DERIVATIVE ANALYSIS OF COMPLEX POLYNOMIAL AMPLITUDE, COMPLEX EXPONENTIAL WITH EXPONENTIAL DAMPING

Sašo Muševič *

Music Technology Group
Universitat Pompeu Fabra
Barcelona, Spain

Jordi Bonada

Music Technology Group
Universitat Pompeu Fabra
Barcelona, Spain

ABSTRACT

In the context of non-stationary sinusoidal analysis two main sinusoidal models prevail: a polynomial phase/polynomial exponential amplitude (*generalised sinusoid*) and polynomial amplitude complex exponential (PACE) model. Recent advances resulted in efficient, robust and accurate parameter estimation algorithms for the *generalised sinusoid* model whereas the high-resolution method (HRM) for PACE model, suffers from high computational complexity. An efficient analysis method for high-degree complex PACE with exponential damping model is presented. Such model better describes sinusoids with sudden drop/increase of amplitude to/from zero, essentially forcing log-amplitude to negative infinity, rendering the generalised sinusoid model analysis unstable. Accuracy and computational efficiency of proposed method are compared to the HRM.

Index Terms— Sinusoidal analysis, Polynomial amplitude, Polynomial Phase, Exponential Sinusoidal Model, Polynomial Amplitude Complex Exponential.

1. INTRODUCTION

The field of non-stationary sinusoidal analysis is primarily concerned with estimating parameters of an amplitude and frequency modulated (AM/FM) sinusoid inside one observation frame. Many methods for the two mentioned AM/FM signal models exist [1, 2, 3, 4, 5, 6, 7, 8, 9]. Current research evolves around 2 families of tractable models: a generalised sinusoid model, sometimes referred to as *exponentially modulated sinusoid model* and the PACE model, which can be generalised to include exponential damping factor (referred to as Exponential Sinusoidal Model [10, 11, 12]). The PACE model is sometimes assumed to have complex rather than a real polynomial amplitude - to avoid ambiguity it will be referred to as complex amplitude-complex exponential model (cPACE) for the purpose of this paper. Extending it further to include the exponential damping factor it will be referred to as cPACE with Exponential Damping - cPACED. Further, a *degree* of the cPACED model will designate the degree of the complex amplitude polynomial.

Percussion sounds can be well modelled using a sinusoid with Gamma-tone amplitude envelope [13], therefore a desirable model for such signals has to include an exponential damping parameter [8] to express the exponential energy loss of a vibrating object without continuous energy supply. Lastly, the main benefit of the complex polynomial amplitude compared to the real one is the ability to encode frequency modulations to some extent, enabling some desirable audio coding properties [14, 5].

Some low-degree cPACE methods (eg: up to second degree polynomial [5]), require initial frequency estimates for all the sinusoids in the signal and thus requires identification of all salient peaks. The benefit is a joint estimation in the least-square sense and enabling the use of a window function to minimise the inter-sinusoid interference. On the other hand, the methods based on rotational invariance require little extra parameters apart from the signal itself. The cPACED model was shown to be the most general one still tractable by methods based on rotational invariance [7]. Many versions of such methods have been successfully used in various audio coding/analysis applications [10, 11], the main advantage being overcoming the time-frequency resolution trade-off, inherent to the Fourier Transform (FT). It will be shown however, that such methods bear significant computational burden. It is therefore desirable to construct an efficient method, able to estimate the parameters of a high degree cPACED model.

The paper is organised as follows: in section 2 the cPACED model and it's relation to exponentially modulated model are outlined. In section 3 a complex pole (the frequency and exponential damping factor) estimator using signal derivatives is proposed and in section 4 the complex polynomial parameters are estimated using the complex pole estimate. Finally, results of comparative tests are presented in section 5, while conclusion and future work suggestions follow in section 6.

*Research was funded by 'Slovene human resources development and scholarship fund' ('Javni sklad Republike Slovenije za razvoj kadrov in štipendije').

