ECE 797: Speech and Audio Processing

Hand-out for Lecture #8

Thursday, March 11, 2004

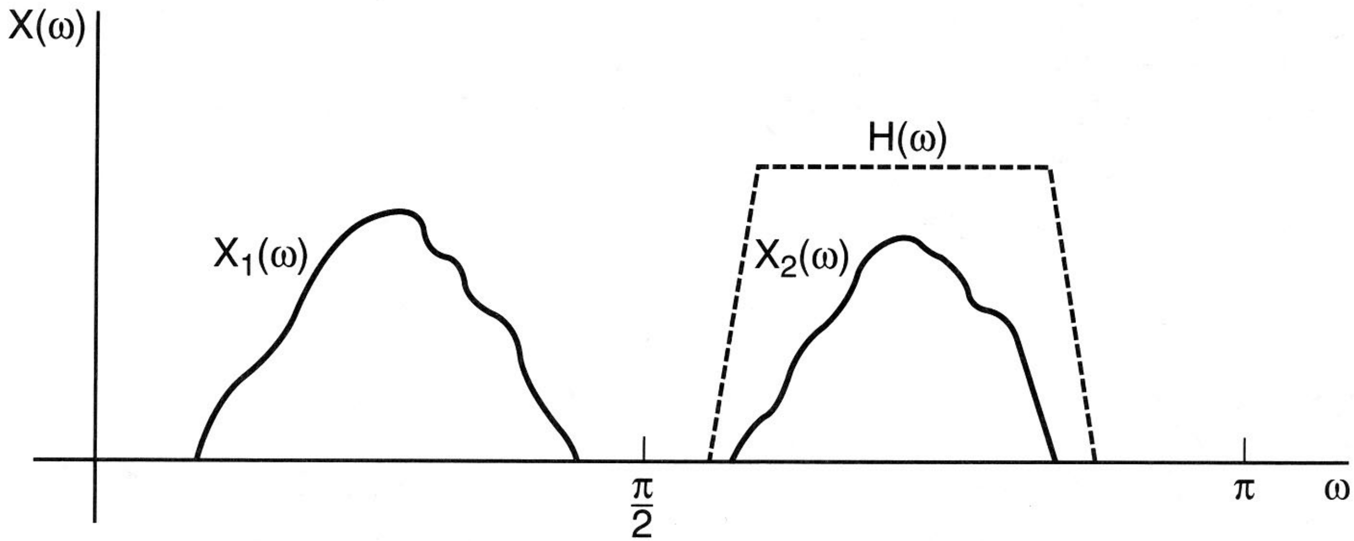


Figure 6.1 Signal with disjoint low- and high-frequency spectra $X_1(\omega)$ and $X_2(\omega)$.

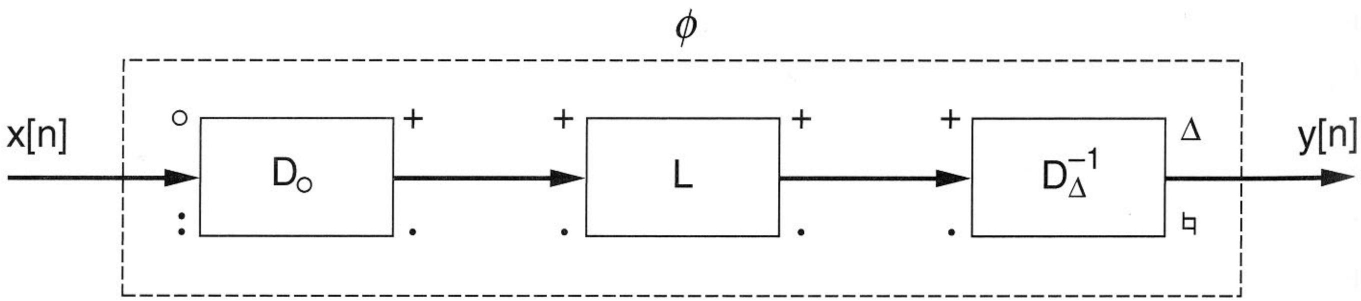


Figure 6.3 Canonical formulation of a homomorphic system.

SOURCE: A.V. Oppenheim and R.W. Schafer, *Discrete-Time Signal Processing* [13]. ©1989, Pearson Education, Inc. Used by permission.

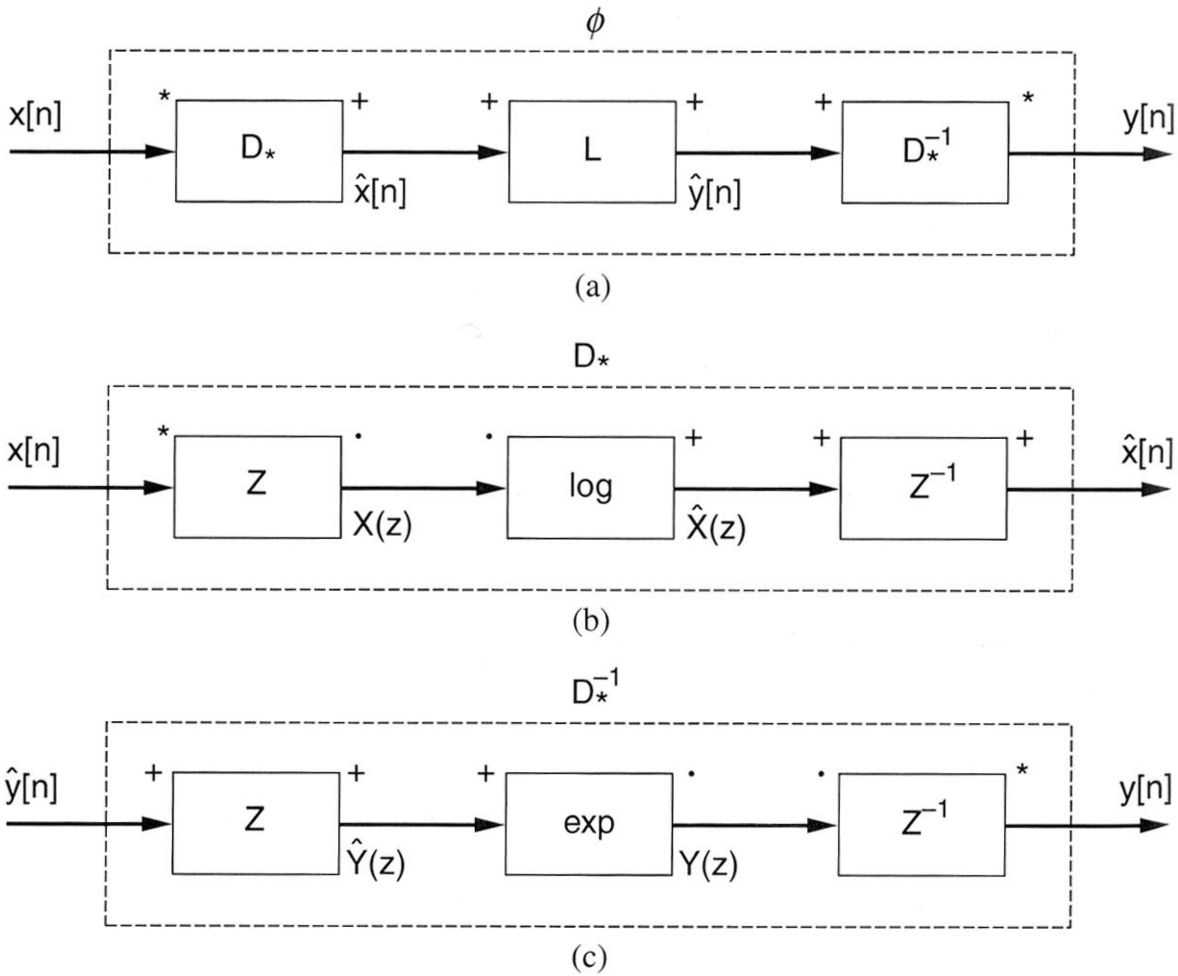


Figure 6.4 Homomorphic system for convolution: (a) canonical formulation; (b) the subsystem D_*; and (c) its inverse.

