signal $s(t)$ with a filter bank $\psi_{a,b}(t)$, in our case, built from the Morlet's Wavelet shown in Eq. 1

$$\psi(t) = C e^{-\frac{t^2}{2\sigma^2}} \left( e^{j\omega_0 t} - e^{-\frac{\omega_0^2}{2}} \right) \approx C e^{-\frac{t^2}{2\sigma^2}} e^{j\omega_0 t} \quad (1)$$

The wavelet coefficients can be calculated as follows:

$$W_s(a,b) = \int_{-\infty}^{\infty} s(t) \psi_{a,b}^*(t) dt \quad (2)$$

where $*$ is the complex conjugate and $\psi_{a,b}$ is the *mother wavelet* scaled by a factor $a$ and shifted by a factor $b$.

The filterbank $\psi_{a,b}$ is generated by shifting the mother wavelet in time with the parameter $b$ and dilating (or contracting) it by the scale parameter $a$. The shapes in the frequency domain of the individual bandpass filters of the filterbank are simple Gaussians of width $\sigma$ and each one centered at $a\omega_0$. Due to the mother wavelet $\psi(t)$ is complex, the obtained complex coefficients can be represented in terms of modulus and phase. In [2] it can be found an extended explanation of CCWT and Morlet's Wavelet.

This work is mainly related with R. Kronland-Martinet's et al. [3] [4] [5] and R. Carmona's et al. [6] [7] works. They found that the minimal information needed to characterize a signal from its wavelet coefficients is located around the points of the time-frequency half-plane where the module of the coefficients has a maximum (the *ridge* of the transform). At these points the phase of the coefficients can be easily obtained. This phase information is closely related with the instantaneous frequency [8] of the original signal.

In this work a slightly different approach is made. Our objective is not only to find the instantaneous frequency but also the instantaneous amplitude of the original signal. In this way a comparison can be made between the theoretical instantaneous amplitude and instantaneous frequency of the canonical pair of the signal and the results obtained with our algorithm.

This paper has been divided as follows. In Sec. 2 the canonical pair of a sinusoidal modelled signal is presented. Only the sum of two sinusoids is detailed but it can be generalized applying the method recursively. In Sec. 3 our analysis and synthesis algorithm is briefly explained. Three kinds of signals have been processed in Sec. 4: a sum of cosines with different frequency distance between them, a sum of two linear chirps and a sum of two exponential chirps. The main conclusions are presented in Sec. 5.

## 2. THE CANONICAL PAIR OF A WAVE

### 2.1. Instantaneous amplitude and phase: the analytic signal

Let's suppose that we are working with a signal $s(t)$ composed by the sum of two general cosine waves:

$$s(t) = A_1(t) \cos[\varphi_1(t)] + A_2(t) \cos[\varphi_2(t)]. \quad (3)$$

To find the canonical pair of this signal, we can work with the related *analytic signal* [3] of every cosine of 3. Then, the signal $s(t)$ is the real part of the sum of the analytic signals:

$$s(t) = \Re \left[ A_1(t) e^{j\varphi_1(t)} + A_2(t) e^{j\varphi_2(t)} \right]. \quad (4)$$

Let's suppose that the amplitudes $A_1(t)$ and $A_2(t)$ satisfy the *asymptotic condition* [3]. That means that the signal amplitude varies slowly enough with respect to the signal phase:

$$\left| \frac{1}{A_i(t)} \frac{dA_i(t)}{dt} \right| \ll \left| \frac{d\varphi_i(t)}{dt} \right| . \tag{5}$$

In such a case, the signal $s(t)$ can be written as:

$$s(t) = \|s(t)\| cos\left\{ \varphi_1(t) + \Phi(t) \right\} . \tag{6}$$

Naming $\Delta(t) = \varphi_2(t) - \varphi_1(t)$ to the difference of the individual phases, the modulus of $s(t)$ can be written as:

$$\|s(t)\| = \sqrt{A_1^2(t) + A_2^2(t) + 2A_1(t)A_2(t)\cos\left[\Delta(t)\right]}, \tag{7}$$

and $\Phi(t)$ in Eq. 6 is:

$$\Phi(t) = \arctan\left\{ \frac{A_2(t)\sin\left[\Delta(t)\right]}{A_1(t) + A_2(t)\cos\left[\Delta(t)\right]} \right\} . \tag{8}$$

