ON THE CONCEPT OF INSTANTANEOUS FREQUENCY

Paulo M. Oliveira (*) and Victor Barroso (**)

(*) Escola Naval - DAEl, Base Naval de Lisboa, Alfeite, 2800 Almada, Portugal
(**) Instituto Superior Técnico ISR/DEEC, Torre Norte, Piso 7 Av. Rovisco Pais, P-1096 Lisboa Codex Portugal

ABSTRACT

The concept of Instantaneous Frequency is still not clearly defined. Current operational definitions give rise to physical paradoxes, difficulting proper interpretation of the obtained results. In this paper, we discuss why those paradoxes appear, and show how they can be avoided. We introduce a new definition of Instantaneous Frequency, which yields physically consistent results. This is confirmed with the help of several examples.

1. INTRODUCTION

The concept of Instantaneous Frequency (IF_t, where the subscript emphasizes a possible time-dependency) is still evading proper definition [2]. To start with, the concept of frequency (based on pure sinusoids, extending from $-\infty$ to $+\infty$) is not easily combined with words such as *instantaneous.* This difficulty can be satisfactorily solved if we fall back to the physical notion of frequency as the rate of change of the phase. This notion poses additional problems: **(i)** determining which phase to use, and **(ii)** defining the phase of a real signal. For complex signals, the phase function is normally used. For real signals, there is a growing tendency of using the phase of the associated analytic signal. However, these choices produce several paradoxical results [2], [4]. In this paper, we will explain when and why this definition of IF_t produces unacceptable results. Our point is that the derivative of the phase of a complex signal cannot generally be used as IF_t in any physically meaningful way. An alternative definition of IF_t is proposed, and shown to produce physically consistent results. The discussion will be restricted to analytic signals, but the results will be directly applicable to any complex signals.

2. INSTANTANEOUS FREQUENCY AND THE ANALYTIC SIGNAL

In [3], Gabor introduced the concept of an analytic signal associated with the real signal $s(t)$, defined by $z(t) = s(t) + j\mathcal{H}[s(t)]$, where $\mathcal{H}[\circ]$ is the Hilbert Transform operator. The analytic signal $z(t)$ can also be obtained by

$$z(t) \quad = \quad A(t)e^{j\varphi(t)} = 2\int_0^\infty S(f)\, e^{j2\pi ft}df, \qquad (1)$$

where $S(f)$ is the Fourier Transform of the signal $s(t)$, and $A(t)$ and $\varphi(t)$ are real functions of time. This extension to the complex plane provides a simple way of associating, to

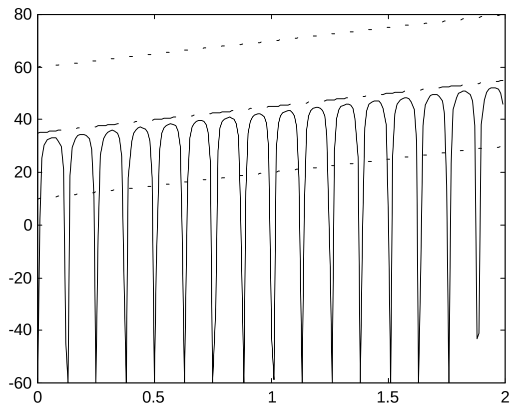

Figure 1: Two chirps. Different amplitudes

the real signal under analysis, a phase function $\varphi(t)$. However, interpreting $\varphi'(t)$ as the IF_t leads to physically unacceptable results in all but a few special cases [2]. Recent work by Loughlin and Tacer [5] has come to shed additional light on the inappropriateness of that definition of IF_t.

• **Problems associated with $\varphi'(t)$.** The main difficulty associated with the definition $IF_t = \varphi'(t)$ is that the obtained IF_t may range outside the band [2]. This imposes severe difficulties when computing the IF_t of general multicomponent signals. To illustrate this point, we may consider a real signal consisting of two chirping cosines. The corresponding analytic signal can be approximated by

$$z(t) = A_1 e^{j(k_1 t^2 + \omega_1 t)} + A_2 e^{j(k_2 t^2 + \omega_2 t)} = A(t)e^{j\varphi(t)}.$$

Defining $IF_t = \varphi'(t)$ results in the solid line in Figure 1, for $k_1 = 5$, $k_2 = 5$, $A_1 = -1.2$, $A_2 = 1$, $\omega_1 = 10$, and $\omega_2 = 60$. In this figure, we also present (dashed lines) the values of IF_t of each one of the chirps individually. We can see that the resulting IF_t is not physically acceptable, not even as a mean of any kind. For a mean IF_t, one would expect a value somewhere in between the frequencies of the two components. As proved in [4], only in the case $A_1 = A_2$ will the obtained IF_t be restricted to that region. These results are also applicable to signals with components of generic IF_t laws [4].

