

# Estimating the Instantaneous Frequency of Sinusoidal Components Using Phase-Based Methods

Mathieu Lagrange and Sylvain Marchand

June 13, 2007

## Abstract

The robust estimation of the frequency of some sinusoidal components is a major prerequisite for many applications, such as in sinusoidal sound modeling, where the estimation has often to be done with a low complexity, on short-term spectra. Among the estimators proposed in the literature, we focus in this paper on a class known as the “phase-based” estimators. In this paper, we prove that five of these estimators are equivalent, at least in theory. In-depth practical experiments demonstrate that these estimators perform roughly similarly in practice, although small differences remain, differences which are most probably due to numerical properties of the mathematical operators used in their implementation.

## 1 Introduction

Among numerous other applications, sinusoidal sound modeling [1, 2] requires the estimation of the frequencies of sinusoidal components. This first step is of great importance because the remaining elements of the sinusoidal analysis chain strongly depend on the precision of these estimates.

If the number of sinusoidal components that should be considered were exactly known, parametric methods such as MUSIC [3] or ESPRIT [4] could be used to achieve virtual infinite frequency precision and resolution (if no noise is added). Unfortunately, the inherent complexity of the analyzed sound signals together with the need for real-time applications often impose the use of estimators based on short-term spectra, such as those obtained with the Short-Time Fourier Transform (STFT) [5], often implemented using the Fast Fourier Transform (FFT) [6]. The frequency resolution of this

transform is bounded by the well-known time/frequency resolution trade-off.

On the other hand, the frequency precision can be enhanced by considering some estimators based on the spectral properties of sinusoidal components. A first class of FFT-based estimators considers some values of the power spectrum around a frequency component to fit some analytic function – e.g. a polynomial. The location of the maximum of this function gives the precise frequency of the sinusoidal component [7, 8, 2, 9]. A second class explicitly uses the phase of the FFT to estimate the frequency, and includes the reassignment estimator [10, 11] and the difference estimator [12] (the latter is commonly used in the phase-vocoder approach [13, 14]). In this paper, we extend the results published in [15], where the consideration of the relationship between the signal and its derivative lead to the proposal of another estimator, known as the derivative estimator.

As it was guessed in [16], and partly demonstrated by Hainsworth during his Ph.D. [17], this so-called derivative estimator is theoretically equivalent to the reassignment one, and so are its discrete version [18] and the difference estimator. However, this does not ensure that these estimators perform equally in practice. For example, a loss of precision of the derivative estimator as well as a bias with the discrete version of the reassignment estimator have been noticed in [19]. The evaluation of the performance of each estimator is a practical issue that should be considered to determine which estimator is relevant for a given application.

The paper is organized as follows: The sinusoidal model and some elements of FFT-based sinusoidal analysis are presented in Section 2. Three phase-based estimators are reviewed in Section 3 and their theoretical equivalence is demonstrated. We show that the loss of precision of the derivative estimator is due to the numerical properties of the considered mathematical operator. This can be avoided without overloading complexity by considering improved estimators, called the trigonometric and arctan estimators, introduced in [20] and [21], respectively.

Section 4 is dedicated to the evaluation of the performance of these estimators from a practical point of view. The estimation of the frequency of complex tones is first considered in Section 4.1. Since musical applications consider real tones rather than complex ones, the performance of these estimators in the real case is next studied in Section 4.2. When compared to the results obtained in the complex case, these experiments show that the performance of the estimators is significantly diminished due to the spectral properties of real signals. To achieve better performance, a pre-processing step that converts the input real signal into an analytic – complex equiva-

lent – signal using the Hilbert transform is also considered in Section 4.3. The results of all these experiments are finally summarized and discussed in Section 5.

## 2 Sinusoidal Modeling

Additive synthesis (see [13]) is the original spectrum modeling technique. It is rooted in Fourier’s theorem, which states that any periodic function can be modeled as a sum of sinusoids at various amplitudes and harmonic frequencies. For stationary pseudo-periodic sounds, these amplitudes and frequencies continuously evolve slowly with time, controlling a set of pseudo-sinusoidal oscillators commonly called *partials*. In the analog domain, it has been used for designing artificial sounds using devices such as the Teleharmonium or the Hammond organ. In the digital domain, additive synthesis has been used by Risset [22] for musical purposes. Many works have later been published in the speech and audio processing areas, and the reader is invited to refer to [23] for proper history.

### 2.1 Additive Model

The audio signal  $s$  can be calculated from the additive parameters using Equations 1 and 2:

$$s(t) = \sum_{p=1}^P A_p(t) e^{j\phi_p(t)} \quad (1)$$

$$\phi_p(t) = \phi_p(0) + 2\pi \int_0^t \omega_p(u) du \quad i.e. \quad \omega_p(t) = \frac{1}{2\pi} \frac{d}{dt} \phi_p(t) \quad (2)$$

where  $P$  is the number of partials and the (real) functions  $\omega_p$ ,  $A_p$ , and  $\phi_p$  are the instantaneous frequency (normalized), amplitude, and phase of the  $p$ -th partial, respectively.

In the real case, the signal  $s$  consists of a sum of (real) sinusoids:

$$s(t) = \sum_{p=1}^P A_p(t) \cos(\phi_p(t)). \quad (3)$$

In fact, each real sinusoid consists of two complex exponentials, since we have:

$$\cos(x) = (e^{+jx} + e^{-jx}) / 2. \quad (4)$$

## 2.2 Elements of Sinusoidal Analysis

The basic method used for estimating the model parameters is the Short-Time Fourier Transform (STFT), where a sliding analysis window  $w$  is used to obtain short-term spectra:

$$S_w(\omega, t) = \int_{-\infty}^{+\infty} s(\tau) w(\tau - t) e^{-j2\pi\omega(\tau - t)} d\tau. \quad (5)$$

Then the sinusoidal components are searched in each spectrum, using specific estimators.

In practice, the signal is discrete. If  $N$  samples are used to compute each spectrum, a rectangular window of size  $N$  is implicitly used:

$$w_r[n] = \begin{cases} 1 & \text{for } 0 \leq n < N \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

Its (discrete) spectrum  $W_r$ , plotted in Figure 1, is given by:

$$W_r[k] = \frac{\sin(k\pi)}{\sin(k\pi/N)} e^{-jk\pi(N-1)/N} \quad (7)$$

where  $k$  is the Discrete Fourier Transform (DFT) bin number (index). A multiplication in the temporal domain leads to a convolution in the spectral domain. The spectrum of each exponential of  $s$  is therefore a frequency-shifted, amplitude-scaled, and phase-rotated version of the spectrum of the analysis window. If the size of the DFT is chosen so that the absolute difference of the frequencies of two components is at least  $F_s/N$  (in Hz), each sinusoidal component  $p$  will then give rise to a local maximum in the power spectrum located at the DFT index  $k_p$  so that:

$$\frac{k_p - 0.5}{N} \leq \omega_p \leq \frac{k_p + 0.5}{N}. \quad (8)$$

In a first attempt, the frequency of each component can then be approximated by:

$$\hat{\omega}_p^{\text{DFT}} = \frac{k_p}{N}. \quad (9)$$

The precision of this rough estimate can be enhanced as desired by using a Zero-Padded (ZP) Fourier Transform, see Table 1. This can be done by adding trailing zeros to the temporal samples before the DFT computation.

Indeed, as shown in [24, 25], when the number of trailing zeros approaches infinity, the resulting estimator is the Maximum Likelihood estimator when

the rectangular window is used. Ideally, the use of a non-rectangular window should then be avoided because it induces a bias in the estimation. In the specific case of the estimation of the frequency of a known number of sinusoids in Additive White Gaussian Noise (AWGN), some methods [4] achieve close to optimal results without any windowing.

