DEFINITION OF INSTANTANEOUS FREQUENCY ON REAL SIGNALS

Sokol Saliu
Chalmers Lindholmen University College
Dept. of Electrical and Computer Engineering
P.O. Box 8873, SE 402 72, Göteborg, SWEDEN
Tel: + 46 31 772 5744; fax +46 31 772 5731
email: sokol.salius@chl.chalmers.se

ABSTRACT

This paper introduces a definition of the instantaneous frequency (IF) on real signals. Basically, the definition $\omega_t = \dot{x}_t/\sqrt{1-x_t^2}$ avoids some pitfalls associated with the corresponding definition on analytic signals. It readily applies to the class of monocomponent frequency modulated signals and also accommodates the class of amplitude and frequency modulated signals. An interpretation of the IF is given for the latter and an extension of it to instantaneous parameters is discussed.

1 INTRODUCTION

The first definition of *instantaneous frequency* (IF) is attributed to Carson and Fry [4] who defined a complex frequency modulated (FM) signal and the IF as the derivative of its phase.

Van der Pol [8], from the simple harmonic motion,

$$\ddot{x} + \omega^2 x = 0 \tag{1}$$

proceeds to the definition of IF as the derivative of the phase $\phi(t)$ of the signal $x(t) = A\cos(\phi(t))$, where

$$\phi(t) = \int_0^t \omega_i(u)du + \theta \tag{2a}$$

$$\omega_i(t) = \dot{\phi}(t) \tag{2b}$$

Cohen [6] generalizes (1) to an equation with non constant coefficients

$$\ddot{x} - 2\Gamma(t)\dot{x} + \eta^2(t)x = 0 \tag{3}$$

The special case

$$\ddot{x} - \frac{\dot{\omega}_i}{\omega_i}\dot{x} + \omega_i^2 x = 0 \tag{4}$$

has as a solution an FM signal with arbitrary IF $\omega_i(t)$.

The prevalent definition of IF is the one given by Gabor and Ville [9, 13]

$$\omega_i(t) = \frac{d}{dt}[\arg z(t)] = \dot{\phi}(t) \tag{5a}$$

where

$$z(t) = A(t)\exp[\mathrm{i}\phi(t)] = x(t) + \mathrm{i}\mathcal{H}[x(t)] \tag{5b}$$

is the analytic signal associated with the real signal $x(t)$ and $\mathcal{H}[x(t)]$ defines the Hilbert transform of $x(t)$.

Given a bilinear time-frequency distributions $P(\omega, t)$, the average frequency at a particular time is given by

$$\langle\omega\rangle_t = \int P(\omega|t)d\omega \tag{6}$$

where $P(\omega|t) = P(\omega,t)/P(t)$ is the distribution of frequency at a given time t and $P(t) = \int P(\omega,t)d\omega$ the marginal distribution of time.

Cohen [7] shows that under some constrains on the kernels of distributions $P(\omega, t)$

$$\langle\omega\rangle_t = \dot{\phi}(t) \tag{7}$$

$$\langle\omega^2\rangle_t = \dot{\phi}^2(t) + \sigma_\omega^2(t) \tag{8}$$

where $\sigma_\omega(t) = \dot{A}(t)/A(t)$ defines the standard deviation of frequency at a particular time.

A related definition is the definition of the complex IF [10]

$$\beta(t) = \frac{d}{dt}ln[A(t)e^{\mathrm{i}\phi(t)}] = \frac{\dot{A}(t)}{A(t)} + \mathrm{i}\dot{\phi}(t) \tag{9}$$

where the imaginary part defines the IF and the real part the instantaneous radial velocity.

Presently there is a controversy about the concept and the definition of IF. In equations (2), the phase of the signal is not explicitly available. The definition (5), as pointed out by Cohen [7], is contradictory—the concept of IF is local while to obtain the analytic signal (5b) the whole signal is required by Hilbert transform $\mathcal{H}$. Furthermore, as discussed in [2], there are some inherent problems with definition (5a): (i) although the real and the derived analytic signals have equal spectra for positive frequencies they are different, (ii) as various complex signals can be constructed from a single real signal, the IF is not uniquely defined, (iii) some processes e.g. zero-crossing demodulation of FM signals, can not be described in terms of analytic signals, and (iv) $\omega_i(t)$ in (5) or $\langle\omega\rangle_t$ in (6) may not represent a physical quantity [5]; for some signals it may even wander outside the frequency range of the signal [11, 12]. However, Boashash [2] considers such signals to be multi-component and outside the scope of the IF definition.

