# ECE 797: Speech and Audio Processing

Hand-out for Lecture #7

Thursday, March 4, 2004

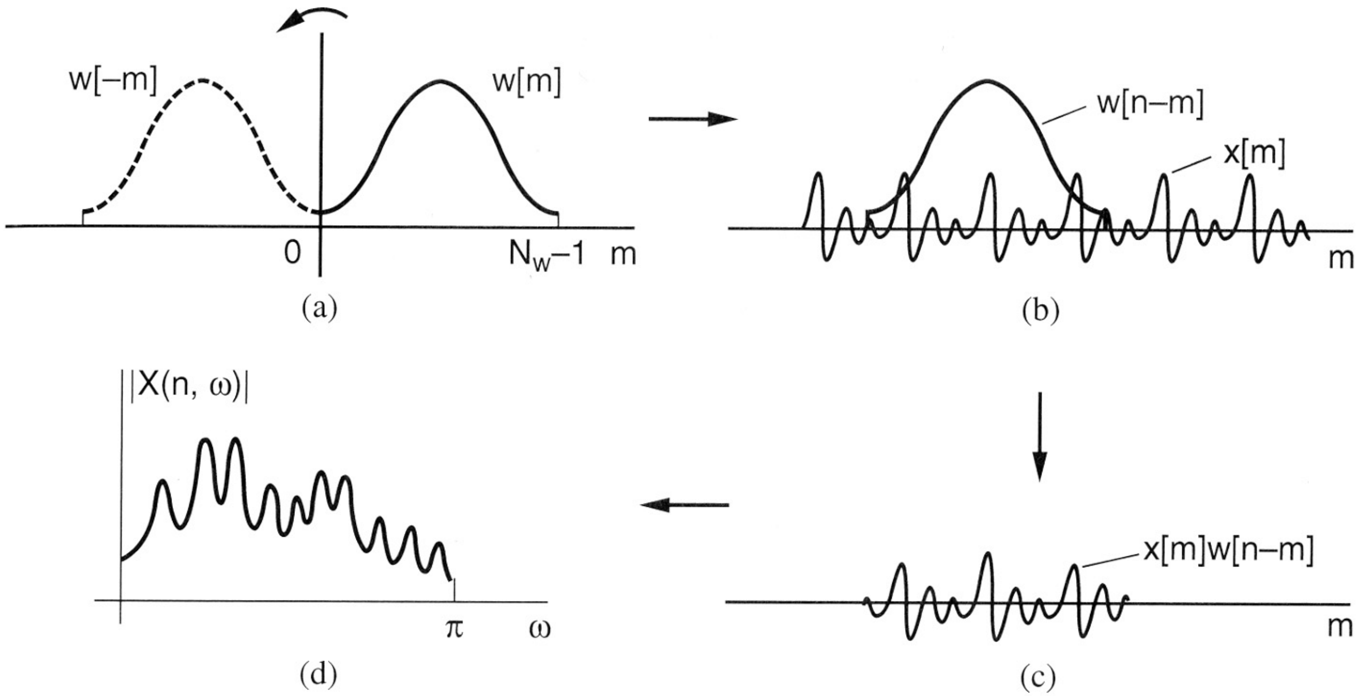


**Figure 7.1** Series of operations required to compute a short-time section and STFT: (a) flip window; (b) slide the window sample-by-sample over the sequence (*Note:* $w[-(m-n)] = w[n-m]$); (c) multiply the sequence by the displaced window; (d) take the Fourier transform of the windowed segment.

