On the Structure of the Phase around the Zeros of the Short-Time Fourier Transform

Peter Balazs[1], Dominik Bayer[2], Florent Jaillet[3]

[1] *Acoustics Research Institute, 1040 Wien, Austria, Email: peter.balazs@oeaw.ac.at*

[2] *Acoustics Research Institute, 1040 Wien, Austria, Email: bayerd@kfs.oeaw.ac.at*

[3] *Acoustics Research Institute, 1040 Wien, Austria, Email: florent@kfs.oeaw.ac.at*

Introduction

The short-time Fourier transform (STFT) is a time-frequency representation widely used in audio signal processing. Two conventions are commonly used for the definition of the short-time Fourier transform:

$$V(f,g)(\tau,\omega) = \int f(t)\overline{g(t-\tau)}e^{-i\omega t}\,dt \tag{1}$$

and

$$W(f,g)(\tau,\omega) = \int f(t)\overline{g(t-\tau)}e^{-i\omega(t-\tau)}\,dt. \tag{2}$$

With convention (1), the transform uses a fixed phase reference at $t = 0$, whereas in convention (2), the phase reference at $t = \tau$ is sliding with the window. The two conventions are thus related by a simple phase term following the relation $W(f,g)(\tau,\omega) = e^{i\omega\tau}V(f,g)(\tau,\omega)$. In the following, we will use convention (2).

The interpretation of the modulus of the STFT is relatively easy, considering the fact that the spectrogram (defined as the square absolute value of the STFT) can be interpreted as a time-frequency distribution of the signal energy. This interpretation led to the important success of STFT in signal processing. In particular, it has been widely used for applications in speech processing and acoustics as a graphical tool for signal analysis.

The interpretation of the phase of the STFT is less obvious, and has thus hardly been considered in applications. Nevertheless the phase can be of particular interest for certain applications, as illustrated by important applications such as phase vocoder [1] or reassignment [3]. Furthermore, for applications requiring the perfect reconstruction of a signal from STFT coefficients, phase information is essential. For this type of applications, in particular for applications using Gabor multipliers [4], which motivated the present study, better understanding of the structure of the phase is necessary to improve the processing possibilities.

The phase derivative over time is of particular interest for analysis of signals containing sinusoidal components, as often encountered in acoustics [1]. This can be seen when considering the case of a single (complex) frequency defined by:

$$f(t) = e^{i\omega_0 t}, \tag{3}$$

for which the STFT is then given by:

$$W(f,g)(\tau,\omega) = e^{i\omega_0\tau}\overline{\hat{g}(\omega_0-\omega)}. \tag{4}$$

The derivative of the phase over time is then

$$\frac{\partial}{\partial\tau}\arg(W(f,g)(\tau,\omega)) = \omega_0, \tag{5}$$

which gives direct access to the value of the angular frequency of the signal.

Numerical Experiments for White Noise

We conducted numerical experiments on random signals, modelled as Gaussian white noise. In this case, only statistical properties of the phase were known so far. We were interested in general structural properties of the phase derivative.

Our experiments gave surprising results, as illustrated by the figures below. As can be seen, the time-frequency distribution of the values appears to be highly structured and in particular, the values of the phase derivative with high absolute values are concentrated around the zeros of the transform when looking at the modulus. Furthermore, the shape of the phase derivative seems to display a regularly recurring behavior in the neighborhood of the zeros, with a typical pattern repeating at each zero. When going from low to high frequencies, it presents a negative peak followed by a positive one. Choice of

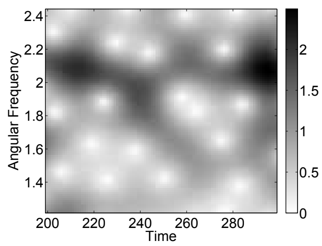

Figure 1: Observation for a Gaussian white noise, using a Gaussian window: modulus of the STFT

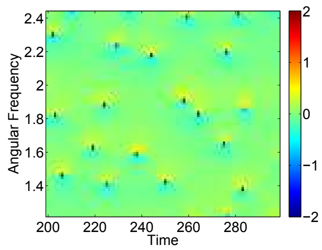

Figure 2: Gaussian white noise with Gaussian window: derivative over time of the phase of the STFT (1)

different windows yields qualitatively the same results.

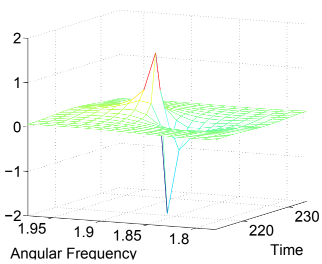

Figure 3: Gaussian white noise with Gaussian window: derivative over time of the phase of the STFT (2)

The phase derivative always shows the same characteristic pattern around the zeros of the transform.

