Journal of Zhejiang University SCIENCE A
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Continuous frequency and phase spectrograms: a study of their 2D and 3D capabilities and application to musical signal analysis

Laurent NAVARRO[1], Guy COURBEBAISSE[†‡2], Jean-Charles PINOLI[1]

([1]*Ecole Nationale Supérieure des Mines de Saint-Etienne, Centre Ingénierie et Santé (CIS) and LPMG, UMR 5148 CNRS, 158 cours Fauriel, 42023 Saint-Etienne cedex 2, France*)

([2]*Medical Imaging Research Center (CREATIS-LRMN), UMR 5220 CNRS, U630 INSERM, Université Lyon 1, INSA Lyon, Bat. Blaise Pascal, 7 Av. Jean Capelle, 69621 Villeurbanne, France*)

[†]E-mail: guy.courbebaisse@creatis.insa-lyon.fr

Received June 22, 2007; revision accepted Sept. 19, 2007; published online Jan. 11, 2008

Abstract: A new lighting and enlargement on phase spectrogram (PS) and frequency spectrogram (FS) is presented in this paper. These representations result from the coupling of power spectrogram and short time Fourier transform (STFT). The main contribution is the construction of the 3D phase spectrogram (3DPS) and the 3D frequency spectrogram (3DFS). These new tools allow such specific test signals as small slope linear chirp, phase jump and small frequency jump to be analyzed. An application case of musical signal analysis is reported. The main objective is to detect small frequency and phase variations in order to characterize each type of sound attack without losing the amplitude information given by power spectrogram.

Key words: Frequency spectrogram (FS), Phase spectrogram (PS), Time-frequency representations, Musical signals
doi:10.1631/jzus.A072140 **Document code:** A **CLC number:** TN911.6

INTRODUCTION

Time-frequency representations are well adapted to the analysis of non-stationary signals (Cohen, 1989). However, the classical time-frequency methods are not efficient for the detection of small frequency modulations, such as the Spectrogram (Qian and Chen, 1996), which is subjected to the Heisenberg- Gabor uncertainty principle (Oliveira and Barroso, 2000), or the Wigner-Ville Distribution and its inherent interference terms (Ville, 1948; Cohen, 1989). Some authors tried to eliminate these limitations, in particular like Auger and Flandrin (1995). Their main idea of the reassignment method is to take into account the phase information in order to improve the accuracy of the time-frequency representations to estimate the instantaneous frequency.

It is interesting to see the phase content of the short time Fourier transform (STFT) (Oppenheim and Lim, 1981). However, this is not directly graphically accessible due to the large amount of information. Léonard (2005) presented an original method that allows representing the phase information in a relevant way that he called "phase spectrogram" (PS). At the same time he defined the "frequency spectrogram" (FS), which is a time-frequency representation computed only from the phase of STFT.

The present paper brings a new lighting on FS and PS, in particular concerning the continuous form of these representations. This paper also introduces new 3D time-frequency representations called the "3D frequency spectrogram" (3DFS) and the "3D phase spectrogram" (3DPS), which allow exploring the time-frequency plane in an original way. Capabilities of 3DFS and 3DPS are presented with the help of specific test signals and musical acoustics signals, since quasi-periodical signals and phase information are of main importance (Wu and Perio, 1979) in mu-

[‡] Corresponding author

sical signal analysis. Three musical acoustic signals have been chosen to represent the three types of string instruments: classical acoustic guitar signal for pinched strings, piano for struck strings and violin for rubbed strings.

Section 2 of this paper will present the method of calculating FS and PS and their continuous forms. Section 3 will report some results obtained with specific tests signals. Section 4 is dedicated to the introduction of 3DFS and 3DPS, and highlights their improvements compared with FS and PS. In Section 5, the two methods will be applied to musical acoustic signals for illustrating their relevance in practical situations. The concluding part will be a synthesis on these new representations.

