

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of Helsinki University of Technology's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org.

By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

On Line Spectral Frequencies

W. Bastiaan Kleijn, *Fellow, IEEE*, Tom Bäckström, and Paavo Alku

Abstract—The commonly used line spectral frequencies form the roots of symmetric and antisymmetric polynomials constructed from a linear predictor. In this letter, we provide a new, simpler proof that the symmetric and antisymmetric polynomials can be regarded as optimal constrained predictors that correspond to predicting from the low-pass and high-pass filtered signal, respectively.

Index Terms—Linear prediction, line spectral frequency, line spectral pair.

I. INTRODUCTION

LINEAR prediction is ubiquitous in speech coding (e.g., [1]). In linear-prediction-based speech coders, the predictor is generally encoded as side-information and interpolated between updates. It has been found that the line spectral frequency (LSF) [2] representation of the predictor is particularly well suited for quantization (e.g., [3]) and interpolation (e.g., [3], [4]). From a theoretical viewpoint, this can be motivated by the fact that the sensitivity matrix relating the LSF-domain squared quantization error to the perceptually relevant log spectrum is diagonal [5], [6]. Furthermore, the minimum-phase property of the quantized predictor is easily guaranteed [7].

Despite the extensive use of the LSFs, insight in their behavior and meaning is less developed than that of other predictor representations, although some physical interpretations have been provided [7], [8]. The goal of this letter is to provide additional understanding of the LSFs by showing that they can be determined from particular optimal constrained predictors.

We show that, for even predictor order, the LSFs can be interpreted as the roots of two particular predictor polynomials. These polynomials (and corresponding predictors) are symmetric, which is consistent with the LSFs being on the unit circle. The predictors are optimal constrained predictors that correspond to prediction of the original signal from a particular high-pass and a particular low-pass filtered signal, respectively. The predictor polynomials are identical to the LSF polynomials of the unconstrained predictor (the symmetric and antisymmetric components of the unconstrained predictor polynomial), if the trivial roots of these LSF polynomials at $+1$ and -1 are removed. This basic equivalence is outlined in Fig. 1. The nature of the constraints indicates that the sequence

of interlaced roots of the symmetric and antisymmetric polynomials will generally start (at the low-frequency end) with a root of the symmetric polynomial. The new interpretation also has an interesting link to fast computational methods for predictors that exploit symmetric polynomials [9]. We note that the constrained predictors are similar to those discussed in [10] and [11]. The present proof is different and more compact than that in [12].

We continue this letter in Section II with the decomposition of the conventional linear prediction polynomial into a symmetric and an antisymmetric LSF polynomial. In Section III, we show that the same symmetric and antisymmetric polynomials are obtained for certain constrained predictors. We discuss the significance of these results in Section IV.

We will, where that does not lead to confusion, refer to the prediction-residual filter as “the predictor.” The notation $\mathbf{a}^{(k)}$ indicates a vector of dimension k , and $A^{(k)}(z)$ is the z transform of the corresponding sequence. Subscripts $+$ and $-$ denote, respectively, symmetry and antisymmetry of a vector or sequence. Superscripts $+$ and $-$ indicate low-pass and high-pass filtering by $H^+(z) = 1 + z^{-1}$ and $H^-(z) = -1 + z^{-1}$, respectively.

II. CONVENTIONAL LINEAR PREDICTION

Let us consider a stationary signal with a $k \times k$ covariance matrix $\mathbf{R}^{(k)}$. The optimal order- p linear predictor $[-a_1^{(p+1)}, \dots, -a_p^{(p+1)}]^T$ for this signal can be determined from the extended normal equations

$$\mathbf{R}^{(p+1)}\mathbf{a}^{(p+1)} = [\sigma_{(p)}^2, 0, \dots, 0]^T \quad (1)$$

where $\mathbf{a}^{(p+1)} = [1, a_1^{(p+1)}, \dots, a_p^{(p+1)}]^T$, and $\sigma_{(p)}^2$ is the prediction-residual variance. The elements of the vector $\mathbf{a}^{(p+1)}$ correspond to the coefficients of the p th-order prediction-error filter, often referred to as the inverse filter.