2. CPACED MODEL

The cPACED model is defined as follows:

$$s(t) = a(t)e^{(\mu_0 + j\omega_0)t}, \tag{1}$$

$$a(t) = p(t) + jq(t) = \sum_{k=0} (p_k + jq_k)t^k, \tag{2}$$

where $a(t)$ is the complex polynomial amplitude with real polynomials $p(t), q(t)$ and its real coefficients p_k, q_k respectively, while μ_0, ω_0 are the exponential damping and frequency parameter respectively, referred to as *pole* when combined into a complex number $\mu_0 + j\omega_0$. Such model covers all the parameters of a Gamma-tone envelope except the exact transient time inside the observed time frame as it is described in [8].

Since the polynomial coefficients are complex, they affect AM as well as FM. Transforming the polynomial into the polar form yields the exponential AM and FM separately:

$$s(t) = \sqrt{p(t)^2 + q(t)^2} \exp\left((\mu_0 + j\omega_0)t + j \arctan \frac{q(t)}{p(t)}\right), \tag{3}$$

$$= \exp\left((\mu_0 t + \alpha(t)) + j(\omega_0 t + \phi(t))\right), \text{where} \tag{4}$$

$$\alpha(t) = \frac{1}{2}\log(p(t)^2 + q(t)^2), \tag{5}$$

$$\phi(t) = \arctan \frac{q(t)}{p(t)}. \tag{6}$$

The Taylor series of $\phi(t), \alpha(t)$ suggest a certain degree of ambiguity is expected. Denoting the phase and log-amplitude power series respectively

$$M_\alpha(t) = \sum_{k=0}^{\infty} \frac{\alpha^{(l)}(0)}{l!} t^l, \tag{7}$$

$$M_\phi(t) = \sum_{k=0}^{\infty} \frac{\phi^{(l)}(0)}{l!} t^l, \tag{8}$$

$$\tag{9}$$

the actual linear phase parameter (i.e. frequency) is a sum of ω_0 and $\phi'(0)$ and the actual exponential damping parameter is a sum of μ_0 and $\alpha'(0)$. It is crucial to recognise this duality when assessing the accuracy of the algorithm. Figure 1 shows an example of such duality. Evaluating estimations of amplitude polynomial and exponential damping separately shows significant discrepancies, however the cumulative amplitude envelope is much more accurate.

3. POLE ESTIMATOR USING DERIVATIVES

A Fourier Transform of signal $s(t)$ at frequency ω, using a window function $w(t)$ is defined as an inner product:

$$S_w(\omega) = \langle s, w\Psi_{j\omega} \rangle, \tag{10}$$

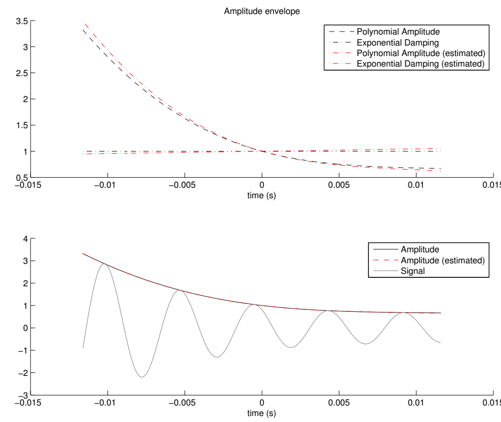

Fig. 1. Amplitude polynomial and exponential damping estimates separately (above) and cumulative (below).

where $\Psi_x = \exp(xt)$. The FT of signal time derivative is designated as

$$S'_w(\omega) = \langle s', w\Psi_{j\omega} \rangle = \tag{11}$$

$$- \langle s, w'\Psi_{j\omega} \rangle + j\omega \langle s, w\Psi_{j\omega} \rangle, \tag{12}$$

where the second equality follows from the distribution derivative rule [17] and implies $w'(-\frac{T}{2}) = w'(\frac{T}{2}) = 0$, where T is the length of the observation window. Higher signal derivatives can be easily derived using higher window derivatives, and further restrictions on window function apply: for $l-th$ signal derivative $w^{(l)}(-\frac{T}{2}) = w^{(l)}(\frac{T}{2}) = 0$. The FT of the derivative of cPACED model follows:

