

The Computation of Line Spectral Frequencies Using Chebyshev Polynomials

PETER KABAL AND RAVI PRAKASH RAMACHANDRAN

Abstract—Line spectral frequencies provide an alternate parameterization of the analysis and synthesis filters used in linear predictive coding (LPC) of speech. In this paper, a new method of converting between the direct form predictor coefficients and line spectral frequencies is presented. The system polynomial for the analysis filter is converted to two even-order symmetric polynomials with interlacing roots on the unit circle. The line spectral frequencies are given by the positions of the roots of these two auxiliary polynomials. The response of each of these polynomials on the unit circle is expressed as a series expansion in Chebyshev polynomials. The line spectral frequencies are found using an iterative root finding algorithm which searches for real roots of a real function. The algorithm developed is simple in structure and is designed to constrain the maximum number of evaluations of the series expansions. The method is highly accurate and can be used in a form that avoids the storage of trigonometric tables or the computation of trigonometric functions. The reconversion of line spectral frequencies to predictor coefficients uses an efficient algorithm derived by expressing the root factors as an expansion in Chebyshev polynomials.

I. INTRODUCTION

IN many speech coders, the parameters of the all-zero predictor filter or the corresponding all-pole synthesis filter are coded and sent as part of the information stream. Recently, there has been a growing interest in the use of line spectral frequencies (LSF's) to code the filter parameters for linear predictive coding (LPC) of speech [1]–[4]. LSF's are an alternative to the direct form predictor coefficients or the lattice form reflection coefficients for representing the filter response.

The direct form coefficient representation of the LPC filters is not conducive to efficient quantization. Instead, nonlinear functions of the reflection coefficients (e.g., log-area ratio or inverse sine of the reflection coefficient) are often used as transmission parameters [5]. These parameters are preferable because they have a relatively low spectral sensitivity.

Line spectral frequencies are an alternate parameterization of the filter with a one-to-one correspondence with the direct form predictor coefficients. The concept of an

LSF was introduced by Itakura [6]. LSF's encode speech spectral information more efficiently than other transmission parameters [2]–[4], [7]. This can be attributed to the intimate relationship between the LSF's and the formant frequencies. Accordingly, LSF's can be quantized taking into account spectral features known to be important in perceiving speech signals. In addition, LSF's lend themselves to frame-to-frame interpolation with smooth spectral changes because of their frequency domain interpretation.

The emphasis of this work is on the efficient computation of line spectral frequencies. This will involve an iterative root finding algorithm for a series representation in Chebyshev polynomials. The algorithm developed is simple in structure and constrains the maximum number of function evaluations. These considerations are important if LSF's are to be used in a real-time environment. The reconversion of LSF's to predictor coefficients is based on reconstructing the expansion in Chebyshev polynomials from the root factors.

II. LINE SPECTRAL FREQUENCIES

The starting point for deriving the LSF's is the response of the prediction error filter with P coefficients

$$A(z) = 1 - \sum_{k=1}^P a(k)z^{-k}. \quad (1)$$

The $\{a(k)\}$ are the direct form predictor coefficients. The corresponding all-pole synthesis filter is $1/A(z)$. A minimum phase prediction error filter (i.e., one with all its roots within the unit circle) has a corresponding synthesis filter which is stable.

A symmetric polynomial $F_1(z)$ and an antisymmetric polynomial $F_2(z)$ related to $A(z)$ are formed by adding and subtracting the time-reversed system function

$$\begin{aligned} F_1(z) &= A(z) + z^{-(P+1)} A(z^{-1}), \\ F_2(z) &= A(z) - z^{-(P+1)} A(z^{-1}). \end{aligned} \quad (2)$$

The roots of these two auxiliary polynomials determine the line spectral frequencies. The two polynomials also have the interpretation of being the system polynomials for a $P + 1$ coefficient predictor derived from a lattice structure. The first P stages of the lattice have the same

Manuscript received November 22, 1985; revised February 15, 1986.

P. Kabal is with the Department of Electrical Engineering, McGill University, Montreal, P.Q., Canada H3A 2A7, and INRS-Télécommunications, Université du Québec, Verdun, P.Q., Canada H3E 1H6.

R. P. Ramachandran is with the Department of Electrical Engineering, McGill University, Montreal, P.Q., Canada H3A 2A7.

IEEE Log Number 8610167.

response as the original P stage predictor. An additional stage is added with reflection coefficients equal to $+1$ or -1 to give the response $F_1(z)$ or $F_2(z)$, respectively [2].

Soong and Juang [2] have shown that if $A(z)$ is minimum phase, then 1) the roots of $F_1(z)$ and $F_2(z)$ are on the unit circle; and 2) the roots are simple and separate each other. In addition, any procedure which determines an ordered set of LSF's can be used to construct a minimum phase prediction error filter.