In this paper we give a definition of IF on real signals and discuss its implications.

2 THEORY

The concept of the *period* is connected to the repetitiveness of phenomena or events. It is an unambiguous measure independent of the signal waveform. Coupled with it is the *rate* defined as the inverse of the period.

A periodic signal of special interest is the *harmonic signal* $A\sin(\omega t+\theta)$ or in complex form $x(t) = C\exp(\mathrm{i}\omega t)$, with C, ω, A and θ constants. The importance of the *harmonic* signal is due to the Fourier series. Connecting the concept of *frequency* specifically to the rate of the harmonic signal allows us to represent any periodic signal as a sum of harmonics on a *frequency set*. By doing so, we lose the unique and waveform-independent definition of the rate to the waveform-dependent *set of Fourier frequencies*.

To determine the rate of a periodic signal one needs to observe the signal over the whole period, that is the rate is not a *local* feature of the signal. Applying this interpretation to the harmonic signal would exclude the notion of local or *instantaneous* frequency, and hence the apparent 'paradox' of IF. Nevertheless, the rate parameter can obtain a locality feature if (a) we single out a reference periodic function, (b) which is smooth—i.e., the function and its derivatives of all orders are continuous, and (c) is a member of an elementary family of functions—in the sense that any other periodic function of some other family[1] can be expressed as a composition of functions from the elementary family.

2.1 Instantaneous frequency of the harmonic signal

The harmonic signal

$$x(t) = A\sin(\omega t+\theta) \tag{10}$$

where ω is the angular frequency, θ an initial phase and A the amplitude, satisfies the three conditions above. It is smooth, elementary, and agreed upon as a reference function. One can locally determine the parameter of interest, in this case the frequency ω, by observing the signal at two close and distinct time instants t_1 and t_2

$$\begin{cases} x_{t_1} = A\sin(\omega t_1+\theta) \\ x_{t_2} = A\sin(\omega t_2+\theta) \end{cases}$$

[1] a harmonic signal with a fractional frequency can be written as a series of harmonic signals with integer frequencies

The basic assumption here is that the parameter ω changes slower than the signal it parameterizes.

Using the equality $\arcsin x = \int_0^x dy/\sqrt{1-y^2}$ for $|x| < 1$, we solve for ω the system of equations above

$$\omega = \frac{1}{\Delta t}\int_{a_{t_1}}^{a_{t_2}} \frac{1}{\sqrt{1-y^2}}dy \approxeq \pm\frac{1}{\sqrt{1-a_t^2}}\frac{\Delta a_t}{\Delta t} \tag{11}$$

where $a_t = x_t/A < 1$, $\Delta t = t_2 - t_1 > 0$, $\Delta a_t = a_{t_2} - a_{t_1}$, and $t \in \Delta t$. Note that the expression (11) is not defined for values $a_t = 1$.

The sign of ω can be determined by solving for the phase θ; in sequel the conventional positive sign of ω is maintained.

The approximation in (11) holds due to the smoothness of $x(t)$. In limit, as $\Delta t \to 0$, we obtain the IF

$$\omega(t) = \frac{\dot{a}(t)}{\sqrt{1-a(t)^2}} \tag{12a}$$

and for $A = 1$,

$$\omega(t) = \frac{\dot{x}(t)}{\sqrt{1-x(t)^2}} \tag{12b}$$

We have attained a definition of instantaneous frequency that is based on the derivative of the signal amplitude weighted by a 'shape' factor—here the inverse of in-quadrature signal $\sin(\omega t+\theta+\pi/2)$.

Dividing the numerator and denominator of (12a) with $a(t)$ (for $a(t) \neq 0$), we rewrite (12)

$$\omega(t) = r_{\dot{a}}(t)\zeta(t) \tag{13a}$$

with

$$r_{\dot{a}}(t) = \frac{\dot{a}(t)}{a(t)};\ \zeta(t) = \frac{1}{\sqrt{\eta^{-2}(t)-1}};\ \eta(t) = a(t) \tag{13b}$$

When $a(t) \to 0$, $r_{\dot{a}}(t) \to \infty$ and $\zeta(t) \to 0$ but $\omega(t)$ remains finite (12a).