• **Why $\varphi'(t)$?** It will now be instructive to discuss the reasons why $\varphi'(t)$ has enjoyed such widespread acceptance as IF_t, despite the paradoxical results it provides.

(a) To start with, its average value is the mean frequency (if we elect the Fourier spectrum as an adequate one, even for the non-stationary case) of the signal [7]:

$$< \omega >^f = \int_{-\infty}^{\infty} \omega \, |S(\omega)|^2 \, d\omega = < \varphi^{'}(t) >^t, \qquad (2)$$

where $< \circ >^t$ and $< \circ >^f$ stand for time and frequency averages, respectively. It thus seems natural to accept this quantity, whose mean is the mean frequency, as the instantaneous value of frequency.

(b) Secondly, we have that the spread of a signal in frequency is given by [2]:

$$\sigma_\omega^2 = \int_{-\infty}^{\infty} A^{'}(t)^2 dt + \int_{-\infty}^{\infty} \left(\varphi^{'}(t) - < \omega >^f\right)^2 A^2(t)\, dt. \quad (3)$$

While the first term in (3) represents the spread in frequency due to $A(t)$, the second term depends on the deviation of $\varphi^{'}(t)$ from the mean frequency. This again suggests that $\varphi^{'}(t)$ should be interpreted as the IF_t.

(c) Another argument in favor of this definition of IF_t came from the theory of bilinear time-frequency distributions. These distributions represent the signal's energy in a bidimensional time-frequency plane, thus providing an Instantaneous Spectrum for each moment in time ($S_t(f)$ or, simply, S_t). Many of these distributions have one property in common [2]: the mean frequency of S_t is, at all time instants, $\varphi^{'}(t)$. This has strengthened the feeling that $\varphi^{'}(t)$ could be interpreted as the IF_t

• **When is $\varphi^{'}(t)$ acceptable?** Now we will discuss why $\varphi^{'}(t)$ sometimes agrees with our physical notion of IF_t, and, in other cases, does not. As we shall see, the answer seems to lie in the symmetry of S_t. Consider an ideal time-frequency distribution $\rho_s(t, f)$ of the energy of $s(t)$. When dealing with non-stationary signals, we intuitively expect a generalized frequency shift property to hold: if $m(t) = s(t) \cdot e^{j\phi(t)}$, then $\rho_m(t, f) = \rho_s(t, f - \phi^{'}(t))$. In this context, the analytic signal (1) is just a heterodyned version of $A(t)$, $\varphi(t)$ being the heterodination law. When we compute $\varphi^{'}(t)$, we are hoping to determine the frequency around which the spectrum of $A(t)$ is centered. It seems physically sound to take this central frequency as the IF_t. $A(t)$ is, however, a real signal and, as such, must have a symmetric power spectrum. Computing $\varphi^{'}(t)$ is, thus, not determining the frequency around which S_t is centered, but instead the frequency around which S_t is symmetric. Therefore, we can expect a physically reasonable solution if and only if S_t is, in fact, symmetric around a central frequency. In all other cases, the problem is ill-posed, and the results provided by $\varphi^{'}(t)$ are meaningless.

Figure 1 is a typical example. The ideal time-frequency distribution of the two linear chirps will be constituted by two different amplitude parallel ridges, shifting in frequency at a constant rate. There are thus no frequencies around which the several S_t are symmetric. The paradoxical results given by $\varphi^{'}(t)$ are to be expected. If both chirps had the same amplitude, then, for all t, S_t would be symmetric around the central frequency, and we could expect $\varphi^{'}(t)$ to give acceptable results. This equal amplitude case can be seen as the dash-dot line of Figure 1.

The problem of lack of symmetry of S_t is mainly felt when dealing with multicomponent signals, since in most cases of signals with a single narrowband component S_t is symmetric around the peak. This is why $\varphi^{'}(t)$ usually gives proper results when analyzing single component signals.

In the following sections we will see how to extend the concept of IF_t to the general multicomponent case.