PHASE AND FREQUENCY SPECTROGRAMS

Frequency spectrogram (FS)

The discrete FS and PS have been introduced recently by Léonard (2005). The present paper's contribution is linked to the continuous formulation for PS and FS representations. FS is presented first since it is the initial step for both FS and PS representations (PS is computed starting from FS). All computing actions are done using Matlab® software and the ISIS Time-Frequency Toolbox (Auger *et al*., 1998). The first step of the method is the calculation of STFT (Qian and Chen, 1996) defined by:

$$STFT_x(t,f)=\int_{-\infty}^{+\infty}x(\tau)h^*(\tau-t)\mathrm{e}^{-2\mathrm{i}\pi f\tau}\mathrm{d}\tau, \qquad (1)$$

where $h(t)$ is a window function (* denotes the conjugate symbol), t the time, f the frequency and $x(t)$ the signal to be transformed. *STFT* is a complex- valued function allowing the representation of both phase and magnitude parts of the signal over time and frequency.

The phase of STFT is then defined by:

$$\phi_{STFT_x}(t,f)=\arctan\left(\frac{\mathrm{Im}(STFT_x(t,f))}{\mathrm{Re}(STFT_x(t,f))}\right). \qquad (2)$$

Then, FS is a time-derivative of the phase of STFT:

$$FS_x(t,f)=\frac{1}{2\pi T}\frac{\mathrm{d}\phi_{STFT_x}(t,f)}{\mathrm{d}t}, \qquad (3)$$

where T is the sampling period.

Indeed, this can be linked with the well-known relation between phase and frequency:

$$f=\frac{1}{2\pi}\frac{\mathrm{d}\phi(t)}{\mathrm{d}t}, \qquad (4)$$

which can be interpreted in terms of instantaneous frequency (IF), and also related to the relation between angular position and rotation frequency in mechanics.

In order to control the number of "angular frequency rotations" in the time-frequency plane for a given frequency modulation signal, FS is multiplied before the wrapping operation by an integer number called the "gain parameter" and denoted P. The wrapping operation is realized to adjust each frequency rotation pattern in the $[-\pi, \pi]$ range. In addition, in order to normalize the frequency, the result must be multiplied by $1/(2\pi TP)$. So, FS can be expressed as:

$$FS_{x,P}(t,f)=\frac{1}{2\pi TP}\underline{P\frac{\mathrm{d}\phi_{STFT_x}(t,f)}{\mathrm{d}t}}, \qquad (5)$$

where the underline notation corresponds to the phase-wrapping operation (Léonard, 2005).

Finally, a threshold of FS is performed conditionally to the power spectrogram energy in order to only keep and show the information for the relevant amplitudes.

Let us recall that the power spectrogram for a signal $x(t)$ is defined by

$$S_x(t,f)=\left|STFT_x(t,f)\right|^2.$$

Phase spectrogram (PS)

The PS is computed by integration starting from the FS. It is based on the same relation between angular position and rotation frequency but in reverse direction as explained hereafter.

First, the FS is indexed on a reference time denoted *tr* and unwrapped (*Uw* corresponds to the phase unwrapping operation):

$$FS_{x,P,tr}(t,f)=\frac{1}{2\pi TP}\left[Uw\left(\underline{P\phi'(t,f)}\right)-Uw\left(\underline{P\phi'(tr,f)}\right)\right] \qquad (6)$$

with

$$\phi'(t,f)=\frac{\mathrm{d}\phi_{STFT_x}(t,f)}{\mathrm{d}t}. \quad (7)$$

Then, the PS referenced to time t=0 can be computed by integration:

$$\Phi_{x,P,tr}(t,f)=2\pi T\int_0^t FS_{x,P,tr}(\tau,f)\mathrm{d}\tau \quad (8)$$

and by combining Eqs.(6) and (8):

$$\Phi_{x,P,tr}(t,f)=\frac{1}{P}\int_0^t\left[Uw\left(\underline{P\phi'(\tau,f)}\right)-Uw\left(\underline{P\phi'(tr,f)}\right)\right]\mathrm{d}\tau$$
$$=\frac{1}{P}\left[\int_0^t Uw\left(\underline{P\phi'(\tau,f)}\right)\mathrm{d}\tau-K(f)t\right] \quad (9)$$

with

$$K(f)=Uw\left(\underline{P\phi'(tr,f)}\right).$$

Note that $K(f)$ does not depend on t.