A conceptually simple method for solving (1) is to solve first

$$\mathbf{R}^{(p+1)}\tilde{\mathbf{a}}^{(p+1)} = [1, 0, \dots, 0]^T \quad (2)$$

for $\tilde{\mathbf{a}}^{(p+1)}$. We then have $\sigma_{(p)}^2 = 1/\tilde{a}_0^{(p+1)}$ and $\mathbf{a}^{(p+1)} = \sigma_{(p)}^2 \tilde{\mathbf{a}}^{(p+1)}$.

Next, we define the symmetric polynomial and the antisymmetric polynomial of the LSF decomposition using a vector notation. Let the superscript $\#$ denote the sequence-reversal operation, i.e., $\tilde{\mathbf{a}}^{(p+1)\#} = [\tilde{a}_p^{(p+1)}, \dots, \tilde{a}_0^{(p+1)}]^T$. (The vector $\tilde{\mathbf{a}}^{(p+1)\#}$ corresponds to the so-called backward predictor.) We can form $(p+2)$ -dimensional symmetric $\tilde{\mathbf{a}}_+^{(p+2)}$ and antisymmetric $\tilde{\mathbf{a}}_-^{(p+2)}$ components of $\tilde{\mathbf{a}}^{(p+1)}$ as follows:

$$\tilde{\mathbf{a}}_{\pm}^{(p+2)} = \begin{bmatrix} \tilde{\mathbf{a}}^{(p+1)} \\ 0 \end{bmatrix} \pm \begin{bmatrix} 0 \\ \tilde{\mathbf{a}}^{(p+1)\#} \end{bmatrix}. \quad (3)$$

Manuscript received May 23, 2002; revised September 4, 2002. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Chuang (Mike) He.

W. B. Kleijn is with KTH (Royal Institute of Technology), Department of Signals, Sensors and Systems, 10044 Stockholm, Sweden (e-mail: bastiaan.kleijn@s3.kth.se).

T. Bäckström and P. Alku are with the Helsinki University of Technology (HUT), Laboratory of Acoustics and Audio Signal Processing, Finland.

Digital Object Identifier 10.1109/LSP.2003.809035

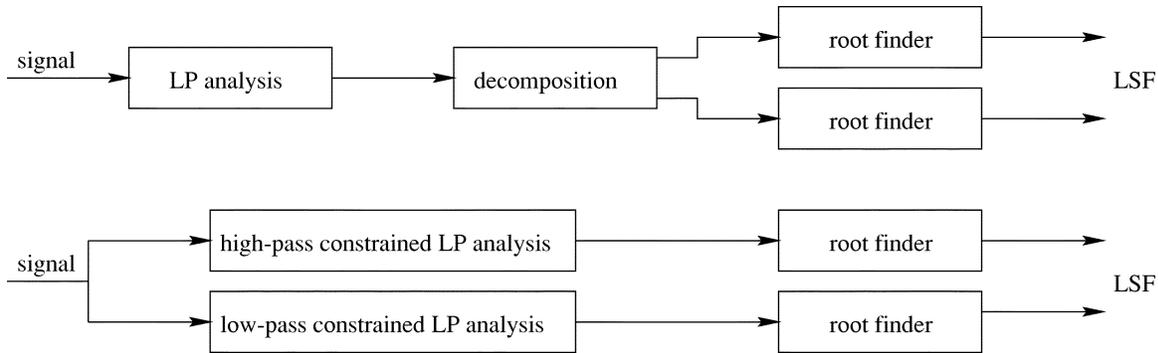


Fig. 1. Block diagram outlining the equivalence shown in this letter for even predictor order. The top diagram shows the conventional method for obtaining the LSFs based on a decomposition of the predictor polynomial into symmetric and antisymmetric polynomials. The bottom diagram shows a method based on prediction from particular high-pass and low-pass filtered signals that renders the symmetric and antisymmetric polynomials as predictor polynomials (ignoring the trivial roots).

The polynomials $\tilde{A}_+^{(p+2)}(z)$ and $\tilde{A}_-^{(p+2)}(z)$ are the symmetric and antisymmetric polynomials, respectively, of the LSF decomposition (usually denoted as $P(z)$ and $Q(z)$, respectively), except for the scaling factor σ_p^2 . Given the (anti)symmetry, it is easy to see from (2) that $\tilde{\mathbf{a}}_+^{(p+2)}$ and $\tilde{\mathbf{a}}_-^{(p+2)}$ obey

$$\mathbf{R}^{(p+2)}\tilde{\mathbf{a}}_{\pm}^{(p+2)} = [\mu, 0, \dots, 0, \pm\mu]^T \quad (4)$$

where μ is a constant.

