On the Duality Between Line-Spectral Frequencies and Zero-Crossings of Signals

Ramdas Kumaresan and Yudong Wang

Abstract—Line spectrum pairs (LSPs) are the roots (located in the complex-frequency or z-plane) of symmetric and antisymmetric polynomials synthesized using a linear prediction (LPC) polynomial. The angles of these roots, known as line-spectral frequencies (LSFs), implicitly represent the LPC polynomial and hence the spectral envelope of the underlying signal. By exploiting the duality between the time and frequency domains, we define analogous polynomials in the complex-time variable \( \zeta \). The angles of the roots of these polynomials in \( \zeta \)-plane now correspond to zero-crossing points. Analogous to the fact that the line-spectral frequencies represent the spectral envelope of a signal, these zero-crossing locations can be used to represent the temporal envelope of bandpass signals.

I. INTRODUCTION

Line spectrum pairs (LSPs) are the roots of symmetric and antisymmetric polynomials synthesized using a linear prediction (LPC) polynomial. These roots lie in the complex-frequency or the z-plane, are simple, and have unit magnitude. LSPs were originally proposed as an alternative representation of LPC coefficients by Itakura [1]. The angles of these roots, called line-spectral frequencies (LSFs), are widely used for speech coding, synthesis, and recognition. They implicitly represent the spectral envelope of the underlying signal. They have been shown to possess statistical properties that are suitable for quantization and coding [2].

In this correspondence, we point out the analogy between the complex-time domain (called the \( \zeta \)-domain) and the complex-frequency domain (the traditional z-domain), and thereby show that the dual of the angles of LSPs are the zero-crossings of certain signals. These zero-crossing locations may then be used to represent the temporal envelope of bandpass signals. In Section II, we summarize the well-known [1], [3] properties of the LSPs. In Section III, we describe the dual relationships in the complex-time domain. Two examples are provided to illustrate the main point. The first is a synthetic example. In the second example a \( T \) second segment of bandpass filtered speech signal’s envelope is represented using the said zero-crossings.

II. LINE-SPECTRAL PAIRS

Let \( A_m(z) \) denote an \( m \)th order linear prediction or inverse filter polynomial [4]

\[
A_m(z) \triangleq a_0 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_m z^{-m}
\]

where \( a_i, i = 0, 1, 2, \cdots, m \) are assumed to be complex valued. Let \( x(n) \) denote a sequence of signal samples. \( A_m(z) \) is obtained [4] by minimizing the linear prediction error energy \( \sum_n |e(n)|^2 \) where

\[
e(n) = x(n) - a_n
\]

where \( \cdot \) denotes linear convolution. \( a_0 \) is constrained to be unity to avoid the trivial solution. Note that

\[
\sum_n |e(n)|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |X(e^{j\omega}) A_m(e^{j\omega})|^2 d\omega.
\]

The resulting polynomial, \( A_m(z) \), can be shown to be minimum-phase, i.e., its roots lie strictly inside the unit circle \( |z| = 1 \) (see, for example, [5]). If \( m \) is sufficiently high, then \( 1/A_m(e^{j\omega}) \) models the spectral envelope of the sequence, \( x(n) \), well. Let \( A_m^*(1/z^*) \) denote the reciprocal polynomial (with roots in reciprocal conjugate locations, i.e., reflected outside the unit circle)

\[
A_m^*(1/z^*) \triangleq a_0^* + a_1^* z + a_2^* z^2 + \cdots + a_m^* z^m.
\]

We define two other polynomials

\[
P(z) \triangleq [A_m(z) + z^{-l} A_m^*(1/z^*)]
\]

\[
Q(z) \triangleq [A_m(z) - z^{-l} A_m^*(1/z^*)].
\]

Note that the coefficients of \( P(z) \) and \( Q(z) \) have conjugate-even and conjugate-odd symmetry, respectively. Two important properties of \( P(z) \) and \( Q(z) \) are well-known [1], [3]: if \( l \geq m \) all the roots of \( P(z) \) and \( Q(z) \) are on the unit circle and interlaced with each other and if \( l < m \), not all the roots of \( P(z) \) and \( Q(z) \) are guaranteed to be on the unit circle.