Considering that Eq. 5 is true, we obtain the instantaneous frequency of the signal $s(t)$ from Eqs. 8 and 6:

$$\nu(t) = \frac{1}{2\pi} \frac{d\left\{ \varphi_1(t) + \Phi\left[ s(t) \right] \right\}}{dt} =$$
$$\frac{1}{2\pi} \left\{ \varphi_1'(t) + \frac{\Delta'(t)\left\{ A_2^2(t) + A_1 A_2 \cos[\Delta(t)] \right\}}{A_1^2(t) + A_2^2(t) + 2A_1(t)A_2(t)\cos[\Delta(t)]} \right\} . \tag{9}$$

If the signal $s(t)$ is composed by three or more sinusoids like 3, we can use this technique recursively to reach more general results.

The envelope and the phase of the canonical pair of $s(t)$ are the instantaneous amplitude and the instantaneous phase of the signal, and they exhibit a time dependence that reflects the amplitude and frequency modulation of the total wave. If the phases of the components are distant enough, there is a clear distinction between carrier wave and modulating wave [8]. This difference becomes more diffuse when the phases approach each other, and there starts to appear intermodulation terms that become into the classical beat movement if the frequencies are close enough. It means that mixing the components of a signal in a modulus-phase representation has a greater physical sense the closer the components are. So Eq. 6 is a valid representation of every signal composed by two or more sinusoids like 3, but useful from a perceptual point of view whenever they are close enough so that intermodulation terms become audible; something similar to the behaviour of the human ear due to the critical bands.

## 2.2. Analysis of a sum of pure cosines

Two signals composed by the sum of two cosines each with constant amplitudes and frequencies have been analyzed. They are shown in Fig. 1. The graphs of instantaneous amplitude and frequency obtained from Eqs. 7 and 9 are shown in Figs. 2 and 3.

It can be seen that the intermodulation terms clearly appear in both signals. In the first one we have two frequencies came apart by the distance of a note. So, this representation, although exact, is not valid in terms of perception. In this representation we are not able to distinguish between the two notes but they are
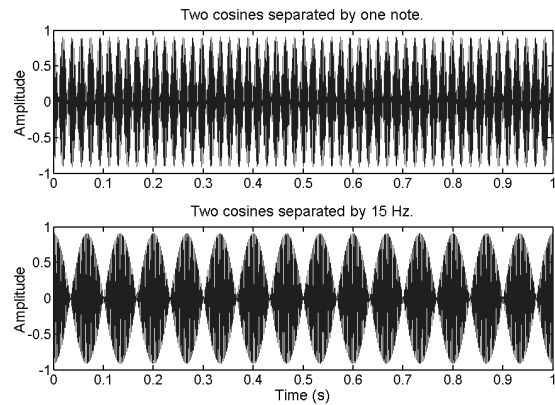
Figure 1: Representation of two waves composed by two cosines. (a) One note apart: frequencies are 440 Hz and 493.88 Hz (notes A4 and B4) and amplitudes 0.5 and 0.499, respectively (b) 15 Hz apart: frequencies are 440 Hz and 455 Hz with the same amplitudes.
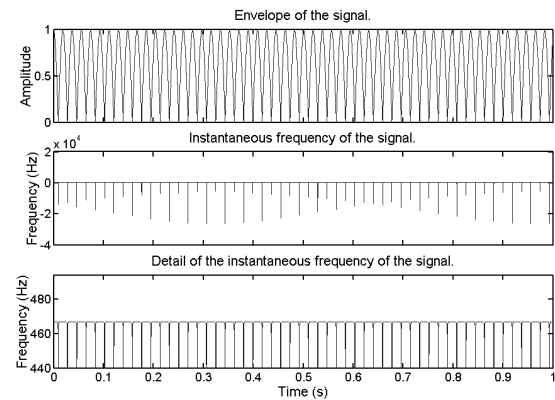
Figure 2: Instantaneous amplitude and frequency of the signal with frequencies 440 Hz and 493.88 Hz.
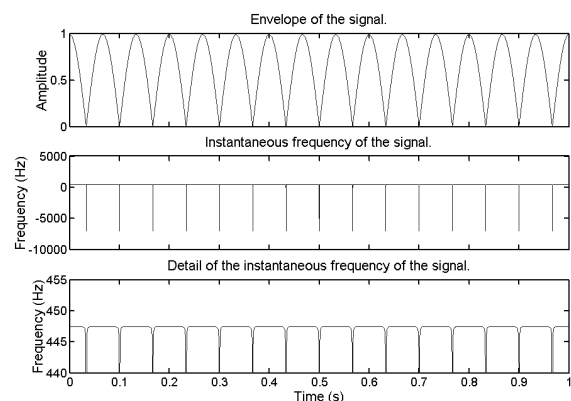
Figure 3: Instantaneous amplitude and frequency of the signal with frequencies 440 Hz and 455 Hz.