SOURCE: A.V. Oppenheim and R.W. Schafer, *Discrete-Time Signal Processing* [13].

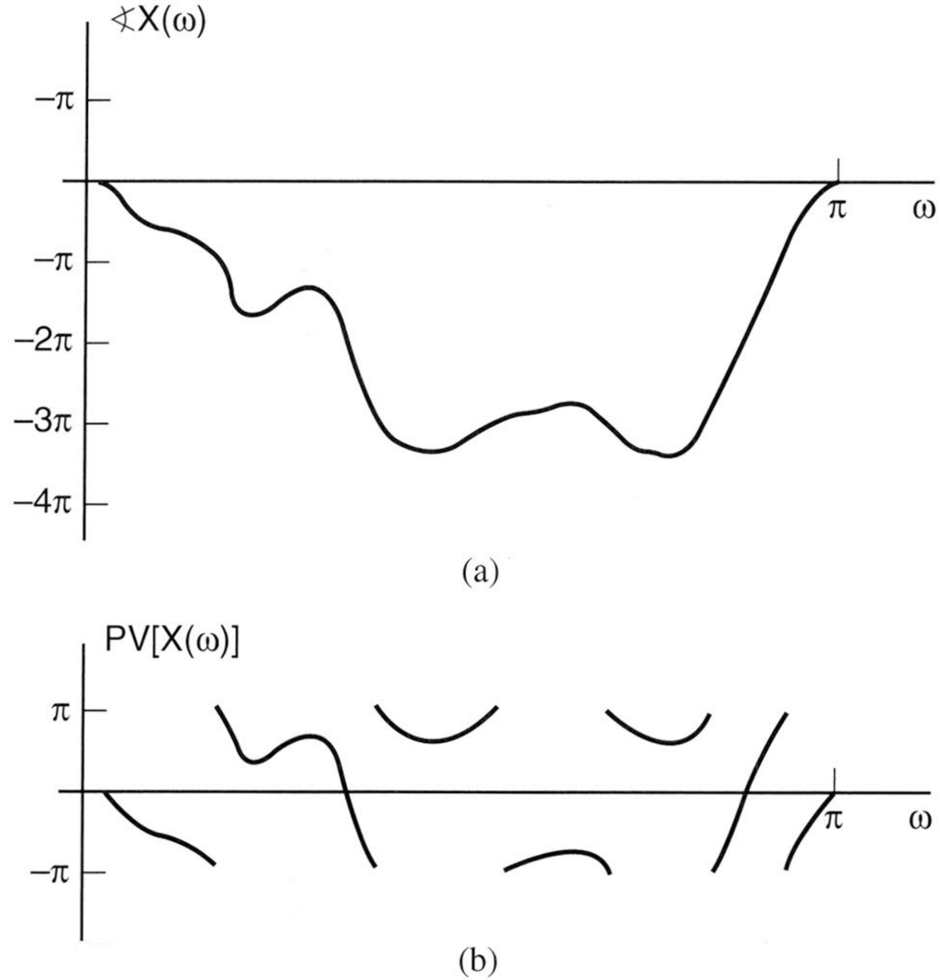


Figure 6.5 Fourier transform phase continuity: (a) typical continuous phase function; (b) its principal value.

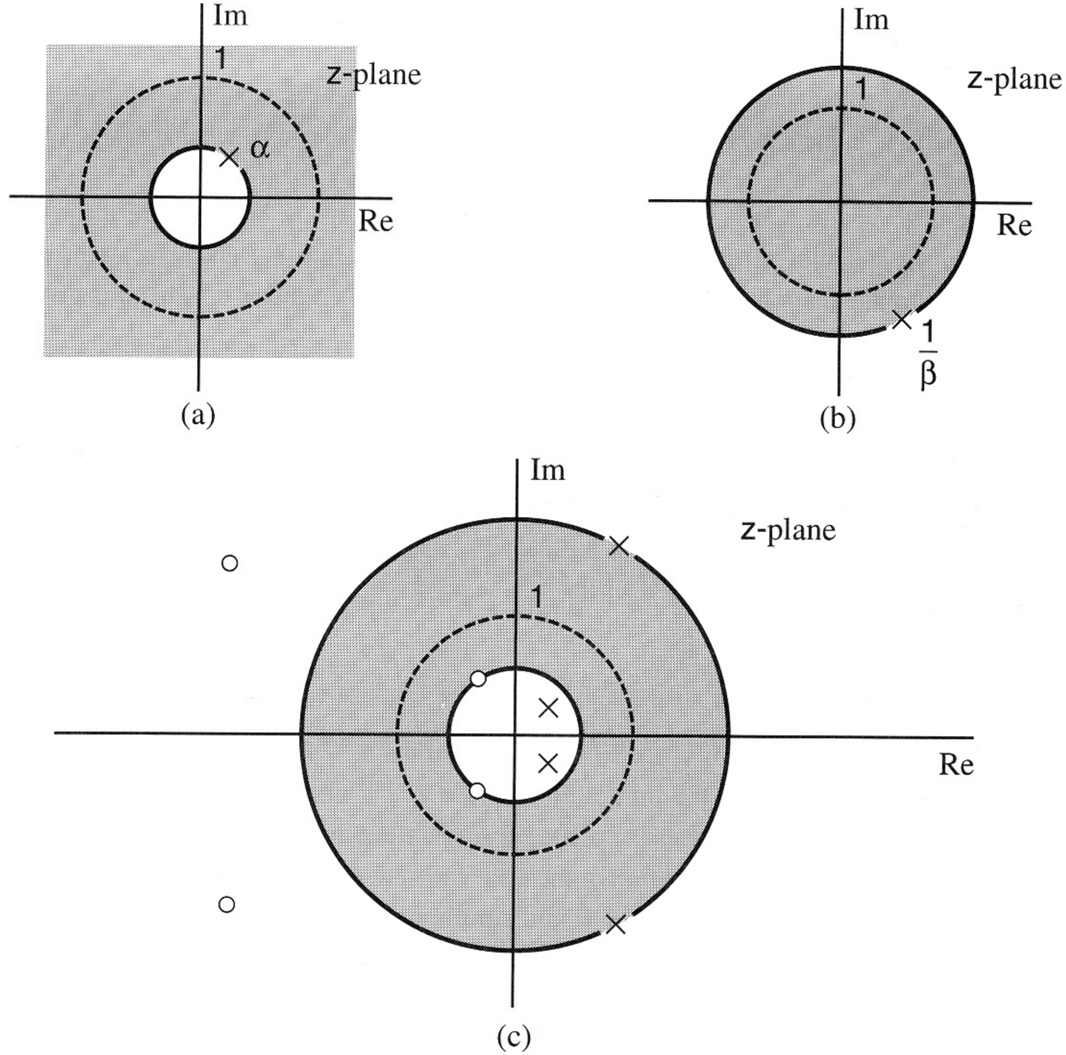


Figure 6.6 Region of convergence (ROC) for (a) $\log(1-\alpha z^{-1})$ and (b) $\log(1-\beta z)$, while (c) shows the annular ROC for a typical rational $X(z)$. For all cases, the ROC includes the unit circle; the unit circle is shown as a dashed line.