The polynomials $F_1(z)$ and $F_2(z)$ being symmetrical and antisymmetrical, respectively, have roots at $z = +1$ and/or $z = -1$ which can be removed by polynomial division

$$G_1(z) = \frac{F_1(z)}{1+z^{-1}} \quad \text{and} \quad G_2(z) = \frac{F_2(z)}{1-z^{-1}}, \quad P \text{ even,}$$

$$G_1(z) = F_1(z) \quad \text{and} \quad G_2(z) = \frac{F_2(z)}{1-z^{-2}}, \quad P \text{ odd.}$$

(3)

These polynomial divisions can be performed by additions and subtractions of the coefficients of $F_1(z)$ and $F_2(z)$. The resulting $G_1(z)$ and $G_2(z)$ are symmetric polynomials of even order. Since the roots occur in complex conjugate pairs, it is only necessary to determine the roots located on the upper semicircle. The roots of interest are $\exp j\omega_i$ for $i = 1, 2, \dots, P$. The line spectral frequencies are the angular positions of the roots $0 < \omega_i < \pi$.

Fig. 1 shows the arrangement of zeros of $F_1(z)$ and $F_2(z)$ for both even and odd P . These plots show the actual root positions for a voiced segment of speech (8 kHz sampling rate). The polynomials $G_1(z)$ and $G_2(z)$ have the same zeros as $F_1(z)$ and $F_2(z)$, respectively, except for the zeros at $z = \pm 1$. It can be noted that for any order, the lowest frequency LSF corresponds to a root of $G_1(z)$. These plots show that for roots of $A(z)$ near the unit circle, a pair of LSF's tends to bracket the angular position of the root of $A(z)$. However, it also indicates that the difference between pairs of LSF's is not necessarily a good indicator of how close a root of $A(z)$ is to the unit circle. Interpretation of the LSF's in terms of formant resonances for P odd is more tenuous due to the influence of the real axis root of $A(z)$.

A stability theorem which uses a form similar to the LSF formulation has been formulated by Schussler [8]. The auxiliary symmetric and antisymmetric polynomials defined by Schussler become the same as $F_1(z)$ and $F_2(z)$ if $A(z)$ is considered to be a polynomial of degree $P+1$ with $a(P+1) = 0$. The root locations of the auxiliary polynomials given by a direct application of Schussler's theorem (i.e., without appending a zero valued coefficient), and those given by an LSF formulation are compared in Appendix A. This examination provides additional insight as to the relationship of the LSF's to the roots of $A(z)$.

The cases of an odd number and an even number of LSF's differ in some details. Let the order of the poly-

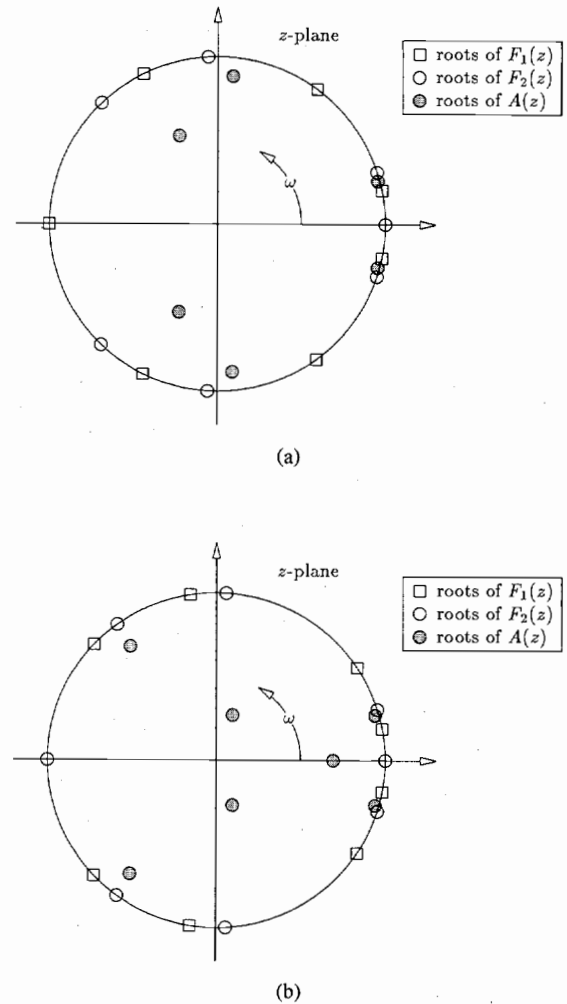


Fig. 1. Root locations. (a) P even (shown for $P = 6$). (b) P odd (shown for $P = 7$).

nomials $G_1(z)$ and $G_2(z)$ be $2M_1$ and $2M_2$, respectively,

$$M_1 = \frac{P}{2} \quad \text{and} \quad M_2 = \frac{P}{2}, \quad P \text{ even,}$$

$$M_1 = \frac{P+1}{2} \quad \text{and} \quad M_2 = \frac{P-1}{2}, \quad P \text{ odd.} \quad (4)$$

Then explicitly showing the symmetry of the polynomial coefficients

$$\begin{aligned} G_1(z) &= 1 + g_1(1)z^{-1} + \dots + g_1(M_1)z^{-M_1} + \dots \\ &\quad + g_1(1)z^{-(2M_1-1)} + z^{-2M_1}, \\ G_2(z) &= 1 + g_2(1)z^{-1} + \dots + g_2(M_2)z^{-M_2} + \dots \\ &\quad + g_2(1)z^{-(2M_2-1)} + z^{-2M_2}. \end{aligned} \quad (5)$$