$$S'_w(\omega) = \frac{\partial}{\partial t} \langle a\Psi_{\beta_0}, w\Psi_{j\omega} \rangle = \tag{13}$$

$$\langle a'\Psi_{\beta_0}, w\Psi_{j\omega} \rangle + \beta_0 S_w(\omega) \Rightarrow \tag{14}$$

$$\langle a'\Psi_{\beta_0}, w\Psi_{j\omega} \rangle = S'_w(\omega) - \beta_0 S_w(\omega), \tag{15}$$

where the pole is designated as $\beta_0 = \mu_0 + j\omega_0$ for compactness. Time derivatives of both hand sides of 15 yield:

$$\langle a''\Psi_{\beta_0}, w\Psi_{j\omega} \rangle + \beta_0 \langle a'\Psi_{\beta_0}, w\Psi_{j\omega} \rangle = S''_w(\omega) - \beta_0 S'_w(\omega) \Rightarrow \tag{16}$$

$$\langle a''\Psi_{\beta_0}, w\Psi_{j\omega} \rangle = S''_w(\omega) - 2\beta_0 S'_w(\omega) + \beta_0^2 S_w(\omega). \tag{17}$$

A general expression can easily be proven:

$$\left\langle a^{(k)}\Psi\beta_0, w\Psi_{j\omega} \right\rangle = \sum_{l=0}^{k} S_w^{(l)} \binom{k}{l} (-\beta_0)^{k-l}. \tag{18}$$

Proof by induction: for $k = 1$ the above expression simplifies to 15. Assuming 18 for some k, its derivative is:

$$\left\langle a^{(k+1)}\Psi_{\beta_0}, w\Psi_{j\omega}\right\rangle + \beta_0 \left\langle a^{(k)}\Psi_{\beta_0}, w\Psi_{j\omega}\right\rangle = \sum_{l=0}^{k} S_w^{(l+1)} \binom{k}{l} (-\beta_0)^{k-l}. \quad (19)$$

Inserting the induction assumption and rearranging the indexes yields:

$$\begin{aligned}\left\langle a^{(k+1)}\Psi_{\beta_0}, w\Psi_{j\omega}\right\rangle &= \\ \sum_{l=0}^{k} S_w^{(l+1)} \binom{k}{l} (-\beta_0)^{k-l} - \sum_{l=0}^{k} S_w^{(l)} \binom{k}{l} (-\beta_0)^{k+1-l} &= \\ \sum_{l=0}^{k+1} S_w^{(l)} (-\beta_0)^{k-l} \left(\binom{k-1}{l} + \binom{k-1}{l-1} \right) &= \\ \sum_{l=0}^{k+1} S_w^{(l)} (-\beta_0)^{k-l} \binom{k}{l}&, \end{aligned} \quad (20)$$

concluding the proof. Assuming the degree of the amplitude polynomial is K and using the fact that for polynomial $p(x)$ of degree D, $\frac{\partial^{D+1} p(x)}{\partial x^{D+1}} = 0$, the following equation can be obtained:

$$\left\langle a^{(K+1)}\Psi_{\beta_0}, w\Psi_{j\omega}\right\rangle = \sum_{l=0}^{K+1} S_w^{(l)} (-\beta_0)^{k-l} \binom{k}{l} = 0, \quad (21)$$

which is a $(K+1)^{th}$ degree polynomial in respect to β_0 and it's K roots $\hat{\beta}_{0,k}, k = 1..K$ are the estimates for the pole β_0. Note that both α_0 and ω_0 are estimated jointly. The K estimates will in general not be equal even in a noise-less case as already outlined in [7]. The *best* estimate can be chosen by $\max_k \left|\left\langle s, e^{j\Im(\hat{\beta}_{0,k})}\right\rangle\right|$.