Musical signals are more complex. An arbitrary high number of sinusoidal components are present, and their frequencies should be estimated reliably and efficiently. To achieve such constraints, the some methods approximate the interpolation of the ZP Fourier Transform (ZPFT) by others means. Polynomial methods combine a ZPFT with reasonable ZP factor and further interpolation of the log-amplitude of the peak using polynomials, to achieve practical computation complexity. Theoretically, a Gaussian window – whose power spectrum is exactly a polynomial – could be used to achieve good performance. However, the Gaussian window should be avoided while analyzing musical signals. An in-depth study of the relationship between the ZP factor and the estimation error of polynomial estimators while considering widely-used windows can be found in [9].

On contrary, the phase-based methods do not require the use of zero-padding, see Table 1. As it will be discussed further in the next section, the precision of the phase-based estimators relies on the integrity of spectral informations, *i.e.* at a given local maximum, the influence of the other sinusoidal components should be negligible. In the most common case where the frequencies  $\omega_p$  are not a multiple of  $1/N$ , the use of a rectangular window induces a non-negligible energy spread so that the above assertion is rarely verified. The multiplication – prior to any DFT computation – of the observed samples with a window function that minimizes the energy spread has to be considered.

The (periodic) Hann window, given by:

$$w_H[n] = \frac{1}{2} \left( 1 - \cos \left( \frac{2\pi n}{N} \right) \right) \quad (10)$$

is a good candidate since it achieves a good asymptotic attenuation of the side lobes as shown in Figure 1.

Its spectrum is expressed by:

$$W_H[k] = \frac{1}{2} W_r[k] - \frac{1}{4} W_r[k-1] - \frac{1}{4} W_r[k+1] \quad (11)$$

where  $W_r[k]$  is the (discrete) spectrum of the rectangular window, see Equation 7. The two last elements of Equation 11 attenuate the side lobes of the first element – the spectrum of a rectangular window – at the expense

ZP	$\hat{\omega}^{\text{DFT}}$			$\hat{\omega}^d$		
	mean	var	max	mean	var	max
1	5.4	9.8	21.5	0.014	0.13	15.4
2	2.7	2.6	20.0	0.015	0.15	15.4
3	1.8	1.2	12.8	0.015	0.12	10.2
4	1.3	0.8	14.6	0.017	0.16	11.9
5	1.1	0.5	11.5	0.014	0.10	8.8
6	0.9	0.5	12.8	0.015	0.12	10.2
7	0.7	0.4	13.8	0.016	0.14	11.2
8	0.6	0.3	11.9	0.014	0.10	9.3
9	0.6	0.3	12.8	0.015	0.12	10.2
10	0.5	0.3	13.5	0.016	0.13	10.9

Table 1: Mathieu Lagrange and Sylvain Marchand. *Mean and variance of the square error of the rough DFT estimator and the difference estimator versus increasing zero-padding factor.*

of the widening of the main lobe, which is not a handicap for phase-based estimators provided that only frequency precision is considered, and not resolution. This window will be used implicitly in the experiments reported in the remainder of the paper.

### 3 Phase-Based Estimators

The (short-term) spectra obtained from the STFT with Equation 5 consist of complex values, which in the polar representation are:

$$S(\omega, t) = A(\omega, t) e^{j\phi(\omega, t)} \quad (12)$$

and in the present study we will use only the phases:

$$\phi(\omega, t) = \angle S(\omega, t) = \Im(\log(S(\omega, t))) \quad (13)$$

where  $\angle x$  and  $\Im(x)$  denote respectively the angle and imaginary part of the complex number  $x$ .

In this article, we focus on the estimation of the frequency, which will be considered as constant during the analysis window  $w$ .

#### 3.1 Reassignment Method

In usual time-frequency representations, the values obtained when decomposing the signal on the time-frequency atoms are assigned to the geometrical



Figure 1: Mathieu Lagrange and Sylvain Marchand. *Power spectra (in dB) of the rectangular and Hann windows plotted versus the normalized frequency, with dashed and solid lines, respectively. The Hann window achieves good asymptotic side-lobe attenuation at the expense of the broadening of the main lobe.*

center of the cells, *i.e.* center of the analysis window and bins of the Fourier transform, see Equation 9. Auger and Flandrin propose in [10] to assign each value to the center of gravity of the cell's energy. The method uses the knowledge of the analytic first derivative  $w'$  of the analysis window  $w$  in order to adjust the frequency inside the DFT bin.

More precisely, if we consider Equations 5 and 13, we can compute:

$$\hat{\omega} = \frac{1}{2\pi} \frac{\partial}{\partial t} \phi(\omega, t) = \frac{1}{2\pi} \Im \left( \frac{\partial}{\partial t} \log(S_w(\omega, t)) \right) \quad (14)$$

$$= \frac{1}{2\pi} \Im \left( \frac{\frac{\partial}{\partial t} \left( \int_{-\infty}^{+\infty} s(\tau) w(\tau - t) e^{-j2\pi\omega(\tau - t)} d\tau \right)}{S_w(\omega, t)} \right) \quad (15)$$

$$= \frac{1}{2\pi} \Im \left( \frac{j2\pi\omega S_w(\omega, t) - S_{w'}(\omega, t)}{S_w(\omega, t)} \right) \quad (16)$$

$$= \omega - \frac{1}{2\pi} \Im \left( \frac{S_{w'}(\omega, t)}{S_w(\omega, t)} \right) \quad (17)$$

where  $S_w$  is the Fourier spectrum of the windowed signal and  $S_{w'}$  is the spectrum of the signal windowed using the derivative of the window. In the special case of a single sinusoid ( $P = 1$  in Equation 1), we have  $\hat{\omega} = \omega_1$ . In the general case of multiple sinusoids, if we consider the spectrum at a frequency  $\omega$  close to  $\omega_p$ , we can neglect the influence of the other frequencies, since the analysis window is band limited, and thus we have  $\hat{\omega} \approx \omega_p$  for  $\omega \approx \omega_p$ . More precisely, their absolute difference should not be higher than the half of the width of the main lobe. If the Hann window is used, we should have  $|\omega - \omega_p| < 2/N$ , see [9, 26] for further references. The frequency  $\hat{\omega}$  gives indeed an excellent estimate of  $\omega_p$  when estimated at the frequency of the DFT bin nearest to  $\omega_p$ .

The reassignment estimator is the first estimator we consider in our study:

$$\hat{\omega}^r = \omega - \frac{1}{2\pi} \Im \left( \frac{S_{w'}(\omega, t)}{S_w(\omega, t)} \right) \quad (18)$$

or more precisely its discrete version:

$$\hat{\omega}_p^r = \frac{k_p}{N} - \frac{1}{2\pi} \Im \left( \frac{S_{w'}[k_p, n]}{S_w[k_p, n]} \right) \quad (19)$$

where  $k_p$  is the DFT bin number of the  $p$ -th local maxima in the magnitude spectrum for the sinusoidal component under consideration, and  $n$  is the sample index corresponding to the time where this  $N$ -point DFT is computed.



## 3.2 Derivative Estimator

We have shown in [27, 15] that it is also possible to greatly improve the precision of the classic Fourier analysis by taking advantage of the first  $d$  derivatives of the signal itself. For  $d = 1$ , this estimator is also known as the derivative algorithm.

### 3.2.1 Continuous Case

In the case of continuous time, from Equation 1, the analytic derivative of the signal  $s$  is given by:

$$s'(t) = \frac{d}{dt}s(t) = \sum_{p=1}^P \frac{d}{dt}A_p(t) e^{j\phi_p(t)} + j \sum_{p=1}^P A_p(t) \frac{d}{dt}\phi_p(t) e^{j\phi_p(t)} \quad (20)$$

If the amplitudes are constant, their derivatives are zero, and since the derivative of the phases are the frequencies, with Equation 2, we have:

$$s'(t) = j2\pi \sum_{p=1}^P A_p \omega_p e^{j\phi_p(t)}. \quad (21)$$

In the special case of a single sinusoid ( $P = 1$ ), we get  $s'(t) = j2\pi\omega_1 s(t)$ , thus  $S'(t) = j2\pi\omega_1 S(t)$  provided that  $\omega_1$  is constant. In the general case of multiple sinusoids, the same considerations as for  $\hat{\omega}^r$  lead to:

$$S'(\omega, t) = j2\pi\omega_p S(\omega, t) \quad \text{for } \omega \approx \omega_p \quad (22)$$

thus we define

$$\hat{\omega}^d = \frac{1}{2\pi} \Im \left( \frac{S'_w(\omega, t)}{S_w(\omega, t)} \right) \quad (23)$$

where  $S'_w$  is the Fourier spectrum of the windowed derivative of the signal.  $\hat{\omega}^d$  is the continuous version of the derivative estimator, proposed in [15].