III. CONSTRAINED LINEAR PREDICTION

We define a predictor $[-b_1^{(p+1)+}, \dots, -b_p^{(p+1)+}]^T$ that predicts a signal sample of a first signal, $x(n)$, from a set of samples $x^+(n), \dots, x^+(n-p+1)$ of a second signal that is obtained by low-pass filtering $x(n)$ with $H^+(z) = 1 + z^{-1}$, i.e., $X^+(z) = H^+(z)X(z)$. The prediction error is

$$e(n) = x(n) + \sum_{i=1}^{i=p} b_i^{(p+1)+} x^+(n-i+1). \quad (5)$$

Similarly, we define a predictor $[-b_1^{(p+1)-}, \dots, -b_p^{(p+1)-}]^T$ that predicts $x(n)$ from samples of the signal $x^-(n)$ obtained by high-pass filtering with $H^-(z) = -1 + z^{-1}$. The extended normal equations defining these predictors are

$$\mathbf{H}^{\pm} \mathbf{R}^{(p+1)} \mathbf{H}^{\pm T} \tilde{\mathbf{b}}^{(p+1)\pm} = [1, 0, \dots, 0]^T \quad (6)$$

where $\tilde{\mathbf{b}}^{(p+1)+}$ and $\tilde{\mathbf{b}}^{(p+1)-}$ are the unnormalized prediction error filters and the matrices \mathbf{H}^+ and \mathbf{H}^- are the $(p+1) \times (p+1)$ Toeplitz matrices

$$\mathbf{H}_{i,j}^{\pm} = \delta_{i,j} \pm \delta_{i-1,j}, \quad i, j \in \{0, \dots, p\}. \quad (7)$$

The matrices $\mathbf{H}_{i,j}^{\pm}$ correspond to a transform where the first row yields the original sample $x(n)$ to be predicted, and the following rows yield samples of $x^{\pm}(n)$ filtered by $H^{\pm}(z)$.

We define the constrained predictors $\tilde{\mathbf{c}}^{(p+1)+} = \mathbf{H}^{+T} \tilde{\mathbf{b}}^{(p+1)+}$ and $\tilde{\mathbf{c}}^{(p+1)-} = \mathbf{H}^{-T} \tilde{\mathbf{b}}^{(p+1)-}$ that correspond to the optimal prediction of the unfiltered signal from, respectively, high-pass and low-pass filtered samples. Note that

the constrained predictors $\tilde{\mathbf{c}}^{(p+1)+}$ and $\tilde{\mathbf{c}}^{(p+1)-}$ have the original signal samples as input: defining $\mathbf{c}^{(p+1)+} = \mathbf{H}^{+T} \mathbf{b}^{(p+1)+}$ (note that, in general, $c_0^{(p+1)}$ is not unity), the prediction error can be written as

$$e(n) = \sum_{i=0}^{i=p} c_i^{(p+1)+} x(n-i). \quad (8)$$

By multiplying (6) by the inverse of \mathbf{H}^{\pm} (the right-hand side can be found with simple forward elimination), it is seen that the predictors $\tilde{\mathbf{c}}^{(p+1)\pm} = \mathbf{H}^{\pm T} \tilde{\mathbf{b}}^{(p+1)\pm}$ satisfy

$$\mathbf{R}^{(p+1)} \tilde{\mathbf{c}}^{(p+1)\pm} = [1, \mp 1, 1, \mp 1, \dots]^T. \quad (9)$$

The same equation can be obtained by minimization of the prediction error subject to a constraint on the structure of $\tilde{\mathbf{c}}^{(p+1)\pm}$. The predictor $\tilde{\mathbf{c}}^{(p+1)+}$ is symmetric ($\tilde{\mathbf{c}}^{(p+1)+} = \tilde{\mathbf{c}}^{(p+1)+\#}$) for p even since $\mathbf{R}^{(p+1)}$ is Toeplitz and symmetric and since the right-hand side of (9) is symmetric. Furthermore, $\tilde{\mathbf{c}}^{(p+1)+}$ is antisymmetric for p odd, since the right-hand side of (9) is antisymmetric. Similarly, $\tilde{\mathbf{c}}^{(p+1)-}$ must be symmetric for p odd and even. The symmetries and antisymmetries can be exploited to reduce computational effort.