perceived as individual tones. In the second example a similar behaviour is shown, but in this case this representation is much more appropriate to our perception experience of beating in amplitude and frequency when the two tones interact. In Fig. 3 the perceived amplitude modulation and the frequency beating can be observed.

### 2.3. Analysis of chirped signals

In Fig. 4 two chirped signals are shown. The sum of two lineal chirps and the sum of two exponential chirps with different values composes them.

$$s_{2lin} = \sum_{i=1}^{2} A_i(t) \cos\left(a_i t^2 + b_i t\right) \qquad (10)$$

$$s_{2exp} = \sum_{i=1}^{2} A_i(t) \cos\left(B_i e^{C_i t}\right) \qquad (11)$$

In this case, for simplicity, we have supposed constant amplitude for each component. Figs. 5 and 6 show the results of equations 7 and 9 for the signals represented in Fig. 4. It can be observed in the detailed view of Figs. 5 and 6 the linear and exponential change in the instantaneous frequencies of each signal.
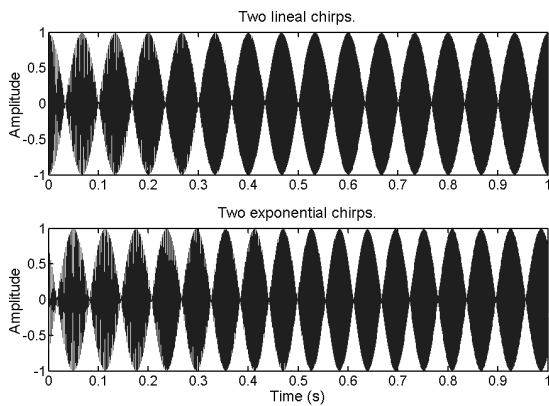


Figure 4: Representation of two chirped waves. (a) Lineal case with parameters: $A_1 = 0.5$, $a_1 = 400$ and $b_1 = 415$, $A_2 = 0.499$, $a_2 = 400$ and $b_2 = 400$. (b) Exponential case with parameters: $A_1 = 0.5$, $B_1 = 414.92$ and $C_1 = 1.074$, $A_2 = 0.499$, $B_2 = 399.97$ and $C_2 = 2.098$.

In these cases the interaction between individual chirps is high and is very difficult to distinguish each isolated one. So, the obtained amplitude and frequency representation describes quite well the signals in terms of perception.

Once described theoretically the signals that will be analyzed we are going to present the developed algorithm in the next section.

### 3. ALGORITHM

An additive synthesis algorithm based on the CCWT has been developed. The algorithm block diagrams can be seen in Figs. 7 and 8.

The main characteristic of the algorithm that differs from the existing ones [3] [4] [5] [6] [7] is that the Q of the filterbank is adjustable. The control parameter is the divisions per octave of the
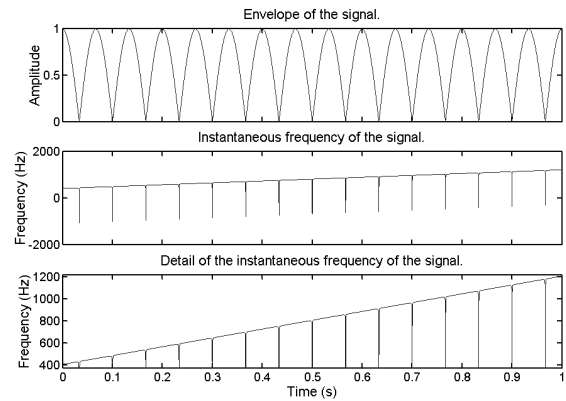