$G_1(z)$ contributes M_1 pairs of conjugate zeros and $G_2(z)$ contributes M_2 pairs of conjugate zeros ($M_1 + M_2 = P$). On the unit circle, the linear phase term can be removed to give two zero phase series expansions in cosines

$$\begin{aligned} G_1(e^{j\omega}) &= e^{-j\omega M_1} G'_1(\omega), \\ G_2(e^{j\omega}) &= e^{-j\omega M_2} G'_2(\omega), \end{aligned} \quad (6)$$

where

$$\begin{aligned} G'_1(\omega) &= 2 \cos M_1 \omega + 2g_1(1) \cos (M_1 - 1)\omega + \cdots \\ &\quad + 2g_1(M_1 - 1) \cos \omega + g_1(M_1), \\ G'_2(\omega) &= 2 \cos M_2 \omega + 2g_2(1) \cos (M_2 - 1)\omega + \cdots \\ &\quad + 2g_2(M_2 - 1) \cos \omega + g_2(M_2). \end{aligned} \quad (7)$$

Various methods to locate the roots of $G'_1(\omega)$ and $G'_2(\omega)$ have been suggested. The procedure developed by Soong and Juang [2] evaluates (7) on a fine grid by using a discrete cosine transform. Sign changes at adjacent grid points isolate intervals containing roots and further bisection of these intervals gives an approximation to the root positions.

Kang and Fransen [7] have proposed two other methods for finding the LSF's. In one method, the autocorrelation functions of the coefficients of $G'_1(\omega)$ and $G'_2(\omega)$ are used to calculate power spectra. The locations of the local minima of the power spectra give the LSF's. The second method uses an all-pass ratio filter

$$R(z) = \frac{z^{-(P+1)} A(z^{-1})}{A(z)}. \quad (8)$$

The phase spectrum of the ratio filter is evaluated and whenever the phase response takes on a value which is a multiple of π , the corresponding frequency is an LSF. All of the above methods require the evaluation of trigonometric functions and may not be appropriate for a real-time environment.

The method proposed in this paper requires no prior storage or calculation of trigonometric functions. The method is introduced in two steps. In the next section, the use of a Chebyshev polynomial expansion is discussed. Subsequently, an efficient numerical algorithm to find the roots with this formulation is established.

III. CHEBYSHEV SERIES FORMULATION

Consider the frequency mapping $x = \cos \omega$. Then

$$\cos m\omega = T_m(x), \quad (9)$$

where $T_m(x)$ is an m th-order Chebyshev polynomial in x . The Chebyshev polynomials satisfy the order recursion

$$T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x), \quad (10)$$

with initial conditions, $T_0(x) = 1$ and $T_1(x) = x$. The series expansions in cosines (7) can now be expressed in terms of Chebyshev polynomials

$$\begin{aligned} G'_1(x) &= 2T_{M_1}(x) + 2g_1(1)T_{M_1-1}(x) + \cdots \\ &\quad + 2g_1(M_1 - 1)T_1(x) + g_1(M_1), \\ G'_2(x) &= 2T_{M_2}(x) + 2g_2(1)T_{M_2-1}(x) + \cdots \\ &\quad + 2g_2(M_2 - 1)T_1(x) + g_2(M_2). \end{aligned} \quad (11)$$

Once the roots $\{x_i\}$ of $G'_1(x)$ and $G'_2(x)$ are determined, the corresponding LSF's are given by $\omega_i = \arccos x_i$. The

mapping $x = \cos \omega$ maps the upper semicircle in the z -plane to the real interval $[-1, +1]$. Therefore, all the roots x_i lie between -1 and $+1$, with the root corresponding to the lowest frequency LSF being the one nearest $+1$.

The Chebyshev polynomial series lends itself to an efficient and accurate evaluation which bypasses an expansion in powers of x . Let the series to be evaluated be represented as

$$Y(x) = \sum_{k=0}^{N-1} c_k T_k(x). \quad (12)$$

Consider the backward recurrence relationship

$$b_k(x) = 2xb_{k+1}(x) - b_{k+2}(x) + c_k, \quad (13)$$

with initial conditions $b_N(x) = b_{N+1}(x) = 0$. This recursion is used to calculate $b_0(x)$ and $b_2(x)$. Then $Y(x)$ can be expressed in terms of $b_0(x)$ and $b_2(x)$

$$\begin{aligned} Y(x) &= \sum_{k=0}^{N-1} [b_k(x) - 2xb_{k+1}(x) + b_{k+2}(x)] T_k(x) \\ &= \frac{b_0(x) - b_2(x) + c_0}{2}. \end{aligned} \quad (14)$$

The benefit of this formulation is that errors in the evaluation of $b_0(x)$ and $b_2(x)$ tend to cancel [9]. This results in a numerically stable evaluation of the Chebyshev polynomial series. Neglecting the factor of 2 which does not affect the root locations, each evaluation can be computed with about N multiplies and $2N$ additions.

IV. NUMERICAL SOLUTION FOR THE LINE SPECTRAL FREQUENCIES

In this section, a numerical algorithm to find the roots corresponding to the line spectral frequencies is developed. The basic task is to isolate the roots of $G'_1(x)$ by searching incrementally for intervals in which the sign changes. The search proceeds backwards from $x = 1$ since $G'_1(x)$ has the root nearest $x = 1$. The location of the root in an interval containing a sign change is refined by successive bisection of the root interval. The function values are determined using the backward recursion given in the previous section to compute the Chebyshev polynomials at a given argument value. In this way, only two function values at a time need be stored. Given the interlacing property of the roots, the search for a root of $G'_2(x)$ starts from the position of the root of $G'_1(x)$ just found. The algorithm continues as before, but interchanges the roles of the functions as each root is found.