Rewriting (5) and (6) in a product form, we have

\[
P(z) = A_m(z) [1 + H(z)]
\]

\[
Q(z) = A_m(z) [1 - H(z)]
\]

where \( H(z) \) is an all-pass function

\[
H(z) = z^{-l-m} \prod_{i=1}^{m} \frac{z_i^* - z^{-1}}{1 - z_i z^{-1}}
\]

where \( z_i \)’s are the roots of \( A_m(z) \), \( z_i = r_i e^{j\phi_i} \), and \( r_i < 1, \theta_0 = \angle(a_0^*/a_0) \). Since \( H(z) \) is an all-pass filter function, we can write

\[
H(e^{j\omega}) = e^{j\phi(\omega)}.
\]

It should be clear from (7) and (8) that \( P(z) \) and \( Q(z) \) have roots at the locations where \( e^{j\phi(\omega)} \) equals \( -1 \) and \( 1 \), respectively.

The phase function \( \phi(\omega) \) can further be expressed as [3], [6]

\[
\phi(\omega) = \theta_0 - (m + (l - m)) \omega - \sum_{i=1}^{m} 2 \tan^{-1} \left( \frac{r_i \sin(\omega - \omega_i)}{1 - r_i \cos(\omega - \omega_i)} \right).
\]

The group delay \( \tau(\omega) \) of \( H(e^{j\omega}) \) is \( -d\phi(\omega)/d\omega \) and is given by [3], [6]

\[
\tau(\omega) = (l - m) + \sum_{i=1}^{m} \frac{1 - r_i^2}{1 + r_i^2 - 2r_i \cos(\omega - \omega_i)}.
\]
If \( l \geq m \), because all \( r_i < 1 \), we conclude that \( \pi(\omega) > 0 \) and \( \phi(\omega) \) is a monotonically decreasing function. Let \( \theta_0 \) denote the phase of \( H(e^{j\omega}) \) at \( \omega = 0 \), i.e., \( \phi(0) = \theta_0 \) and \( \phi(2\pi) = \theta_0 - 2\pi \). Therefore, \( \phi(\omega) \) crosses lines corresponding to each integer multiple of \( \pi \) (odd and even multiple of \( \pi \) for \( P(z) \) and \( Q(z) \), respectively) exactly once, resulting in \( 2l \) crossing points for \( 0 \leq \omega < 2\pi \). Because the solution to \( P(z) = 0 \) or \( Q(z) = 0 \) requires that \( H(z) = \pm 1 \), these crossing points constitute the total \( 2l \) roots of \( P(z) \) and \( Q(z) \) alternately on the unit circle \([1],[3]\). If we choose \( l = m + 1 \), the angles corresponding to the total \( 2l \) roots of the \( l \)th degree polynomials, \( P(z) \) and \( Q(z) \), are called line spectral frequencies (LSFs).

In summary, if all the roots of a polynomial \( A_m(z) \) lie strictly within the unit circle and \( l \geq m \), then all the roots of \( P(z) \) and \( Q(z) \) are simple and lie on the unit circle. However this may not be the case if \( l < m \). Given \( P(z) \) and \( Q(z) \) we can recover \( A_m(z) \) by computing \( (P(z) + Q(z))/2 \). If \( l \geq 2m \) then \( P(z) \) or \( Q(z) \) is sufficient to recover \( A_m(z) \). We shall make use of these results in the next section and apply them to functions in the complex-time domain.

### III. Duality between Zero-Crossings and LSFs

Let us consider periodic signals with period \( T \). Let \( \Omega = 2\pi/T \) denote the fundamental frequency, \( h_m(t) \) denotes a periodic signal, which we shall call the “inverse signal” \([7]\), for reasons that will become apparent in what follows:

\[
h_m(t) \triangleq h_0 + h_1 e^{j2\Omega t} + h_2 e^{j4\Omega t} + \cdots + h_m e^{j2m\Omega t}.
\]