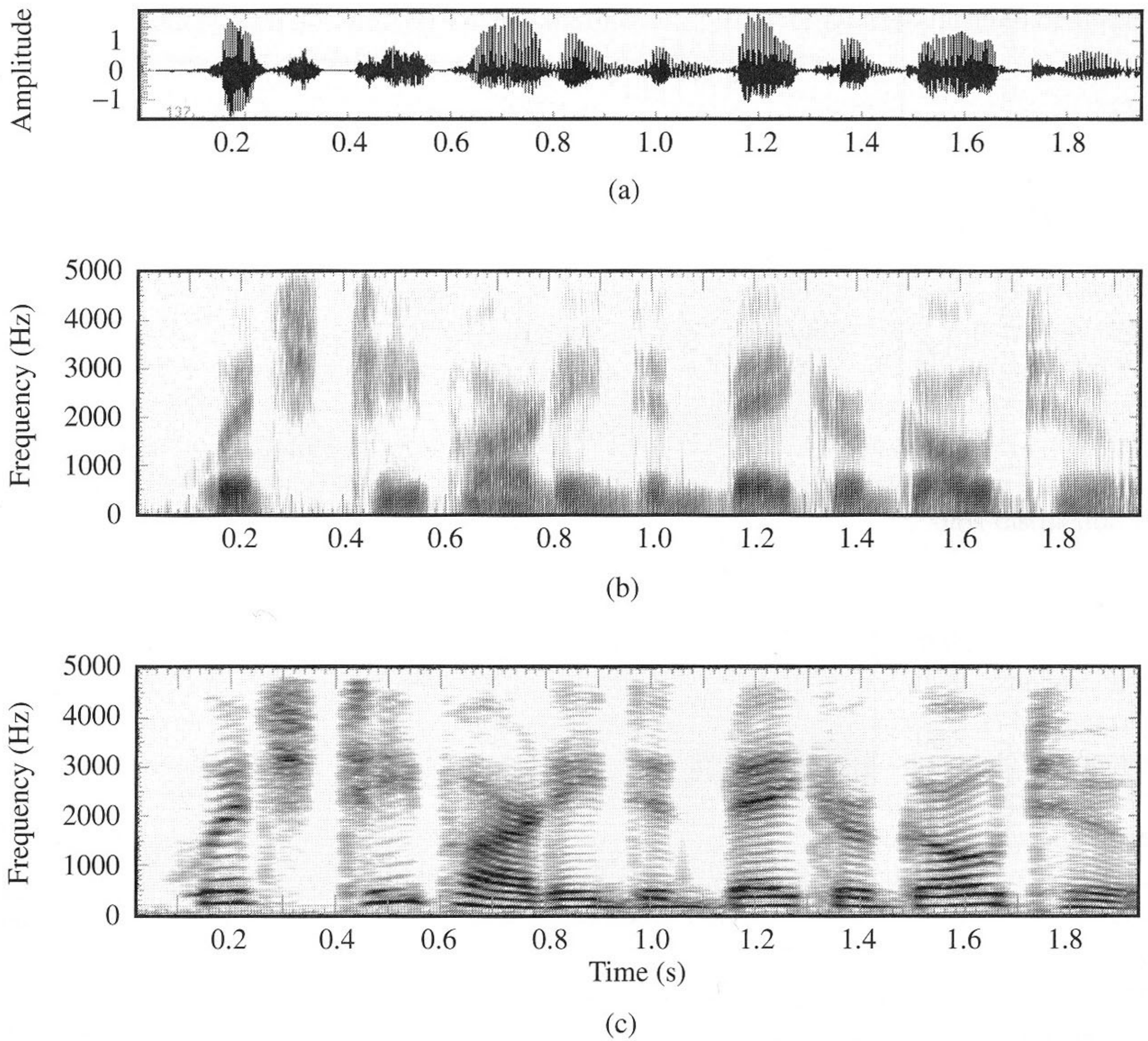


**Figure 3.15** Comparison of measured spectrograms for the utterance, "Which tea party did Baker go to?": (a) speech waveform; (b) wideband spectrogram; (c) narrowband spectrogram.