Two different precisions must be specified for the numerical algorithm. The initial evaluation interval δ must be sufficiently small so that two or more roots of the same function do not occur in the same interval. Let the roots be denoted by $\{x_i\}$ for $i = 1, 2, \dots, P$, and let them be ordered such that $x_i > x_{i-1}$. The roots of $G'_1(x)$ (x_i with i odd) interlace with the roots of $G'_2(x)$ (x_i with i even).

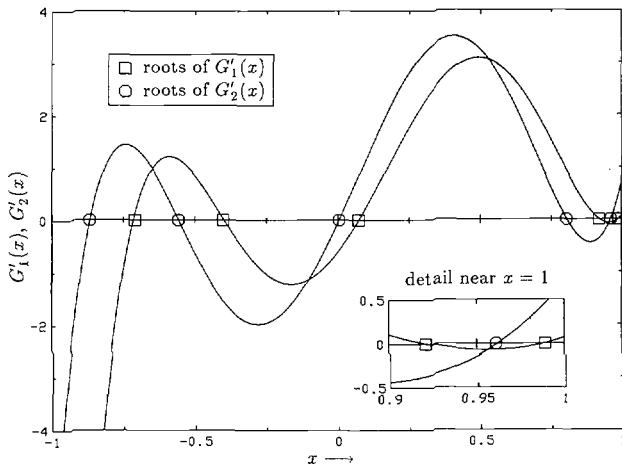


Fig. 2. Plots of $G'_1(x)$ and $G'_2(x)$ ($P = 10$).

Then the initial evaluation interval must satisfy

$$\delta < \min_i (x_i - x_{i-2}). \quad (15)$$

A second increment ϵ specifies the acceptable uncertainty in root position. This value must be small enough so that in switching the search for roots from one function to the other, a root is not missed or roots are not interchanged in order. To guarantee this, ϵ must be smaller than the minimum spacing between pairs of roots, one taken from each function,

$$\epsilon < \min_i (x_i - x_{i-1}). \quad (16)$$

Experiments with speech data were conducted to determine reasonable values for δ and ϵ . Five utterances comprising 10 s of speech sampled at 8 kHz were used. Three utterances were spoken by males and two by females. In all cases, a 20 ms Hamming window was used to perform a 10th-order autocorrelation analysis. This ensures that the predictor is minimum phase [10]. The root locations corresponding to LSF's were determined to a high precision. A plot of $G'_1(x)$ and $G'_2(x)$ for a voiced segment of speech is shown in Fig. 2. It shows a tendency of roots to pair which requires ϵ to be significantly smaller than δ .

Histograms of root differences are shown in Fig. 3. The extreme values of the root differences are summarized in Table 1. These results indicate that $\delta = 0.02$ is sufficiently small to avoid missing sign changes. Each interval of length δ will be bisected to further resolve the root location. From the table, ϵ should be chosen to be less than 0.0015, implying that 4 bisections will be sufficient (for $\delta = 0.02$). While the worst case uncertainty in the x -domain is constant, the uncertainty in the ω -domain varies with ω due to the nonlinear relationship between x and ω . For the parameters given above and assuming 8 kHz sampling, the worst case uncertainty in the LSF's varies between 64 Hz at low and high frequencies down to 1.6 Hz at the middle frequencies. However, the uncertainty remains less than 10 Hz for the frequencies between 200 and 3800 Hz. Kang and Franssen [7] suggest a 10 Hz resolution in evaluating LSF's and, furthermore, find that

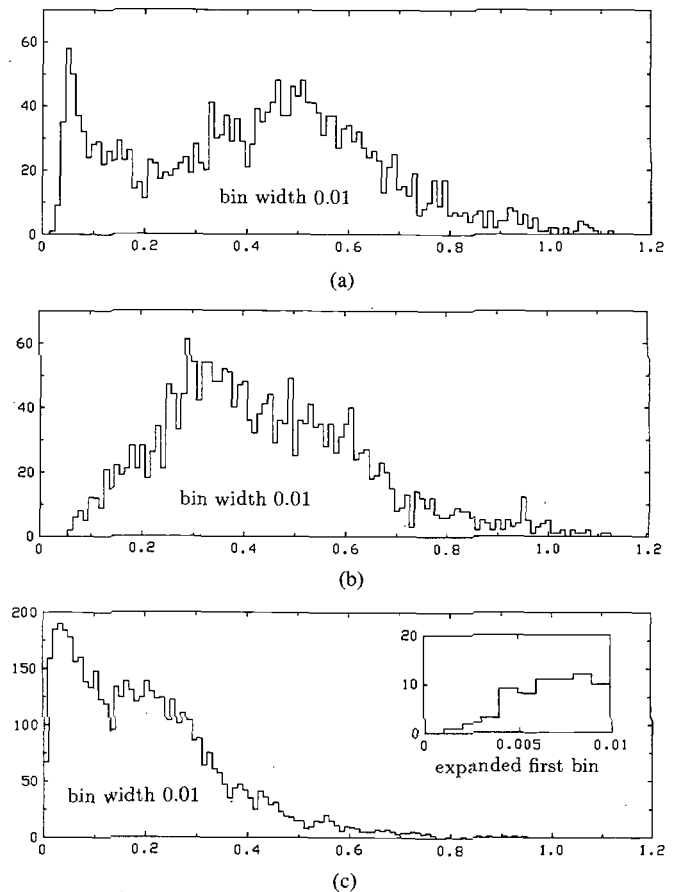


Fig. 3. Histograms of root differences. (a) Differences for roots of $G'_1(x)$. (b) Differences for roots of $G'_2(x)$. (c) Differences between roots of $G'_1(x)$ and $G'_2(x)$.