Next, the phase must be referenced to the time tr in order to create a zero-phase reference:

$$PS_{x,P,tr}(t,f)=\Phi_{x,P,tr}(t,f)-\Phi_{x,P,tr}(tr,f) \quad (10)$$
$$=\frac{1}{P}\left[\int_0^t Uw\left(\underline{P\phi'(\tau,f)}\right)\mathrm{d}\tau-K(f)t-\int_0^t Uw\left(\underline{P\phi'(tr,f)}\right)\mathrm{d}\tau+K(f)tr\right]$$
$$=\frac{1}{P}\left[\int_0^t Uw\left(\underline{P\phi'(\tau,f)}\right)\mathrm{d}\tau-K(f)t-K(f)t+K(f)tr\right]$$
$$=\frac{1}{P}\left[\int_0^t Uw\left(\underline{P\phi'(\tau,f)}\right)\mathrm{d}\tau-2K(f)t+K(f)tr\right]. \quad (11)$$

For t=tr,

$$PS_{x,P,tr}(tr,f)=\frac{1}{P}\left[K(f)tr-2K(f)tr+K(f)tr\right]=0. \quad (12)$$

So, PS is null at t=tr for all frequency-values: a zero-phase reference has thus been created.

TEST SIGNALS ANALYSIS

Case of the sinusoidal signal

As an example, let $x(t)$ be a sinusoid signal composed by a single constant frequency f_0 called the nominal frequency:

$$x(t)=\cos(2\pi f_0 t+\delta). \quad (13)$$

At this frequency, $\phi'(t,f_0)$=$2\pi f_0$ remains constant for every time value t, thus

$$Uw\left(\underline{P\phi'(t,f_0)}\right)=Uw\left(\underline{P\phi'(tr,f_0)}\right). \quad (14)$$

Then using Eqs.(7), (9) and (11),

$$FS_{x,P,tr}(t,f_0)=0,\ \ PS_{x,P,tr}(t,f_0)=0,$$
$$\Phi_{x,P,tr}(t,f_0)=\Phi_{x,P,tr}(tr,f_0).$$

Therefore, for the nominal frequency f_0 of a sinusoidal signal, PS is null: the constant phase progression $2\pi f_0 t$ has been cancelled.

Test signal 1: small slope linear chirp

The frequency and phase spectrograms allow to reveal characteristics that cannot be observed with other time-frequency representations. The first simple example is a signal with a few-varying frequency modulation, more accurately a linear chirp with a very small slope. For a high value of the gain P (here P=30), FS exhibits a small frequency modulation of 1 Hz (Fig.1). The color bar on the right of the representation is a frequency zoom scale that is proportional to the gain P. It displays the frequency shifts of the signal. On the power spectrogram of this signal (Fig.2), the frequency modulation cannot be observed, since it is too small.

Test signal 2: phase jump

A second relevant test signal is a single frequency with a phase jump. The PS is well adapted to observe phase jumps on this type of signal, but it is also very interesting as it allows to quantify these phase jumps. The color bar on the right of the representation is a phase zoom scale that is proportional to the gain P. It displays the phase shifts of the signal. On the PS representation (Fig.3), the phase jump of 1.5 rad is shown. On the power spectrogram of the same signal (Fig.4), such information cannot be shown. Indeed, only a small break is perceptible but it cannot be determined if it is a phase jump or not.

Test signal 3: small frequency jump

Since the link between PS and FS is an integral relation, in general, the information contained in one of these representations can be observed on the second. In fact, this is not clear on FS (Fig.5), but is obvious on PS (Fig.6).