In order to prove the equivalence of this estimator and the reassignment estimator, we introduce  $\rho = \tau - t$  which gives another expression for the STFT:

$$S_w(\omega, t) = \int_{-\infty}^{+\infty} s(t + \rho) w(\rho) e^{-j2\pi\omega\rho} d\rho \quad (24)$$

from which we can derive, as we did for the reassignment to obtain Equation

18:

$$\hat{\omega} = \frac{1}{2\pi} \frac{\partial}{\partial t} \phi(\omega, t) = \frac{1}{2\pi} \Im \left( \frac{\partial}{\partial t} \log(S_w(\omega, t)) \right) \quad (25)$$

$$= \frac{1}{2\pi} \Im \left( \frac{\frac{\partial}{\partial t} \left( \int_{-\infty}^{+\infty} s(t + \rho) w(\rho) e^{-j2\pi\omega\rho} d\rho \right)}{S_w(\omega, t)} \right) \quad (26)$$

$$= \frac{1}{2\pi} \Im \left( \frac{S'_w(\omega, t)}{S_w(\omega, t)} \right) \quad (27)$$

that is exactly Equation 23, thus  $\hat{\omega}^r$  and  $\hat{\omega}^d$  are equivalent, at least in theory, since they are two different mathematical formulations of the same physical quantity.

### 3.2.2 Discrete Case

However, in practice, the signal  $s$  is discrete and the derivative of the signal is unknown and must be approximated. In [15], it is proposed to consider the difference  $s^-$  as an approximation of the derivative  $s'$ . More precisely, we have:

$$s[n] = s(n/F_s) \quad (28)$$

$$s^-[n] = (s[n+1] - s[n])F_s \quad (29)$$

where  $F_s$  is the sampling frequency. In fact,  $s^-$  defines a high-pass filter of the signal  $s$  whose gain is  $2F_s \sin(\pi\omega)$ , and we derived in [18] the derivative estimator in the discrete case:

$$\hat{\omega}_p^d = \frac{1}{\pi} \arcsin \left( \left| \frac{S^-[k_p, n]}{2F_s S[k_p, n]} \right| \right) = \frac{1}{\pi} \arcsin \left( \left| \frac{S[k_p, n+1] - S[k_p, n]}{2S[k_p, n]} \right| \right) \quad (30)$$

This estimator will be the second estimator considered in our study.

### 3.3 Difference Estimator

The difference estimator is used in the classic phase vocoder approach [28, 29, 12]. Considering that the frequency is constant during the time-interval between two successive short-term spectra, with a hop size of  $H$  samples, Equation 2 shows that the frequency can be estimated from the phase difference:

$$\hat{\omega} = \frac{1}{2\pi H} \Delta_\phi. \quad (31)$$

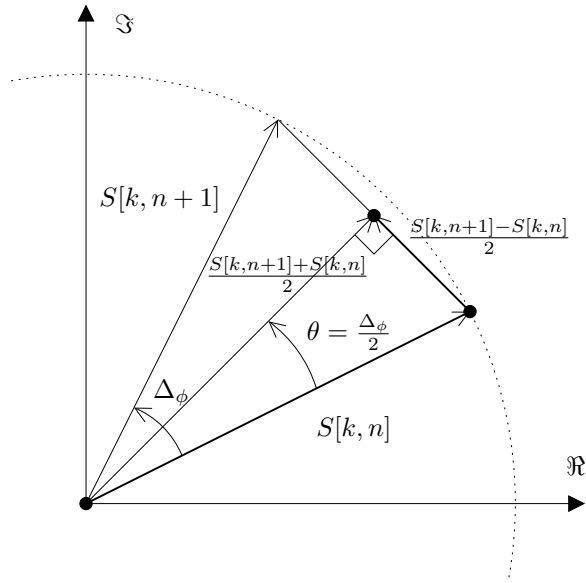


Figure 2: Mathieu Lagrange and Sylvain Marchand. *Vector relationships for phase difference and discrete derivative estimators.*  $S[k, n]$  is the spectrum at DFT bin number  $k$  and time index  $n$ , and  $\Delta\phi$  is the phase difference between two consecutive (short-term) spectra.

Thus, considering that the frequency is constant during the time-interval between two successive short-term spectra, which is especially the case if they are separated by only 1 sample, the frequency is then estimated directly from the phase difference  $\Delta_\phi$  (see Figure 2) by taking care of unwrapping the phase so that this difference is never negative. The resulting estimator, known as the difference estimator, is the third one we consider:

$$\hat{\omega}_p^\Delta = \frac{1}{2\pi} (\angle S[k_p, n+1] - \angle S[k_p, n])_{\text{unwrap}}. \quad (32)$$

### 3.4 Trigonometric Estimator

As shown in [20], the behavior of the derivative estimator is closely linked to the properties of the mathematical function arcsin of Equation 30, which is not a linear transfer function. If the argument gets close to 1, a small error will lead to a non-negligible error on the estimated frequency.

Fortunately, the following identity:

$$\frac{S[k, n+1] + S[k, n]}{2} - S[k, n] = \frac{S[k, n+1] - S[k, n]}{2} \quad (33)$$

and simple geometric considerations lead us to consider a right-angled triangle, see Figure 2 (lower triangle, delimited by 3 dots). Another of its angles measures  $\theta = \Delta_\phi/2$  radians. Simple trigonometric considerations give:

$$\sin(\theta) = \left| \frac{S[k, n+1] - S[k, n]}{2} / S[k, n] \right| \quad (34)$$

$$\cos(\theta) = \left| \frac{S[k, n+1] + S[k, n]}{2} / S[k, n] \right| \quad (35)$$

that is

$$\Delta_\phi = 2 \arcsin \left( \left| \frac{S[k, n+1] - S[k, n]}{2S[k, n]} \right| \right) \quad (36)$$

$$\Delta_\phi = 2 \arccos \left( \left| \frac{S[k, n+1] + S[k, n]}{2S[k, n]} \right| \right). \quad (37)$$

By considering again Equation 31 with  $H = 1$  and a particular frequency component  $p$ , but this time with Equations 36 and 37 to compute  $\Delta_\phi$ , we define:

$$\hat{\omega}_p^- = \frac{1}{\pi} \arcsin \left( \left| \frac{S[k_p, n+1] - S[k_p, n]}{2S[k_p, n]} \right| \right) \quad (38)$$

$$\hat{\omega}_p^+ = \frac{1}{\pi} \arccos \left( \left| \frac{S[k_p, n+1] + S[k_p, n]}{2S[k_p, n]} \right| \right). \quad (39)$$

The errors of the estimator  $\omega_p^-$  – the derivative one – and the estimator  $\omega_p^+$  behave symmetrically with respect to the frequency of the analyzed tone. More precisely, the error of  $\omega_p^+$  is high in the low frequencies and low in the high frequencies whereas the error of the derivative estimator is low in the low frequencies and grows as the frequency grows. Indeed, the argument of the arccos function of the  $\hat{\omega}_p^+$  estimator gets close to 1 when the frequency is close to 0.