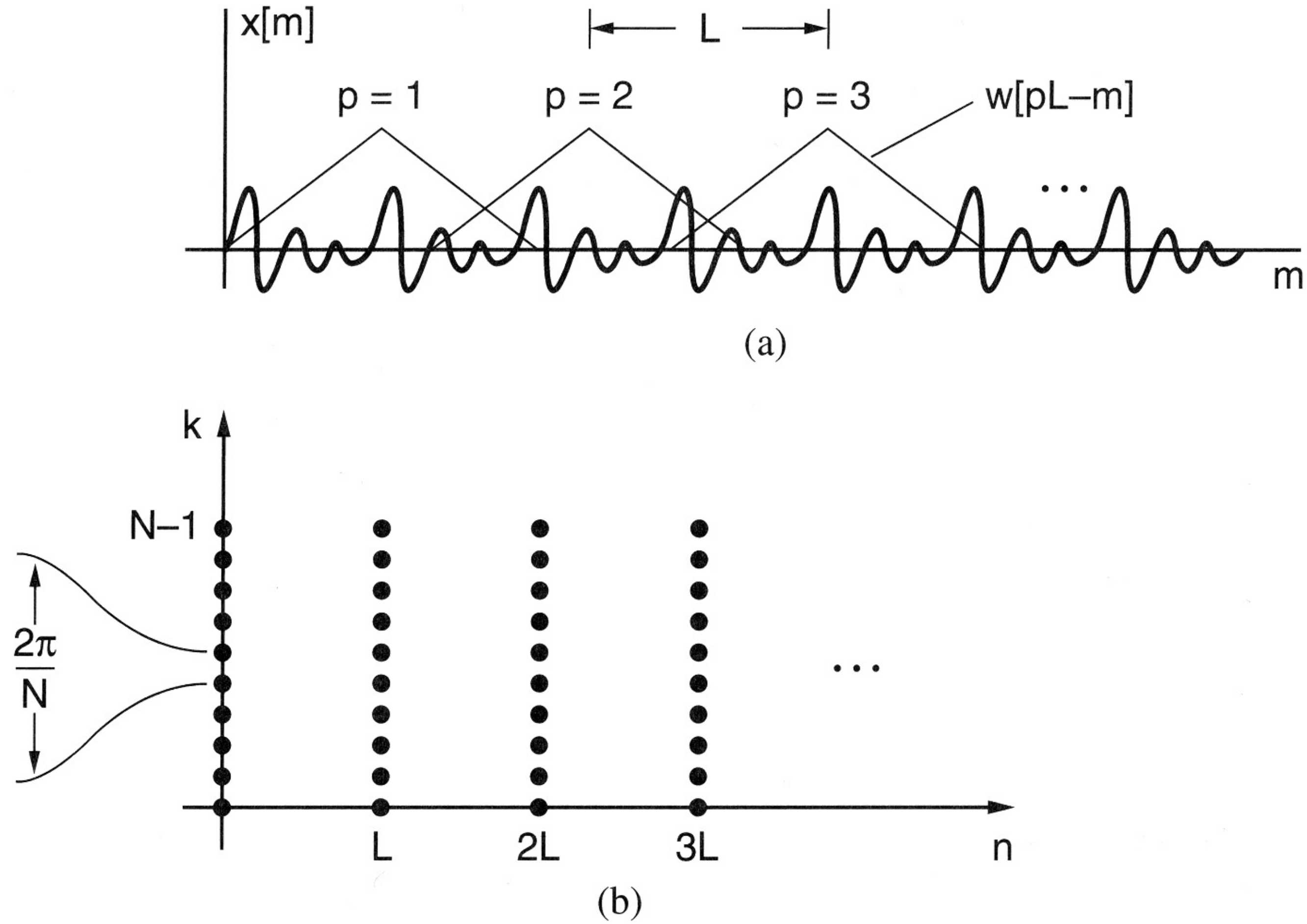


**Figure 7.2** Time and frequency decimation used in computing the discrete STFT $X(nL, k)$: (a) analysis window positions; (b) time-frequency sampling.

SOURCE: S.H. Nawab and T.F. Quatieri, "Short-Time Fourier Transform" [13]. ©1987, Pearson Education, Inc. Used by permission.

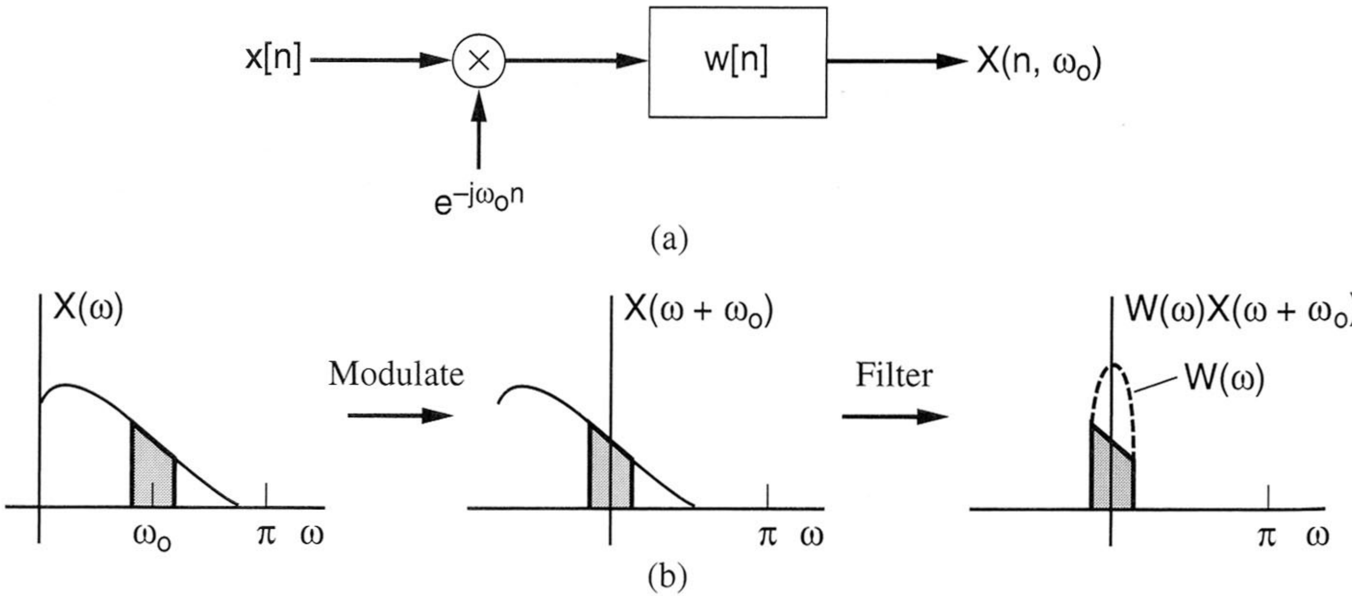


**Figure 7.3** Filtering view of STFT analysis at frequency $\omega_o$: (a) block diagram of complex exponential modulation followed by a lowpass filter; (b) operations in the frequency domain.