We shall define a complex variable \( \zeta \equiv e^{-j\Omega t} \), \( \zeta \), the complex-time variable \([8]\), is analogous to the commonly used complex-frequency variable \( \omega \). Note that, on the unit circle in the \( \zeta \) plane, as the angle \( \Omega t \) traverses from 0 to \( 2\pi \) the time variable \( t \) traverses one period, from 0 to \( T \) seconds. We can rewrite (14) in terms of \( \zeta \), analogous to \( A_m(z) \) in (1), as follows:

\[
H_m(\zeta) = h_0 + h_1 \zeta^{-1} + h_2 \zeta^{-2} + \cdots + h_m \zeta^{-m}.
\]

We also define \( P(\zeta) \) and \( Q(\zeta) \) analogous to \( P(z) \) and \( Q(z) \) as follows:

\[
P(\zeta) = [\zeta^{l/2} H_m(\zeta) + \zeta^{-l/2} H_m(1/\zeta^*)] \quad \text{(16)}
\]

\[
Q(\zeta) = [\zeta^{l/2} H_m(\zeta) - \zeta^{-l/2} H_m(1/\zeta^*)] \quad \text{(17)}
\]

Using the same arguments as in Section II, we note that if the roots of \( H_m(\zeta) \) fall within the unit circle in the \( \zeta \) plane and \( l \geq m \), then the roots of \( P(\zeta) \) and \( Q(\zeta) \) fall on the unit circle, \( |\zeta| = 1 \), and are interlaced. The polynomials \( P(\zeta) \) and \( Q(\zeta) \) are defined such that \( \zeta \) is raised to powers ranging from \(-l/2 \) to \( l/2 \), instead of 0 to 1. \( l \) is assumed even. Since we assume that \( H_m(\zeta) \) has all roots inside the unit circle (i.e., minimum-phase), the log-magnitude of \( H_m(\zeta) \) and phase of \( H_m(\zeta) \) (when evaluated around the unit circle \( |\zeta| = 1 \)) are related by the Hilbert transform \([6],[7]\). That is

\[
H_m(\zeta)|_{\zeta = e^{j\pi}} = e^{j\phi(t)} + j\theta(t) \quad \text{(18)}
\]

where the phase function \( \theta(t) \) is the Hilbert transform of the log-magnitude function \( \phi(t) \). Similarly, evaluating \( H_m^*(1/\zeta^*) \) around the unit circle we have

\[
H_m^*(1/\zeta^*)|_{\zeta = e^{-j\pi}} = e^{j\phi(t)} - j\theta(t). \quad \text{(19)}
\]

Plugging (18), (19) and \( \zeta = e^{-j\Omega t} \) in (16) and (17) we have

\[
p(t) = P(e^{j\Omega t}) = 2e^{j\gamma(t)} \cos \left( \frac{l}{2} \Omega t - \zeta(t) \right). \quad \text{(20)}
\]

Similarly

\[
q(t) = jQ(e^{j\Omega t}) = 2e^{j\gamma(t)} \sin \left( \frac{l}{2} \Omega t - \zeta(t) \right). \quad \text{(21)}
\]

Since \( P(\zeta) \) and \( Q(\zeta) \), each have \( l \) simple roots on the unit-circle, say, at locations \( \zeta_1, \zeta_2, \cdots, \zeta_l \), respectively, the corresponding \( p(t) \) and \( q(t) \) have zero-crossings at \( \zeta_1, \zeta_2, \cdots, \zeta_l \), and \( \zeta_1, \zeta_2, \cdots, \zeta_l \), respectively, where \( \zeta_i/\Omega = t_i, i = 1, 2, \cdots, l \). Just as the LSFs represent the LPC zero-crossings \( t_1, t_2, \cdots, t_l \), and \( t_1, t_2, \cdots, t_l \), represent the inverse signal in (14). The signals (such as \( p(t) \) and \( q(t) \)) that can be reconstructed from their zero-crossing locations only (to within a scale factor) are called real-zero (RZ) signals \([9]\).