The  $\omega_p^+$  estimator can therefore be used in order to improve the precision of the derivative one in the high frequencies. The resulting estimator, recently proposed in [20] and called the trigonometric estimator, is the fourth we consider in this paper:

$$\hat{\omega}_p^t = \begin{cases} \hat{\omega}_p^- & \text{if } k_p/N < 0.25 \\ \hat{\omega}_p^+ & \text{otherwise.} \end{cases} \quad (40)$$

This estimator is of equivalent complexity since it also requires the computation of only two FFTs:  $S[k, n]$  and  $S[k, n + 1]$ .

### 3.5 Arctan Estimator

The third and last way of computing the angle  $\theta$  from the triangle plotted in Figure 2 leads to another estimator proposed in [21]. This estimator, called the arctan estimator, is the fifth and last one we consider in this paper:

$$\hat{\omega}_p^a = \frac{1}{\pi} \arctan \left( \left| \frac{S[k_p, n + 1] - S[k_p, n]}{S[k_p, n + 1] + S[k_p, n]} \right| \right). \quad (41)$$

Together with Equations 30 and 31, and because of the trigonometric relation on  $\theta$  given in Equation 34, Figure 2 shows that the estimators  $\hat{\omega}^d$ ,  $\hat{\omega}^\Delta$ ,  $\hat{\omega}^t$ , and  $\hat{\omega}^a$  are equivalent, at least in theory, since  $\Delta_\phi = 2\theta$ .

Despite the theoretical equivalence of these phase-based estimators, the different mathematical operations used in their discrete versions influence their performance. The next section will study this issue by comparing them on a more practical (implementation) point of view.

## 4 Practical Experiments

To compare the estimators presented in the previous section, we consider in turn a complex or a real signal composed of a periodic part  $x$  with amplitude unity and constant frequency embedded in noise  $y$ . The power of the

noise is chosen to achieve a desired Signal-to-Noise Ratio (SNR) expressed in decibels:

$$\text{SNR} = 10 \log_{10} \left( \frac{\text{var}(x)}{\text{var}(y)} \right). \quad (42)$$

In the experiments, the SNR ranges from  $-20$  dB to  $100$  dB.

We use frames of  $N = 128$  samples ( $F_s = 4$  kHz) and consider 400 different (normalized) frequencies in each considered range.

The lower bound of the limited frequency range is set to  $0.24$  and its upper bound to  $0.26$  (normalized frequencies). The lower bound of the whole frequency range is set to  $0$  and its upper bound to  $0.5$ . These bounds are exclusive, so that the first evaluated frequency in the whole range is  $0.0025$ .

For each frequency, 30 different phases are evaluated from  $0$  to  $2\pi$ . At each evaluation, the noise is randomized. For all the tested methods, the detection picks the greatest local maximum in the power spectrum. Although the estimators can be used for the case of multiple sinusoids, the remainder of the article focuses on the estimation of the frequency of one sinusoid. The index  $p$  is therefore omitted for the sake of clarity.

#### 4.1 The Complex Case

When evaluating the performance of an estimator in terms of variance of the estimation error, an interesting element to compare with is the Cramér-Rao Bound (CRB), defined as the limit to the best possible performance achievable by an unbiased estimator given a dataset [24].

Let us consider a complex sinusoid  $x$  (of amplitude 1) in a Gaussian complex noise  $y$ :

$$x[n] = \exp(2\pi j\omega n + \Phi) \quad (43)$$

$$y[n] = 10^{-\text{SNR}/20} z[n] \quad (44)$$

where  $\omega$  is the frequency and  $z$  is a Gaussian noise of variance 1. The variance of the signal part  $x$  is 1, and the variance of the noise part  $y$  is  $\text{var}(y) = \sigma^2 = 10^{-\text{SNR}/10}$ . The analyzed signal is  $s = x + y$ .

For the case of the estimation of the frequency  $\omega$  of a complex sinusoid in noise, the lower Cramér-Rao bound is [24]:

$$\text{CRB}_c = \frac{6\sigma^2}{a^2 N(N^2 - 1)} = \frac{6}{N(N^2 - 1)} 10^{-\text{SNR}/10} \quad (45)$$

where  $a$  is the amplitude of the sinusoid (here  $a = 1$ ), and the SNR is given by Equation 42. We can easily show that, in the log scales, the CRB in function of the SNR is a line of slope -1.

As discussed in the Section 3.4, the performance of the tested estimators may depend on the frequency of the analyzed sinusoid, whether this frequency is close to the frequency boundaries – 0 or Nyquist frequency – or not. To first evaluate these estimators with a minimal influence of these frequency boundaries, we compare the estimators in the narrow frequency range located around the 0.25 normalized frequency.

As asserted in [19], all the estimators seem to perform similarly at medium SNR, see Figure 3(a). A loss of precision can be observed for  $\hat{\omega}^r$  in the high SNR range. In the special case of only one sinusoid, this bias can be removed [30]. The estimators  $\hat{\omega}^\Delta$  and  $\hat{\omega}^a$  perform slightly better than  $\hat{\omega}^d$  and  $\hat{\omega}^t$ .

When the whole  $]0, 0.5[$  frequency range is considered,  $\hat{\omega}^d$  performs poorly, see Figure 3(b). This can be explained by the lack of precision in the high-frequency region, see Figure 4.

## 4.2 The Real Case

Musical applications consider real sinusoids rather than complex ones. We then consider in this section the signal  $s = x + y$  made of a sinusoid  $x$  (of amplitude 1) in a Gaussian noise  $y$ :

$$x[n] = \sin(2\pi\omega n + \Phi) \quad (46)$$

$$y[n] = \frac{1}{\sqrt{2}} 10^{-\text{SNR}/20} z[n] \quad (47)$$

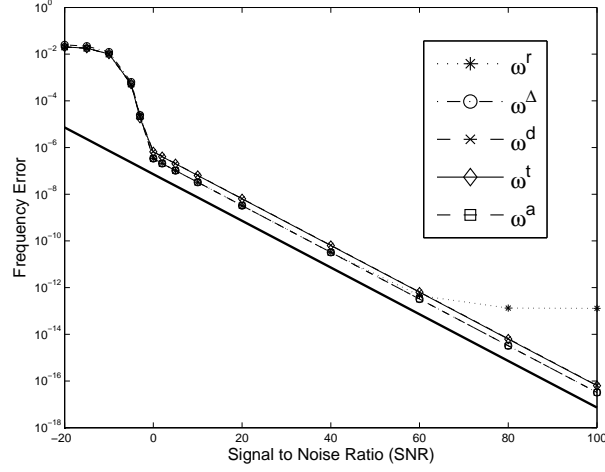
where  $\omega$  is the frequency (in radians per sample) and  $z$  is a Gaussian noise of variance 1. We use the  $1/\sqrt{2}$  normalizing factor to ensure the validity of Equation 42, because in the real case the variance of the sinusoid is  $1/2$ , while we still consider, by definition,  $\text{var}(y) = \sigma^2$ .

For the case of the estimation of the frequency  $\omega$  of a real sinusoid in noise, the lower Cramér-Rao bound is shown to be [31]:

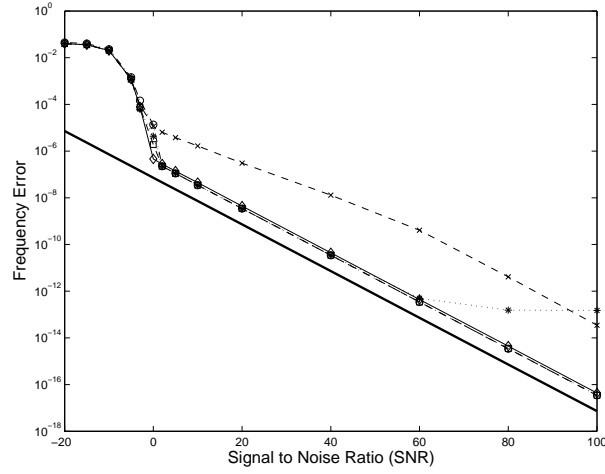
$$\text{CRB}_r = \frac{24\sigma^2}{a^2N(N^2-1)} = \frac{12}{N(N^2-1)} 10^{-\text{SNR}/10} = 2\text{CRB}_c \quad (48)$$

where  $a$  is the amplitude of the sinusoid (here  $a = 1$ ), and the SNR is given by Equation 42.