Analytical Treatment of a Simple Example

Considering the signal given by

$$f(t) = e^{i\omega_1 t} + e^{i\omega_2 t} \tag{6}$$

and using a Gaussian window $g(t) = e^{-\frac{t^2}{2\sigma^2}}$, we can explicitly compute the expression of the STFT, which results in the formula:

$$W(f,g)(\tau,\omega) = \tag{7}$$
$$= \sqrt{2\pi}\sigma \left(e^{i\tau\omega_1} e^{-\frac{\sigma^2(\omega-\omega_1)^2}{2}} + e^{i\tau\omega_2} e^{-\frac{\sigma^2(\omega-\omega_2)^2}{2}} \right).$$

The zeros of this STFT are the points of coordinates (τ_k, ω_m) in the time-frequency plane, with $\omega_m = \frac{\omega_1+\omega_2}{2}$ and $\tau_k = \pi \frac{1+2k}{\omega_1-\omega_2}$ for $k \in \mathbb{Z}$.

The expression of the phase derivative for this signal is (cf.[2])

$$\frac{\partial}{\partial \tau} \arg(W(f,g)(\tau,\omega)) = \tag{8}$$
$$= \omega_m + \delta \tanh(s) \frac{1+\tan^2(\tau\delta)}{1+\tan^2(\tau\delta)\tanh^2(s)}$$

with $\delta = \frac{\omega_2-\omega_1}{2}$ and $s = \sigma^2(\omega-\omega_m)\delta$.

The phase derivative of this function (for values $\omega_1 = 0.5$, $\omega_2 = 1.5$ and $\sigma = 3$) around one of the zeros, depicted in the following figure, shows the same behavior observed before. The results for this analytical example confirm

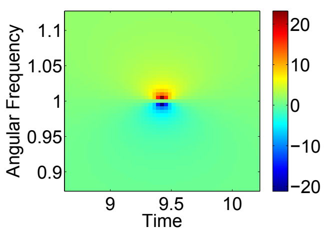

Figure 4: The analytical example: derivative over time of the phase of the STFT (1).

that the structure observed in the experiments are not only due to numerical artefacts, but present also for the continuous transform.

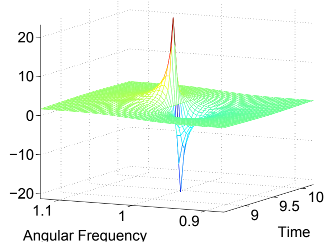

Figure 5: The analytical example: derivative over time of the phase of the STFT (2).

The Mathematical Explanation

The phenomenon is due to a certain analyticity property of the STFT with Gaussian window $g(t) = e^{-t^2/2}$ (cf. [5]):

$$V(f,g)(\bar{z}) = e^{i\tau\cdot\omega/2} \cdot Bf(z) \cdot e^{-|z|^2/4}, \qquad z = \tau + i\omega \in \mathbb{C},$$

with

$$Bf(z) = e^{-z^2/4} \int f(t) \cdot e^{t\cdot z - t^2/2}\, dt,$$

the **Bargmann transform** of f, an *entire* (= holomorphic on the entire complex plane) function.

The zeros of the STFT are the zeros of the Bargmann transform reflected along the real axis:

$$V(f,g)(z) = 0 \quad \Leftrightarrow \quad Bf(\bar{z}) = 0.$$

Apart from (harmless) non-zero factors and the reflection, $V(f,g)$ behaves exactly like the holomorphic function Bf.

THEOREM: The partial derivative of the phase ϕ of a holomorphic function $h = |h| \cdot e^{i\cdot\phi}$ shows the following behavior in a neighborhood of a zero $z_0 = x_0 + iy_0$ of h:

$$\lim_{z\to z_0} \frac{\partial\phi}{\partial x}(z) = \begin{cases} +\infty, & \text{if } z = x_0 + it,\ t \uparrow y_0 \text{ from below} \\ -\infty, & \text{if } z = x_0 + it,\ t \downarrow y_0 \text{ from above} \end{cases}$$

References

[1] Dolson, M.: The Phase Vocoder: a Tutorial. Computer Music Journal 10 (1986), 14-27

[2] Delprat, N., Escudie, B., Guillemain, P., Kronland–Martinet, R., Tchamitchian, P. and Torrésani, B.: Asymptotic Wavelet and Gabor Analysis: Extraction of Instantaneous Frequencies. IEEE Transactions on Information Theory 38 (1992), 644-664

[3] Flandrin, P., Auger, F. and Chassande–Mottin, E.: Time-Frequency Reassignment: from Principles to Algorithms. In: Papandreou-Suppappola, A. (Ed.), Applications in Time-Frequency Signal Processing, CRC Press, 2002, 179-203

[4] Feichtinger, H.G. and Nowak, K.: A First Survey of Gabor Multipliers. In: Feichtinger, H.G. and Strohmer, T. (Eds.), Advances in Gabor Analysis, Birkhäuser, 2003, 99-128

[5] Gröchenig, K.: Foundations of Time-Frequency Analysis. Birkhäuser, 2001