TABLE I
ROOT DIFFERENCES

| | Minimum Difference | Maximum Difference |
|-------------------------|--------------------|--------------------|
| $G'_1(x)$ only | 0.0232 | 1.121 |
| $G'_2(x)$ only | 0.0564 | 1.195 |
| $G'_1(x)$ and $G'_2(x)$ | 0.0015 | 0.946 |

coarse quantization of LSF's below 300 Hz does not affect speech quality.

As a last step, the root position is estimated by linearly interpolating between the function values determined in the last bisection step. This results in an average error which is significantly smaller than the worst case value given by ϵ . The given root finding algorithm has been formulated to limit the number of function evaluations. The initial search for intervals containing roots uses approximately $2/\delta + P$ evaluations. Bisection then uses an additional $(\lceil \log_2(\delta/\epsilon) \rceil)P$ evaluations. For the values of δ and ϵ given above, the number of function evaluations for a 10th-order LPC system is less than 150. The actual number is somewhat smaller than this value, since the search for roots can be terminated when all roots are found.

The algorithm to find the LSF's is extremely simple in structure and keeps the number of function evaluations

relatively constant. The numerical analysis literature abounds with more sophisticated root finding algorithms. These will tend to find the roots with fewer function evaluations on the average. However, most of these algorithms are inappropriate in a real-time environment since the maximum number of function evaluations is indeterminate.

A slightly more complex root finding algorithm which combines bisection with inverse parabolic interpolation [11] merits consideration for some applications. It can be used to refine the root position when an interval containing a root has been identified. The worst case number of function evaluations for this algorithm is two or three times that for simple bisection, but the average number is smaller.

V. CONVERSION OF LSF'S TO PREDICTOR COEFFICIENTS

The conversion of LSF's to predictor coefficients is less computationally intensive than deriving LSF's from predictor coefficients. Each LSF ω_i gives rise to a second-order polynomial factor of the form $1 - 2 \cos \omega_i z^{-1} + z^{-2}$. These can be multiplied together to form the auxiliary polynomials directly. In this section, an alternate reconstruction process using the Chebyshev series representation will be formulated. This leads to an efficient reconstruction process which takes symmetries in the auxiliary polynomials into account.

The polynomials $G_1'(x)$ and $G_2'(x)$ are reconstructed from their roots by successive polynomial multiplication of the appropriate first-order LSF polynomials

$$\begin{aligned} G_1'(x) &= \prod_{k=1}^{M_1} 2(x - x_{2k-1}), \\ G_2'(x) &= \prod_{k=1}^{M_2} 2(x - x_{2k}). \end{aligned} \tag{17}$$

However, it is not the coefficients of the powers of x that are desired, but the coefficients of the Chebyshev polynomial terms. Consider an N th-order polynomial expressed as a Chebyshev series

$$Y_N(x) = \sum_{k=0}^N c_k T_k(x). \tag{18}$$

Adding one more root factor to form an $N + 1$ st-order Chebyshev representation,

$$\begin{aligned} Y_{N+1}(x) &= 2(x - x_r) Y_N(x) \\ &= \sum_{k=-1}^{N+1} [c_{k-1} - 2x_r c_k + c_{k+1}] T_k(x). \end{aligned} \tag{19}$$

This expression has been put in this form by applying the relation $2xT_k(x) = T_{k-1}(x) + T_{k+1}(x)$ and defining c_k to be zero for $k < 0$ and $k > N$. In addition, note that $T_{-1}(x) = T_1(x)$, which means that the term for $k = -1$ should be combined with the term for $k = 1$. The bracketed term in (19) is the coefficient of the k th Chebyshev polynomial

in the representation of $Y_{N+1}(x)$. This equation defines one step in the recursion to determine the coefficients of the Chebyshev representation from the root factors.

The coefficients for $G_1(z)$ and $G_2(z)$ can be determined directly from $G_1'(x)$ and $G_2'(x)$. This involves applying a factor of $\frac{1}{2}$ to all but one of the coefficients. In fact, multiplication by this factor can be avoided if the recursion is modified to directly produce the coefficients of $G_1(z)$ and $G_2(z)$.

As the penultimate step, $G_1(z)$ and $G_2(z)$ must be multiplied by the polynomial terms with roots at ± 1 to give $F_1(z)$ and $F_2(z)$. This can be carried out on half of the total number of coefficients in these auxiliary polynomials (using symmetry) and involves only additions and subtractions. Finally, the coefficients of the prediction error filter are determined from

$$A(z) = \frac{F_1(z) + F_2(z)}{2}. \tag{20}$$

The reconstruction procedure described requires about $P^2/4 - P/2$ multiplications and $P^2/2 + 2P - 4$ additions.¹ This is less than $\frac{1}{4}$ of the number of multiplications and additions cited for the reconstruction procedure suggested by Kang and Fransen [7].