Due to the spectral properties of real signals, the performance of the estimators are significantly worse than those obtained with complex signals. Indeed, the spectrum of a real sinusoid is made of two Dirac's impulses, located at frequencies  $\omega$  and  $-\omega$ , see Equation 4. Since the spectrum of the sampled signal is periodic, the more the frequency of the analyzed sinusoid



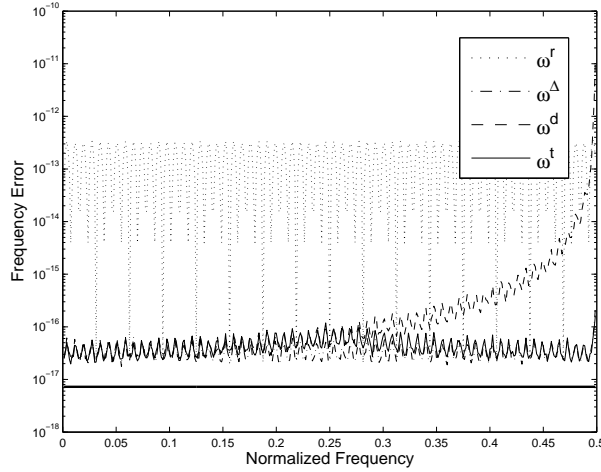
(a) Complex case: narrow frequency range.



(b) Complex case: whole frequency range.

Figure 3: Mathieu Lagrange and Sylvain Marchand. *Performance of the tested estimators for the analysis of a complex sinusoid signal with frequency lying in the  $(0.24, 0.25)$  normalized frequency range (a), and in the  $(0, 0.5)$  range (b). The CRB is plotted with a double solid line.*





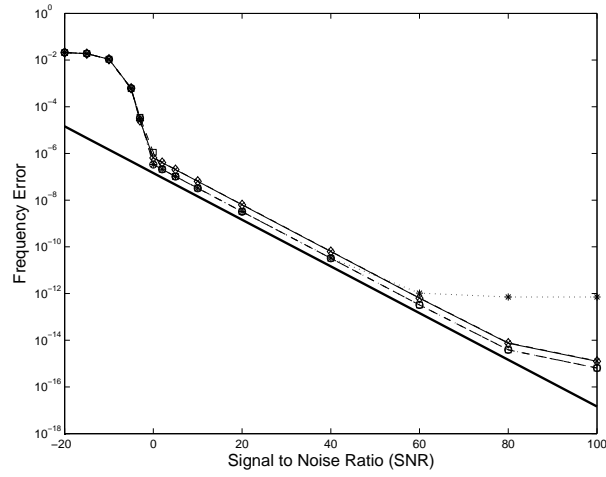
(a) Complex case: whole frequency range.

Figure 4: Mathieu Lagrange and Sylvain Marchand. *Performance of the tested estimators at SNR=100 dB versus the frequency of the analyzed complex sinusoid. The CRB is plotted with a double solid line.*

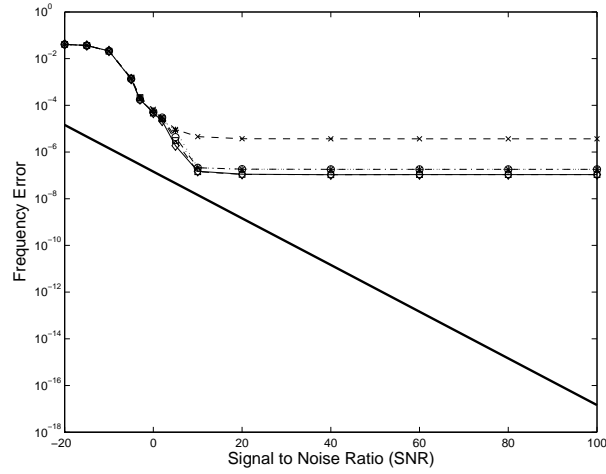
is close to 0 or the Nyquist frequency, the more the interference between the two complex exponentials is pronounced.

If the narrow frequency range is considered, this interference phenomenon is therefore lightly pronounced and the results are roughly equivalent to the complex case, see Figure 5(a). On the other hand, the precision is limited by the interference phenomenon if the whole frequency range is considered, so that the squared error is held asymptotically constant at SNR higher than 10 dB, see Figure 5(b). The estimators  $\hat{\omega}^d$  and  $\hat{\omega}^r$  perform equally, slightly worse than  $\omega^t$  and  $\hat{\omega}^a$  estimators. This is due to the fact that the first DFT bin ( $S[0, n]$ ) is by definition purely real when the signal is real. Since the two last estimators only consider the magnitudes of the spectra, their performance are not affected. On the other hand, the difference and reassignment estimators will always estimate a frequency zero if the frequency of the analyzed sinusoid falls into the first bin, *i.e.*  $\omega < 1/N$ .

These experiments show that the influence of the negative frequencies greatly disturbs the phase-based estimators.



(a) Real case: narrow frequency range.



(b) Real case: whole frequency range.

Figure 5: Mathieu Lagrange and Sylvain Marchand. *Performance of the tested estimators for the analysis of a real sinusoidal signal with frequency lying in the  $(0.24, 0.25)$  normalized frequency range (a), and in the  $(0, 0.5)$  range (b).*

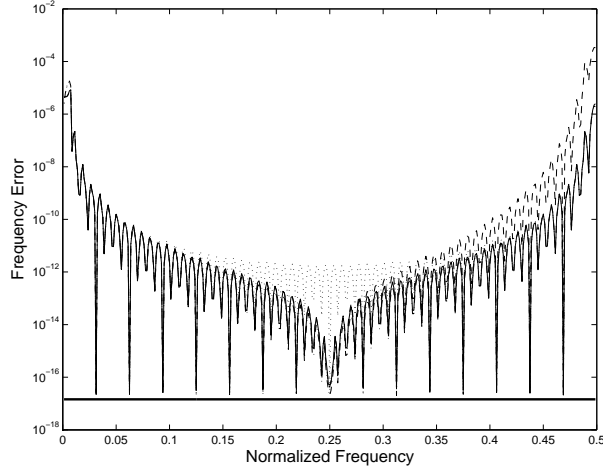


Figure 6: Mathieu Lagrange and Sylvain Marchand. *Performance of the tested estimators at SNR=100 dB versus the frequency of the analyzed real sinusoid.*

### 4.3 The Analytic Case

We now propose to bring down the influence of the negative frequencies over the positive part of the spectrum by considering the “analytic” signal instead of the real input signal.

An analytic signal is a signal which has no negative frequency. For example, a real sinusoid may be converted into a positive frequency complex sinusoid by adding a quarter-cycle shifted imaginary part, since:

$$\cos(2\pi\omega t + \Phi) + j \sin(2\pi\omega t + \Phi) = e^{j(2\pi\omega t + \Phi)} \quad (49)$$

For more complex signals, we can use a “Hilbert transform filter” which shifts each sinusoidal component by a quarter cycle. Ideally, this filter has magnitude 1 at all frequencies and introduces a phase shift of  $-\pi/2$  at each positive frequency and  $\pi/2$  at each negative frequency. The impulse response of such an ideal  $\pi/2$  radians phase shifter is [32, 33]:

$$h[n] = \frac{1 - \cos(\pi n)}{\pi n} \text{ for } n \in \left[ -\frac{N-1}{2}, \frac{N-1}{2} \right] \quad (50)$$