Defining a $(p+2)$ -dimensional symmetric component $\tilde{\mathbf{c}}_+^{(p+2)+}$ of the vector $\tilde{\mathbf{c}}^{(p+1)+}$ and a $(p+2)$ -dimensional antisymmetric component $\tilde{\mathbf{c}}_-^{(p+2)-}$ of $\tilde{\mathbf{c}}^{(p+1)-}$ as in (3), it follows from (9) that, for p even

$$\mathbf{R}^{(p+2)} \tilde{\mathbf{c}}_{\pm}^{(p+2)\pm} = [\gamma^{\pm}, 0, \dots, 0, \pm\gamma^{\pm}]^T \quad (10)$$

where γ^+ and γ^- are constants. (We recall that superscripts $+$ and $-$ indicate low-pass and high-pass filtering, whereas subscripts $+$ and $-$ indicate symmetry and antisymmetry.) We recognize that (10) and (4) are identical and that, therefore, for even p

$$\tilde{\mathbf{c}}_{\pm}^{(p+2)\pm} = \frac{\gamma^{\pm}}{\mu} \tilde{\mathbf{a}}_{\pm}^{(p+2)}. \quad (11)$$

This implies that $C_{\pm}^{(p+2)\pm}(z)$ and $A_{\pm}^{(p+2)}(z)$ (the z transforms corresponding to $\tilde{\mathbf{c}}_{\pm}^{(p+2)\pm}$ and $\tilde{\mathbf{a}}_{\pm}^{(p+2)}$) are identical except for their scaling.

Next, let us consider odd values of p . In this case, we have $\tilde{\mathbf{c}}^{(p+1)\pm} = \mp \tilde{\mathbf{c}}^{(p+1)\pm\#}$. The antisymmetric component of $\tilde{\mathbf{c}}^{(p+1)+}$ is obtained by subtraction

$$\tilde{\mathbf{c}}_-^{(p+2)+} = \begin{bmatrix} \tilde{\mathbf{c}}^{(p+1)+} \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ \tilde{\mathbf{c}}^{(p+1)+\#} \end{bmatrix}. \quad (12)$$

For p odd we then obtain

$$\tilde{\mathbf{c}}_-^{(p+2)+} = \frac{\gamma^+}{\mu} \tilde{\mathbf{a}}_-^{(p+2)}. \quad (13)$$

However, for p odd there is no equation equivalent to (11) for the symmetric case (i.e., for $\mathbf{a}_+^{(p+2)}$).

We now summarize our results. For any p and with proper scaling, the antisymmetric components of the conventional and constrained predictors are equal. Furthermore, for even values of p and with proper scaling, the symmetric component of the conventional and constrained predictors are also equal. Some implications of these results are discussed in Section IV.

IV. DISCUSSION

We saw that, for p even, the constrained predictor $\mathbf{c}^{(p+1)+}$ has the same $(p+2)$ -dimensional symmetric component as the unconstrained predictor $\mathbf{a}^{(p+1)}$, ignoring a scaling. Similarly, again for p even, the constrained predictor $\mathbf{c}^{(p+1)-}$ has the same $(p+2)$ -dimensional antisymmetric component as the unconstrained predictor $\mathbf{a}^{(p+1)}$, except for a scaling. It then follows that

$$\frac{\gamma^\pm}{\mu} \tilde{\mathbf{a}}_\pm^{(p+2)} = \tilde{\mathbf{c}}_\pm^{(p+2)+} = \begin{bmatrix} \tilde{\mathbf{c}}^{(p+1)\pm} \\ 0 \end{bmatrix} \pm \begin{bmatrix} 0 \\ \tilde{\mathbf{c}}^{(p+1)\pm} \end{bmatrix} \quad (14)$$

since both $\mathbf{c}^{(p+1)+}$ and $\mathbf{c}^{(p+1)-}$ are symmetric. This results in $\tilde{A}_\pm^{(p+2)}(z) = (1 \pm z^{-1}) \tilde{C}^{(p+1)\pm}(z)(\mu/\gamma^\pm)$, rendering the well-known fact that $A_+^{(p+2)}(z)$ has a trivial root at $z = -1$ and that $A_-^{(p+2)}(z)$ has a trivial root at $z = +1$.