The frequency definition (12) or (13) represents the relative speed of change of the signal's *instantaneous amplitude*(IA) $r_{\dot{a}}(t)$ or $r_{\dot{x}}(t)$—or when written equivalently as $r_{\dot{a}}(t) = d[ln(a(t)]/dt$, the speed of logarithmic IA—at any given time instant t weighted by the shape factor $\zeta(t)$ based on the relative IA $\eta(t)$. Therefore, the IF is not directly connected to the period of the sinus wave, but to the very reference of the sinusoidal signal as standard function. The derivation leads naturally to the concept of instantaneous frequency. Uniqueness of the IF derives from the sinusoidal function being elementary.

At time instants t_i, when $a(t_i) = 1$ and $\dot{a}(t_i) = 0$, i.e. the relative speed of IA is zero and the shape factor infinite, the frequency ω cannot be determined. This problem does not arise in the case of complex or analytic harmonic signals, $\sin(\omega t+\theta)+\mathrm{i}\cos\sin(\omega t+\theta)$, due to the existence of the quadrature imaginary part. Anyhow, the difficulty with singular points remains, reflected on the phase discontinuity.

2.2 Instantaneous frequency of AM-FM signals

So far, the IF (13) is defined for a sinusoidal signal. It needs to remain coupled to the harmonic signal in order to preserve the *uniqueness property* but on the other hand, in order to be a meaningful concept the IF needs to cover a broader class of signals.

Let us consider the class of amplitude and frequency-modulated (AM-FM) signals.

$$x(t) = A(t)\sin(\omega t + \theta) \tag{14}$$

Setting $m(t) = x(t)/A(t)$, $a(t) = x(t)/A$, where A being here the maximum value of $A(t)$, and evaluating

$$\dot{m}(t) = \frac{1}{A(t)}\left[\dot{x}(t) - x(t)\frac{\dot{A}(t)}{A(t)}\right] \tag{15}$$

we obtain

$$\omega(t) = \left[r_{\dot{a}}(t) - r_{\dot{A}}(t)\right]\zeta(t) \tag{16a}$$

where, similarly to (13b),

$$r_{\dot{a}}(t) = \frac{\dot{a}(t)}{a(t)} = r_{\dot{x}}(t); \quad r_{\dot{A}}(t) = \frac{\dot{A}(t)}{A(t)} \tag{16b}$$

$$\zeta(t) = \frac{1}{\sqrt{\eta^{-2}(t) - 1}}; \quad \eta(t) = \frac{x(t)}{A(t)} = m(t)$$

These results could also be obtained if one modulates the signal $sin(\phi(t))$ in amplitude, maintaining that $\phi(t)$ in (2) is the phase of the resulting signal, and takes the derivative of $\phi(t) = \arcsin[x(t)/A(t)]$.

When the signal has constant amplitude A, i.e. $r_{\dot{A}}(t) = 0$, the equations (16) are reduced to equation (12), that is to the IF definition in (12a). In this case the IF can be interpreted as the IF of a harmonic signal that locally fits an FM signal.

When $r_{\dot{A}}(t) \neq 0$, cf. $r_{\dot{A}}(t)$ to the local spread $\sigma_\omega(t)$ of IF in (8), the interpretation above is no longer valid. Anyhow, when $|r_{\dot{A}}(t)| \ll |r_{\dot{a}}(t)|$ then $\omega(t) \approxeq r_{\dot{a}}(t)\zeta(t)$, which would approximate the IF of the FM signal (12a).

Note that the terms $r_{\dot{a}}(t)$ and $r_{\dot{A}}(t)$ in equation (16a) are nor indepedent, the amplitude modulation $A(t)$ effects both terms. The first term $r_{\dot{x}}(t)\zeta(t)$ cannot be considered separately as the IF of FM signal (12a). The amplitude modulation makes the signal non-elementary and blurs the concept of IF.