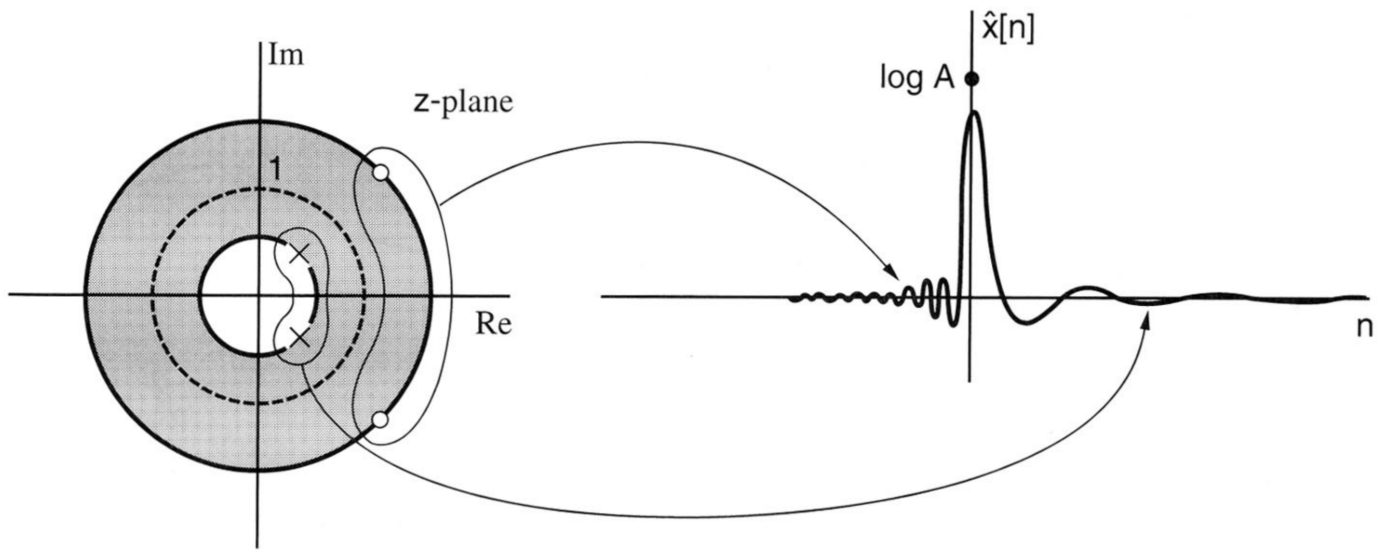


Figure 6.7 Schematized example illustrating right- and left-side contributions to the complex cepstrum.

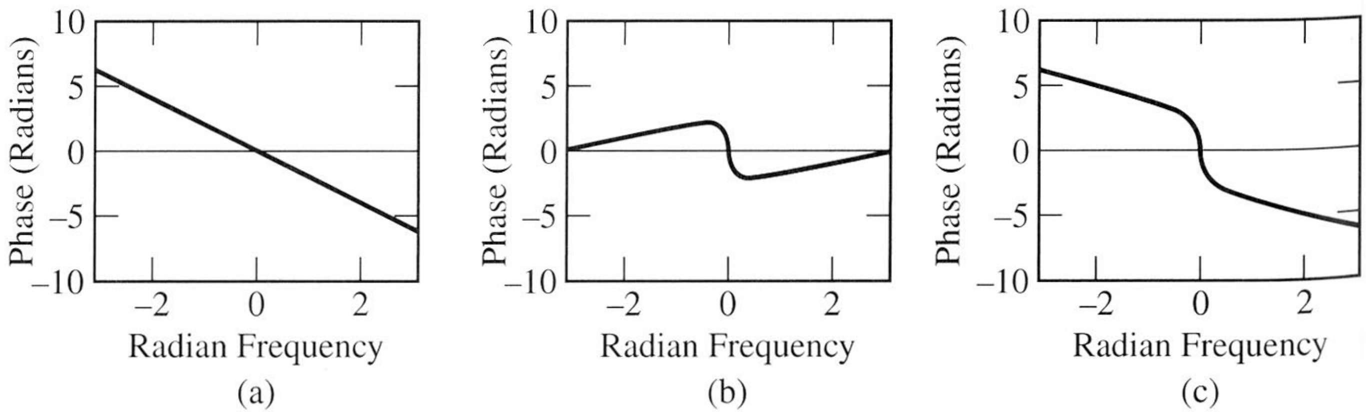


Figure 6.8 Illustration of phase contributions of Example 6.5: (a) unwrapped phase of the linear phase contribution; (b) unwrapped phase of the pole component; (c) sum of (a) and (b). The value of r in z^{-r} is negative (-2) and the pole occurs on the real z axis (0 Hz).

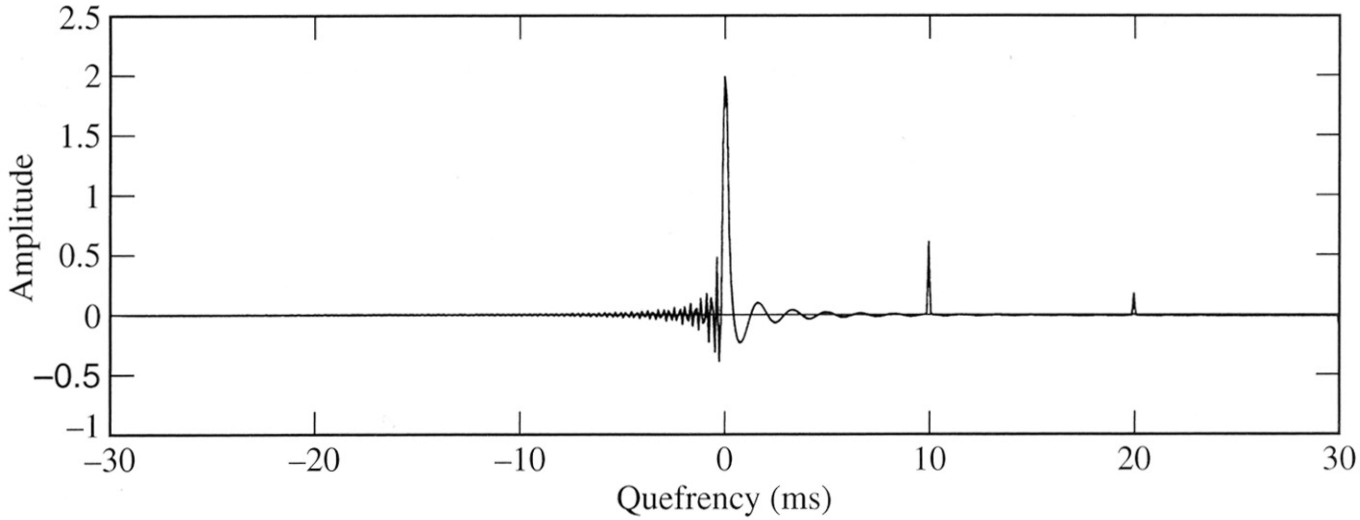


Figure 6.9 Complex cepstrum of $x[n] = p[n] * h[n]$ of Example 6.6. The sequence $p[n]$ is minimum-phase and $h[n]$ is mixed phase (zeros inside and poles outside the unit circle).