A minimum phase prediction error filter results. This follows directly from the fact that the reconstruction procedure is the step-by-step inverse of the procedure to find the LSF's. As long as the LSF's are distinct and $F_1(z)$ and $F_2(z)$ are formed from alternating roots, the minimum phase property of the reconstructed prediction error filter is guaranteed.

An alternative to converting the LSF's to predictor coefficients is the use of filter structures that use the LSF's directly as parameters. This kind of structure implements $F_1(z)$ and $F_2(z)$ directly as cascaded second-order sections and can be used as the basis of both the analysis (prediction error) filter and the corresponding synthesis filter [7]. However, this form of filter requires more arithmetic operations per sample than a direct form filter using the predictor coefficients. The tradeoff is then between this extra computation which occurs for each sample of data processed and the computation required to convert LSF's to predictor coefficients. Kang and Fransen [7] show that for reasonable frame sizes in an LPC coder, conversion to predictor coefficients and the use of a direct form filter structure results in a lower operations count than the use of an LSF based filter structure. This conclusion is strengthened by the more efficient procedure to convert to direct form coefficients described here.

The reconstruction procedure given in this section starts with the roots in the x -domain and thus complements the algorithm to derive them from the predictor coefficients. In any transmission system, the LSF's must be quantized and coded. In order to avoid conversion to the ω -domain,

¹These counts apply for P even. For P odd, the number of operations is slightly smaller than given by these formulas. Note also that the counts do not include P multiplications by the factor $\frac{1}{2}$ and P multiplications by the factor 2.

the quantization procedure must be modified to work directly on the x_i values. For scalar quantization in the ω -domain, an equivalent scheme in the x -domain can be formulated by simply applying the cosine nonlinearity to the quantizer boundaries and output values. Kang and Fransen [7] describe a scheme to code the center frequency and difference frequency of a pair of LSF's. This procedure can be viewed in two dimensions as a quantizer with (rotated) rectangular boundaries. In the x -domain, the equivalent quantizer boundaries are curved, but locally, can be well approximated by straight lines. With straight line boundaries, the x -domain quantizer can be implemented with only a small increase in complexity and storage over the ω -domain quantizer. Carrying out the quantization in the x -domain eliminates the need to evaluate transcendental functions in the conversion to or from predictor coefficients.

VI. SUMMARY AND CONCLUSIONS

This paper has reported a method for converting predictor coefficients to a set of line spectral frequencies which can be used for both even- and odd-order LPC systems. The proposed method with the given interval parameters is highly accurate. The accuracy can be further increased by performing more bisections within the root interval, of course at the cost of more function evaluations. The use of an expansion in Chebyshev polynomials obviates the calculation of trigonometric functions and/or the storage of trigonometric tables. The evaluation of these expansions makes use of an efficient and numerically stable algorithm. The root finding algorithm which determines the LSF's has been structured to limit the maximum number of function evaluations for a given accuracy constraint.

The reconversion of the LSF's to predictor coefficients is formulated in terms of a recursive calculation of the coefficients of the Chebyshev expansion. This gives a computationally efficient algorithm which takes into account inherent symmetries in the auxiliary polynomials. If the LSF's are expressed in the cosine domain, trigonometric computations can be avoided altogether.

As a test of the overall procedure, LSF's were found using the procedure described in Section IV for the speech data used previously. The analysis conditions are the same as specified earlier. The 10 LSF's for each frame of speech were reconverted to predictor coefficients by the procedure described in Section V. The maximum difference between a reevaluated predictor coefficient and the original predictor coefficient was 3.8×10^{-5} .

APPENDIX A

RELATIONSHIP BETWEEN LSF'S AND THE PREDICTOR ROOTS

The relationship between the LSF's and the roots of the prediction error filter $A(z)$ is explored in this appendix. In the main text, examples for real speech data have shown that there is indeed a tendency for the LSF's to cluster around the angular positions corresponding to roots of $A(z)$

when these are close to the unit circle. Some insight into the clustering phenomenon can be obtained by examining another formulation related to that for the LSF's.

Schussler [8] has given a stability theorem for a polynomial $A(z)$. The stability condition is expressed in terms of two auxiliary polynomials

$$\begin{aligned}\tilde{F}_1(z) &= A(z) + z^{-P} A(z^{-1}), \\ \tilde{F}_2(z) &= A(z) - z^{-P} A(z^{-1}).\end{aligned}\quad (\text{A.1})$$

The polynomial $A(z)$ has all its roots within the unit circle if and only if: 1) the roots of $\tilde{F}_1(z)$ and $\tilde{F}_2(z)$ are on the unit circle; 2) the roots are simple and separate each other; and 3) $|a(P)| < 1$.² These auxiliary polynomials differ from those used for LSF's by being of order P instead of order $P + 1$. As a result, the roots of Schussler's polynomials cannot be used to uniquely reconstruct $A(z)$. An additional quantity must be specified. This could be the coefficient $a(P)$, which in addition is known to have magnitude less than unity. As will be seen, the formulation derived from Schussler's theorem has drawbacks as a pseudo-LSF representation.