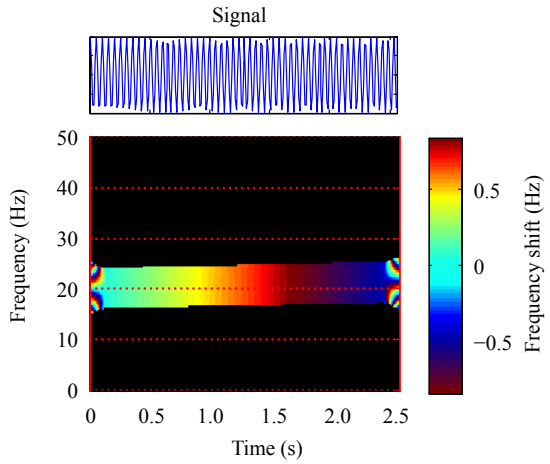

Fig.1 Frequency spectrogram (FS) of a linear chirp signal with a small slope (1 Hz/2.5 s). *P*=30, sampling rate: 100 Hz. Computed with a Hanning window of 31 points

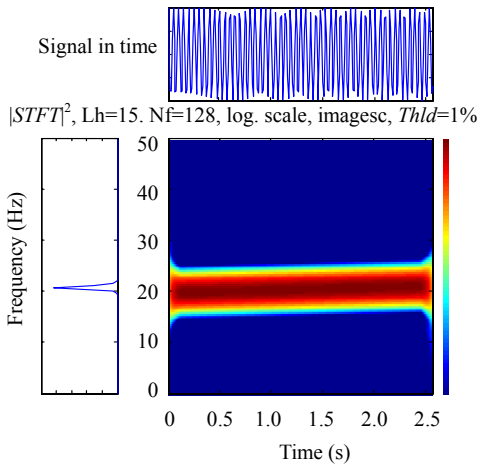

Fig.2 Power spectrogram (PS) of a linear chirp signal with a small slope (1 Hz/2.5 s). Sampling rate: 100 Hz. Computed with a Hanning window of 31 points

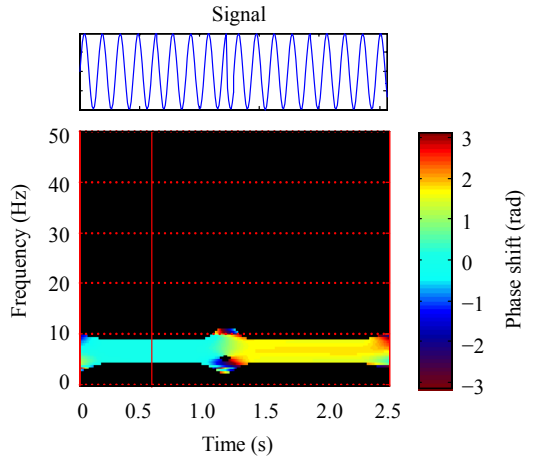

Fig.3 Power spectrogram (PS) of a signal containing a phase jump (1.5 rad). *P*=1, sampling rate: 100 Hz. Computed with a Hanning window of 61 points

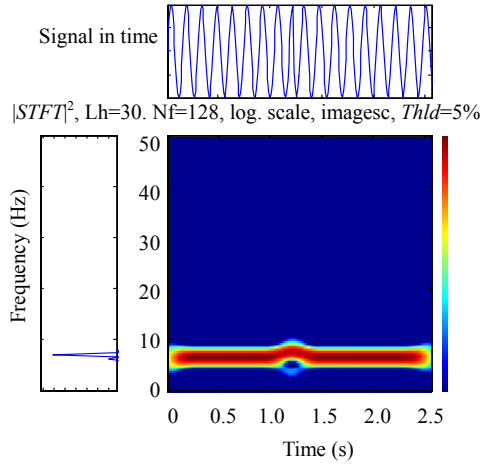

Fig.4 Power spectrogram (PS) of a signal containing a phase jump (1.5 rad). Sampling rate: 100 Hz. Computed with a Hanning window of 61 points

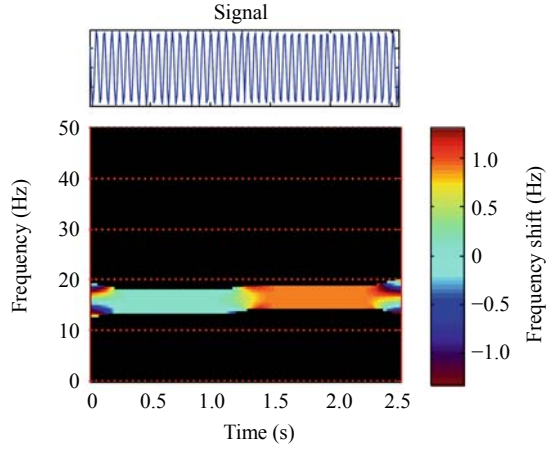

Fig.5 FS of a signal containing a small frequency jump (Δ*f*=0.8 Hz). *P*=19, sampling rate: 100 Hz. Computed with a Hanning window of 61 points