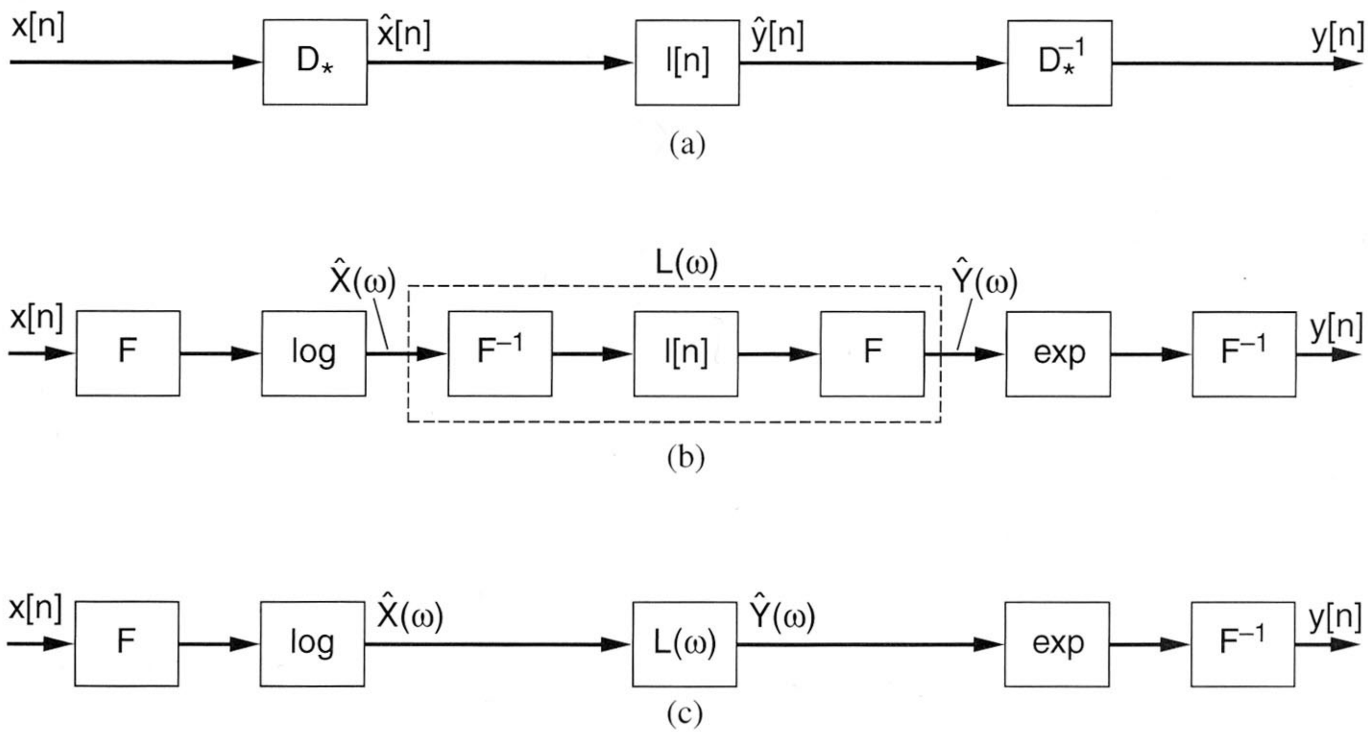


Figure 6.10 Homomorphic filtering interpreted as a linear smoothing of $\log[X(\omega)]$: (a) quefrency-domain implementation; (b) expansion of the operations in (a); (c) frequency-domain interpretation.

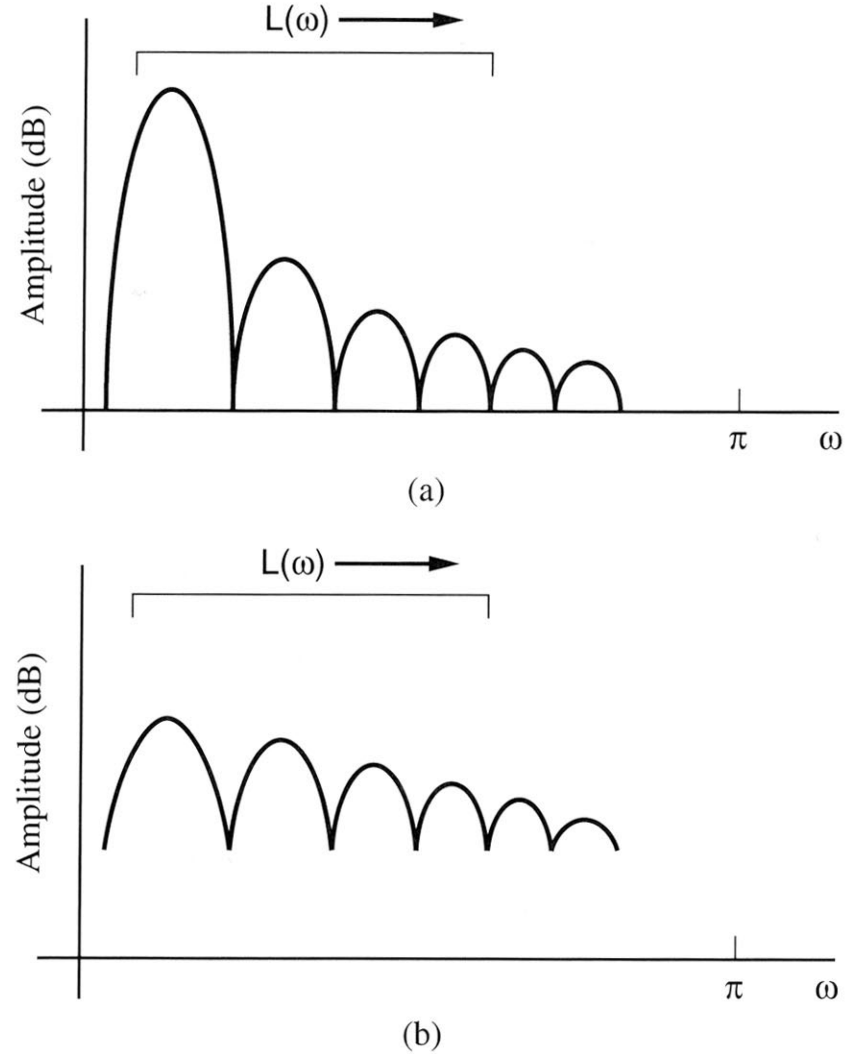
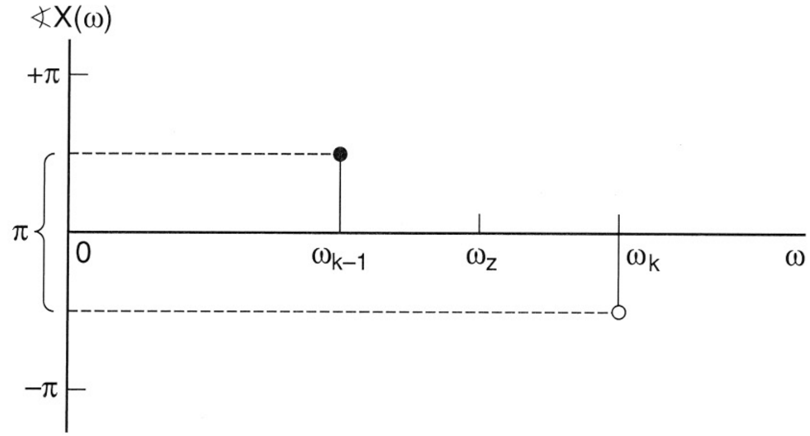
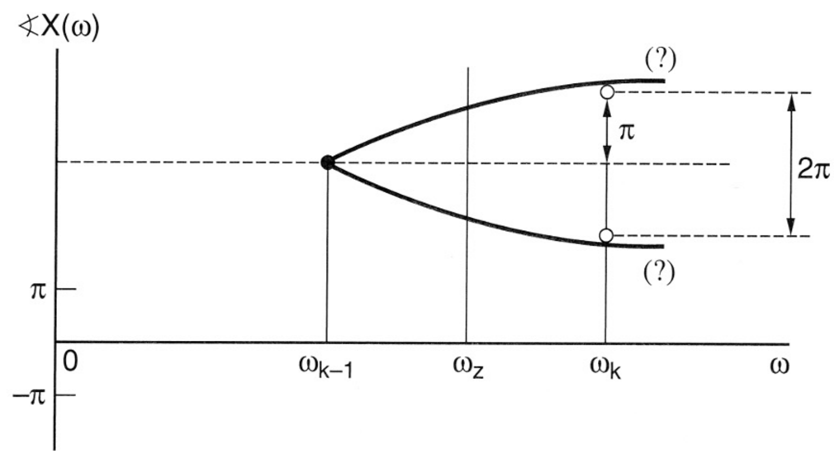


Figure 6.11 Schematic of smoothing (a) a harmonic spectrum in contrast to (b) the logarithm of a harmonic spectrum.



(a)



(b)

Figure 6.12 Phase unwrapping ambiguity: (a) natural unwrapped phase change of π across two DFT frequencies, with zero close to unit circle; (b) two possible phase values in performing phase unwrapping.

SOURCE: J.M. Tribolet, "A New Phase Unwrapping Algorithm" [19]. ©1977, IEEE. Used by permission.