Figure 7 shows how the system can be used to compute the analytic signal  $s_a[n]$  defined as:

$$s_a[n] = s_r[n] + js_h[n] \quad (51)$$

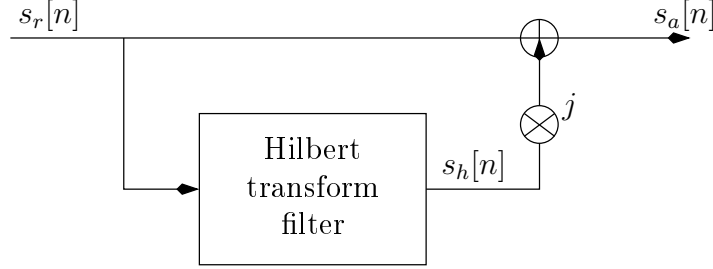


Figure 7: Mathieu Lagrange and Sylvain Marchand. *Block diagram representation of the creation of the analytic signal  $s_a[n]$  given the real signal  $s_r[n]$  and its Hilbert transform  $s_h[n]$ .*

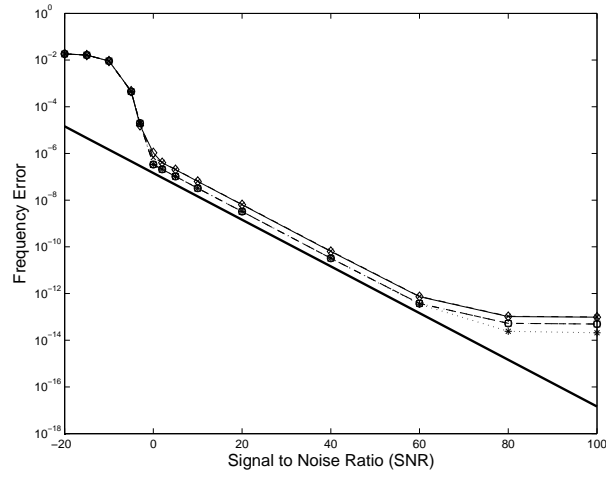
where  $s_r[n]$  is the analyzed real signal and  $s_h[n]$  is the output of the Hilbert transform filter. Given the impulse response of the filter given by Equation 50 or computed using practical Hilbert transform filter design methods [34, 35], the Hilbert transform is efficiently computed using an FFT-based convolution.

The real signal is now converted into an analytic signal before the frequency estimation process. If an ideal filter could be considered, the performance of the estimators should be similar to those of the complex case. Unfortunately, the magnitude of this filter is not exactly unity due to the ripple of the filter in the passing band. This globally degrades the performance of the estimators in the short frequency range, except for the  $\hat{\omega}^r$  which obtains slightly better results. This enhancement is due to a reduction of the estimation bias as it will be explained in Section 5.1. The  $\hat{\omega}^d$  and  $\hat{\omega}^t$  estimators perform similarly with slightly lower performance than the two other estimators, see Figure 8(a).

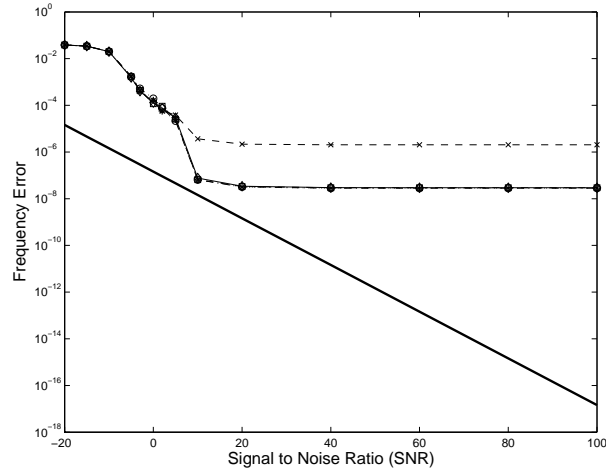
If the whole frequency range is considered, all the estimators achieve similar performance, see Figure 8(b). Due to the lower influence of the negative frequencies, the performance of the estimators are globally enhanced with respect to the ones obtained in the real case as discussed in the next section.

## 5 Discussion

The previous section detailed the performance behavior of the phase-based estimators for three signal types: complex, real, or analytic. We now globally



(a) Analytic case: short frequency range.



(b) Analytic case: whole frequency range.

Figure 8: Mathieu Lagrange and Sylvain Marchand. *Performance of the tested estimators for the analysis of a Hilbert transformed real sinusoid with frequency lying in the  $(0.24, 0.25)$  normalized frequency range (a), and in the  $(0, 0.5)$  range (b).*

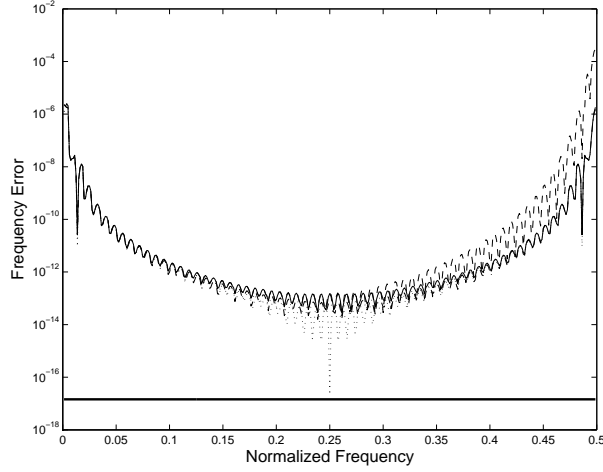


Figure 9: Mathieu Lagrange and Sylvain Marchand. *Performance of the tested estimators at SNR=100 dB after applying a Hilbert transform to the real sinusoidal signal versus its frequency.*

compare these estimators depending on the type of signal, frequency range, and SNR range. The estimation bias is studied in a first part and the variance of the estimation error in a second part.

### 5.1 Bias of the Error

We compare the estimation bias of the evaluated estimators by considering the log-bias. The log-bias is defined as the base-10 logarithm of the absolute value of the mean error achieved at a given frequency so that this error is maximal in the considered frequency range:

$$b_{\log} = \log_{10} (\max_{\omega} |\hat{\omega} - \omega|). \quad (52)$$

For each type of signal, the two frequency ranges are considered: the narrow frequency range around 0.25 normalized frequency and the whole frequency range (0,0.5). Two SNR ranges are also considered: a high SNR range  $\{20, 40, 60, 80, 100\}$  and a low SNR range  $\{-20, -10, 0, 10\}$ .

As can be seen in Table 2,  $\hat{\omega}^r$  exhibits a higher bias than the others estimators, if not masked by the interference phenomenon when real signals are considered. However, the delta between the bias of  $\hat{\omega}^r$  and the bias of the others is consequently diminished when considering the analytic signal.

Signal	Complex		Real		Analytic	
Range	Narrow	Whole	Narrow	Whole	Narrow	Whole
difference $\hat{\omega}^\Delta$	<b>-6.86</b> (-2.77)	-6.82 (-1.96)	<b>-6.81</b> <b>(-2.93)</b>	-2.57 <b>(-1.88)</b>	-6.18 <b>(-2.91)</b>	-3.15 <b>(-1.59)</b>
reassignment $\hat{\omega}^r$	-5.80 <b>(-2.85)</b>	-5.82 <b>(-2.41)</b>	-5.85 (-2.91)	-2.57 (-1.83)	<b>-6.38</b> (-2.88)	-3.15 (-1.58)
derivative $\hat{\omega}^d$	-6.72 (-2.81)	-4.49 (-1.70)	-6.70 (-2.73)	-1.95 (-1.50)	-6.06 (-2.79)	-1.97 (-1.37)
trigonometric $\hat{\omega}^t$	-6.73 (-2.79)	<b>-6.84</b> <b>(-2.41)</b>	-6.74 (-2.84)	-2.81 (-1.57)	-6.09 (-2.79)	-3.20 (-1.46)
arctan $\hat{\omega}^a$	-5.78 (-1.81)	-5.76 (-1.66)	-5.70 (-1.81)	<b>-3.79</b> (-1.60)	-5.42 (-1.81)	<b>-4.08</b> (-1.55)