4. COMPLEX POLYNOMIAL AMPLITUDE ESTIMATOR

Pole estimates can be used to construct a simple linear system $Ax = b$:

$$A = \begin{bmatrix} \left\langle t^K \Psi_{\hat{\beta}_0}, w\Psi_{j\omega_1}\right\rangle & \cdots & \left\langle t\Psi_{\hat{\beta}_0}, w\Psi_{j\omega_1}\right\rangle & \left\langle \Psi_{\hat{\beta}_0}, w\Psi_{j\omega_1}\right\rangle \\ \left\langle t^K \Psi_{\hat{\beta}_0}, w\Psi_{j\omega_2}\right\rangle & \cdots & \left\langle t\Psi_{\hat{\beta}_0}, w\Psi_{j\omega_2}\right\rangle & \left\langle \Psi_{\hat{\beta}_0}, w\Psi_{j\omega_2}\right\rangle \\ \vdots & & \vdots & \vdots \\ \left\langle t^K \Psi_{\hat{\beta}_0}, w\Psi_{j\omega_R}\right\rangle & \cdots & \left\langle t\Psi_{\hat{\beta}_0}, w\Psi_{j\omega_R}\right\rangle & \left\langle \Psi_{\hat{\beta}_0}, w\Psi_{j\omega_R}\right\rangle \end{bmatrix} \quad (22)$$

$$x = \begin{bmatrix} a_K \\ a_{K-1} \\ \vdots \\ a_1 \\ a_0 \end{bmatrix} b = \begin{bmatrix} S_w(\omega_1) \\ S_w(\omega_2) \\ \vdots \\ S_w(\omega_{R-1}) \\ S_w(\omega_R) \end{bmatrix}, \quad (23)$$

where $\hat{\beta}_0 = \hat{\alpha}_0 + j\hat{\omega}_0$ is the pole estimation acquired as described in section 3. Solutions of above linear system give estimates for coefficients of the complex amplitude polynomial. The window function does not have to be the same as the one used for pole estimation - restriction on edge values does not apply. Each row of matrix A corresponds to an arbitrary frequency, the most reasonable choice being the ones carrying most of the energy of the sinusoid in question, eg: as close to magnitude peak frequency as possible [15]. An efficient algorithm implementation can utilise FFT bin values and zero-padding to avoid costly computation of DTFT at specific frequencies and to adjust inter-bin frequency difference. Matrix A need not be square, many times an overdetermined system is desired: the number of estimation frequencies R can be larger than the number of unknowns K. Such systems can solved on least-square basis via Moore-Penrose matrix pseudo-inverse A^+:

$$A^+ = (A^*A)^+ A^* = (A^*A)^{-1} A^*, \quad (24)$$

where A^* designates a conjugate transpose of matrix A and $A^{-1} = A^+$ if A is square.

5. TESTS AND RESULTS

A polynomial amplitude of degree 4 was studied, the polynomial denoted as: $[a_3, a_2, a_1, a_0] = [p_3 + jq_3, p_2 + jq_2, p_1 + jq_1, p_0 + jq_0]$. The test values for p_3, p_2, p_1 were chosen so all the terms of the amplitude polynomial have equal impact on the final value:

$$p_3 \in \left[-\left(\frac{fs}{2T}\right)^3, \left(\frac{fs}{2T}\right)^3 \right] \quad (25)$$

$$p_2 \in \left[-\left(\frac{fs}{2T}\right)^2, \left(\frac{fs}{2T}\right)^2 \right] \quad (26)$$

$$p_1 \in \left[-\frac{fs}{2T}, \frac{fs}{2T} \right]. \quad (27)$$

Exact same value sets were used for the imaginary part of the polynomial $q(t)$. A Hann[3] window function of length 512 samples was used for pole estimation and Hann window for the complex polynomial coefficients estimation. Damping factor was varied in the bounds [-100,100] and only one frequency of 1000Hz was considered. For each parameter except frequency, only the 2 extreme values and 0 have been tested in order to keep computational time reasonable. The comparison to a 4^{th} degree (i.e. 4 poles and amplitudes) simple high-resolution method implementation from DESAM Toolbox [16] (section 5.1.2.) without whitening was conducted. The signal tested is the real part of the complex cPACED signal, reflecting the real world scenario when analytical signal isn't available.