3. SPLITTING THE PHASE FUNCTION

As was stated earlier, for many bilinear time-frequency distributions, the derivative of the phase of a complex signal is also the center of mass of S_t. Considering the wild behavior of this derivative, we should ask why does the center of mass in these distributions range outside the frequency band where the energy of S_t is located. The answer lies in the negative values that the distributions with that property always contain [1]. In an attempt to circumvent these difficulties, Loughlin and Tacer recently proposed a different approach to obtain (and, in fact, a different definition of) IF_t. In their proposal, IF_t is obtained as the center of mass of one of the positive time-frequency distributions of the Cohen-Posh class [5]. Since these distributions are always non-negative and have strong finite spectral support ($S(f) = 0 \Longrightarrow \rho(t, f) = 0$), their center of mass will always be contained in the frequency band of the signal, and does not necessarily coincide with $\varphi^{'}(t)$. The concept of IF_t in this proposal is thus separated from $\varphi^{'}(t)$, a necessary condition to obtain an IF_t meaningful in the general multicomponent case. When we perform coherent demodulation of the signal with the obtained "correct" IF_t, we end up, in general, with a complex signal [5]. This is in accordance with our previous discussion, reinforcing the fact that, in the model for IF_t extraction

$$s(t) = C(t)\, e^{j \int_{-\infty}^{t} IF_t \, dt},$$

$C(t)$ should be allowed to be complex. For any signal under analysis (namely the analytic signal), only part of the phase function should be differentiated to obtain IF_t. This part of the phase function will correspond to the heterodination law. The rest of the phase function is needed to account for the lack of symmetry of S_t. We will refer to these two phase terms by $\varphi_{IF}(t)$ and $\varphi_A(t)$, respectively, ending up with the following model:

$$s(t) = C(t)\, e^{j\varphi_{IF}(t)} = A(t)\, e^{j\varphi_A(t)}\, e^{j\varphi_{IF}(t)},$$

with $IF_t = \varphi^{'}_{IF}(t)$. If S_t is symmetric around some central frequency, then all the phase of the signal can be attributed to the heterodination law (and, thus, differentiated to obtain IF_t), since there is no lack of symmetry to account for. These are the cases where $\varphi^{'}(t)$ gives proper results.

Due to the separation of the phase of the signal in two terms, equations (2) and (3) must be generalized to the equally meaningful identities:

$$< \omega >^f = < \varphi^{'}_A(t) >^t + < \varphi^{'}_{IF_t}(t) >^t . \qquad (4)$$

$$\sigma_\omega^2 = \int_{-\infty}^{\infty} A^{'}(t)^2 dt + \int_{-\infty}^{\infty} \varphi^{'}_A(t)^2 A^2(t)\, dt$$

$$+\int_{-\infty}^{\infty}\left(\varphi'_{IF_t}(t)-<\omega>^f\right)^2 A^2(t)\,dt. \qquad (5)$$

Both equations (4) and (5) have now an extra term, depending on the phase of the baseband signal $C(t)$. As expected, the mean of $\varphi'_A(t)$ concurs to the global mean frequency, and the second moment of $\varphi'_A(t)$ concurs to the overall bandwidth.

4. MAXIMIZING THE SYMMETRY

• **Redefining IF_t.** The separation of $\varphi(t)$ into two terms ($\varphi'_A(t)$ and $\varphi'_{IF}(t)$) is totally arbitrary. The only real guideline to follow lies in the physical validation of the IF_t that will emerge. From the previous sections, we know that $\varphi'(t)$ produces physically acceptable results if and when S_t is symmetric around the IF_t. This suggests that the center of symmetry of S_t may be an acceptable definition of IF_t. For each frequency υ, let us then consider the even (S_t^e) and the odd (S_t^o) components of S_t:

$$S_{t,\upsilon}^e(f)=\frac{1}{2}\left[S_t(f)+S_t(2\upsilon-f)\right]$$

$$S_{t,\upsilon}^o(f)=\frac{1}{2}\left[S_t(f)-S_t(2\upsilon-f)\right].$$

We can now define IF_t as being the value of υ that maximizes the energy of the even (symmetric) component or, equivalently, that minimizes the energy of the odd (antisymmetric) component. That is, the value of υ that minimizes (S_t is real valued)

$$\int_{-\infty}^{\infty}\left(S_t(f)-S_t(2\upsilon-f)\right)^2 df. \qquad (6)$$

This definition produces an IF_t with a behavior substantially different than the one obtained in [5]. They will produce the same results for symmetric S_t, since in these cases the center of symmetry will also be the center of mass. This type of definition, based on symmetry considerations, allows us the use of the (eventually negative) bilinear distributions, thus avoiding the time-consuming task of obtaining a positive distribution.