Figure 1(a) shows the IF of a linearly frequency-modulated harmonic signal(0–0.1 Hz). The IF is calculated according to the definition (12b) with the differentiation implemented as central finite difference(CFD), $\dot{a}(t) \approx [a(n+1) - a(n-1)]/2$. When the linear FM range was increased up to 0.5 Hz, we observed a saturation of IF at at about 0.16 Hz (six samples per period) followed by a decrease towards zero frequency, caused mainly by the approximation of signal derivative by CFD. Recall that the IF formulation (12b) is not meant to be an estimator—the reader could refer to [3] for an overview of IF estimation methods and to [1] for frequency estimation from few measurements.

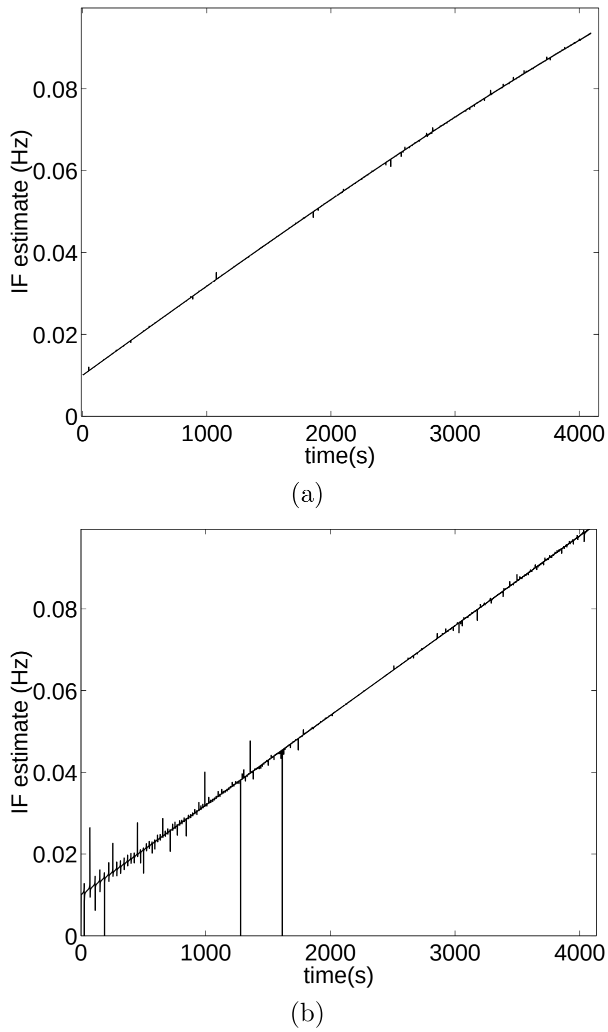

Figure 1: Instantaneous frequency of (a) linearly frequency modulated signal and (b) linearly frequency and amplitude modulated signal

2.3 Instantaneous amplitude

The IF definition given in sections 2.1 and 2.2 assumes that the amplitude $A(t)$ is known, which is not generally the case. Nevertheless, following a similar approach to the derivation of IF in section 2.1, the amplitude can be obtained by observing the signal on three neighboring time instants

$$\begin{cases} x(t-\delta) & = A_t \sin[\omega_t(t-\delta) + \theta] \\ x(t) & = A_t \sin[\omega_t t + \theta] \\ x(t+\delta) & = A_t \sin[\omega_t(t+\delta) + \theta] \end{cases} \tag{17}$$

where in addition to $\omega(t)$, $A(t)$ is assumed to be constant in the interval $[t-\delta, t+\delta]$, i.e. it changes slower than $x(t)$—the later assumption is less substantiated than the one on ω in section 2.1.

One can solve the nonlinear equations above and obtain the parameters ω_t, A_t, and θ. A simpler solution can be obtained by considering an arbitrary phase [1], i.e. an arbitrary initial time $t = 0$. In this case the IA $A(t)$ and IF $\omega(t)$ are given by

$$A(t) = \frac{[x^2(t) - x(t-\delta)x(t+\delta)]^{1/2}}{\sin(\omega_t)} \tag{18a}$$

$$\omega_t = \arccos \frac{x(t-\delta) + x(t+\delta)}{2x(t)} \tag{18b}$$

When $A(t) = A$, the obtained solution is accurate, as shown in section 2.1. By being able to determine instantaneously the amplitude A we validate the definition (12) as the definition of IF.