To measure accuracy, the commonly used Signal-to-Residual-Ratio (SRR) metric was used:

$$SRR = \frac{\langle s, ws\rangle}{\langle s - \hat{s}, w(s - \hat{s})\rangle}, \quad (28)$$

where $s, \hat{s}$ are the original signal (without noise) and the estimated signal respectively, and w the Hann window. The Signal-to-Noise-Ratio (SNR) range from [50,-20] was studied. The total computation times for both methods follow:

Proposed method	28s
High-resolution	5400s

(29)

Since HRM involves singular value decomposition of correlation matrix of size $N/2 \times N/2$ the computation cost is significantly higher as the proposed method only requires $K-1$ FFTs for the pole and K DTFTs for the complex polynomial estimates.

Classic Cramer-Rao bounds (CRBs) parameter-by-parameter comparison would total to 10 plots (8 for the real/imaginary polynomial coefficients and 2 for the pole), overcomplicating the results and obscuring the overall accuracy. A more intuitive approach would involve only one SRR/SNR plot, thus a different upper accuracy bound is required. A *near perfect* estimator can be constructed by substituting the pole estimates with actual poles and solving the linear system 22. Mean and variance of the SRR, computed with the aforementioned estimator represents a good upper SRR bound. Figure 2 depicts the mean and variance of the baseline estimator, proposed method and HRM. At low SNR the methods perform roughly the same, HRM reaching the upper bound while proposed method performing 2dB below. At high SNR both methods reach a plateau, however the proposed method's plateau is about 10dB higher.

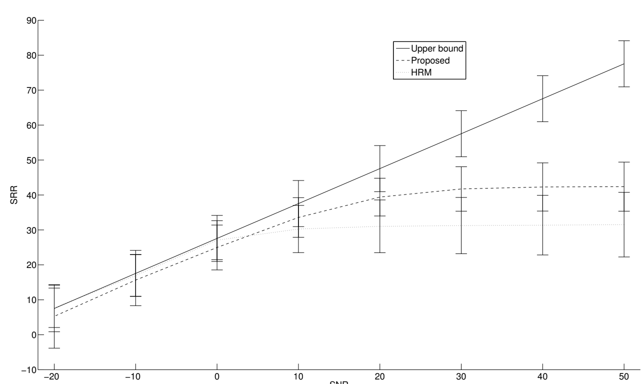

Fig. 2. SRR: Mean and variance

6. CONCLUSION AND FUTURE WORK

In this paper a novel method for analysing cPACED signals was represented, tested and compared to the HRM. In conducted tests the HRM performs marginally better in high noise cases, while the proposed method performs significantly better in low-noise cases. More rigorous testing, involving multicomponent signals is required, but is out of scope of this document.

While HRM is a computationally intense method, designed to jointly estimate parameters of multiple cPACED sinusoids in the entire frequency range, the proposed method focuses on a narrow frequency range to estimate a single cPACED sinusoid. The flexibility of HRM is a huge overkill for the tests conducted, which is reflected in substantially higher (about 2 degrees of magnitude) computational costs. However the proposed method could be invoked on different frequency ranges, effectively covering the whole spectrum. The advantage of the proposed method in this case is the ability to process only certain frequency regions of interest, reducing the final computational cost.

The pole estimator is inspired by the generalised reassignment method [2], imposing significant requirement on the number and type of the window function. Such window set usually exhibits suboptimal time-frequency properties and quickly reduces the condition number of the resulting linear system to a value too low to handle even using very high precision computation. The distribution derivative method [17] circumvents this problem by constructing the linear system using FT values at different frequencies, rather than using a single FT value, but using different windows. A version of proposed method, that does not impose the window requirement would enable an extremely accurate, very high degree cPACED analysis and would, analogously, correspond to the distribution derivative method. cPACE model is well suited for analysis of close frequency non-stationary sinusoids, as the amplitude beating function can be much better approximated with a polynomial than a generalised sinusoid, thus coding such signals is expected to be of higher accuracy compared to the generalised reassignment.