Consider rewriting $\tilde{F}_1(z)$ and $\tilde{F}_2(z)$ as

$$\begin{aligned}\tilde{F}_1(z) &= A(z) [1 + \tilde{R}(z)], \\ \tilde{F}_2(z) &= A(z) [1 - \tilde{R}(z)],\end{aligned}\quad (\text{A.2})$$

where the ratio filter $\tilde{R}(z)$ is defined as

$$\tilde{R}(z) = \frac{z^{-P} A(z^{-1})}{A(z)}.\quad (\text{A.3})$$

Note that the only difference between this ratio filter and that for the LSF formulation [see (8)] is an extra z^{-1} delay term. The auxiliary polynomials have roots at those points on the unit circle at which the phase of the all-pass ratio filter passes through multiples of π .

A simple example will point out some of the ramifications of the extra delay term associated with the LSF formulation. Consider an $A(z)$ which has conjugate pairs of roots near the unit circle. The LSF's coalesce as the roots of $A(z)$ approach the unit circle. By contrast, the formulation that arises from Schussler's stability theorem gives single roots at the angular position of the roots of $A(z)$ but, in addition, roots appear midway between these positions. Fig. 4 shows a phase plot of both $R(e^{j\omega})$ (LSF formulation) and $\tilde{R}(e^{j\omega})$ (Schussler's formulation) for the case of an $A(z)$ which has 3 pairs of conjugate roots, each with magnitude 0.99. The angular positions of the roots correspond to ω equal to $\pi/4$, $\pi/2$, and $3\pi/4$. Symbols are used to mark the places where the phase angle crosses a multiple of π . In the vicinity of the roots of $A(z)$, the phase undergoes an excursion through nearly 2π rad for both formulations. However, the phase offset due to the linear phase component in the LSF formulation is enough to substantially change the positions of the roots of the LSF

²The last condition was added by Gnanasekaran [12]. Conditions (1) and (2) by themselves also hold if $A(z)$ has all its roots outside the unit circle.

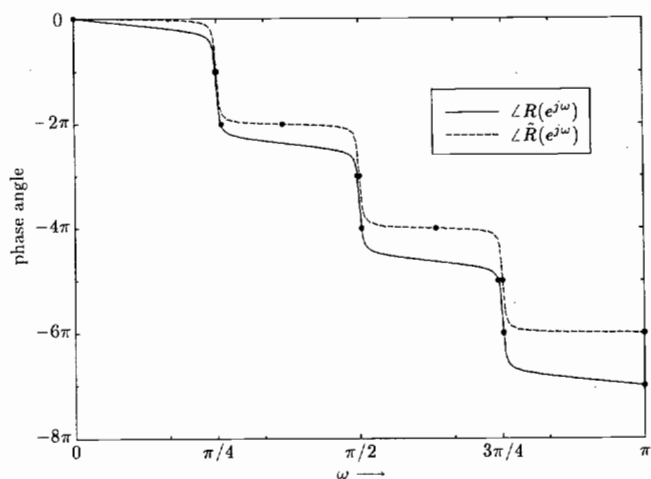


Fig. 4. Phase responses of $R(e^{j\omega})$ and $\hat{R}(e^{j\omega}) - A(z)$ has roots at $0.99e^{\pm j\pi/4}$, $0.99e^{\pm j\pi/2}$, and $0.99e^{\pm j3\pi/4}$.

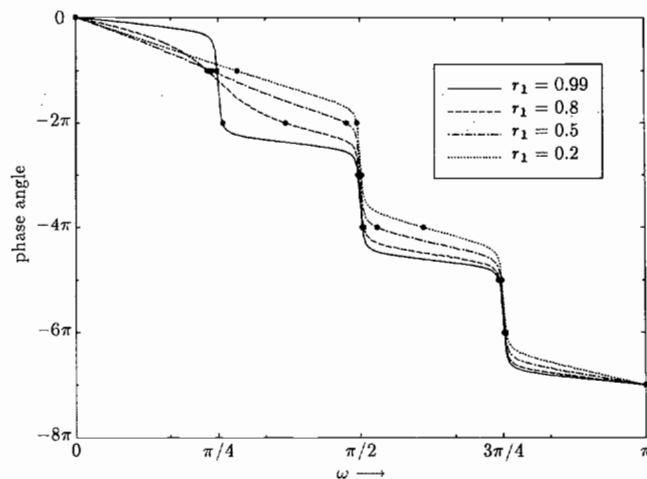


Fig. 5. Phase response of $R(e^{j\omega}) - A(z)$ has roots at $r_1 e^{\pm j\pi/4}$, $0.99e^{\pm j\pi/2}$, and $0.99e^{\pm j3\pi/4}$.

auxiliary polynomials from those for Schussler's formulation.

Kang and Fransen [4] point out that the group delay of $R(z)$ increases in the neighborhood of a root of $A(z)$. This clearly also applies to $\hat{R}(z)$. However, the previous example has shown that peaking of the group delay or the equivalent rapid change in phase angle is not sufficient in itself for a pair of LSF's to occur close together.

