ON THE TEAGER-KAISER ENERGY OPERATOR "LOW FREQUENCY ERROR"

M.J. Lipsey and J.P. Havlicek

School of Electrical and Computer Engineering
University of Oklahoma, Norman, OK 73019-1023 USA
{lipsdogg, joebob}@ou.edu

ABSTRACT

We investigate frequency-dependent frequency estimation errors in the DESA-2 energy separation algorithm associated with the discrete Teager-Kaiser energy operator (TKEO). The TKEO and DESA-2 algorithm associate instantaneous amplitude and frequency functions to a real-valued discrete-time signal. It has been observed informally that there seems to be a frequency dependent error in the instantaneous frequency estimates delivered by the approach, where the error is more pronounced at lower frequencies. By studying quadratic phase signals, we demonstrate that this TKEO "low frequency error" does in fact exist and is related to both the magnitude frequency and the signal chirp rate.

1. INTRODUCTION

The continuous-time Teager-Kaiser energy operator (TKEO) applied to a real-valued signal $x(t)$ is given by [1,2]

$$\Psi_c[x(t)] = \dot{x}^2(t) - \ddot{x}(t)x(t). \tag{1}$$

This operator and its discrete counterpart were introduced by Teager and systematically studied by Kaiser and others [1–3]. For a pure sinusoid $x(t) = A\cos(\omega_0 t)$, where $A, \omega_0 \in \mathbb{R}$ are constants, application of the TKEO results in

$$\Psi_c[x(t)] = A^2\omega_0^2. \tag{2}$$

The quantity in (2) is known as the *Teager energy* of the signal; it is proportional to the energy required to generate the displacement $x(t)$ in a mass-spring harmonic oscillator.

More generally, consider a real signal $x(t)$ with joint amplitude and frequency modulation given by

$$x(t) = a(t)\cos[\varphi(t)], \tag{3}$$

where $a, \varphi : \mathbb{R} \to \mathbb{R}$. Signals of this type are referred to as *AM-FM signals*. The function $a(t)$ is called the AM function of $x(t)$, while $\dot{\varphi}(t)$ is called the FM function of $x(t)$. Although it is often useful to make a further constraint in the model (3) by requiring $a(t)$ to be positive semidefinite, we will not do so in this paper because we will be concerned primarily with frequency modulation.

If the modulating functions $a(t)$ and $\dot{\varphi}(t)$ in the AM-FM signal (3) are locally smooth in a certain sense [4,5], or, alternatively, if the AM and FM modulation indices are not too large [1,2], then application of the TKEO yields

$$\Psi_c[x(t)] \approx a^2(t)\dot{\varphi}^2(t) \tag{4}$$

with an approximation error that is generally small. The quantity $a^2(t)\dot{\varphi}^2(t)$ on the RHS of (4) defines the Teager energy of the general AM-FM signal (3).

The discrete TKEO is defined by [3]

$$\Psi_d(x[n]) = x^2[n] - x[n+1]x[n-1]. \tag{5}$$

Analogous to the continuous case, application of the discrete operator (5) to a discrete-time AM-FM signal

$$x[n] = a[n]\cos(\varphi[n]) \tag{6}$$

yields

$$\Psi_d(x[n]) \approx a^2[n]\sin^2(\dot{\varphi}[n]) \tag{7}$$

with an error that is small under appropriate smoothness assumptions on the local variations in the AM and FM functions [1,2,4,5]. The notation $\dot{\varphi}[n]$ in (7) indicates that the derivative operates symbolically on the discrete variable n as though it were a continuous variable; an equivalent interpretation is that $x[n]$ in (6) contains the samples of $x(t)$ in (3) while $\dot{\varphi}[n]$ contains the samples of $\dot{\varphi}(t)$ (where a unity sampling interval may be assumed for simplicity WLOG). The RHS of (7) defines the discrete Teager energy for the signal (6). Like the continuous case, there is no approximation error in (7) when $x[n]$ is a pure sinusoid. Also, the approximation errors tend to increase as $a[n]$ and $\varphi[n]$ increasingly deviate from constant and linear behaviors, respectively.