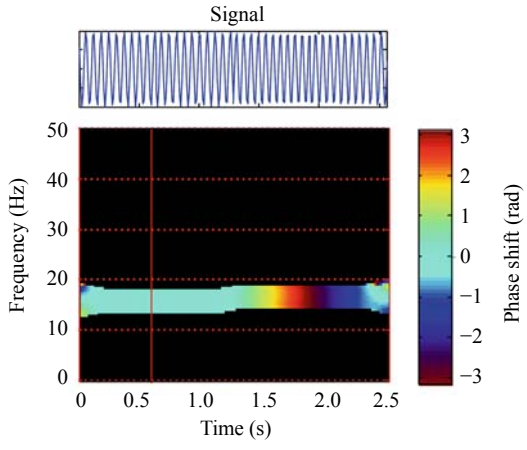

Fig.6 PS of a signal containing a small frequency jump (Δ*f*=0.8 Hz). *P*=1, sampling rate: 100 Hz. Computed with a Hanning window of 61 points

Let a signal be composed of two parts: the first part is a single frequency component and the second part is also a single frequency component but with a slightly higher frequency. This signal allows the frequency jump to be quantified. On the second part, the delay Δt between two fringes lasts 1.25 s. That corresponds to a frequency shift of

$$\Delta f = \Delta\phi(t)/(2\pi\Delta t) = 0.8 \text{ Hz},$$

where Δt=1.25 s and $\Delta\Phi$=2π.

This value can also be read on the frequency color bar of FS (Fig.5). So the differential frequency can be also precisely quantified with the PS.

3D PHASE AND FREQUENCY SPECTROGRAMS

It has been shown that a threshold must be performed in order to only keep the information contained for the relevant amplitudes of the power spectrogram.

3DFS and 3DPS are based on the same calculus as that of FS and PS, except that no threshold is performed, and therefore all information is remained. The principle is to map FS and PS on the power spectrogram. The result is a 3D representation with time, frequency, amplitude, and small frequency modulations or time, frequency, amplitude, and phase shift of the signal. This representation allows a new reading of the time-frequency plan. In the case of 3DPS, it is possible to define a time-frequency-phase surface. With classical time-frequency representations, time frequency trajectories are analysed. With this new representation, it is possible to analyse geodesic trajectories that can be computed with the Bresenham's algorithm (Bresenham, 1965) for instance.

Fig.7 is an example of a 3DFS of the third test signal (small frequency jump, Δf=0.8 Hz). The readability is improved, since all amplitudes are taken into account but not only the crests. Edges phenomena appear clearly with a succession of horizontal fringes. They can be compared with impulses, and a "frequency distortion" appears at the bottom of the representation (low energy) when the frequency jump occurs. The 3DPS of the same signal (Fig.8) exhibits some phase distortion at the bottom of the representation, on the second part of the signal between 1.5 and 2.5 s.

With the threshold operation, this phase distortion cannot be seen, and the information that it contains is lost. The zero-phase phenomena for all frequency values at the phase reference can also be observed directly on the 3DPS.

APPLICATION TO MUSICAL ACOUSTIC SIGNALS

The need of time-frequency representations has been necessary since the beginning of musical signal analysis. As a proof, the power spectrogram used to be called the sonagram (Qian and Chen, 1996).

3DFS and 3DPS are well adapted to the analysis of musical acoustic signals, since typical characteristics such as few varying frequency modulations are often present in this kind of signals.

Three musical signals of the same musical note (fundamental frequency: *A3*=220 Hz) associated with three musical instruments have been processed, which are classical acoustic guitar, piano and violin. Only the attack part of these signals have been kept in order to differentiate their frequency and phase variations.