Table 2: Mathieu Lagrange and Sylvain Marchand. *Mean log-bias of the phase-based estimators for the SNR values expressed in dB:  $\{20, 40, 60, 80, 100\}$ . For the mean log-bias in parentheses, the SNR values are:  $\{-20, -10, 0, 10\}$ . The estimators are tested for several signal types. From left to right, the analyzed signal is a complex tone in a narrow frequency range and in the whole frequency range, so as for the real tone and the analytic tone. Bold numbers indicate – column-wise – the minimal bias achieved by the tested estimators for a given signal type, frequency range, and SNR range.*

## 5.2 Variance of the Error

We compare the variance of the error of each estimator by considering the log-efficiency. The log-efficiency of an estimator is defined as the base-10 logarithm of the ratio between the variance of the error [36] versus the CRB for a given SNR:

$$e_{\log} = \log_{10} \left( \frac{\text{var}(\hat{\omega} - \omega)}{\text{CRB}} \right). \quad (53)$$

As in the previous section, the two frequency ranges are considered for each type of signal: the narrow frequency range around 0.25 normalized frequency and the whole frequency range. A limited range  $]1/N, 0.5[$  is also considered for the real case to compare the estimators without favoring the trigonometric and arctan estimators, see Section 4.2. Two SNR ranges are considered: a high SNR range  $\{20, 40, 60, 80, 100\}$  and a low SNR range  $\{-20, -10, 0, 10\}$ , in parentheses in Table 3. For each of these ranges, the results reported are the mean value of the log-efficiency over the considered range.

The results of the experiments reported in the last section show that the estimation bias is not negligible and thus strongly influences the variance of the error. Therefore, a correlation between the results of the two Tables 2 and 3 is observed.

In the case of a complex signal, the difference and arctan estimators achieve the best performance, so as for the narrow frequency range in the case of real signals. The trigonometric and arctan estimators achieve the best performance for the whole frequency range in the real case and the reassignment performs best in the analytic case thanks to the reduction of the estimation bias.

Considering the analytic signal instead of the real one enhances the performance if the whole frequency range is considered. The use of the analytic signal computed with the Hilbert transform filter of Equation 50 is then relevant for improving the precision of the estimation of the frequency using phase-based estimators since this pre-processing step could also be achieved efficiently in the spectral domain.

## 5.3 Comparison with Other Classes of Estimators

The phase-based estimators are now compared to other classes of estimators. This evaluation is only provided to roughly rank those approaches and do not aim at a fair and in-depth comparison. The Grandke estimator [7] belongs to the polynomial class and requires only one FFT computation per frame.



Signal	Complex		Real			Analytic	
Range	Narrow	Whole	Narrow	Whole	Limited	Narrow	Whole
difference	<b>0.65</b> (2.27)	<b>0.68</b> (2.81)	<b>0.62</b> <b>(1.94)</b>	6.10 (2.84)	<b>5.10</b> <b>(2.60)</b>	1.24 (1.90)	<b>5.30</b> (2.84)
reassignment	1.72 <b>(2.23)</b>	1.77 (2.66)	1.79 (1.94)	6.10 (2.83)	<b>5.10</b> (2.64)	<b>1.09</b> <b>(1.90)</b>	<b>5.30</b> (2.82)
derivative	0.95 (2.38)	3.41 (3.19)	0.93 (2.09)	7.41 (3.20)	7.40 (3.12)	1.54 (2.05)	7.15 (3.24)
trigonometric	0.94 (2.37)	0.78 <b>(2.44)</b>	0.91 (2.08)	<b>5.87</b> <b>(2.79)</b>	5.15 (2.63)	1.53 (2.10)	5.32 (2.81)
arctan	<b>0.65</b> <b>(2.23)</b>	<b>0.68</b> (2.57)	0.63 (2.07)	<b>5.87</b> <b>(2.79)</b>	5.15 (2.65)	1.24 <b>(1.90)</b>	5.32 <b>(2.78)</b>

Table 3: Mathieu Lagrange and Sylvain Marchand. *Mean log-efficiency of the phase-based estimators for the SNR values expressed in dB:  $\{20, 40, 60, 80, 100\}$ . For the mean log-efficiency in parentheses, the SNR values are:  $\{-20, -10, 0, 10\}$ . The estimators are tested for several signal types. From left to right, the analyzed signal is a complex tone in a narrow frequency range and in the whole frequency range, so as for the real tone and the analytic tone. Bold numbers indicate – column-wise – the best performance achieved for a given signal type, frequency range, and SNR range.*

Signal	Complex		Real		Analytic	
Range	Narrow	Whole	Narrow	Whole	Narrow	Whole
$\hat{\omega}^{\text{DFT}}$	7.78 (3.04)	7.86 (3.27)	7.47 (2.75)	7.57 (3.20)	7.47 (2.72)	7.56 (3.28)
Grandke	0.84 (2.33)	0.83 (2.48)	0.54 (2.03)	6.12 (2.86)	0.78 (2.00)	5.58 (2.84)
$\hat{\omega}^{\Delta}$	0.65 (2.27)	0.68 (2.81)	0.62 (1.94)	6.10 (2.84)	1.24 (1.90)	5.30 (2.84)
ESPRIT	0.41 (2.54)	0.41 (3.02)	0.11 (1.85)	0.71 (2.50)	3.10 (1.74)	5.36 (2.65)

Table 4: Mathieu Lagrange and Sylvain Marchand. *Mean log-efficiency of several classes of estimators for the SNR values:  $\{20, 40, 60, 80, 100\}$  dB and in parentheses:  $\{-20, -10, 0, 10\}$  dB.*

The difference estimator  $\hat{\omega}^{\Delta}$  belongs to the phase-based class and require two FFT computation per frame. The ESPRIT algorithm [4] belongs to the High-Resolution class and requires knowledge of the number of sinusoids to achieve the best performance. This algorithm is based on the subspace decomposition of the autocorrelation matrix of the signal and requires two eigenvalues decompositions. This approach is therefore much more complex. In the complex and analytic case, the number of complex sinusoids to be found is set to 1. In the real case, the number of complex sinusoids is set to 2.

According to the results presented in Table 4, the Grandke and  $\hat{\omega}^{\Delta}$  estimators provide a considerable improvement over the rough FFT-based estimator.  $\hat{\omega}^{\Delta}$  behaves better, but according to [9], the polynomial estimators can be improved as desired by using a convenient ZP factor at the expense of some computational complexity.

The ESPRIT algorithm takes advantages of the properties of the analyzed signal which should be a sum of a known number of complex sinusoids with AWGN. Using this prior knowledge, this class of algorithms outperforms the Fourier-based methods in terms of resolution and precision. This leads to a significant improvement in the case of a real sinusoid in the whole frequency range. However, this estimator achieves poor results in the analytical case, probably due to the fact that the signal is filtered and therefore cannot be purely explained as a sum of sinusoids with AWGN and also because the Hilbert transform does not suppress all the energy of the negative frequency components.

## 6 Conclusion

In this article, we have studied the use of phase-based estimators to estimate the frequency of complex and real tones.

We have first demonstrated their theoretical equivalence. More precisely, the reassignment estimator was proved to be equivalent to the derivative one in the continuous case. The discrete version of this last estimator was then shown to be equivalent to the difference and the trigonometric estimators.

This equivalence was guessed in practice by Keiler and Marchand in [16], partly demonstrated in theory by Hainsworth [17] during his Ph.D., and recently explained in [37].

The numerical experiments reported in this article globally confirm these results although slight differences were observed, which are most probably due to the numerical imprecision of mathematical operators. The difference estimator achieves the most stable results, except in the case of real signals with very low frequency components. In this case, the trigonometric or arctan estimators should be preferred.

Finally, the Hilbert transform can be used as a pre-processing step to minimize the influence of the negative frequencies in the case of real signals. Experiments show that the overall performance of these estimators is enhanced, especially in the case of the reassignment estimator.