The nonlinear operators Ψ_c and Ψ_d are interesting because they are computationally simple, temporally localized, and capable under appropriate signal constraints of accurately tracking the instantaneous amplitude a and instantaneous frequency $\dot{\varphi}$. Clearly, it is the discrete operator

0-7803-7523-8/02/$17.00 ©2002 IEEE

Ψ_d that is of greater practical interest in modern DSP applications. Ψ_d has been used extensively in digital speech processing and analysis with significant results [2, 6–9].

In Section 2 we briefly review energy separation algorithms (ESA's) that use the TKEO to estimate individually the AM and FM functions of a signal. On several occasions we have heard it remarked informally that, relative to the frequency estimation error that is incurred at high instantaneous frequency, there seems to be an increased error in the estimates of $\dot{\varphi}[n]$ that occurs when the instantaneous frequency is low. We have also observed this unexpected "low frequency error" in our own laboratory. This paper presents preliminary results from the first systematic study of the TKEO "low frequency error," where our investigation is limited to discrete-time quadratic phase signals (chirps). We present empirical results in Section 3 that clearly demonstrate the existence of a consistent, repeatable frequency-dependent frequency estimation error that increases as the instantaneous frequency $\dot{\varphi}[n]$ approaches zero. Conclusions are reserved for Section 4.

2. ENERGY SEPARATION

It was shown in [2] that, for the signal $x(t)$ in (3), application of the TKEO to the differentiated signal $\dot{x}(t)$ gives

$$\Psi_c[\dot{x}(t)] \approx a^2(t)\dot{\varphi}^4(t) \tag{8}$$

with approximation errors that may be expected to be small under conditions similar to those that lead to small errors in (4). Together, (4) and (8) immediately imply the ESA [2]

$$|a(t)| \approx \frac{\Psi_c[x(t)]}{\sqrt{\Psi_c[\dot{x}(t)]}} \tag{9}$$

$$|\dot{\varphi}(t)| \approx \sqrt{\frac{\Psi_c[\dot{x}(t)]}{\Psi_c[x(t)]}} \tag{10}$$

for estimating the magnitudes of the AM and FM functions individually.

Discrete energy separation algorithms (DESA's) applicable to the signal $x[n]$ in (6) were also developed in [2]. By passing the continuous-time derivatives in (9), (10) to discrete asymmetric backward differences, one obtains the DESA-1a algorithm

$$|a[n]| \approx \sqrt{\frac{\Psi_d(x[n])}{1-\left(1-\frac{\Psi_d(x[n]-x[n-1])}{2\Psi_d(x[n])}\right)^2}} \tag{11}$$

$$|\dot{\varphi}[n]| \approx \arccos\left(1-\frac{\Psi_d(x[n]-x[n-1])}{2\Psi_d(x[n])}\right) \tag{12}$$

and the DESA-1 algorithm

$$|a[n]| \approx \sqrt{\frac{\Psi_d(x[n])}{1-\left(1-\frac{\Psi_d(y[n])+\Psi_d(y[n+1])}{4\Psi_d(x[n])}\right)}} \tag{13}$$

$$|\dot{\varphi}[n]| \approx \arccos\left(1-\frac{\Psi_d(y[n])+\Psi_d(y[n+1])}{4\Psi_d(x[n])}\right), \tag{14}$$

where $y[n]=x[n]-x[n-1]$. By alternatively passing the continuous-time derivatives in (9), (10) to discrete symmetric differences, the DESA-2 algorithm

$$|a[n]| \approx \frac{2\Psi_d(x[n])}{\sqrt{\Psi_d(x[n+1]-x[n-1])}} \tag{15}$$

$$|\dot{\varphi}[n]| \approx \arcsin\left(\sqrt{\frac{\Psi_d(x[n+1]-x[n-1])}{4\Psi_d(x[n])}}\right) \tag{16}$$

is obtained.