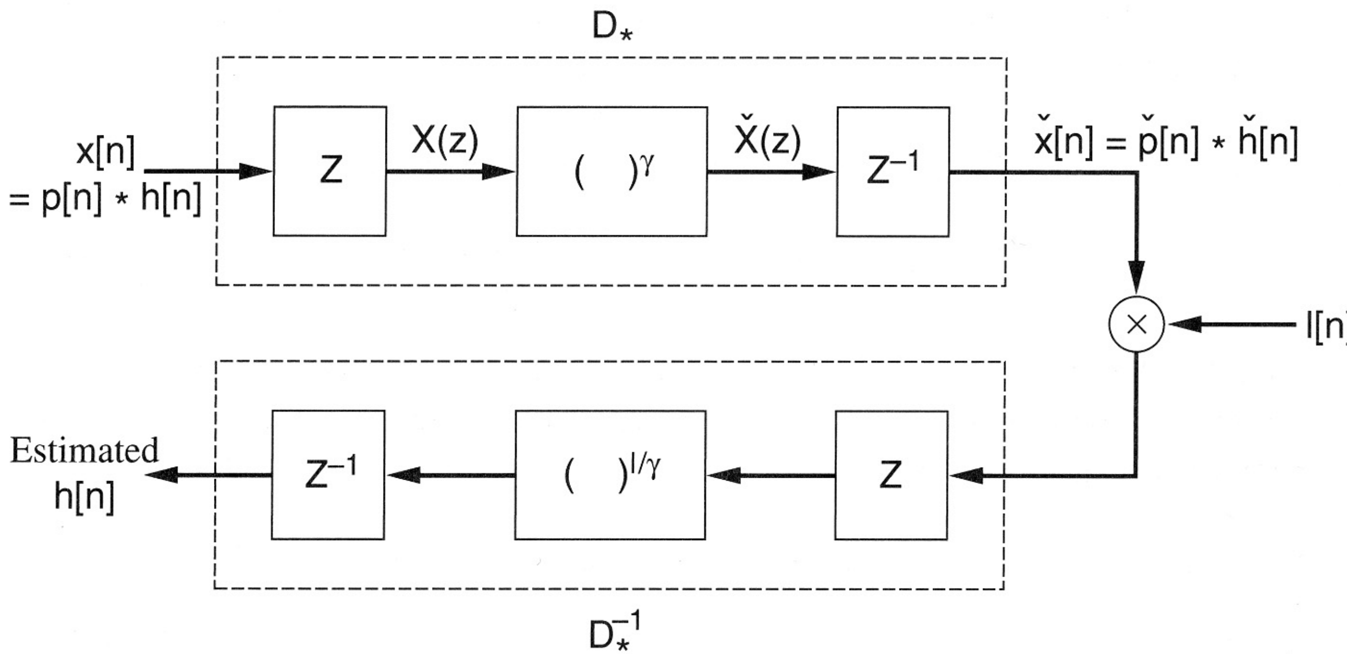


Figure 6.13 Spectral root homomorphic filtering.

SOURCE: J.S. Lim, "Spectral Root Homomorphic Deconvolution System" [5]. ©1979, IEEE. Used by permission.

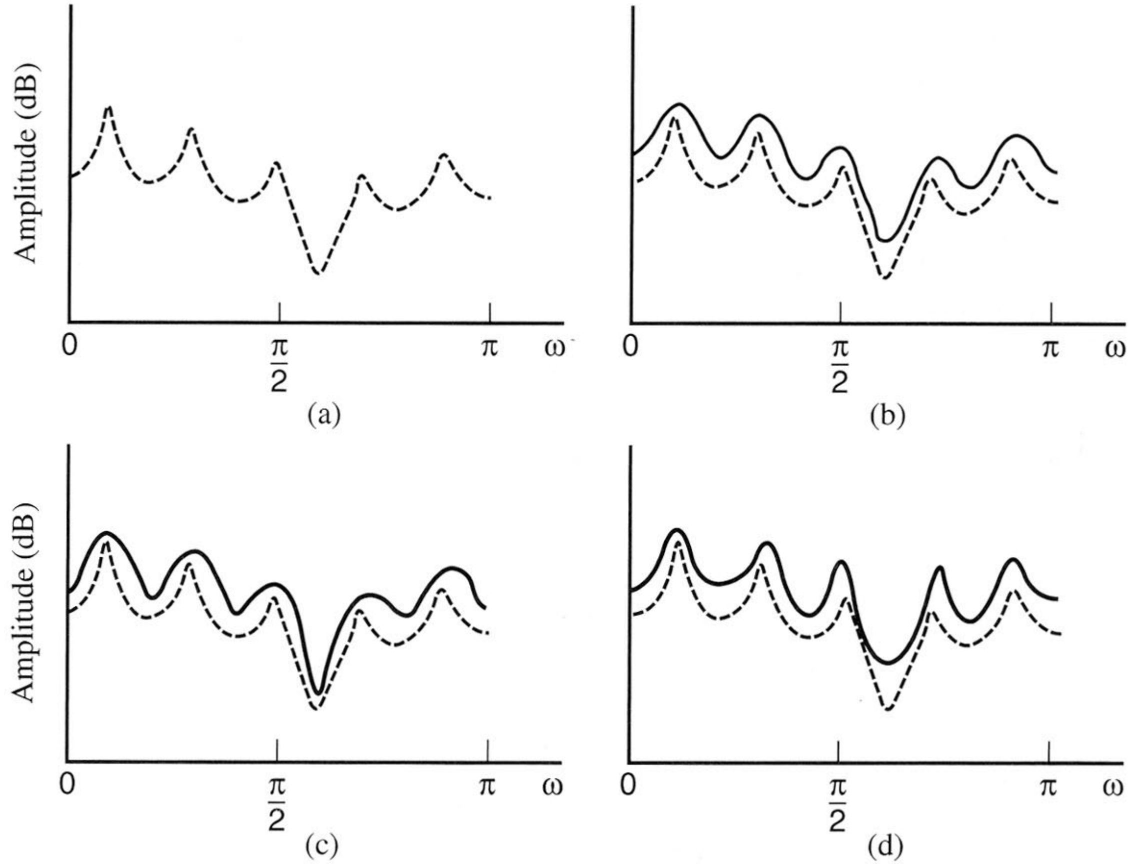


Figure 6.14 Example of spectral root homomorphic filtering on synthetic vocal tract impulse response: (a) log-magnitude spectrum of impulse response; (b) estimate of log-magnitude spectrum of $h[n]$ derived from low-time gating real cepstrum; (c) log-magnitude spectral estimate derived from low-time gating spectral root cepstrum with $\gamma = +0.5$; (d) same as (c) with $\gamma = -0.5$.

SOURCE: J.S. Lim, "Spectral Root Homomorphic Deconvolution System" [5]. ©1979, IEEE. Used by permission.

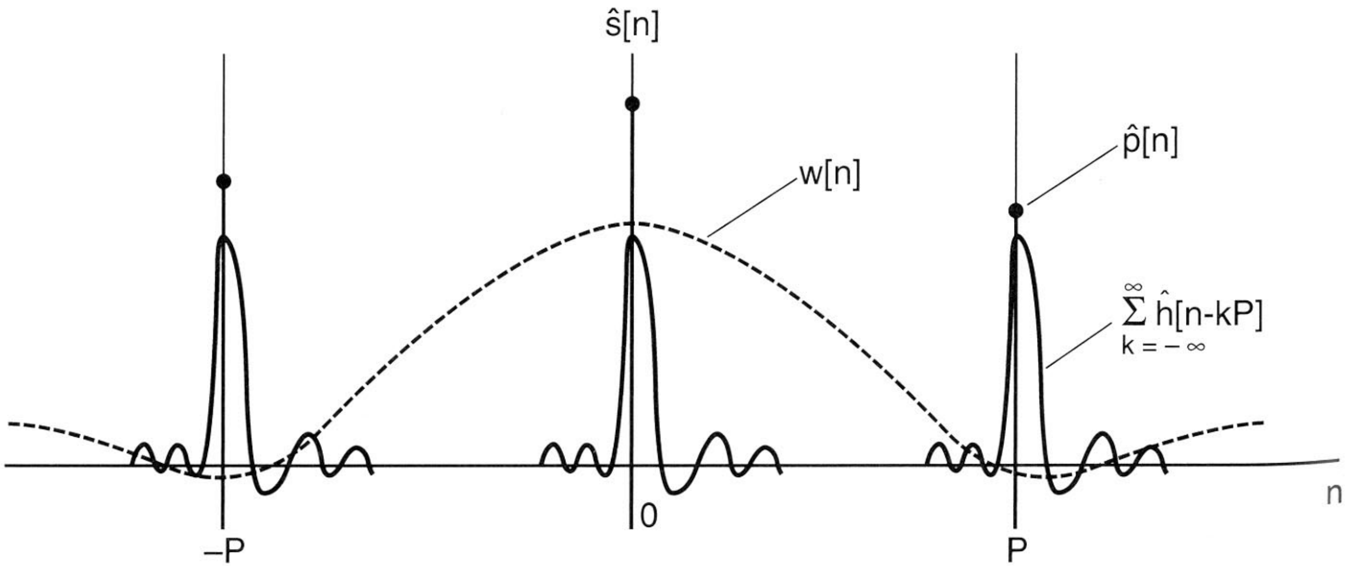


Figure 6.15 Schematic of complex cepstrum of windowed periodic sequence.