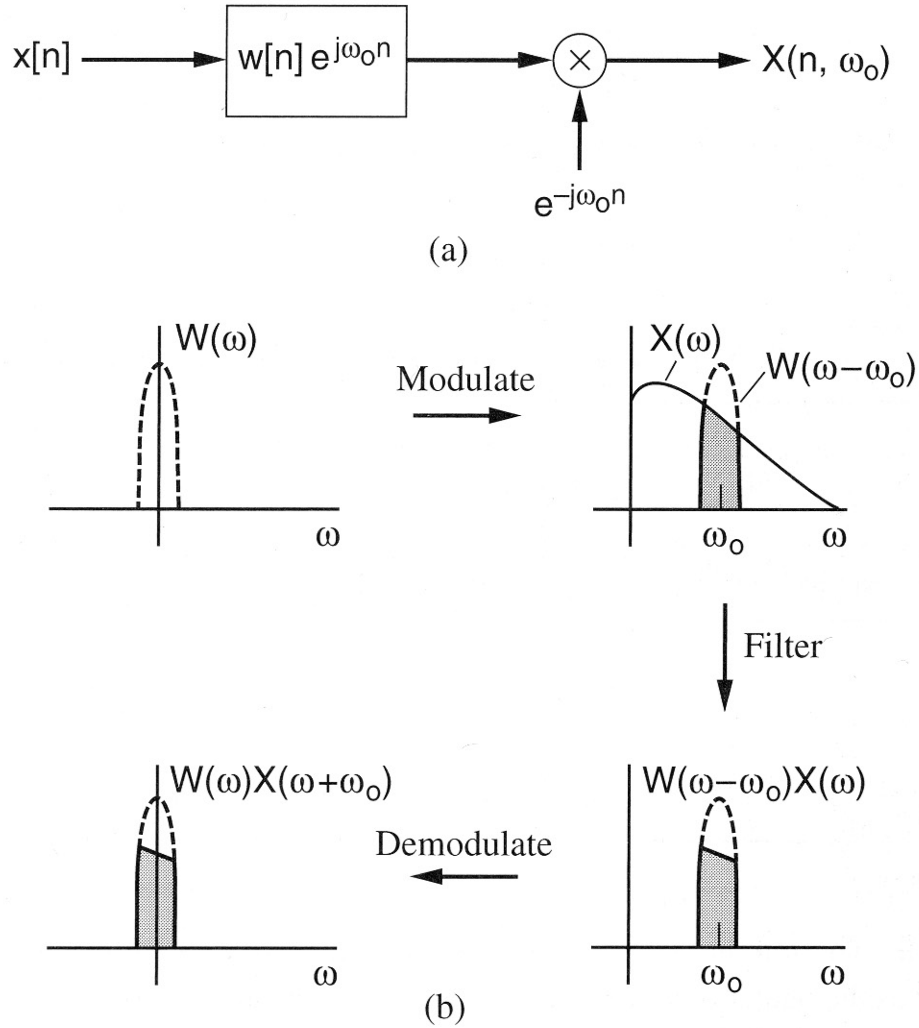


**Figure 7.4** Alternative filtering view of STFT analysis at frequency $\omega_o$: (a) block diagram of bandpass filtering followed by complex exponential modulation; (b) operations in the frequency domain.