## References

- [1] Robert J. McAulay and Thomas F. Quatieri, “Speech Analysis/Synthesis Based on a Sinusoidal Representation,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 4, pp. 744–754, 1986.
- [2] Xavier Serra, *Musical Signal Processing*, chapter Musical Sound Modeling with Sinusoids plus Noise, pp. 91–122, Studies on New Music Research. Swets & Zeitlinger, Lisse, the Netherlands, 1997.
- [3] Ralph O. Schmidt, “Multiple Emitter Location and Signal Parameter Estimation,” *IEEE Transactions on Antennas and Propagation*, vol. 4, no. 34, pp. 276–280, 1986.
- [4] Robert Roy, Arogyaswami Paulraj, and Thomas Kailath, “ESPRIT – A Subspace Rotation Approach to Estimation of Parameters of Cisoids in Noise,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 5, no. 34, pp. 1340–1344, 1986.

- [5] James L. Flanagan and Roger M. Golden, “Phase Vocoder,” *Bell System Technical Journal*, vol. 45, pp. 1493–1509, 1966.
- [6] Michael R. Portnoff, “Implementation of the Digital Phase Vocoder Using the Fast Fourier Transform,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 3, pp. 243–248, 1976.
- [7] Thomas Grandke, “Interpolation Algorithms for Discrete Fourier Transforms of Weighted Signals,” *IEEE Transactions on Instruments and Measurements*, vol. 32, no. 2, pp. 350–355, 1983.
- [8] Malcolm D. Macleod, “Fast Nearly ML Estimation of the Parameters of Real or Complex Single Tones or Resolved Multiple Tones,” *IEEE Transactions on Signal Processing*, vol. 46, no. 1, pp. 141–148, 1998.
- [9] Mototsugu Abe and Julius O. Smith, “Design Criteria for Simple Sinusoidal Parameter Estimation based on Quadratic Interpolation of FFT Magnitude Peaks,” in *117th Convention of the Audio Engineering Society*, San Francisco, October 2004, Preprint 6256.
- [10] François Auger and Patrick Flandrin, “Improving the Readability of Time-Frequency and Time-Scale Representations by the Reassignment Method,” *IEEE Transactions on Signal Processing*, vol. 43, pp. 1068–1089, 1995.
- [11] Kelly Fitz and Lippold Haken, “On the Use of Time/Frequency Reassignment in Additive Sound Modeling,” *Journal of the Audio Engineering Society*, vol. 50, no. 11, pp. 879–893, 2002.
- [12] Daniel Arfib, Florian Keiler, and Udo Zölzer, *DAFx – Digital Audio Effects*, J. Wiley and Sons, 2002, chapter 9, pp. 299–372.
- [13] James A. Moorer, “Signal Processing Aspects of Computer Music – A Survey,” *Computer Music Journal*, vol. 1, no. 1, pp. 4–37, 1977.
- [14] James A. Moorer, “The Use of the Phase Vocoder in Computer Music Applications,” *Journal of the Audio Engineering Society*, vol. 26, no. 1/2, pp. 42–45, 1978.
- [15] Myriam Desainte-Catherine and Sylvain Marchand, “High Precision Fourier Analysis of Sounds Using Signal Derivatives,” *Journal of the Audio Engineering Society*, vol. 48, no. 7/8, pp. 654–667, 2000.

- [16] Florian Keiler and Sylvain Marchand, “Survey on Extraction of Sinusoids in Stationary Sounds,” in *Proceedings of the Digital Audio Effects (DAFx) Conference*, Hamburg, Germany, September 2002, pp. 51–58.
- [17] Stephen W. Hainsworth, *Techniques for the Automated Analysis of Musical Audio*, Ph.D. thesis, University of Cambridge, United Kingdom, Signal Processing Group, Department of Engineering, December 2003.
- [18] Sylvain Marchand, *Sound Models for Computer Music (analysis, transformation, synthesis)*, Ph.D. thesis, University of Bordeaux 1, France, LaBRI, December 2000.
- [19] Stephen W. Hainsworth and Malcom D. Macleod, “On Sinusoidal Parameter Estimation,” in *Proceedings of the Digital Audio Effects (DAFx) Conference*, London, United Kingdom, September 2003, pp. 151–156.
- [20] Mathieu Lagrange, Sylvain Marchand, and Jean-Bernard Rault, “Improving Sinusoidal Frequency Estimation Using a Trigonometric Approach,” in *Proceedings of the Digital Audio Effects (DAFx) Conference*, Madrid, Spain, September 2005, pp. 110–115.
- [21] Michaël Betser, Patrice Collen, Gaël Richard, and Bertrand David, “Review and Discussion on Classical STFT-Based Frequency Estimators,” in *120th Convention of the Audio Engineering Society*, Paris, May 2006.
- [22] Jean-Claude Risset and Max V. Mathews, “Analysis of Musical-Instrument Tones,” *Physics Today*, vol. 22, no. 2, pp. 23–30, 1969.
- [23] Thomas F. Quatieri and Robert J. McAulay, “Audio Signal Processing Based on Sinusoidal Analysis/Synthesis,” in *Applications of DSP to Audio & Acoustics*, Mark Kahrs and Karlheinz Brandenburg, Eds., pp. 343–416. Kluwer Academic Publishers, Boston/Dordrecht/London, 1998.
- [24] David C. Rife and Robert R. Boorstyn, “Single-Tone Parameter Estimation from Discrete-Time Observations,” *IEEE Transactions on Information Theory*, vol. 20, pp. 591–598, 1974.
- [25] David. C. Rife and Robert R. Boorstyn, “Multiple Tone Parameter Estimation from Discrete-Time Observations,” *Bell System Technical Journal*, vol. 55, no. 3, pp. 1389–1410, 1976.

- [26] Saman S. Abeysekera and Kabi P. Padhi, “An Investigation of Window Effects on the Frequency Estimation Using the Phase Vocoder,” *IEEE Transactions on Acoustics, Speech, and Language Processing*, vol. 14, no. 4, pp. 1432–1440, 2006.
- [27] Sylvain Marchand, “Improving Spectral Analysis Precision with an Enhanced Phase Vocoder using Signal Derivatives,” in *Proceedings of the Digital Audio Effects (DAFx) Conference*, Barcelona, Spain, November 1998, pp. 114–118.
- [28] Mark Dolson, “The Phase Vocoder: A Tutorial,” *Computer Music Journal*, vol. 10, no. 4, pp. 14–27, 1986.
- [29] Miller Puckette, “Phase-locked vocoder,” in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, New York, October 1995.
- [30] Stephen W. Hainsworth and Malcom D. Macleod, “Time-Frequency Reassignment: Measures and Uses,” in *Proceedings of the Cambridge Music Processing Colloquium*, 2003, pp. 36–39.
- [31] Steven M. Kay, *Fundamentals of Statistical Signal Processing – Estimation Theory*, Signal Processing Series. Prentice Hall, 1993.
- [32] Alan V. Oppenheim and Ronald W. Schaffer, *Discrete Time Signal Processing*, Prentice Hall, 1989.
- [33] Julius O. Smith, *Mathematics of the Discrete Fourier Transform (DFT)*, W3K Publishing, 2003.
- [34] Larry R. Rabiner and Bernard Gold, *Theory and Application of Digital Signal Processing*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1975.
- [35] Michael Z. Komodromos, Steve F. Russel, and Ping Tak Peter Tang, “Design of FIR Hilbert Transformers and Differentiators in the Complex Domain,” *IEEE Transactions on Circuits and Systems*, vol. 45, no. 1, pp. 64–67, 1998.
- [36] Roland Badeau, *Méthodes à haute résolution pour l’estimation et le suivi de sinusoïdes modulées – Application aux signaux de musique*, Ph.D. thesis, École Nationale Supérieure des Télécommunications, Paris, France, April 2005, in French.

- [37] Sylvain Marchand and Mathieu Lagrange, “On the Equivalence of Phase-Based Methods for the Estimation of Instantaneous Frequency,” in *Proceedings of the 14th European Conference on Signal Processing (EUSIPCO)*, Florence, Italy, September 2006.