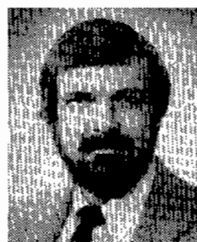
The previous example can be explored further. For this discussion, only the LSF formulation will be considered. Consider changing the radial position of the first pair of roots of $A(z)$ (at angular position $\omega = \pi/4$) while keeping the other root positions constant. As the radius of this pair of roots is reduced to 0.8, the phase change occurs more gently. The lowest frequency LSF's are now spread apart (see Fig. 5). When the radius is reduced to 0.5, the phase change is even more gradual and effects the overall phase in the vicinity of $\omega = \pi/2$. There are now three closely spaced LSF's near $\omega = \pi/2$. As the radius is further reduced to 0.2, the LSF's take on a different configuration. The lowest LSF has moved to just above $\omega = \pi/4$, the second and third LSF's occur close together near $\omega = \pi/2$, and the fourth LSF lies midway between these LSF's and the two paired LSF's near $\omega = 3\pi/4$. This root configuration shows that the sharp resonance at $\omega = \pi/2$ is signaled by a close spacing between the second and third LSF's. In the coding scheme proposed by Kang and Fransen [7], LSF's are coded in pairs specified by a center frequency and an offset frequency. In this case, the coding would be applied to two pairs of LSF's, each of which has a relatively large offset. After quantization of the offset, this coding scheme may not adequately represent the fact that the upper LSF from one pair and the lower LSF from another pair are closely spaced.

This appendix has shown examples that point out that the relationship between the positions of the roots of $A(z)$ and the LSF configuration is more fragile than the literature would lead one to believe. Nonetheless, this does not invalidate the use of LSF's in speech coding. Clearly, the extensive perceptual testing carried out by Kang and

Fransen shows that in spite of hurdles in the interpretation of the LSF's in terms of the spectral features or roots of $A(z)$, LSF's do efficiently represent the perceptually important features of speech spectra.

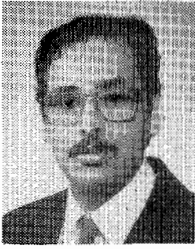
REFERENCES

- [1] H. Wakita, "Linear prediction voice synthesizers: Line spectrum pairs (LSP) is the newest of several techniques," *Speech Technol.*, Fall 1981.
- [2] F. K. Soong and B.-W. Juang, "Line spectrum pair (LSP) and speech data compression," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, San Diego, CA, Mar. 1984, pp. 1.10.1-1.10.4.
- [3] J. R. Crosmer and T. P. Barnwell, III, "A low bit rate segment vocoder base on line spectrum pairs," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, Tampa, FL, Apr. 1985, pp. 7.2.1-7.2.4.
- [4] G. S. Kang and L. J. Fransen, "Application of line spectrum pairs to low bit rate speech encoders," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, Tampa, FL, Apr. 1985, pp. 7.3.1-7.3.4.
- [5] J. D. Markel and A. H. Gray, Jr., *Linear Prediction of Speech*. New York: Springer-Verlag, 1976.
- [6] F. Itakura, "Line spectrum representation of linear predictor coefficients of speech signals," *J. Acoust. Soc. Amer.*, vol. 57, S35(A), 1975.
- [7] G. S. Kang and L. J. Fransen, "Low bit rate speech encoders based on line spectrum frequencies (LSFs)," Naval Res. Lab., Rep. 8857, Nov. 1984.
- [8] H. W. Schussler, "A stability theorem for discrete systems," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, pp. 87-89, Feb. 1976.
- [9] L. Fox and I. B. Parker, *Chebyshev Polynomials in Numerical Analysis*. London, England: Oxford University Press, 1968.
- [10] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1978.
- [11] R. P. Brent, "An algorithm with guaranteed convergence for finding a zero of a function," *The Comput. J.*, vol. 14, pp. 422-425, 1971.
- [12] R. Gnanasekaran, "A note on the new 1-D and 2-D stability theorems for discrete systems," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 1211-1212, Dec. 1981.



Peter Kabal received the B.A.Sc., M.A.Sc., and Ph.D. degrees in electrical engineering from the University of Toronto, Toronto, Ont., Canada.

He is an Associate Professor in the Department of Electrical Engineering at McGill University, Montreal, P.Q., Canada, and a Visiting Professor at INRS-Télécommunications (a research institute affiliated with the Université du Québec), Verdun, P.Q., Canada. His current research interests focus on digital signal processing as applied to speech coding, adaptive filtering, and data transmission.



Ravi Prakash Ramachandran was born in Bangalore, India. He received the B.Eng. degree with great distinction from Concordia University, Montreal, P.Q., Canada, in 1984, and the M.Eng. degree from McGill University, Montreal, P.Q., Canada, in 1986. In high school he received a medal from the University of Waterloo for being on the Canadian Honor List in the Junior Mathematics Competition.

During his undergraduate studies he held a Concordia University Undergraduate Fellowship.

When studying for the M.Eng. degree he held a Natural Sciences and Engineering Research Council of Canada (NSERC) Postgraduate Fellowship. He plans to pursue his doctoral studies at McGill University, and has been awarded another NSERC Postgraduate Fellowship for the Ph.D. degree. His main research interests are in speech coding, data communications, and digital signal processing.

Mr. Ramachandran received the Order of Engineers of Quebec Student Award in 1984, the John H. Chapman Award (presented by Spar Aerospace for scholarship in communications), the Electrical Engineering Medal as the Most Outstanding Student in Electrical Engineering, and the Morris Chait Medal as the Highest Ranking Student in the B.Eng. Program.