7. REFERENCES

[1] Sylvain Marchand and Philippe Depalle, "Generalization of the derivative analysis method to Non-Stationary sinusoidal modeling," *Digital Audio Effects, 2008. Proceedings of the 11th Int. Conference on (DAFx-08)*, 2008.

[2] Mark Sandler Xue Wen, "Notes on model-based non-stationary sinusoid estimation methods using derivative," in *Digital Audio Effects, 2009. Proceedings of the 12th Int. Conference on (DAFx-09)*, 2009.

[3] B. Hamilton and P. Depalle, "A unified view of non-stationary sinusoidal parameter estimation methods using signal derivatives," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, march 2012, pp. 369 –372.

[4] S. Muševič and J. Bonada, "Comparison of non-stationary sinsoid estimation methods using reassignment and derivatives," in *Proc. 7th Sound and Music Computing Conf.*, Barcelona, Spain, July 2010.

[5] Y. Pantazis, O. Rosec, and Y. Stylianou, "Adaptive am/fm signal decomposition with application to speech analysis," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 19, no. 2, pp. 290 –300, feb. 2011.

[6] G.P. Kafentzis, Y. Pantazis, O. Rosec, and Y. Stylianou, "An extension of the adaptive quasi-harmonic model," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, march 2012, pp. 4605 –4608.

[7] R. Badeau, B. David, and G. Richard, "High-resolution spectral analysis of mixtures of complex exponentials modulated by polynomials," *Signal Processing, IEEE Transactions on*, vol. 54, no. 4, pp. 1341 – 1350, april 2006.

[8] M.G. Christensen and S. van de Par, "Efficient parametric coding of transients," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 14, no. 4, pp. 1340 –1351, july 2006.

[9] Roland Badeau, *Méthodes à haute résolution pour l'estimation et le suivi de sinusoïdes modulées. Application aux signaux de musique*, Ph.D. thesis, École Doctorale d'Informatique, Télécommunications et Électronique de Paris, 2005.

[10] J. Jensen, R. Heusdens, and S.H. Jensen, "A perceptual subspace approach for modeling of speech and audio signals with damped sinusoids," *Speech and Audio Processing, IEEE Transactions on*, vol. 12, no. 2, pp. 121 – 132, march 2004.

[11] Kris Hermus, Werner Verhelst, Philippe Lemmerling, Patrick Wambacq, and Sabine Van Huffel, "Perceptual audio modeling with exponentially damped sinusoids," *Signal Processing*, vol. 85, no. 1, pp. 163 – 176, 2005.

[12] J. Nieuwenhuijse, R. Heusens, and Ed.F. Deprettere, "Robust exponential modeling of audio signals," in *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*, may 1998, vol. 6, pp. 3581 –3584 vol.6.

[13] Simon Scholler and P. Purwins, "Sparse approximations for drum sound classification," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, pp. 933 – 940, 2011 2011.

[14] Maciej Bartkowiak, "A complex envelope ssinusoidal model for audio codig," in *Proc. 10th Int. Digital Audio Effects*, 2007.

[15] M. Betser, P. Collen, G. Richard, and B. David, "Estimation of frequency for AM/FM models using the phase vocoder framework," *Signal Processing, IEEE Transactions on*, vol. 56, no. 2, pp. 505–517, 2008.

[16] Roland Badeau, Bertrand David, Nancy Bertin, Jose Echeveste, Mathieu Lagrange, Olivier Derrien, and Sylvain Marchand, "Desam toolbox v1.1," .

[17] M. Betser, "Sinusoidal polynomial parameter estimation using the distribution derivative," *Signal Processing, IEEE Transactions on*, vol. 57, no. 12, pp. 4633 –4645, dec. 2009.