All signals have been recorded at 16 bits, 44.1 kHz using an Appex 170 pencil microphone, a Phonic MM1002 console and Audacity software. The sampling rate has been reduced to 2205 Hz after recording, in order to facilitate the computing.

Classical acoustic guitar signal: pinched strings

The 3DFS of this signal (Fig.9) exhibits no frequency variation on both fundamental frequency and harmonics, in spite of an important amplitude variation. Horizontal fringes can be practically observed from the minimum to the maximum frequency, which means that the attack's behaviour mimics a Heaviside function. An example of what can be considered as a Heaviside function can be observed at the end of the 3DFS, when the signal is brutally cut: the same phenomenon of horizontal fringes still appears.

The 3DPS (Fig.10) of the classical guitar signal exhibits an interesting feature: it seems that there is some phase variations on the fundamental frequency whereas no one appears on harmonics. It could be a limitation of this kind of representation. Indeed, the mapping operation stretches the phase information. So the small variations appear very wide for the highest amplitudes.

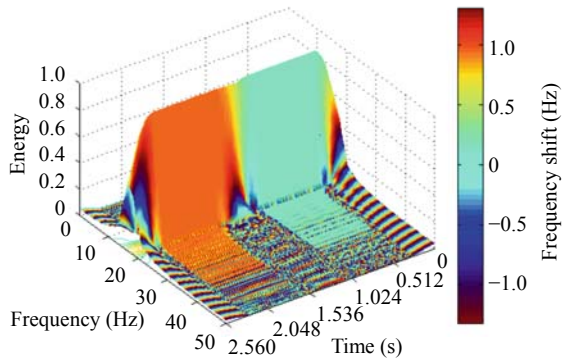

Fig.7 3DFS of a signal containing a small frequency jump (Δf=0.8 Hz). P=19, sampling rate: 100 Hz. Computed with a Hanning window of 61 points

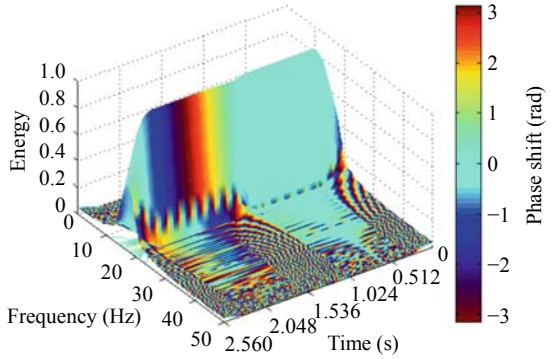

Fig.8 PS of a signal containing a small frequency jump (Δf=0.8 Hz), with the same phase reference as Fig.6. P=1, sampling rate: 100 Hz. Computed with a Hanning window of 61 points

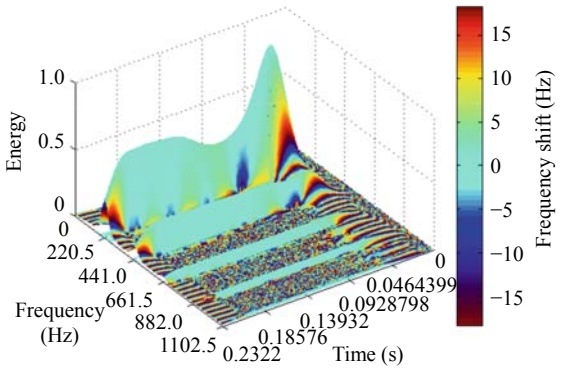

Fig.9 3DFS of a classical guitar signal. P=30, sampling rate: 2205 Hz. Computed with a Hanning window of 121 points

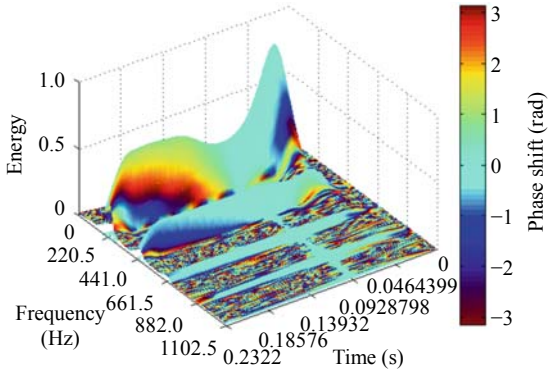

Fig.10 3DPS of a classical guitar signal. P=1, sampling rate: 2205 Hz. Computed with a Hanning window of 121 points