Figure 5: Instantaneous amplitude and frequency of the double lineal chirp signal.
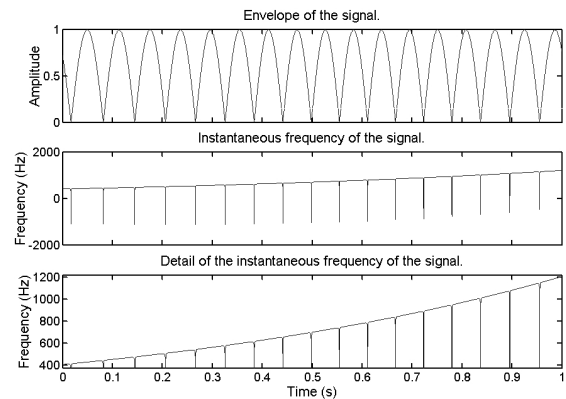


Figure 6: Instantaneous amplitude and frequency of the double exponential chirp signal.

filterbank. This behaviour is obtained modifying appropriately the parameter $\sigma$ on Eq. 1 while maintaining the proper covering of the whole frequency axis. The mathematical analysis behind this property was presented in [9]. A very flexible algorithm which can be fitted to perform a well-localized signal analysis either in frequency or time is obtained.

As can be seen in Fig. 8 a two-pass scheme has been implemented. The stationary analysis is performed with a great number of divisions per octave (more than 16) and the transient analysis is performed with a few divisions per octave (less than 6). Those numbers can be adjusted for every need. In this way we obtain a compact additive model of the signal with the same structure for the stationary and for the non-stationary part of the audio signal.

In Figs. 7 and 8 a partial is the set of wavelet coefficients around a maximum of the modulus of the transform grouped when the energy of the modulus of the transform is high enough. To detect a partial the modulus of the coefficients are calculated and summed in the time axis. The result is the *scalegram* of the signal. A partial is the number of consecutive bands in which the value of the scalegram is above a certain limit. Then, the sum of the complex wavelet coefficients involved in each partial from its lower frequency limit to its upper one is done. The modulus of these
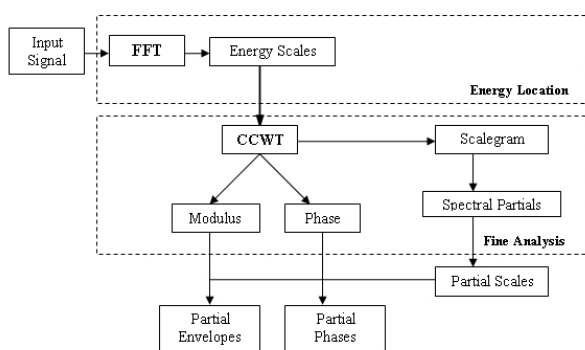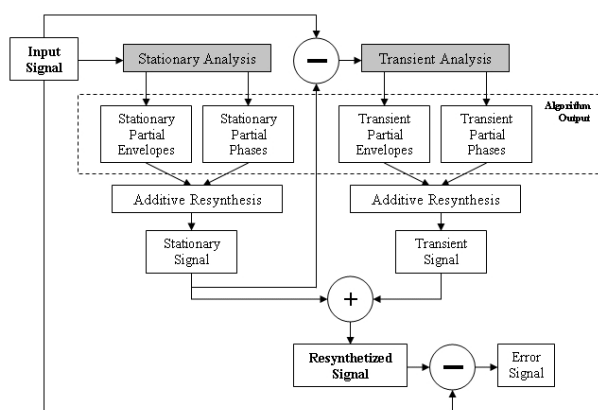
Figure 7: Analysis algorithm block diagram.



Figure 8: Stationary plus transient algorithm block diagram.

summed coefficients is the instantaneous amplitude of the partial and the phase of the grouped coefficients is the instantaneous phase of the partial. A more theoretical view of this work has been presented in [10]. Unlike Serra's work [11] the error signal we obtain (see Fig. 8) is not modelled.

The algorithm output is, then, the canonical pair of every detected partial. This algorithm is very suitable for detecting the amplitudes and frequencies variations presented in Sec. 2.

## 4. BANDPASS FILTERING RESULTS

Trying to cover the possibilities detailed in Sec. 2, a set of five different signals has been generated to test the algorithm and to find its limits. For computing purposes we have chosen the number of divisions per octave equal to 16 for all the analyzed cases below and the presented results correspond merely to the stationary analysis of Fig. 8.