Let us restate these findings. We have found that the symmetric polynomial $A_+^{(p+2)}(z)/(1+z^{-1})$ is the optimal predictor $C^{(p+1)+}(z)$ that is constrained to predict from the low-pass filtered signal (except for a scaling). This predictor generally retains a low-pass character. Furthermore, we have found that the symmetric polynomial $A_-^{(p+2)}(z)/(-1+z^{-1})$ is the optimal predictor $C^{(p+1)-}(z)$ that is constrained to predict from the high-pass filtered signal (except for a scaling). This predictor

generally retains a high-pass character. When p is odd, which is less relevant in practice, we find similarly that $A_-^{(p+2)}(z)/(1+z^{-1})$ equals $C^{(p+1)+}(z)$ (except for a scaling), but no equivalent result for $A_+^{(p+2)}(z)$ exists.

A structure similar to the one given in (4) is used in the Levinson recursive computation of the predictor. In the Levinson algorithm, the forward and backward predictors of (3) are multiplied by factors so as to render the right-hand side of (4) of the form $[1, 0, 0, \dots]^T$ [13]. Moreover, the symmetric version of (9) has appeared earlier in the context of a split Levinson–Durbin recursion [9]. This algorithm takes advantage of both the symmetry and recursion properties of $\tilde{\mathbf{c}}^{(p+1)+}$ to lower the computational requirements for finding an optimal predictor (its numerical properties make it less attractive [14]).

REFERENCES

- [1] R. V. Cox, “Speech coding standards,” in *Speech Coding and Synthesis*, W. B. Kleijn and K. K. Paliwal, Eds. Amsterdam, The Netherlands: Elsevier, 1995, pp. 49–78.
- [2] F. Itakura, “Line spectrum representation of linear predictive coefficients,” *J. Acoust. Soc. Amer.*, vol. 57 Suppl., no. 1, p. S35, 1975.
- [3] K. K. Paliwal and W. B. Kleijn, “Quantization of LPC parameters,” in *Speech Coding and Synthesis*, W. B. Kleijn and K. K. Paliwal, Eds. Amsterdam, The Netherlands: Elsevier, 1995, pp. 433–466.
- [4] J. Erkelens, “Autoregressive modeling for speech coding: Estimation, interpolation and quantization,” Ph.D. dissertation, Delft Univ. Technol., The Netherlands, 1996.
- [5] J. Erkelens and P. Broersen, “On the statistical properties of line spectrum pairs,” in *Proc. ICASSP*, Detroit, MI, 1995, pp. 768–771.
- [6] W. R. Gardner and B. D. Rao, “Optimal distortion measures for the high rate vector quantization of LPC parameters,” in *Proc. ICASSP*, Detroit, MI, 1995, pp. 752–755.
- [7] F. K. Soong and B. Huang, “Line spectrum pair (LSP) and speech data compression,” in *Proc. ICASSP*, San Diego, CA, 1984, pp. 1.10.1–1.10.4.
- [8] M. Hasegawa-Johnson, “Line spectral frequencies are poles and zeros of the glottal driving-point impedance of a discrete matched-impedance vocal tract model,” *J. Acoust. Soc. Amer.*, vol. 108, pp. 457–460, 2000.
- [9] H. Krishna, “New split Levinson, Schur, and lattice algorithms for digital signal processing,” in *Proc. ICASSP*, New York, 1988, pp. 1640–1642.
- [10] P. Alku and S. Varho, “A new linear predictive method for compression of speech signals,” in *Proc. Int. Conf. Spoken Lang. Process.*, Sydney, Australia, 1998, pp. 2563–2566.
- [11] S. Varho and P. Alku, “Separated linear prediction—A new all-pole modeling technique for speech analysis,” *Speech Commun.*, vol. 24, pp. 111–121, 1998.
- [12] T. Bäckström, P. Alku, and W. B. Kleijn, “A time-domain reformulation of linear prediction equivalent to the LSP decomposition,” in *Proc. ICASSP*, vol. I, Orlando, FL, 2002, pp. 661–664.
- [13] P. Delsarte and Y. V. Genin, “The split Levinson algorithm,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 470–478, 1986.
- [14] Y. Wang, H. Krishna, and B. Krishna, “Split Levinson algorithm is weakly stable,” in *Proc. ICASSP*, Glasgow, U.K., 1989, pp. 1215–1218.