uniquely from the LSP parameters, approximations of these polynomial can be computed at the receiver. From these, we can identify an approximation of $A_N(z)$ because

$$A_N(z) = (P(z) + Q(z))/2.$$

The speech segment can then be reconstructed from the stable filter $1/A_N(z)$ as described in Section 5.6. Several features of the LSP coding scheme are worth noting.

1. *Stability preserved*. As long as the ordering (e.g., Eq. (7.39)) is preserved in the quantization, the reconstructed version of $A_N(z)$ is guaranteed to have all zeros in $|z| < 1$. Thus, stability of $1/A_N(z)$ is preserved despite quantization as in the case of lattice coefficient quantization.
2. *Perceptual spectral interpretation*. Unlike the lattice coefficients, the LSP coefficients are perceptually better understood. To explain this, recall first that for sufficiently large N, the A$R(N)$ model gives a good approximation of the speech power spectrum $S_{xx}(e^{j\omega})$. This approximation is especially good near the peak frequencies, called the *formants* of speech. Now, the peak locations correspond approximately to the pole angles of the filter $1/A_N(z)$. Near these locations, the phase response tends to change rapidly. The same is true of the phase response $\phi(\omega)$ of the all-pass filter, as shown in Fig. 7.12(b). The LSP coefficients, which are intersections of the horizontal lines (multiples of π) with the plot of $\phi(\omega)$, therefore tend to get *crowded* near the formant frequencies. Thus, the crucial features of the power spectrum tend to get coded into the *density-information* of the LSP coefficients at various points on the unit circle. This information allows us to perform bit allocation among the LSP coefficients (quantize the crowded LSP coefficients with greater accuracy). For a given bit rate, this results in perceptually better speech quality, as compared with the quantization of lattice coefficients. Equivalently, for a given perceptual quality, we can reduce the bit rate; in early work, an approximately 25% saving has been reported using this idea, and more savings have been obtained in a series of other papers. Furthermore, as the bit rate is reduced, the speech degradation is found to be more gradual compared with lattice coefficient quantization.
3. *Acoustical tube models*. It has been shown in speech literature that the lattice structure is related to an acoustical tube model of the vocal tract (Markel and Gray, 1976). This is the origin of the term *reflection coefficients*. The values $k_{N+1} = \pm 1$ in Fig. 7.9 indicate, for example, situations where one end of the tube is open or closed. Thus, the LSP frequencies are related to the open- and close-ended acoustic tube models.
4. *Connection to circuit theory*. In the theory of passive electrical circuits, the concept of *reactances* is well-known (Balabanian and Bickart, 1969). Reactances are input impedances of electrical LC networks. It turns out that reactances have a pole-zero alternation property similar to the

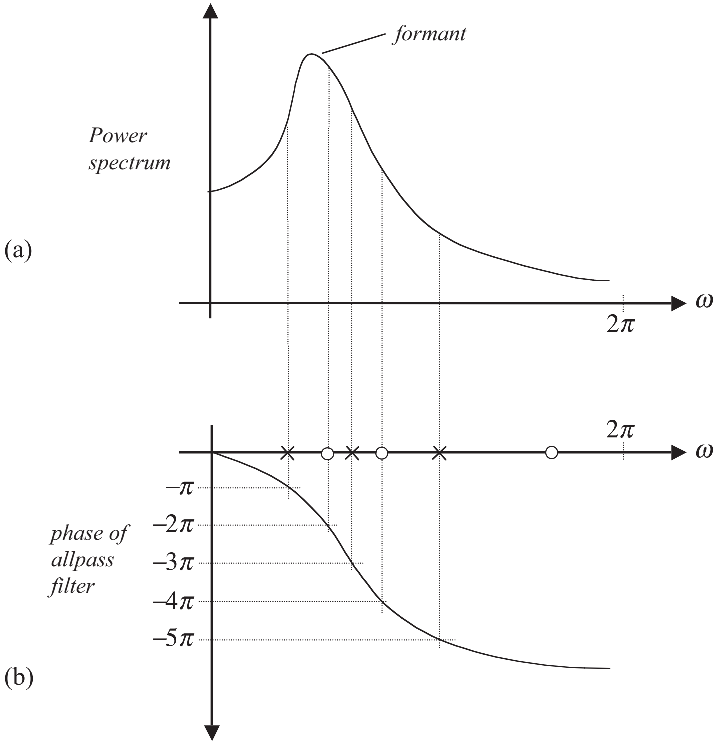

FIGURE 7.12: Explanation of how the LSP coefficients tend to get crowded near the formant regions of the speech spectrum. (a) A toy power spectrum with one formant and (b) the phase response of the all-pass filter $z^{-(N+1)}\widetilde{A}_N(z)/A_N(z)$, where $A_N(z)$ is the prediction polynomial (see text).

alternation of the zeros of the LSP polynomials $P(z)$ and $Q(z)$. In fact, the ratio $P(z)/Q(z)$ is nothing but a discrete time version of the reactance function.

7.9 CONCLUDING REMARKS

In this chapter, we presented a rather detailed study of line spectral processes. The theory finds applications in the identification of sinusoids buried in noise. A variation of this problem is immediately applicable in array signal processing where one seeks to identify the direction of arrival of a signal using an antenna array (Van Trees, 2002). Further applications of the concepts in signal compression, using LSPs, was briefly reviewed.

• • • •