### 4.1. The sum of pure cosines

We have analyzed three signals, each one composed by the sum of two cosines. In the first one the frequencies of the two cosines are 440 Hz and 880 Hz and the amplitudes are 0.5 and 0.499, respectively (two A notes separated by one octave). In the second signal we have chosen the same amplitudes and two constant frequencies

of 440 Hz and 493.88 Hz (notes A and B). Finally, the third signal has constant amplitudes of 0.5 and 0.499 and constant frequencies of 440 Hz and 455 Hz, respectivelly. These signals are shown in Fig. 9.
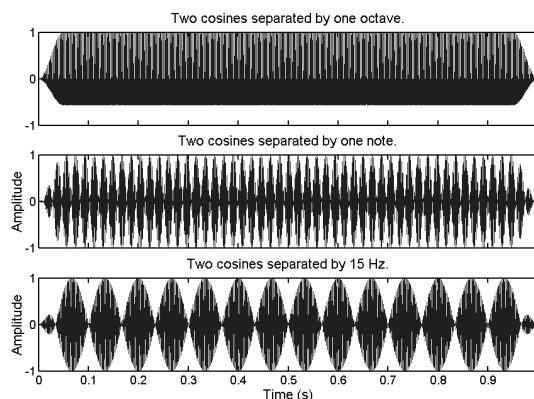


Figure 9: (a) Two cosines with frequencies of 440 Hz and 880 Hz, (b) frequencies of 440 Hz and 493.88 Hz (notes A and B) and (c) frequencies of 440 Hz and 455 Hz.

For the signal in top of Fig. 9 our algorithm finds two main partials. The obtained instantaneous amplitudes and frequencies of each partial are shown in Fig. 10. The bandpass filter used has resolution enough to find each component separately. Then we obtain the canonical pair of each component instead of obtaining the canonical pair of the whole signal. It can be seen a smoothing envelope in the upper graph of Fig. 10 due to a windowing process applied to the signal. This smoothing envelope has been applied to all the analyzed signals to avoid fast transient effects.
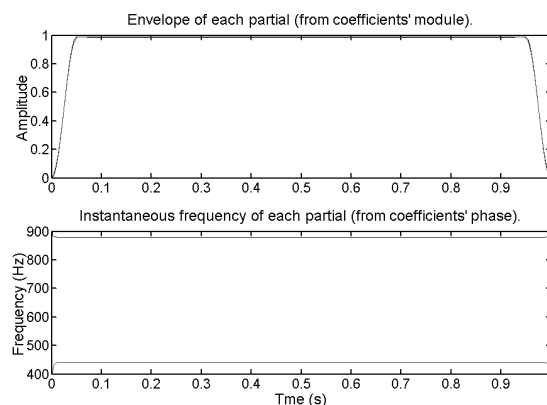


Figure 10: Instantaneous amplitude and frequency of the partials found for the top wave of Fig. 9.

The results of the analysis of the central wave of Fig. 9 are shown in Fig. 11. In this case the filterbank has not resolution enough to find each component completely isolated from the other. So, each partial mutually influences the other. The instantaneous amplitude oscillates and the oscillation in amplitude causes a frequency oscillation, as can be observed in Fig. 11. This fluctuation is not very large because the amount of energy of each sinu-

soidal component that contributes to the opposite band is only a part of the total energy of the component. This sensation is often described as *roughness*.
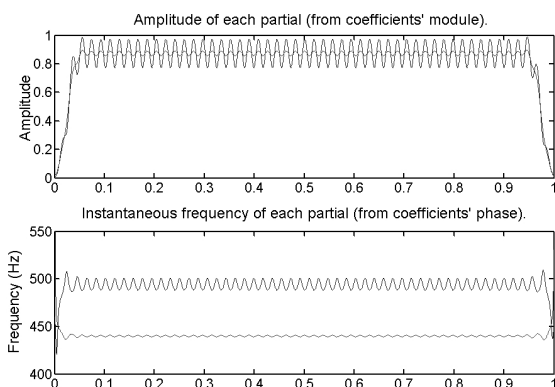


Figure 11: Instantaneous amplitude and frequency of the partials found for the central wave of Fig. 9.