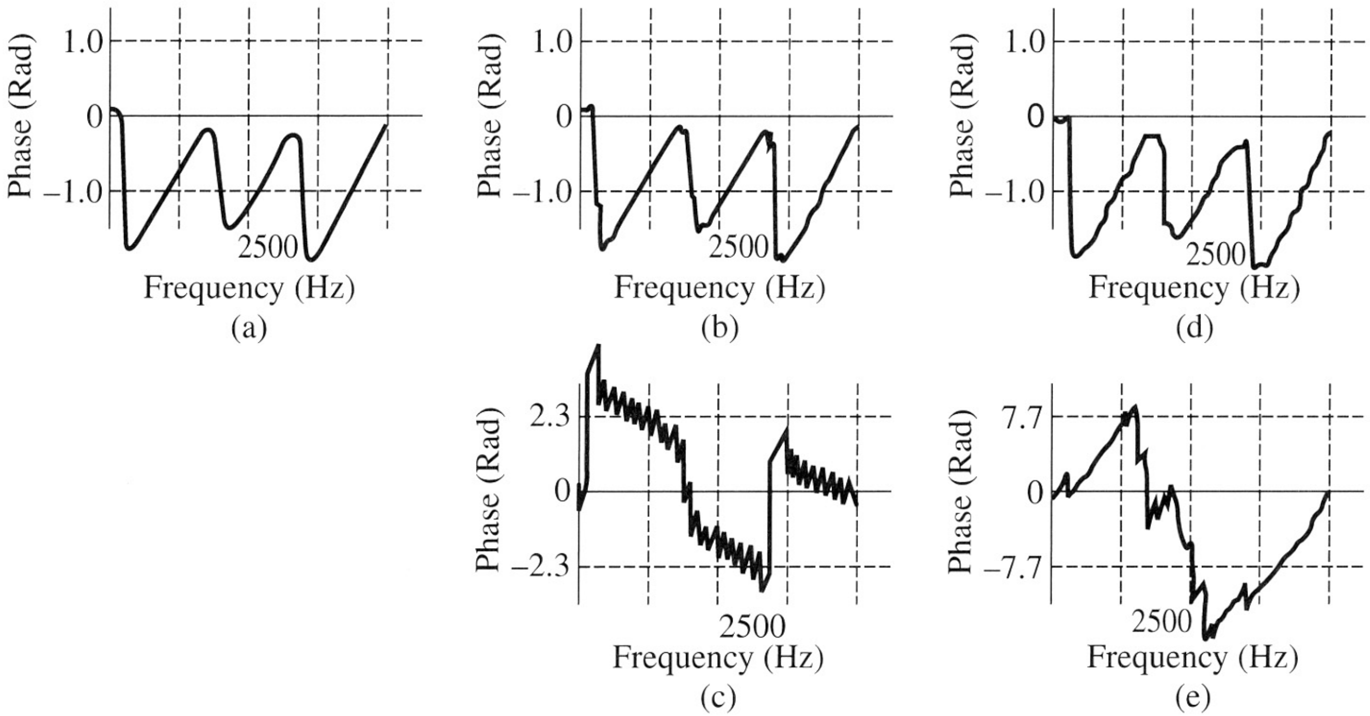


Figure 6.16 Sensitivity of system phase estimate to the analysis window in Example 6.10: (a) unwrapped phase of artificial vocal tract impulse response; (b) unwrapped phase of a periodic waveform with window $w[n] = [\frac{\sin(\pi n/P)}{\pi n)}]^2$ and time-aligned; (c) same as (b) with window displacement; (d) same as (b) with Hamming window 2 pitch periods in length; (e) same as (d) with Hamming window 3.9 pitch periods in length.

SOURCE: T.F. Quatieri, "Minimum- and Mixed-Phase Speech Analysis/Synthesis by Adaptive Homomorphic Deconvolution" [14]. ©1979, IEEE. Used by permission.

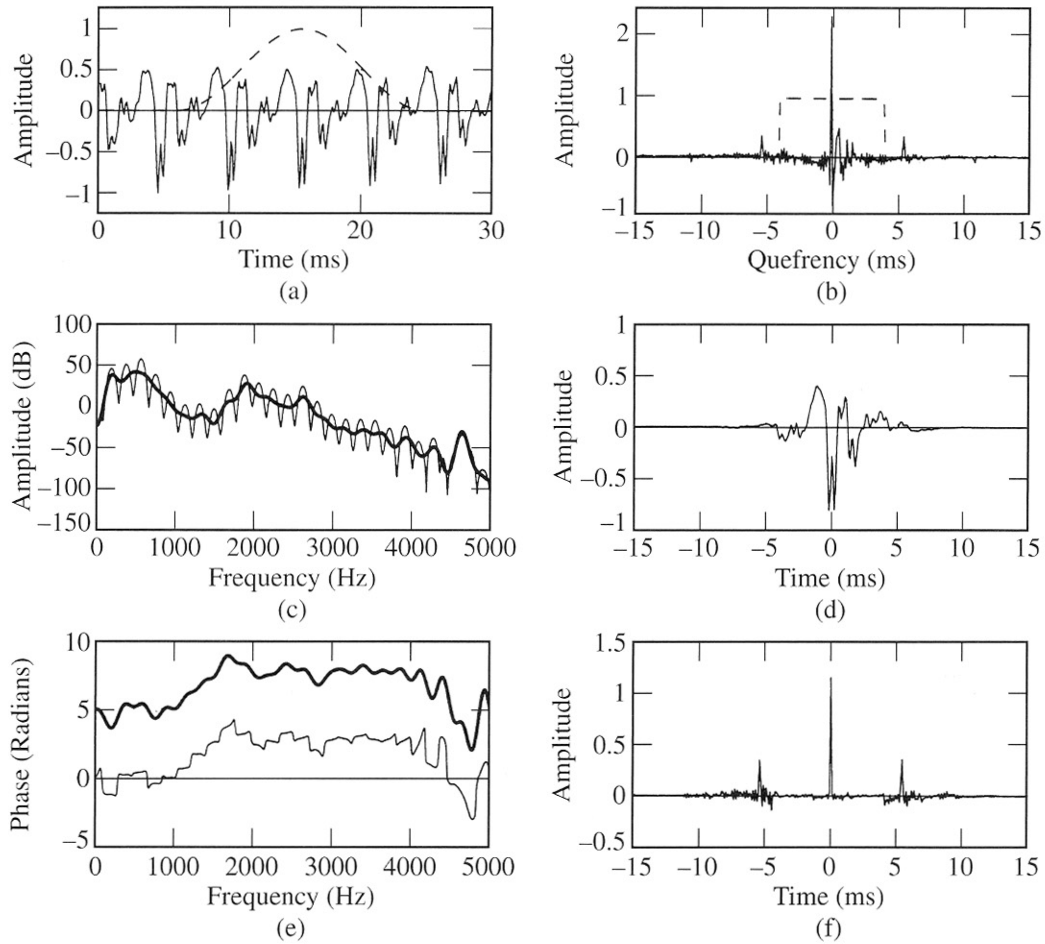


Figure 6.17 Homomorphic filtering of voiced waveform from female speaker: (a) waveform (solid) and aligned analysis window (dashed); (b) complex cepstrum of windowed speech signal $s[n]$ (solid) and low-quefrency lifter (dashed); (c) log-magnitude spectrum of $s[n]$ (thin solid) and of the impulse response estimate (thick solid); (d) impulse response estimate from low-quefrency liftering; (e) spectral unwrapped phase of $s[n]$ (thin solid) and of the impulse response estimate (thick solid) (The smooth estimate is displaced for clarity.); (f) estimate of windowed impulse train source from high-quefrency liftering.