Let us then restrict our scope to the bilinear distributions of Cohen's class [2], and define the *Instantaneous Autocorrelation (R_t)*, as being, at each time instant, the function whose Fourier Transform is S_t. For the distributions in this class, $R_t(\tau)$ is then given by

$$R_t(\tau)=\int_{-\infty}^{\infty} s(t_1+\frac{\tau}{2})\,s^*(t_1-\frac{\tau}{2})\,\gamma(t-t_1,\tau)\,dt_1,$$

with

$$\gamma(t-t_1,\tau)=\int_{-\infty}^{\infty}\Phi(\theta,\tau)\,e^{-j2\pi\theta(t-t_1)}\,d\theta,$$

where $\Phi(\theta,\tau)$ is the so called kernel function. Each choice of $\Phi(\theta,\tau)$ will originate a different distribution [2] and, thus, a different version of $R_t(\tau)$. The Wigner-Ville Distribution (WD), [7] has an Instantaneous Autocorrelation given by

$$R_t^{WD}(\tau)=s(t+\frac{\tau}{2})\,s^*(t-\frac{\tau}{2}),$$

while the Instantaneous Power Spectrum (IPS) provides a slightly (in appearance) different version [6]:

$$R_t^{IPS}(\tau)=\frac{1}{2}[s(t)\,s^*(t-\frac{\tau}{2})+s^*(t)\,s(t+\frac{\tau}{2})].$$

The absolute unconstrained minimum of (6) can be shown to be obtained for $IF_t=\Omega(\tau)/2\pi\tau$, where $\Omega(\tau)$ is the phase of $R_t(\tau)$. The value of IF_t can not, however, be a function of the auxiliary variable τ. To satisfy this constraint, we can expand the unconstrained solution in a Taylor series around zero, and retain only its linear term, resulting in:

$$IF_t=\frac{\Omega'(0)}{2\pi}. \qquad (7)$$

This definition of IF_t is, for both IPS and WD, exactly equivalent to the traditional definition $IF_t=\varphi'(t)$. In fact, it can be shown that the same can be said for all distributions belonging to Cohen's class with kernel functions of the form

$$\Phi(\theta,\tau)=\sum_{i=1}^{N}A_i e^{j\theta f_i(\tau)},$$

for all real functions f_i satisfying $f_i(0)=0$. If, furthermore, $f'_i(0)=0$, then

$$\left.\frac{\partial\Phi(\theta,\tau)}{\partial\tau}\right|_{\tau=0}=0, \qquad (8)$$

which is the condition for $\varphi'(t)$ to be the mean frequency [2]. For these distributions, $\varphi'(t)$ is, simultaneously, the center of mass of the distribution, and a first order approximation to its center of symmetry. The approximation (7) will be exact if $R_t(\tau)$ has linear phase. For both IPS and WD, $R_t(\tau)$ will have linear phase and even modulus if the signal has linear phase. For the WD, this will also be true even for signals with quadratic phase. Since both these distributions have kernel functions that satisfy (8), we can see that $\varphi'(t)$ will, for these signals, provide physically acceptable results.

• **Conditions for consistency.** The results obtained by defining the Instantaneous Frequency as the frequency that maximizes the symmetry of S_t will depend on the choice of a particular distribution, but are easily made to comply with the (applicable) conditions that have been deemed necessary for physical consistency [5]:

1. If $S(f)$ is bandlimited to some range of frequencies, IF_t should also be limited to that range;
2. If the signal has constant amplitude and constant frequency ω_0, its IF_t should be ω_0;
3. If the signal is scaled in amplitude, its IF_t should remain the same.

Satisfaction of the first condition requires the used time-frequency distribution to have weak finite spectral support. The condition on the distribution's kernel function for this property to hold is [2]:

$$\int_{-\infty}^{\infty}\Phi(\theta,\tau)\;e^{-j\tau\omega}\;d\tau=0\quad for\quad |\theta|\le 2\,|\omega|\,.$$

Both WD and IPS have this property. The Choi-Williams distribution (CWD), which we will also be using, does not, but can easily be made to approximate it [2]. For the second condition to hold, the used distribution must respect the frequency marginal, that is:

$$\int_{-\infty}^{\infty} \rho(t, f)\ dt = |S(\omega)|^2 .$$

The condition on the distribution's kernel is $\Phi(0, \tau) = 1$, a condition satisfied namely by the WD, IPS, and the CWD. Satisfaction of the third condition is guaranteed if we use any one of the distributions of Cohen's class, which again is the case of the ones we will be using (WD, IPS and CWD)

