Appl. Comput. Harmon. Anal. 40 (2016) 610–621

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Letter to the Editor

The pole behavior of the phase derivative of the short-time Fourier transform

Peter Balazs [a], Dominik Bayer [a,*], Florent Jaillet [b], Peter Søndergaard [c]

[a] *Acoustics Research Institute, Austrian Academy of Sciences, Vienna A-1040, Austria*
[b] *Aix Marseille Université, CNRS, LIF UMR 7279, 13288, Marseille, France*
[c] *Oticon A/S, 2765 Smørum, Denmark*

A R T I C L E I N F O

Article history:
Received 14 April 2014
Received in revised form 12 September 2015
Accepted 6 October 2015
Available online 22 October 2015
Communicated by Richard Gundy

A B S T R A C T

The short-time Fourier transform (STFT) is a time–frequency representation widely used in audio signal processing. Recently it has been shown that not only the amplitude, but also the phase of this representation can be successfully exploited for improved analysis and processing. In this paper we describe a rather peculiar pole phenomenon in the phase derivative, a recurring pattern that appears in a characteristic way in the neighborhood around any of the zeros of the STFT, a negative peak followed by a positive one. We describe this phenomenon numerically and provide a complete analytical explanation.

1. Introduction

The *short-time Fourier transform* (STFT) [5,13] is a time–frequency representation widely used in audio signal processing. A common definition of the STFT[1] is

$$V(f,g)(x,\omega) = \int f(t)\overline{g(t-x)}e^{-2\pi i\omega t}\,dt. \tag{1}$$

The STFT $V(f,g)(x,\omega)$ provides information about the frequency content of the signal f at time x and frequency ω. The analyzing window g determines the resolution in time and frequency.

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 the 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 [21].

* Corresponding author.
E-mail address: bayerd@kfs.oeaw.ac.at (D. Bayer).

[1] This is the *frequency-invariant* STFT.

http://dx.doi.org/10.1016/j.acha.2015.10.001
1063-5203/

But the interpretation of the phase of the STFT is less obvious, and was thus hardly considered in applications for some time. The phase can be of particular interest for certain applications, as illustrated by important applications such as phase vocoder [12,7] or reassignment [18,1]. In digital image processing it is well known that the phase information of the discrete Fourier transform is at least as important as the amplitude information. In [19] it is shown that as long as the phase of the discrete Fourier transform of an image is retained and the amplitude is set to 1, the image can still be recognized. Similar effects can also be shown for acoustic signal depending on the parameters of the STFT [4].

For applications modifying the STFT coefficients, phase information is essential again. For these types of applications, in particular for the applications using Gabor frame multipliers [9,3] which motivated the present study, better understanding of the structure of the phase is necessary to improve the processing possibilities.

The phase of the STFT is usually not considered directly. In fact, it is more interesting to consider the phase derivative over time or frequency. Indeed, these quantities appear naturally in the context of reassignment [1] and manipulations of phase derivative over time is the idea behind the phase vocoder [7]. Their interpretation is easier, as the derivative of phase over time can be interpreted as local instantaneous frequency while the derivative of the phase over frequency can be interpreted as a local group delay.

To numerically compute the local instantaneous frequency, an unwrapping of the phase is needed to avoid discontinuities. This is the classical method used in [7,18]. Another method was found in [1]:

$$\frac{\partial}{\partial x}\arg(V(f,g)(x,\omega)) = \mathrm{Im}\left(\frac{V(f,g')(x,\omega)\overline{V(f,g)(x,\omega)}}{|V(f,g)(x,\omega)|^2}\right), \tag{2}$$

with $g'(t) = \frac{dg}{dt}(t)$. The benefit of this method is that is does not require unwrapping, instead the phase derivative is computed by pointwise operations using a second STFT based on the derivative of the window.

To understand the phase of the STFT more thoroughly, in particular for applications dealing with multipliers, see for example [22,23,20], we conducted related extensive numerical experiments. In the process we observed a rather peculiar phenomenon in the phase derivative, a recurring pattern that appears in a similar way in the neighborhood around any of the zeros of the STFT. The behavior of the phase derivative close to the singularity always shows the same characteristic shape, i.e., a negative peak followed by a positive one. We describe this phenomenon and provide a complete analytical explanation.

This paper is organized as follows: In Section 2 we report the numerical results. In Section 3 we give a short, instructive, analytical example for this behavior. In Section 4 we give the full analytical results.

Results in this paper have partly been reported at a conference [17], and a preprint of this paper has already been cited in [2].

2. Numerical observations

For noise, naturally only statistical properties of the phase are accessible. Some interesting results for the phase derivative have been shown in the context of reassignment. In [8], the following result is given: We consider a zero-mean Gaussian analytic white noise f such that

$$\mathrm{E}[\mathrm{Re}(f(t))\cdot \mathrm{Re}(f(s))] = \mathrm{E}[\mathrm{Im}(f(t))\cdot \mathrm{Im}(f(s))] = \frac{\sigma^2}{2}\delta(t-s) \tag{3}$$

and $\mathrm{E}[f(t)f(s)] = 0$ for any $(t,s)\in\mathbb{R}^2$, with its real and imaginary parts a Hilbert transform pair. Using a Gaussian window given by $g(t) = e^{-\pi\frac{t^2}{2\sigma^2}}$, the phase derivative over time of $V(f,g)$ is a random variable with distribution of the form:

$$\rho(v) = \frac{1}{2(1+v^2)^{\frac{3}{2}}}. \tag{4}$$