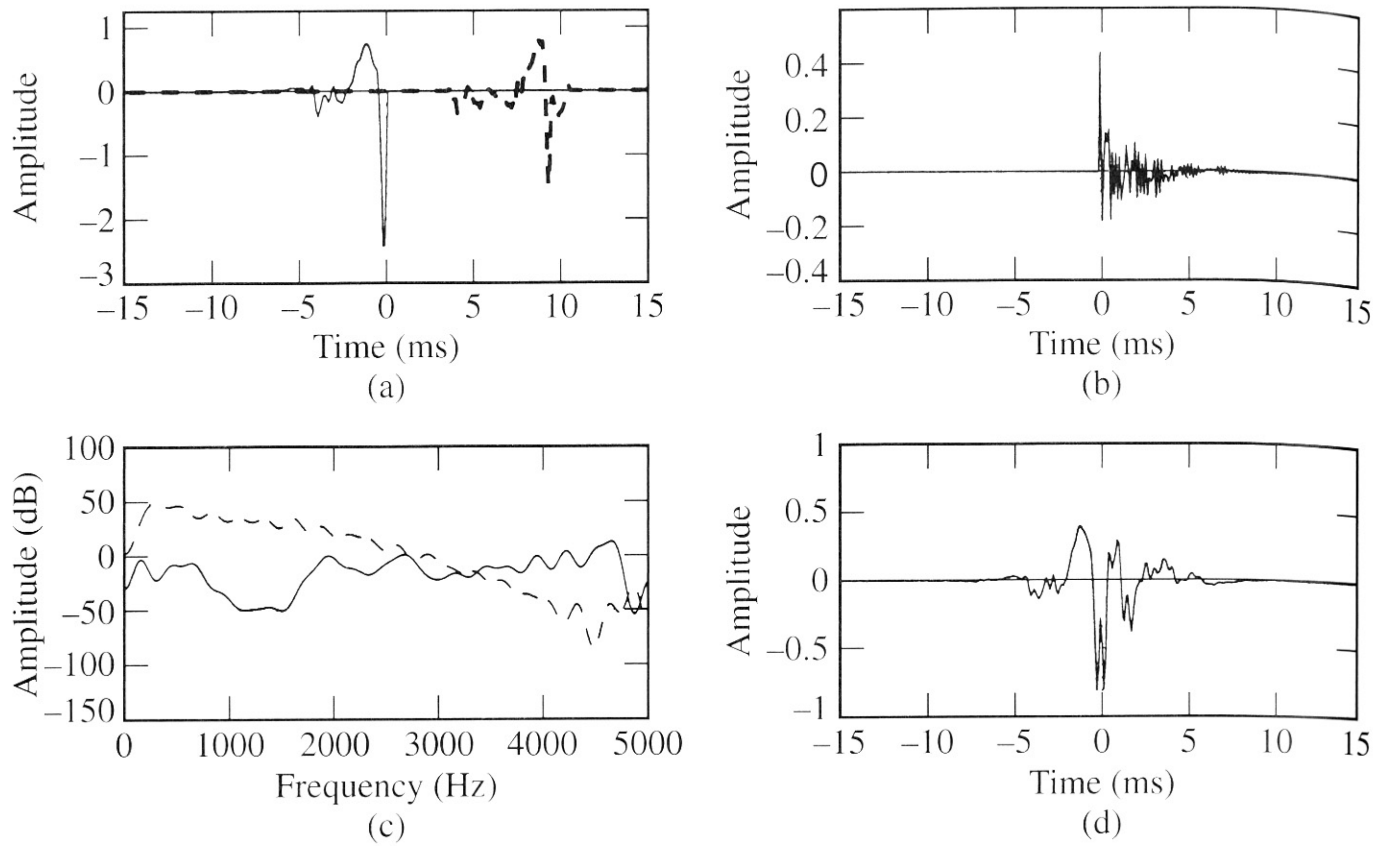


Figure 6.18 Deconvolved maximum-phase component (a) (solid) and minimum-phase component (b) in Example 6.11. The convolution of the two components (d) is identical to the sequence of Figure 6.17d. Panel (c) shows the log-magnitude spectra of the maximum-phase (dashed) and minimum-phase (solid) component. The maximum-phase component in panel (a) and its log-magnitude spectrum in panel (c) resemble those of a typical glottal flow derivative sequence. For reference, the dashed sequence in panel (a) is the glottal flow derivative derived from the pole/zero-estimation method of Section 5.7.2.

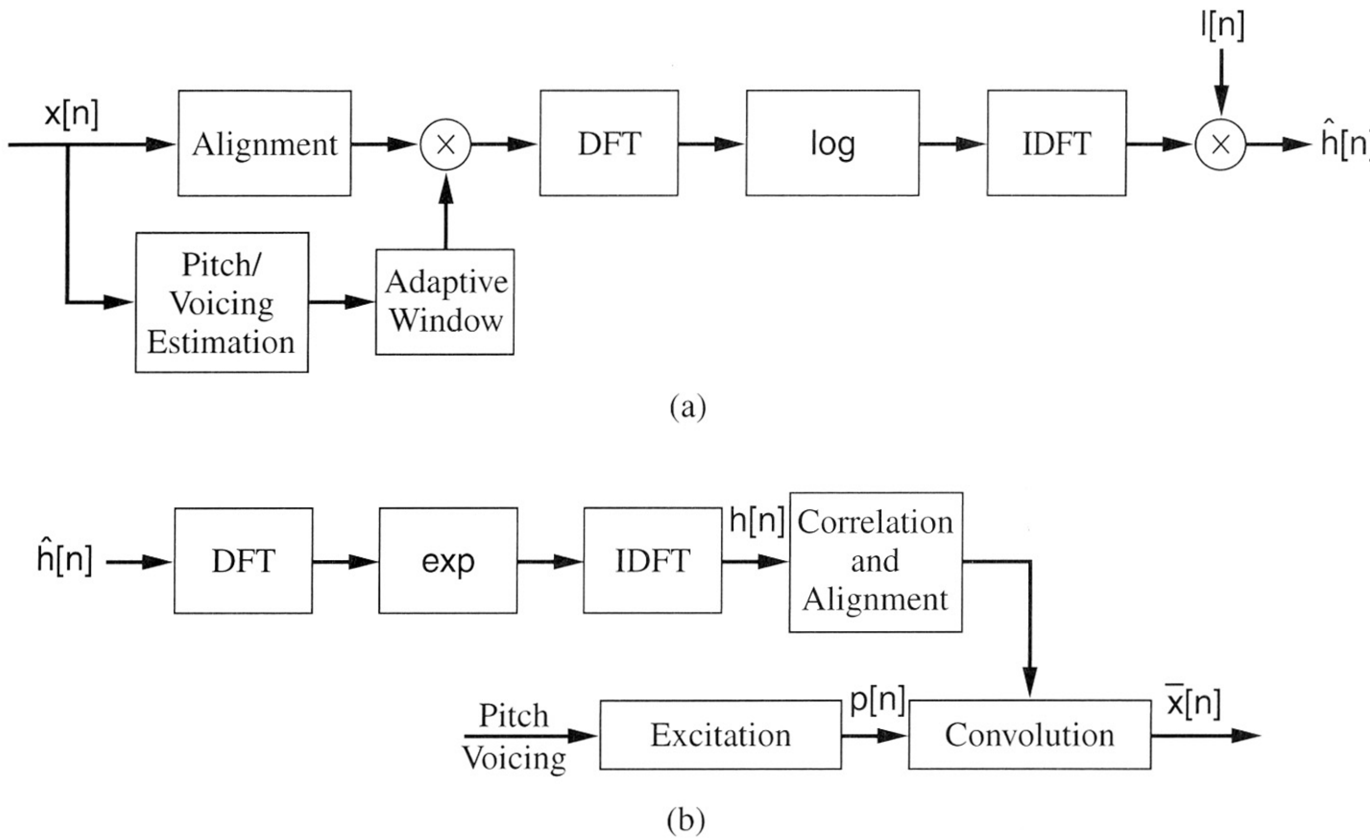


Figure 6.19 General framework for homomorphic analysis/synthesis: (a) analysis; (b) synthesis. The alignment and correlation operations are used for mixed-phase reconstruction.