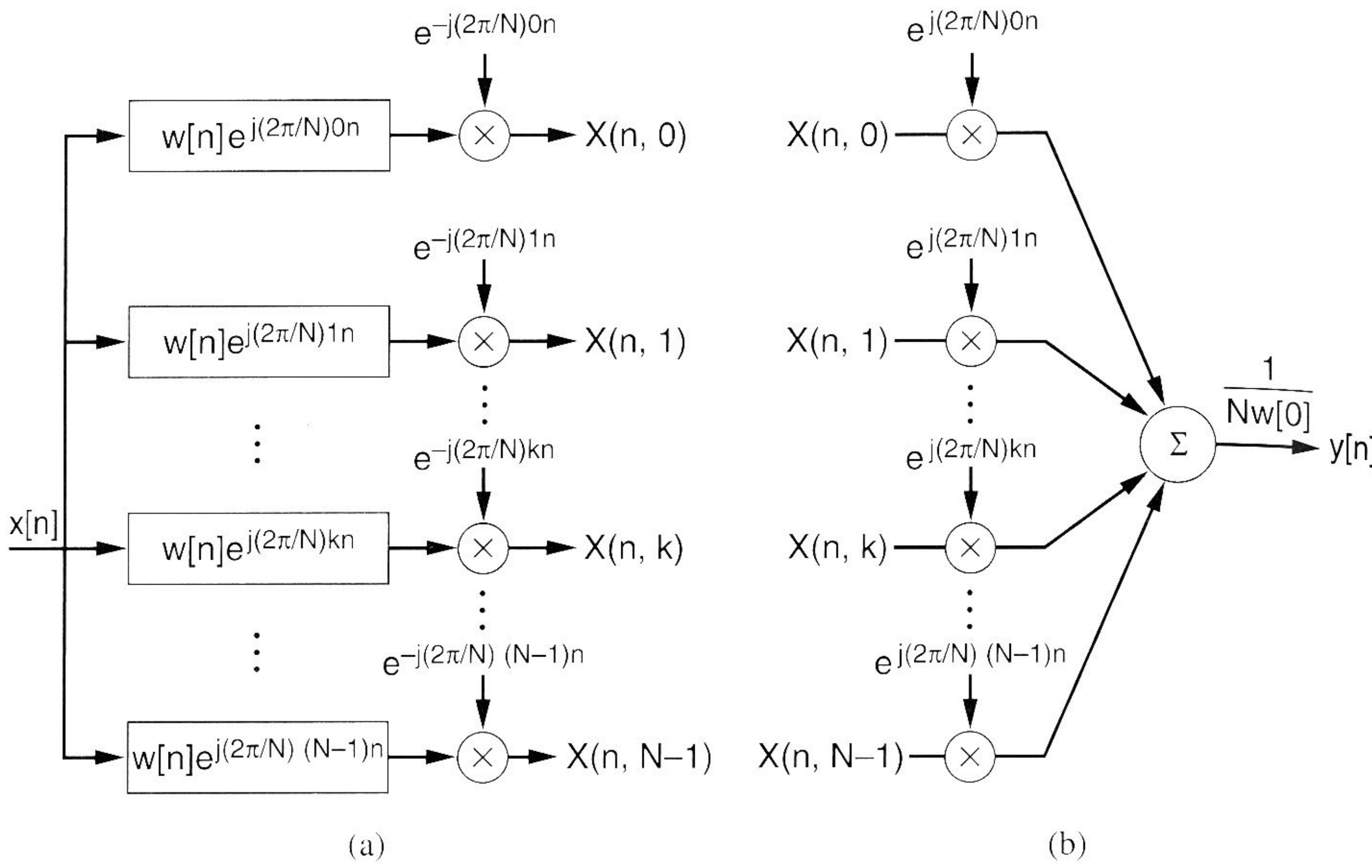


**Figure 7.5** The filtering view of analysis and synthesis with the discrete STFT: (a) the discrete STFT (analysis) as the output of a filter bank consisting of bandpass filters; (b) filter bank summation procedure for signal synthesis from the discrete STFT.

SOURCE: S.H. Nawab and T.F. Quatieri, "Short-Time Fourier Transform" [13]. 