### IV. Computing the Inverse Signal \([10]\)

Consider a real-valued bandpass signal \( s(t) \) observed over an interval of 0 to \( T \) seconds. We can express \( s(t) \) in terms of its Fourier coefficients, \( b_k \), as follows:

\[
s(t) = \sum_{k=-N}^{N} b_k e^{jK\theta_kt}. \quad \text{(22)}
\]

Since \( s(t) \) is real-valued \( b_{-k} = b_k^* \), \( \omega_k = K\Omega \) is the high frequency band-edge of the signal. (Refer to the example in Fig. 1(a)). Further, let \( s(t) \) be such that the Fourier coefficients \( b_{-K+1}, \cdots, b_0, \cdots, b_{K-1} \) are equal to zero for some \( K < N \). In other words, \( s(t) \), is a bandpass signal. \( \omega_k = K\Omega \) represents the lower band-edge of the signal. Let us define an error signal, \( e(t) \), over 0 to \( T \) seconds as follows:

\[
e(t) = s(t)h_m(t). \quad \text{(23)}
\]

where \( h_m(t) \) is defined in (14), \( h_0 = 1 \). Our goal is to find an inverse signal, \( h_m(t) \), (i.e., choose the coefficients, \( h_i \)) such that the energy in the error signal \( e(t) \) is minimized. Plugging in the expression for \( s(t) \) from (22) into the error-energy expression we get

\[
\frac{1}{T} \int_0^T |e(t)|^2 dt = \frac{1}{T} \int_0^T |s(t)h_m(t)|^2 dt = \sum_{m=-\infty}^{\infty} |g_m|^2 \quad \text{(24)}
\]

where \( g_m = h_m * h_m \) (* denotes linear convolution). Note the obvious duality between the expressions in (24) and (3). The inverse signal \( h_m(t) \) attempts to eliminate the temporal envelope variations in the error signal \( e(t) \), just as the inverse filter tends to flatten the spectral envelope of the error sequence \( e(n) \). The inverse signal coefficients, \( h_i \), can be determined by solving linear equations. In earlier work \([7],[10],[11]\), we called this method Linear Prediction in Spectral Domain (LPSD). The \( h_m(t) \) obtained by minimizing the error energy in (24) is not only a minimum phase signal, but further, \( 1/h_m(t) \) also gives an accurate estimate of the Hilbert envelope of \( s(t) \) (i.e., \( |s(t) \pm \hat{s}(t)| \)), although it is computed directly from the real-valued \( s(t) \) \([7],[10]\).

However, for the latter to be true, \( s(t) \) must have sufficient number of zero-valued Fourier coefficients in the low frequency region, i.e., \( m \) must be less than or equal to \( 2K - 1 \) \([10]\). Once \( h_m(t) \) is computed, then \( p(t) \) and \( q(t) \) can be calculated (for some \( l \geq m \) using the expressions in Section III. The zero-crossings of \( p(t) \) and \( q(t) \) can then be determined and they implicitly represent the temporal envelope of \( s(t) \).
V. EXAMPLES

We provide two examples to illustrate the above ideas. The first example uses a real-valued synthetic signal and the second example is based on bandpass filtered speech signal.

A. Synthesized Bandpass Signal

We synthesized an analytic periodic signal, \( s_a(t) \), using \( 7 (M = 6) \) Fourier coefficients as follows:

\[
s_a(t) = e^{i\omega_L t} \sum_{k=0}^{M} a_k e^{i\lambda_k t}
\]

where \( \omega_L \), the low frequency band-edge, is an integer multiple of \( \Omega \), i.e., \( \omega_L = K\Omega, K = 15; \Omega = 2\pi/T \) and \( T = 16 \text{ ms} \); \( a_k \) are the complex amplitudes of the sinusoids \( e^{i\lambda_k t} \). \( \lambda_0 = 1, \lambda_1 = -0.6024 - 3.2827i, \lambda_2 = -5.6441 + 1.5835i, \lambda_3 = -0.1454 + 7.4390i, \lambda_4 = -6.4822 - 1.1832i, \lambda_5 = -4.6306 - 6.7388i, \lambda_6 = 1.0737 + 2.7369i \). The polynomial \( S_a(\zeta) \) corresponding to the signal \( s_a(t) \), has two (non-trivial) zeros inside the unit circle and four outside. Let \( s(t) \) be the real part of \( s_a(t) \) and is sampled at 16000 Hz. \( s(t) \) is plotted in Fig. 1(b).