Figure 1(b) shows the IF, calculated according to (18), of an FM-AM harmonic signal with linear modulation laws—IF increases from 0.01 to 0.1 Hz and IA decreases from 1 to 0.5.

When the assumptions on $A(t)$ (17) are violated, the solution (18) is approximate, with effects already discussed in section 2.2. In this case, we can determine *instantaneous parameters*(IP) of the signal, without necessary associating $\omega(t)$ with IF.

2.4 Instantaneous parameters

Given a smooth parametric family of functions $x(t, \mathbf{p})$, where $\mathbf{p} = [p_1, p_2, \cdots, p_m]$ is the vector of parameters, with Taylor series

$$x(t, \mathbf{p}) = \sum_{n=0}^{\infty} x^{(n)}(t_0, \mathbf{p})/n! \tag{19}$$

we can define locally the parameters $\mathbf{p}$—see [5] for an operator approach. From a single value $x(t_0)$ and signal derivatives $x^{(n)}(t)$ at the time instant t_0 we can derive—for a given waveform— $m-1$ other values of $x(t)$ in the vicinity of t_0. A system of m equations

$$x_i = x(t_i, \mathbf{p}), \quad i = 0, 1, \ldots, m-1. \tag{20}$$

can be used to solve for the m unknown parameters. Practically, we can measure m amplitude values of $x(t, \mathbf{p})$ and solve (20) for $\mathbf{p}$. The existence of Taylor series ensures that the limit $t_{m-1} \to t_{m-2}, \to \cdots \to t_0$ exists. The instantaneous parameters (IP) of a smooth parametric function can therefore be seen as a generalization of the IF.

3 CONCLUSION

A definition of the instantaneous frequency of real signals is given. The definition avoids the apparent IF paradox related to the analytic signal and has an intuitive interpretation as the derivative of instantaneous amplitude. The class of monocomponent FM signals is found a legitimate class for the definition. The effect of AM modulation is discussed in the context of instantaneous signal parameters.

REFERENCES

[1] R. M. Adelson. Frequency estimation from few measurement. *Dig. Signal Proc.*, 7:47–54, 1997.

[2] B. Boashash. Estimating and interpreting the instantaneous frequency of a signal–part 1: Fundamentals. *Proc. IEEE*, 80(4):520–538, 1992.

[3] B. Boashash. Estimating and interpreting the instantaneous frequency of a signal–part 2: Algorithms and applications. *Proc. IEEE*, 80(4):520–538, 1992.

[4] J. Carson and T. Fry. Variable frequency electric circuit theory with application to the theory of frequency modulation. *Bell System Tech. J.*, 16:513–540, 1937.

[5] L. Cohen. Instantaneous "anything". In *IEEE Int. Conf. Acoust, Speech and Signal Proc.*, volume 4, pages 105–108, 1993.

[6] L. Cohen. The frequency and scale content of biological signals. In *Records of Twenty-Eighth Asilomar Conf. on Signals, Systems and Computers*, volume 2, pages 1541–45, 1994.

[7] L. Cohen and C. Lee. Instantaneous frequency and time-frequency distributions. In *IEEE Int. Conf. Circuit and Systems*, volume 2, pages 1231–1234, 1989.

[8] B. V. der Pol. The fundamental principles of frequency modulation. *Proc. IEE*, 93(III):153–158, 1946.

[9] D. Gabor. Theory of communication. *Proc. IEE*, 93(III):429–457, 1946.

[10] S. Hahn. Complex variable frequency electric circuit theory. In *Proc. IEEE*, volume 52, pages 736–737, 1964.

[11] P. J. Loughlin and B. Tacer. Comments on the interpretation of instantaneous frequency. *IEEE Signal Procesing Letters*, 4(5):123–125, 1997.

[12] P. M. Oliviera and V. Barroso. On the concept of instantaneous phase and frequency. In *Proc. IEEE SP Int. Symposium on Time-Frequency and Time-scale Analysis*, pages 105–108, 1998.

[13] J. Ville. Theorie et application de la notion de signal analytique. *Cables et Transmissions*, 93(III):153–158, 1948.