In Fig. 12 the instantaneous amplitude and instantaneous frequency of the bottom wave of Fig. 9 is shown. In this case, both components are inside the same analysis band. They are, then, indistinguishable as different signals. It can be seen that the instantaneous amplitude and frequency obtained with our algorithm is quite similar to the canonical pair behaviour proposed in Sec. 2 (see Fig. 3). It is a logical consequence of the complex band-pass filtering: when the filter cannot separate the components, it tends to find a global expression for the whole input signal, like the canonical pair does. We have found a representation of these two cosines more related with our perceptual ability than the one given by a high-resolution spectral analysis method.
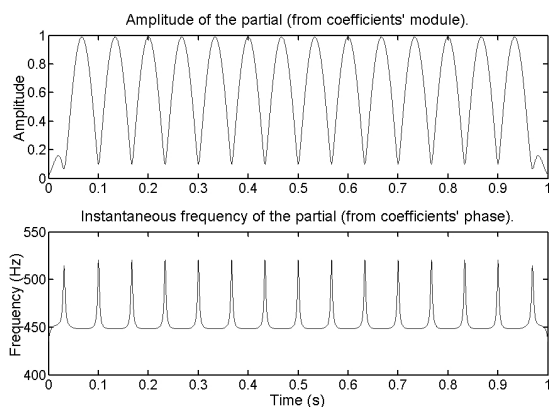


Figure 12: Instantaneous amplitude and frequency of the partials found for the bottom wave of Fig. 9.

**4.2. Chirps analysis**

Finally, we have generated and analyzed two different signals composed, respectively, by the sum of two lineal chirps and two exponential chirps like the ones proposed in Eqs. 10 and 11. In

these cases $A_1(t)$ and $A_2(t)$ are windowed amplitudes of maximum value 0.5 and 0.499, respectively. The corresponding signals have been depicted in Fig. 13. The values of the parameters in the linear case are: $a_1 = 400$, $b_1 = 415$ and $a_2 = 400$, $b_2 = 400$. In the exponential case the chosen parameters are: $B_1 = 414.92$, $C_1 = 1.074$ and $B_2 = 399.97$, $C_2 = 2.098$.

In Fig. 14 we can see the instantaneous amplitude and frequency of the sum of the two linear chirps on the top of Fig. 13. A deep amplitude modulation and a frequency beating is observed. Also the constant linear change in frequency is made clear. It can be seen the agreement in the obtained representation with respect to the results presented in Fig. 5.
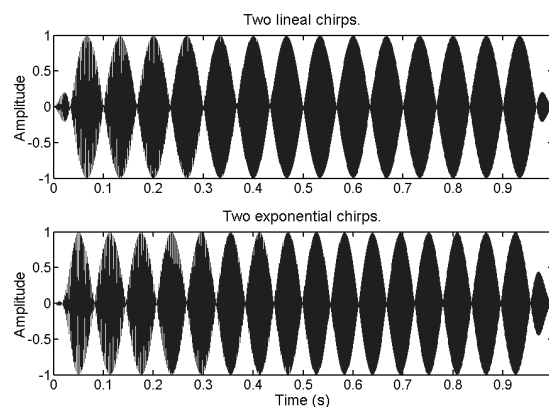


Figure 13: (a) Signal composed by the sum of two linear chirps. (b) Signal composed by the sum of two exponential chirps.
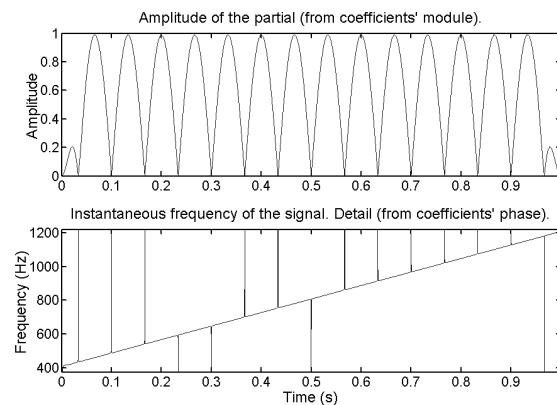


Figure 14: Instantaneous amplitude and frequency of the double lineal chirp signal.