Its magnitude spectrum is shown in Fig. 1(a). Using \( s(t) \), we computed \( h_m(t) \) by minimizing the error in (24) (its discrete-time version) by using the LPSD algorithm [7], [10]. The corresponding \( h_m(t) \) is always a minimum phase signal for any order \( m \), i.e., all the zeros of \( H_m(\zeta) \) are inside the unit circle. \( m = 7 \) in this example. The estimated envelope \( 1/[h_m(t)] \) is shown (dashed line) in Fig. 1(b) along with the true Hilbert envelope \( |s_a(t)| \) (dash-dotted line). They almost perfectly coincide. This would be the case as long as the envelope of \( s(t) \) (i.e., \( |s_a(t)| \)) does not go to zero in the time interval of 0 to \( T \) seconds and the condition that \( m \leq 2K-1 \) is met [7]. In Fig. 1(d), the RZ functions \( p(t) \) and \( q(t) \) are shown in upper and lower panels, respectively. We have marked the zero-crossing locations of \( p(t) \) and \( q(t) \) by spikes along the time axis in (d). These spike locations alone contain enough information to reconstruct the envelope of \( s(t) \) up to a scale factor.

B. Bandpassed Speech Signal

In this section, we apply our algorithm to a bandpass filtered speech signal to represent its temporal envelope. A speech signal segment from the ISOLET database corresponding to the utterance /e/ by a female was used: isolett/isolet1/fcm00/fcm00-e2-t.dat. It is first downsampled by a factor of two. We then filter the speech signal with a bank of gamma-tone filters. These filters were defined by Patterson and Holdsworth [12] and are believed to be a good model for the cochlear filters. We chose one block output from the bandpass filter centered around 2240 Hz with a 3 dB bandwidth of \( \approx 232 \text{ Hz} \) (centered around the second formant) for further processing. The filtered signal was weighted by a 14.6-ms rectangular window. The magnitude spectrum of this signal is shown in Fig. 2(a) and the signal waveform is shown in Fig. 2(b). Note that the Fourier coefficients in the low-frequency region are not exactly zero (unlike in the synthetic example above,) but their magnitudes are relatively small. Using this signal, we computed the inverse signal \( h_m(t) \) and the associated \( p(t) \) and \( q(t) \) (\( m = 6, l = 14, \))
Fig. 2. Bandpassed speech signal: The bandpassed and time-windowed speech signal and its spectrum are plotted in (b) and (a). The estimated envelope \( p(t) \) and \( q(t) \) (be the same) are plotted in (c). In (d), \( p(t) \) and \( q(t) \) are shown in upper and lower panels respectively. Since they are described fully (to within a scale factor) by their zero-crossings, we can represent \( p(t) \) and \( q(t) \), and hence the envelope of speech signal by only marking their zero-crossing time locations.

and \( K \approx 38 \). As before the \( p(t) \) and \( q(t) \) and its zero-crossings are shown in Fig. 2(d).

We note from this speech example that, even though the Fourier coefficients at the low-frequency region of the bandpassed speech are not exactly zero, we can still approximate its envelope by well.

Then we can represent the signal envelope by the two sets of spikes in Fig. 2(d).

VI. DISCUSSIONS AND CONCLUSION

By invoking the analogy between time and frequency domains we have shown that the envelope of a periodic bandpass signal \( s(t) \) can be represented implicitly by zero-crossings of certain RZ signals \( p(t) \) and \( q(t) \). These RZ signals are derived from \( s(t) \) by inverse signal analysis, i.e., by minimizing the energy in the error signal \( e(t) = s(t) - h_m(t) \). Although these results are valid for periodic signals, they can be applied to nonperiodic signals by appropriately windowing the signal. More recent results [13] have shown that the error signal \( e(t) \) itself is an all-phase signal (dual of all-pass filter) and hence can also be represented by another set of zero-crossings. This leads to the possibility of representing bandpass signals (actually its envelope and phase) by solely using zero-crossing or timing information as originally envisaged by Logan [14].

REFERENCES