The three algorithms DESA-1a, DESA-1, and DESA-2 generally deliver comparable performance with regards to approximation errors for broad classes of signals of general interest in practical applications. Since DESA-2 offers the lowest computational complexity, it has been the most widely utilized among the three; we henceforth focus our attention on DESA-2 exclusively.

3. DESA-2 FREQUENCY ESTIMATION ERRORS

For a discrete-time signal $x[n]$ that is a pure sinusoid (constant AM and linear phase), there are no approximation errors in the DESA-2 algorithm. To investigate the discrete TKEO "low frequency error," we consider the simplest nontrivial case where the AM function $a[n]$ is constant and the instantaneous phase $\varphi[n]$ is a quadratic of the form

$$\varphi[n] = \omega_c n + \omega_m\left(\frac{n^2}{N}-n\right)+\theta \tag{17}$$

defined for $0 \le n \le N$. In (17), $\theta = \varphi[0]$ is a constant phase offset, the constant ω_c may be interpreted as the FM carrier frequency, and the constant ω_m determines the chirp rate of the signal (which will be defined below).

For the AM-FM signal in (6) with constant AM given by $a[n]=1$ and phase given by (17), the FM function (instantaneous frequency) is given by

$$\begin{aligned}\dot{\varphi}[n] &= \omega_c + \omega_m\left(\frac{2n}{N}-1\right)\\ &= \frac{2\omega_m}{N}n + (\omega_c-\omega_m).\end{aligned} \tag{18}$$

Thus, the signal $x[n]$ has an instantaneous frequency that is linear in n. We refer to the slope $\frac{2\omega_m}{N}$ as the *chirp rate*. The discrete Teager energy of this signal is given by $\sin^2(\dot{\varphi}[n])$.

It was shown in [1] that application of the discrete TKEO to this chirp signal yields

$$\Psi_d(x[n]) = \sin^2(\dot{\varphi}[n]) + \sin\left(\frac{\omega_m}{N}\right)\sin\left(2\varphi[n]+\frac{2\omega_m}{N}\right), \tag{19}$$

where the term $D[n] = \sin^2(\dot{\varphi}[n])$ is the desired discrete Teager energy and the term

$$E[n] = \sin\left(\frac{\omega_m}{N}\right)\sin\left(2\varphi[n] + \frac{2\omega_m}{N}\right) \qquad (20)$$

may be regarded as an additive approximation error. Clearly $E[n]$ is itself a chirp with a chirp rate that is twice that of the original signal $x[n]$ and with a constant AM function given by $\sin\left(\frac{\omega_m}{N}\right)$. For $2\omega_m < N\pi$, which is typical, the maximum magnitude of the error term $E[n]$ grows as the chirp rate is increased. This is precisely as one would expect, since increases in ω_m represent increasing departure of the instantaneous phase from a purely linear behavior.

For any given fixed chirp rate, however, (20) does not indicate the presence of an error depending directly on the instantaneous frequency $\dot{\varphi}[n]$ itself. In other words, the existence of the TKEO "low frequency error" is not revealed by (20). Nevertheless the "low frequency error" has been observed informally or suspected in a variety of cases not limited to FM-only chirps, but including more general AM-FM signals generated from analytical models for the AM and FM functions as well as experimentally acquired signals for which the true AM and FM functions are unknown.

For a variety of FM-only chirp signals, the actual estimation errors that occur in the DESA-2 frequency demodulation algorithm (16) are depicted in Fig. 1. The time variable n does not appear explicitly in the figure; rather, the x-axis gives the chirp rate in cycles/sample/sample while the y-axis gives the instantaneous frequency $\dot{\varphi}[n]$ in cycles per sample. The frequency estimation error is depicted along the z-axis, also in units of cycles per sample. Two consistent, systematic trends are immediately apparent in the plotted data. First, as predicted by (20), the magnitudes of the maximum errors that occur in a neighborhood about any fixed frequency $\dot{\varphi}[n]$ grow as the chirp rate increases. Second, and even more pronounced, the TKEO "low frequency error" is clearly visible. At each chirp rate, the magnitude of the maximum frequency errors more than doubles as the instantaneous frequency approaches zero. With regards to the high frequency error behavior, it should be noted that the DESA-2 algorithm is capable of estimating frequencies only up to $\frac{1}{4}$ of the sampling frequency, or 0.25 cycles per sample.