5. EXPERIMENTAL RESULTS

As examples, we will be using three different time-frequency distributions (WD, IPS and CWD). These distributions can be thought of as representatives of the three types of behavior that the bilinear distributions may have, in what refers to the cross-terms, always present when dealing with multi-component signals. In the WD, the cross-terms lie midway between the components of the signal. This fact will, in many instances, originate spurious points of symmetry, and thus seriously degrade the usefulness of the extracted IF_t (this problem can be alleviated with appropriate smoothing). IPS, on the other hand, has cross-terms lying on top of the true components of the signal, originating an amplitude modulation of these terms. These time-varying amplitudes will sometimes force the best symmetry point to be found in the peaks of the different components (preference will be given to the strongest one). The CWD, with its reduced interference kernel, will have its cross-terms spread in the time-frequency plane [2], thus largely ignoring them in the process of detecting the best point of symmetry. These different types of behavior can be seen in Figure 2. In this figure, the values for IF_t obtained with the CWD, IPS and WD are marked with a cross, asterisk, or circle, respectively. For comparison purposes, the same signal of Figure 1 was used.

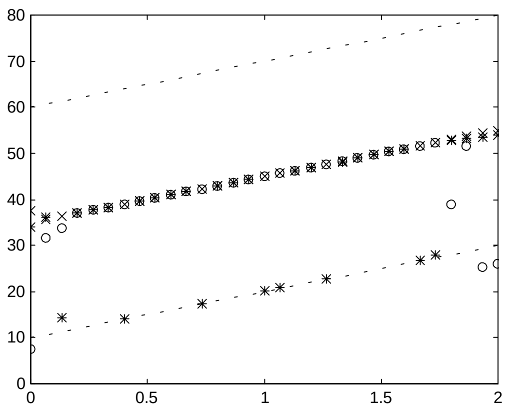

Figure 2: CWD. Two chirps of different amplitudes

In Figure 2, we had a 20% difference in amplitude between the two components. If we had been looking for the center of mass of the several S_t, as in [5] (we could not, in that case, have used these non-positive distributions), this difference in amplitude would have implied a proportional shift in IF_t towards the stronger component. This dependency of IF_t on the ratio of amplitudes of the components may, or may not, be considered undesirable, depending on our concept of Instantaneous Frequency. It doesn't happen when we look for the point of maximum symmetry, which seems to better enable an interpretation of IF_t as the heterodination law of the signal. This constancy of IF_t is, however, typically short lived. For higher ratios of the individual amplitudes, and for all three distributions, we will see IF_t locking quickly on the strongest component. In short: when the amplitudes are similar, this criterion goes for the equidistant locations; when the amplitudes start to differ, the strongest component is selected.

6. CONCLUSION

The traditional definition ($IF_t = \varphi'(t)$) is inappropriate, since its results are paradoxical in most situations. The key reason for the success/failure of this method seems to lie in the symmetry of (S_t). For many types of S_t, the following facts have been ascertained: If symmetry exists, then $\varphi'(t)$ gives acceptable results; when symmetry in S_t doesn't exist, the traditional definition is making a linear approximation to the problem of determining the point of maximum symmetry. A criterion for determining IF_t was proposed, based on maximization of the symmetry of S_t. The conditions under which the results are guaranteed to be physically acceptable were presented, and some examples were given.

7. REFERENCES

[1] Claasen T., and Mecklenbrauker W., "The Wigner distribution - A tool for time-frequency analysis, Parts I and III,", Philips Journ. Res. 35, 1980.

[2] Cohen L., *Time-Frequency Analysis.* Prentice-Hall, NJ, 1995.

[3] Gabor D., "Theory of Communication," *Journal IEE*, vol. 93, pp. 429-457, 1946.

[4] Loughlin Patrick J., and Tacer Berkant, "Instantaneous Frequency And the Conditional Mean Frequency of a Signal," to appear in Signal Processing.

[5] Loughlin Patrick J., and Tacer Berkant, "On the amplitude - and frequency - modulation decomposition of signals," Journal of the Accoustical Society of America, Vol. 100. n^o 3, September, 1996.

[6] Oliveira Paulo M., "Instantaneous Power Spectrum," Engineer's thesis, Dep. Elec. Comput. Eng., Naval Postgraduate School, Monterey, CA, 1989.

[7] Ville J., "Theorie et applications de la notion de signal analytique", *Cables et Transmissions*, vol. 2A, no. 1, pp. 61-74, 1948.