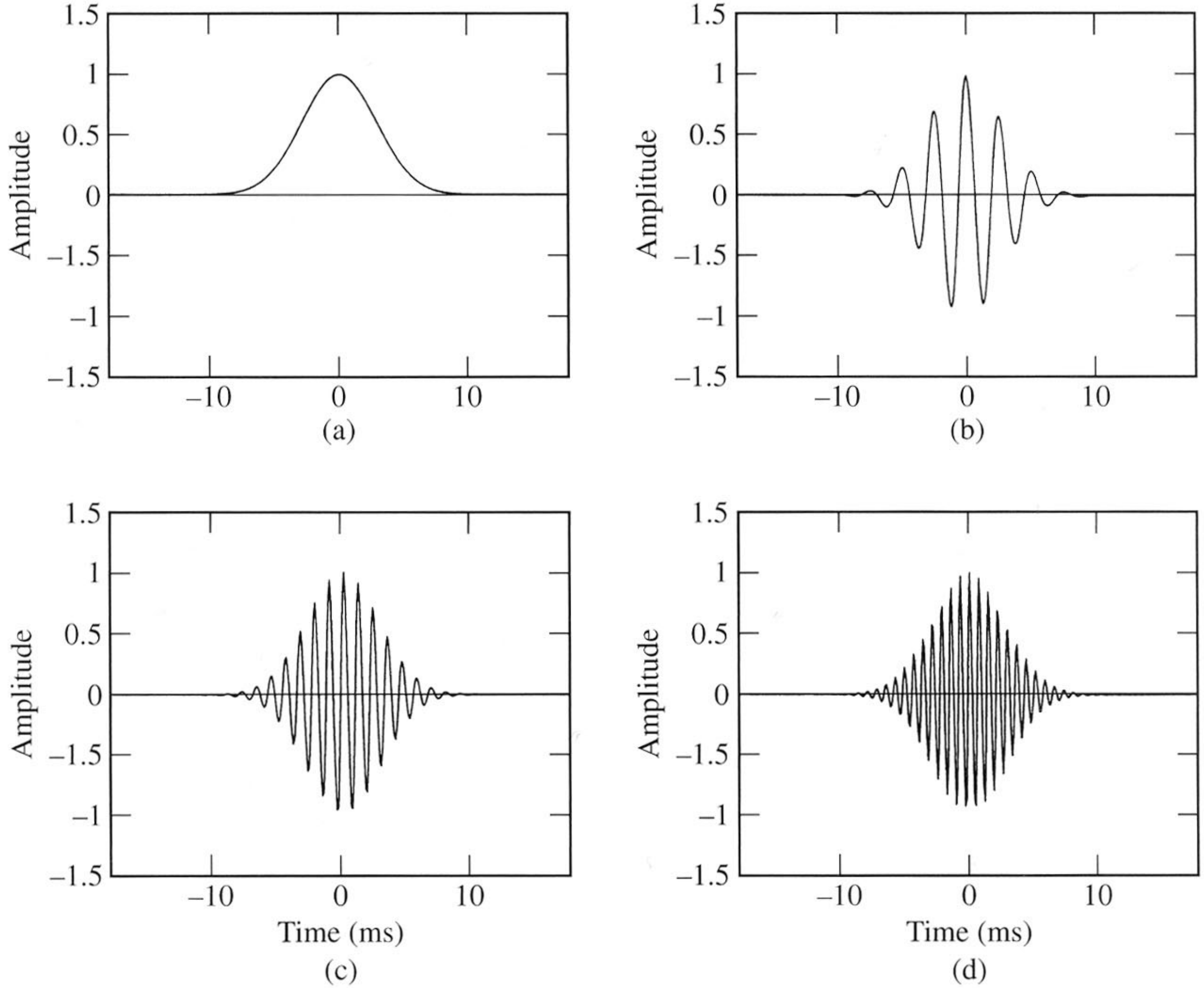


**Figure 7.6** Real part of the bandpass filter outputs with unit sample input $\delta[n]$ for the discrete filter bank of Example 7.2 prior to demodulation: (a) Gaussian window $w[n]$ (also output of filter $k = 0$); (b) discrete frequency $k = 5$; (c) discrete frequency $k = 10$; (d) discrete frequency $k = 20$. Frequency of the output increases with increasing $k$, while the Gaussian envelope remains intact.