Piano signal: struck strings

Contrary to the classical acoustic guitar signal, the piano signal 3DFS (Fig.11) exhibits some frequency variations for both fundamental frequency and harmonics. This fact is interesting, since intuitively struck strings instruments should be more stable in frequency behaviour. The second important thing to be noticed is that the Heaviside-like phenomenon does not appear for the piano sound. That means the attack of the struck string is not so fast as that of the pinched string.

The 3DPS (Fig.12) exhibits some variations of the harmonic components phase. The first harmonic tends to be constant, but the second harmonic's phase decreases and the third one increases. The total difference between the phases of harmonics 2 and 3 reaches 4 rad at the end of the 3DPS.

Violin signal: rubbed strings

The third musical instrument tested is the violin. Its signal is interesting, since it corresponds to the third type of strings instruments: the rubbed strings instruments.

The 3DFS (Fig.13) of this signal exhibits a frequency variation of 10 Hz on the third harmonic. This variation is important, and the amplitude of this harmonic is high, so it will result in an audible variation on the sound. This variation might be due to the player's finger position or the beginning of a vibrato.

Like other rubbed strings instruments, the attack is very soft and the sound increases slowly.

On the 3DPS (Fig.14) of the violin sound, only the phase variation corresponding to the frequency variation of 10 Hz can be observed.

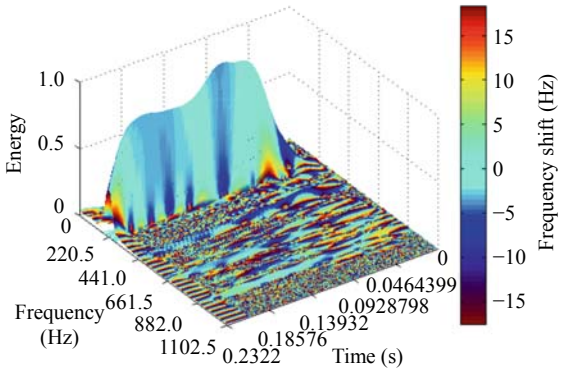

Fig.11 3DFS of a piano signal. *P*=30, sampling rate: 2205 Hz. Computed with a Hanning window of 121 points

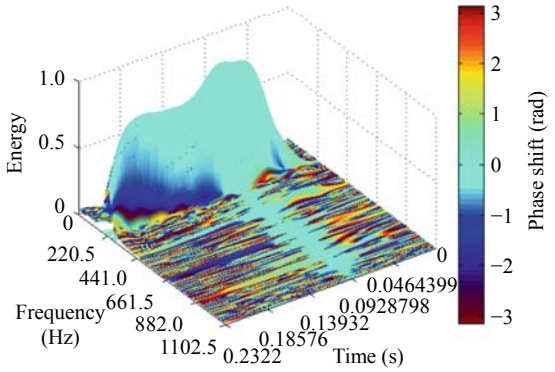

Fig.12 3DPS of a piano signal. *P*=1, sampling rate: 2205 Hz. Computed with a Hanning window of 121 points

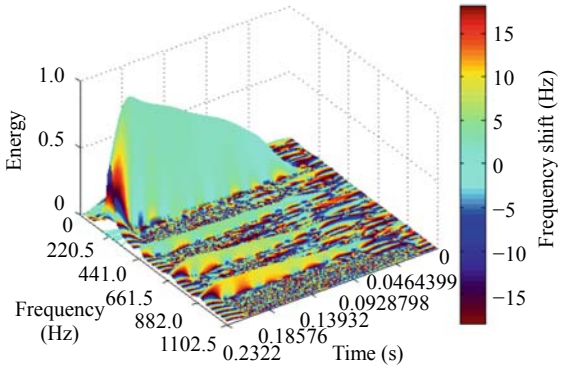

Fig.13 3DFS of a violin signal. *P*=30, sampling rate: 2205 Hz. Computed with a Hanning window of 121 points