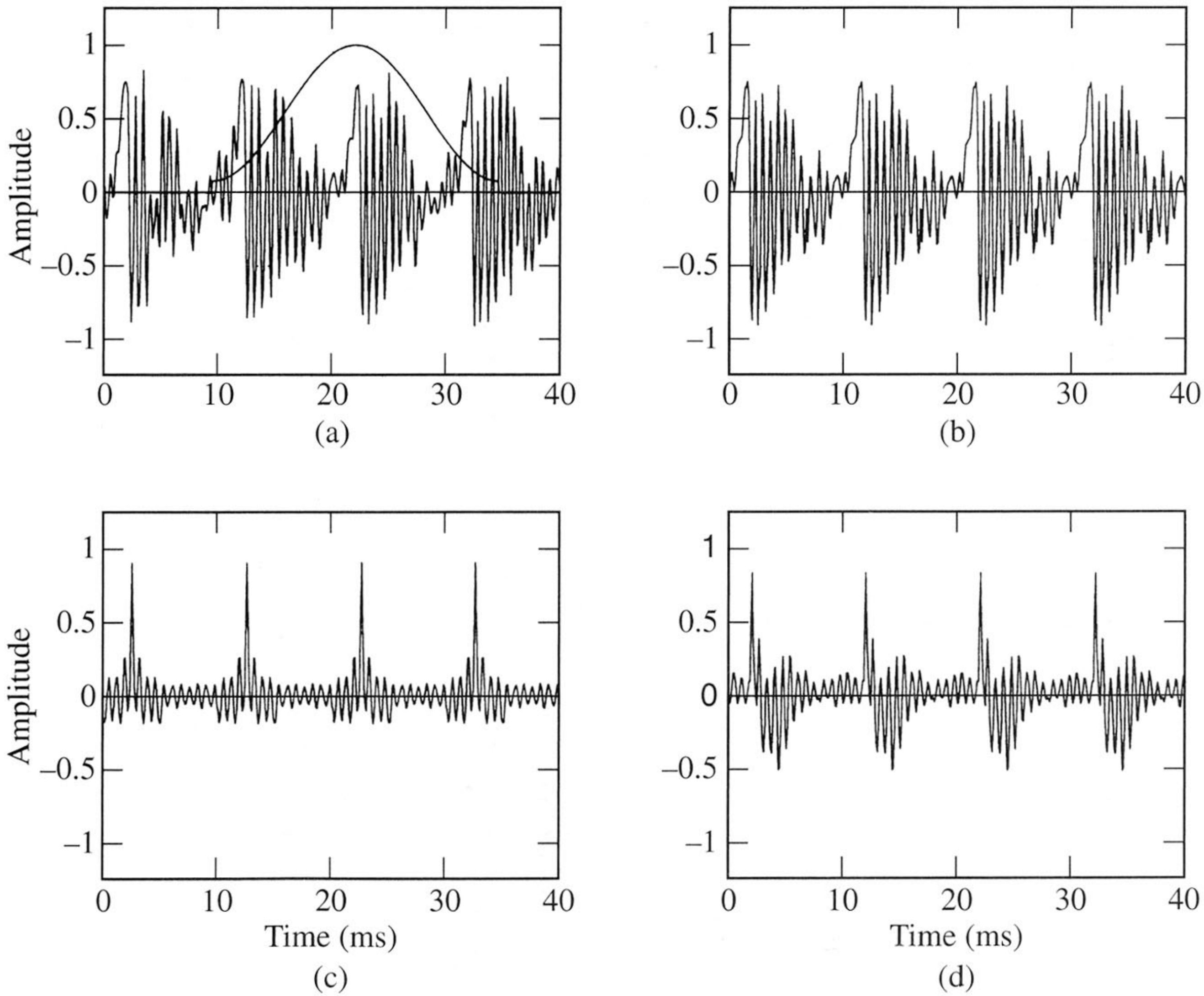


Figure 6.20 Homomorphic synthesis based on (b) mixed-phase, (c) zero-phase, and (d) minimum-phase analysis and synthesis. Panel (a) is the original.

	Linear Prediction	Homomorphic Filtering
Parametric vs non-parametric	Sharp, smooth resonances	Wider, spurious resonances
Rational function representation	All-pole model	Pole-zero model
Phase properties	Minimum-phase	Mixed-phase
Speech synthesis	“Crisper” but more “mechanical”	“Natural” but “muffled”
Effects of pitch	Aliasing problems	Aliasing problems
Effects of windowing	Distortion from values outside the window	Distorts convolutional model
Model order	Number of poles	Length of lifter

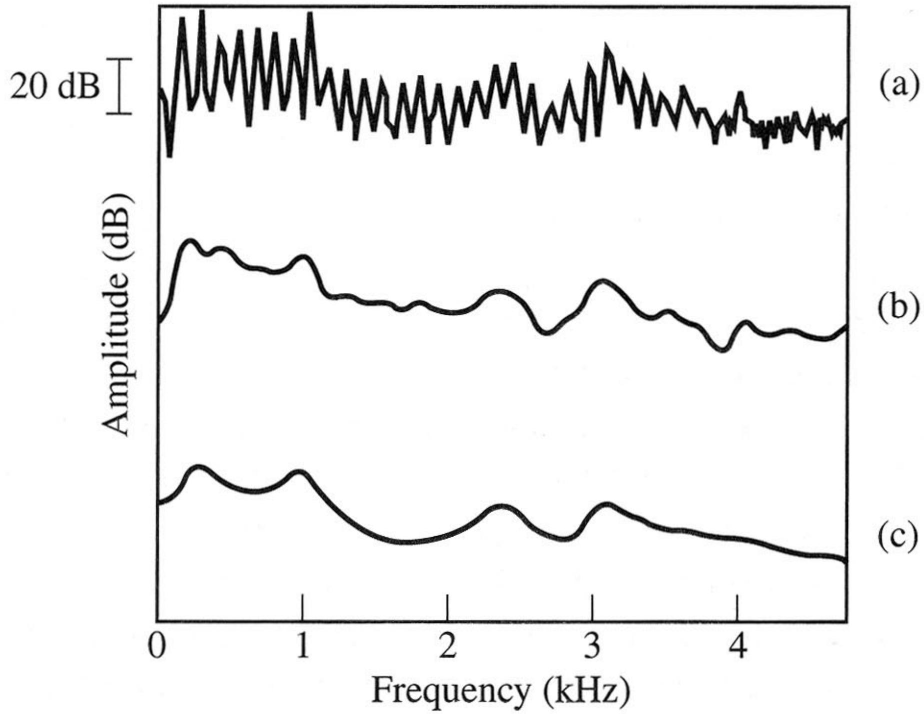


Figure 6.21 Homomorphic prediction applied to a nasalized vowel from /u/ in "m<u>oo</u>n": (a) log-magnitude spectrum of speech signal; (b) log-magnitude spectrum obtained by homomorphic filtering (low-time liftering the real cepstrum); (c) log-magnitude spectrum of 10-pole/6-zero model with zeros from Shanks method.

SOURCE: G.E. Kopec, A.V. Oppenheim, and J.M. Tribolet, "Speech Analysis by Homomorphic Prediction" [4].