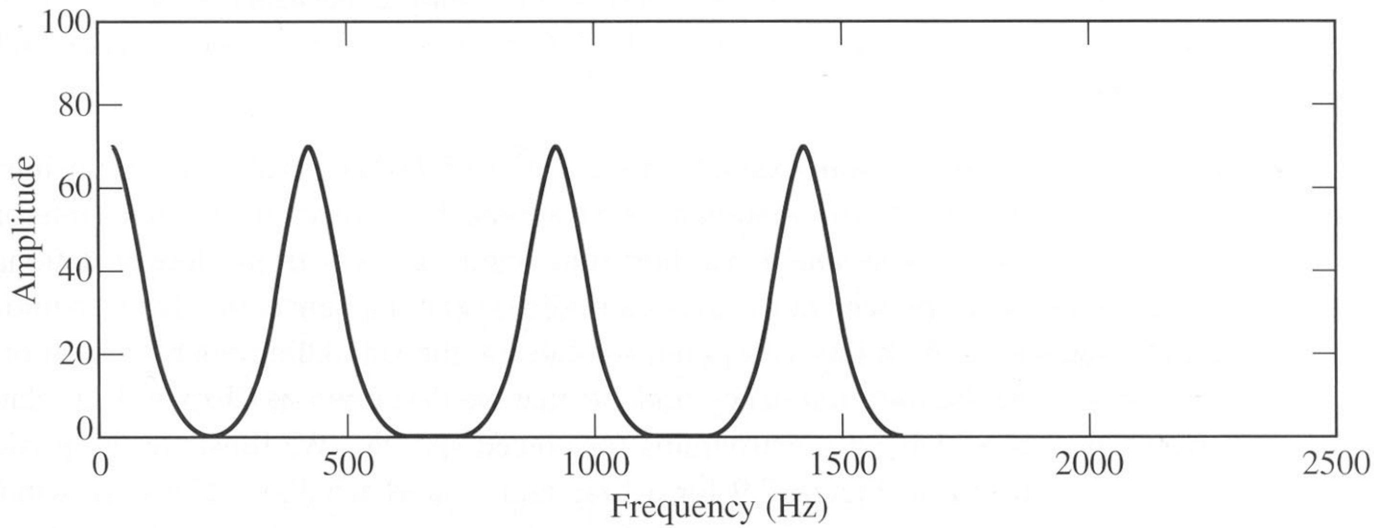


**Figure 7.7** Superimposed spectra of bandpass filter outputs of Figure 7.6.

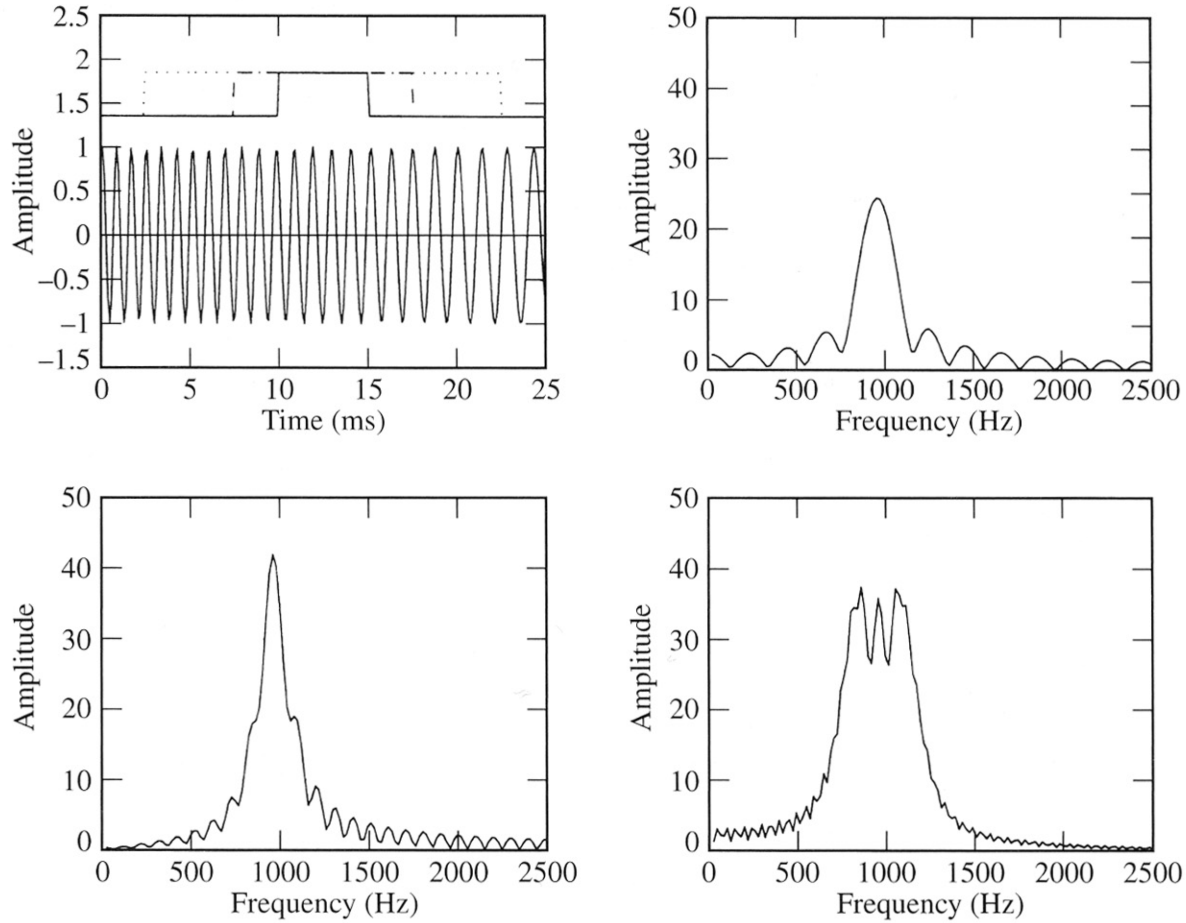


**Figure 7.8** Effect of the length of the analysis window on the discrete Fourier transform of a linearly frequency-modulated sinusoid of length 25 ms whose frequency decreases from 1250 Hz to 625 Hz. The discrete Fourier transform uses a rectangular window centered at 12.5 ms, as illustrated in panel (a). Transforms are shown for three different window lengths: (b) 5 ms [solid in (a)]; (c) 10 ms [dashed in (a)]; (d) 20 ms [dotted in (a)].

SOURCE: S.H. Nawab and T.F. Quatieri, "Short-Time Fourier Transform" [13]. 