Finally, in Fig. 15 the instantaneous amplitude and frequency of the composed exponential chirps at the bottom graph of Fig. 13 are presented. It can be seen the amplitude modulation and the exponential change in frequency. Again a comparison with the graphs presented in Fig. 6 reveals the good agreement between the theoretical analysis and the practical results.
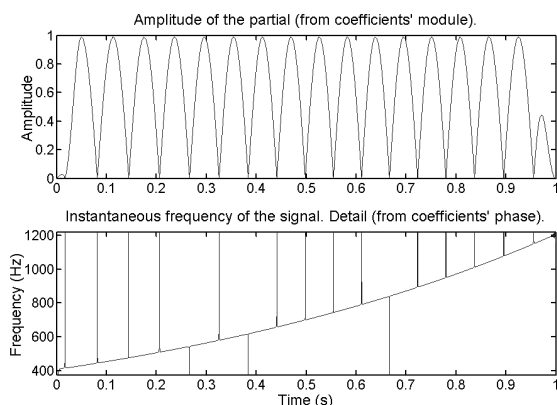
Figure 15: Instantaneous amplitude and frequency of the double exponential chirp.

## 5. CONCLUSIONS

In this work we have obtained a signal representation that includes the modulation in amplitude and frequency that appears when the frequency distance of some interacting tones is under the bandwidth of a bandpass filter. This representation has been motivated by the behaviour of the human ear in terms of the critical bands. A mathematical analysis modelling the interacting sinusoids in terms of the canonical pair of the signals has been made. To obtain such a representation a complex bandpass filtering algorithm has been implemented using the CCWT. A bandwidth adjustable algorithm is presented.

It has been shown that if the components of a composite signal are distant enough in frequency, our complex bandpass-filtering algorithm can detect them as different partials and the canonical pair model of the whole signal is not a good signal representation. As components get closer, the chosen filter size doesn't allow to completely discriminating them, and a representation in instantaneous amplitude and instantaneous phase is found by our algorithm. The bandwidth of the filterbank can be adjusted to simulate the critical bandwidth of the human hearing.

The computed instantaneous amplitudes and frequencies show the same behaviour that the obtained ones by analyzing the signal in terms of the canonical pair if the frequencies involved are close enough.

A question still remains open. Is it really necessary to detect each frequency component, or the obtained model is good enough (or even better) than a high spectral resolution one?

## 6. ACKNOWLEDGEMENTS

## 7. REFERENCES

[1] F. A. Everest, *Handbook for Sound Engineers. The New Audio Cyclopedia*, chapter 2, Psychoacoustics, pp. 25–42, SAMS, 2nd. edition, 1991.

[2] I. Daubechies, *Ten Lectures on wavelets*, vol. 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*, SIAM, 1992.

[3] R. Kronland-Martinet, "The Wavelet Transform for Analysis, Synthesis and Processing of Specch and Music Sounds," *Computer Music Journal*, vol. 12, no. 4, 1988.

[4] P. Guillemain and R. Kronland-Martinet, "Characterization of Acoustic Signals through Continous Linear Time-Frequency Representations," *Proceedings of the IEEE*, vol. 84, no. 4, Apr. 1996.

[5] N. Delprat, B. Escudié, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian, and B. Torrésani, "Asymthotic Wavelet and Gabor Analysis: Extraction of Instantaneous Frequencies," *IEEE Tran. on Inform. Theory*, vol. 38, no. 2, Mar. 1992.

[6] R. A. Carmona, W. L. Hwang, and B. Torrésan, "Characterization of signals by the ridges of their wavelet transforms," *IEEE Transactions on Signal Processing*, vol. 45, no. 10, pp. 2586–2590, Oct. 1997.

[7] R. A. Carmona, W. L. Hwang, and B. Torrésan, "Multiridge detection and time-frequency reconstruction," *IEEE Transactions on Signal Processing*, vol. 47, no. 2, pp. 480–492, Feb. 1999.

[8] B. Boashash, "Estimating and Interpreting the Instantaneous Frequency of a Signal. Part 1: Fundamentals," *Proceedings of the IEEE*, vol. 80, no. 4, Apr. 1992.

[9] J. R. Beltran and F. Beltran, "Additive Synthesis Based on the Continuous Wavelet Transform: a Sinusoidal Plus Transient Model," *Proc. of the 6th International Conference on Digital Audio Effects (DAFx-03)*, 2003.

[10] J. R. Beltran and J. Ponce de Leon, "Analysis and synthesis of sounds through complex filter banks," *Submmited to the 118th Convention of the Audio Engineering Society. (AES'05)*, 2005.

[11] X. Serra and J. O. Smith, "Spectral modelling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition," *Computer Music Journal*, vol. 14, no. 4, pp. 12–24, 1990.