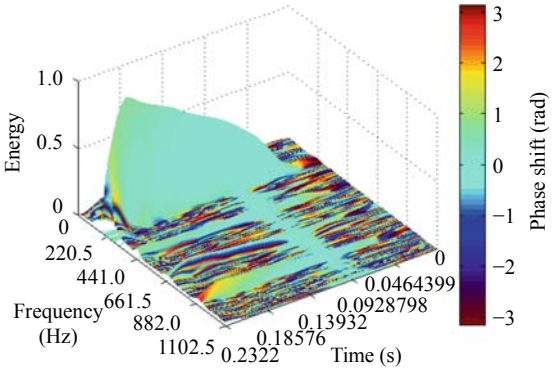

Fig.14 3DPS of a violin signal. *P*=1, sampling rate: 2205 Hz. Computed with a Hanning window of 121 points

CONCLUSION

A continuous expression for PS and FS representations is proposed in this paper. Some properties of the frequency and phase spectrograms have been studied. In particular, it has been shown that the phase spectrogram of a sinusoidal frequency is null for the nominal frequency whatever the time *t*, so that the constant phase progression of a sinusoid is cancelled with this representation.

Examples of the capabilities of the PS to reveal important characteristics of tests signals such as small frequency variations and phase jumps have been presented. It has been shown that the power spectrogram cannot give such results, in particular concerning the phase content.

Finally, 3DFS and 3DPS have been introduced, and these new representations have been tested on musical acoustics signals representing the three types of string instruments: classical acoustic guitar signal for pinched strings, piano for struck strings and violin for rubbed strings. These new representations allow the observation of frequency and phase characteristics that could not be observed simultaneously before. In spite of the fact that they are all harmonic signals, the 3DFS and 3DPS differentiate them. This can be a strong improvement in musical signal analysis, since applications such as pitch detection or instrument synthesis are the important today stake.

Further studies will be focused on other instrument signals, especially these like wind instruments (Wu and Perio, 1981). Biomechanical and biomedical signals, such as GRF (Ground Reaction Forces), heart beat and other biological signals that have stable reference frequencies with less variation, will also be tested.

In fact, this new kind of representation may be considered a tool for all signals whose small frequency and phase variations need to be analysed.

References

Auger, F., Flandrin, P., 1995. Improving the readability of time-frequency and time-scale representations by the reassignment method. *IEEE Trans. on Signal Processing*, **43**:1068-1089. [doi:10.1109/78.382394]

Auger, F., Flandrin, P., Lemoine, O., Gonçalvès, P., 1998. Time-Frequency Toolbox for MATLAB.

Bresenham, J.E., 1965. Algorithm for computer control of a digital plotter. *IBM Systems Journal*, **4**:25-30.

Cohen, L., 1989. Time-frequency distributions—a review. *Proc. IEEE*, **77**:941-981. [doi:10.1109/5.30749]

Léonard, F., 2005. Phase spectrogram and frequency spectrogram as new diagnostic tools. *Mechanical Systems and Signal Processing*, **21**:125-137. [doi:10.1016/j.ymssp.2005.08.011]

Oliveira, P.M., Barroso, V., 2000. Uncertainty in the time-frequency plane. *Statistical Signal and Array Processing*, **1**:607-611. [doi:10.1109/SSAP.2000.870197]

Oppenheim, A., Lim, J., 1981. The importance of phase in signals. *Proc. IEEE*, **69**:529-541.

Qian, S., Chen, D., 1996. Joint Time-Frequency Analysis: Methods and Applications. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Ville, J., 1948. Théorie et applications de la notion de signal analytique. *Câbles et Transmissions*, **1**:61-74 (in French).

Wu, F., Perio, P., 1979. La phase en acoustique musicale. I. Analyse d'un signal quasi-périodique. *Journal de Physique*, **40**:799-810 (in French). [doi:10.1051/jphys:01979004008079900]

Wu, F., Perio, P., 1981. La phase en acoustique musicale. II. Le rayonnement des instruments à vent. *Journal de Physique*, **42**:627-633 (in French). [doi:10.1051/jphys:01981004205062700]