In an effort to gain additional insight into the nature of the "low frequency error," we numerically studied the approximation errors in the numerator of (16). For an FM-only chirp $x[n]$ with a constant AM function and an FM function given by (17), the desired discrete Teager energy of the symmetric difference signal $(x[n+1] - x[n-1])/2$ is $\sin^4(\dot{\varphi}[n])$. Comparing this quantity to the actual value $\Psi_d\{(x[n+1] - x[n-1])/2\}$, we obtained the surprising result that the approximation error in Ψ_d oscillates with a maximum magnitude that rapidly *decreases* as the instantaneous frequency descends toward zero. Thus, independent study of neither the numerator nor the denominator approximation errors in (16) alone is sufficient to explain the "low frequency error." As the frequency decreases, the approximation error in the denominator oscillates with a quadratic phase between extrema of $\pm\sin\left(\frac{\omega_m}{N}\right)$, while the numerator oscillates between extrema that decay rapidly. We believe therefore that the TKEO "low frequency error" must be subtly manifest in the local interplay between the oscillating numerator and denominator approximation errors. Rigorous analysis of this phenomenon is difficult to carry through the radical and transcendental in (16) and is beyond the scope of the preliminary results presented in this paper.

4. DISCUSSION

The data in Fig. 1 clearly demonstrate existence of the TKEO "low frequency error" that has previously been observed informally by us as well as by others. For a discrete AM-FM signal with constant amplitude and quadratic phase, there seem to be two main trends in the frequency estimation errors delivered by the DESA-2 algorithm. First, as predicted by theory, the maximum magnitudes of the errors grow as the chirp rate is increased. Second, for any given chirp rate, there is a dramatic increase that occurs in the maximum magnitudes of the frequency estimation errors as the frequency decreases toward zero. This second effect is surprising and is not predicted by any theoretical results we are aware of.

As we commented in Section 3, independent study of the numerator and denominator errors in (16) is not sufficient to explain the "low frequency error," which must therefore arise from the interaction between these oscillating errors. Although analytical treatment of the error is difficult due to the nonlinear nature of the DESA-2 algorithm, this is an aspect that we will continue to pursue in our ongoing research. There are several additional questions that deserve future investigation. Our study here was limited to the application of DESA-2 to FM-only chirp signals. Informally, we have observed a similar "low frequency error" in the DESA-1a and DESA-1 algorithms, although we have not investigated these systematically (DESA-2 is used most often in practice because it provides an attractive tradeoff between error performance and computational efficiency [2]). It would also be interesting to add a mild cubic component to the instantaneous phase (17) and/or a mild but nontrivial AM function to the signal (6) and see if the "low frequency error" persists in these cases – we believe emphatically that it will.

We close by suggesting two strategies that may be effective for improving the error performance of the DESA-2 algorithm at low frequency. Since DESA-2 is highly localized in the time variable n, it is quite easy to monitor the frequency estimates and make adjustments to the algorithm

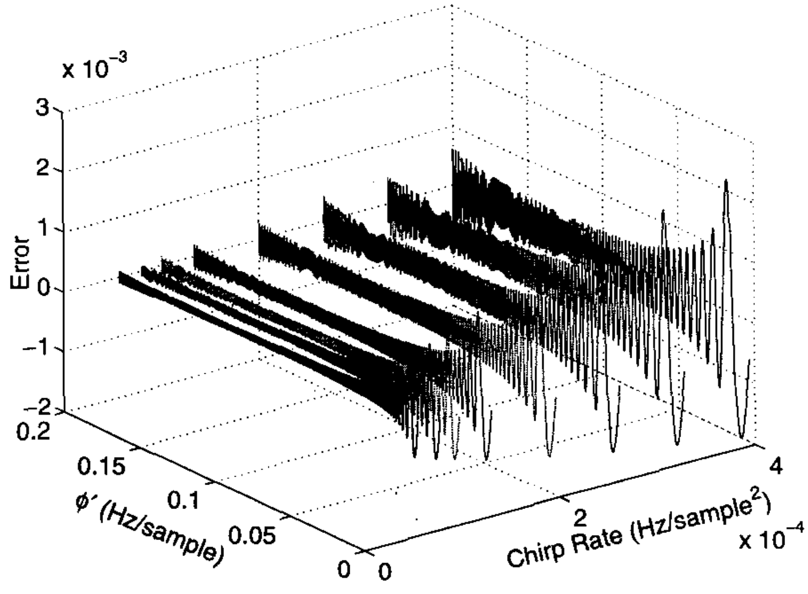

Figure 1: DESA-2 frequency estimation errors for several FM-only chirp signals. The error is depicted as a function of the chirp rate and of the instantaneous frequency $\dot{\varphi}[n]$. The TKEO "low frequency error" is clearly visible at all chirp rates.

on the fly. When low frequencies are detected, it may be possible to elevate the frequency perceived by the algorithm by replacing the central difference $(x[n+1] - x[n-1])/2$ in the numerator of (16) with a quantity such as $(x[n+K] - x[n-K])/(2K)$ for $K > 1$. We refer to this approach as the *wide algorithm*. Another possibility is to elevate the perceived frequency by performing a rational rate conversion to decrease the effective sampling rate. Since DESA-2 is capable of estimating frequencies only up to $\frac{1}{4}$ the sampling rate, care must be taken to avoid aliasing in this case.

5. REFERENCES

[1] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "On amplitude and frequency demodulation using energy operators," *IEEE Trans. Signal Proc.*, vol. 41, no. 4, pp. 1532–1550, April 1993.

[2] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "Energy separation in signal modulations with applications to speech analysis," *IEEE Trans. Signal Proc.*, vol. 41, no. 10, pp. 3024–3051, October 1993.

[3] J. F. Kaiser, "On a simple algorithm to calculate the 'energy' of a signal," in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Proc.*, Albuquerque, NM, April 3-6 1990, pp. 381–384.

[4] A. C. Bovik, J. P. Havlicek, D. S. Harding, and M. D. Desai, "Limits on discrete modulated signals," *IEEE Trans. Signal Proc.*, vol. 45, no. 4, pp. 867–879, April 1997.

[5] J. P. Havlicek, D. S. Harding, and A. C. Bovik, "Multidimensional quasi-eigenfunction approximations and multicomponent AM-FM models," *IEEE Trans. Image Proc.*, vol. 9, no. 2, pp. 227–242, February 2000.

[6] D. A. Cairns and J. H. L. Hansen, "Nonlinear analysis and classification of speech under stressed conditions," *J. Acoust. Soc. Amer.*, vol. 96, no. 6, pp. 3392–3400, December 1994.

[7] J. Foote, D. Mashao, and H. Silverman, "Stop classification using DESA-1 high-resolution formant tracking," in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Proc.*, Minneapolis, MN, April 27-30, 1993, vol. II, pp. 720–723.

[8] H. M. Hanson, P. Maragos, and A. Potamianos, "A system for finding speech formants and modulations via energy separation," *IEEE Trans. Speech and Audio Proc.*, vol. 2, no. 3, pp. 436–443, July 1994.

[9] T. F. Quatieri, C. R. Jankowski, and D. A. Reynolds, "Energy onset times for speaker identification," *IEEE Signal Proc. Let.*, vol. 1, no. 11, pp. 160–162, November 1994.