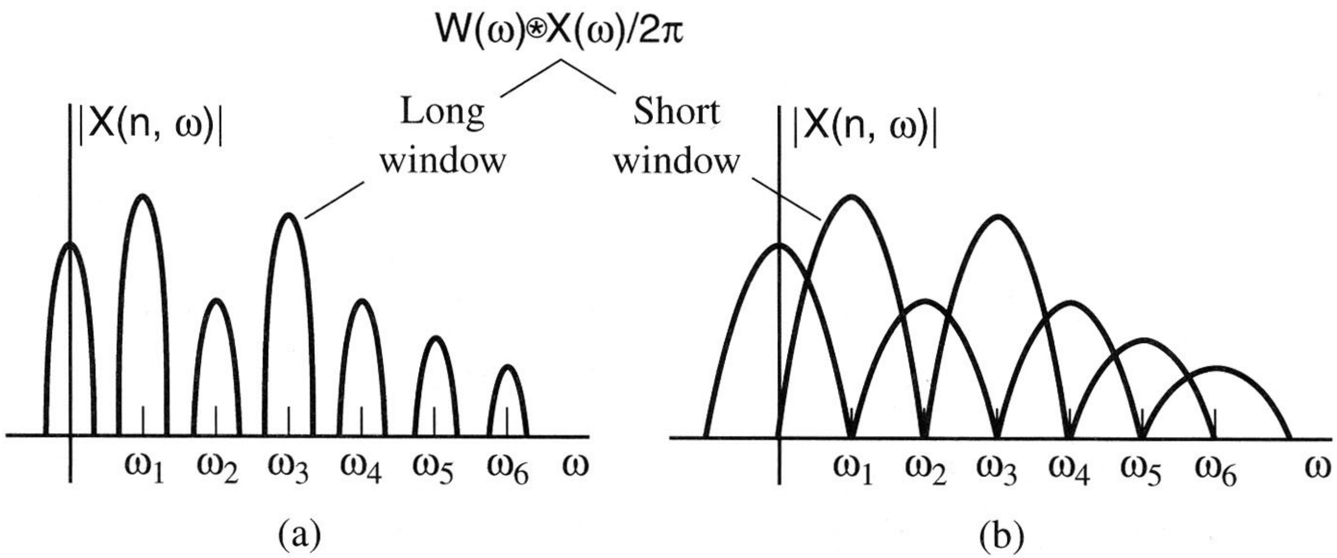


**Figure 7.9** Schematic of convolutional view of time-frequency resolution tradeoff with long and short analysis windows for harmonic spectra: (a) long window; (b) short window.

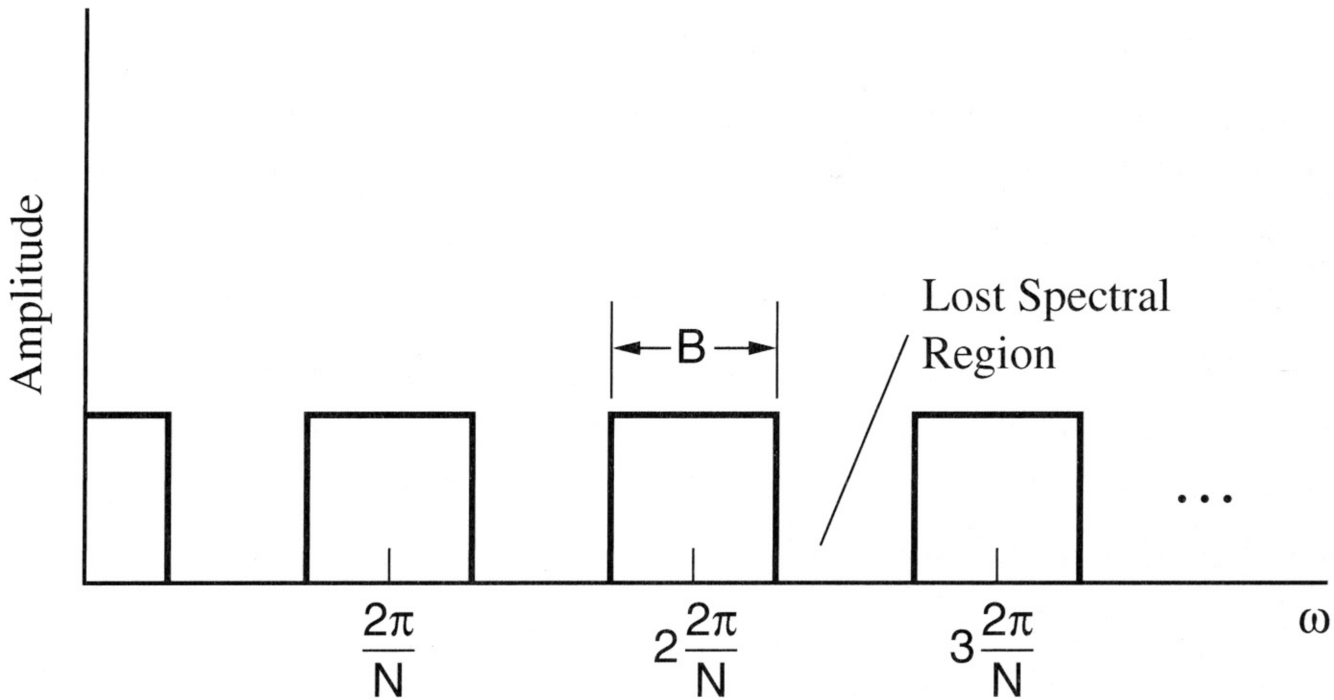


**Figure 7.10** Undersampled STFT when the frequency sampling interval $\frac{2\pi}{N}$ is greater than the analysis-filter bandwidth $B$.

SOURCE: S.H. Nawab and T.F. Quatieri, "Short-Time Fourier Transform" [13].

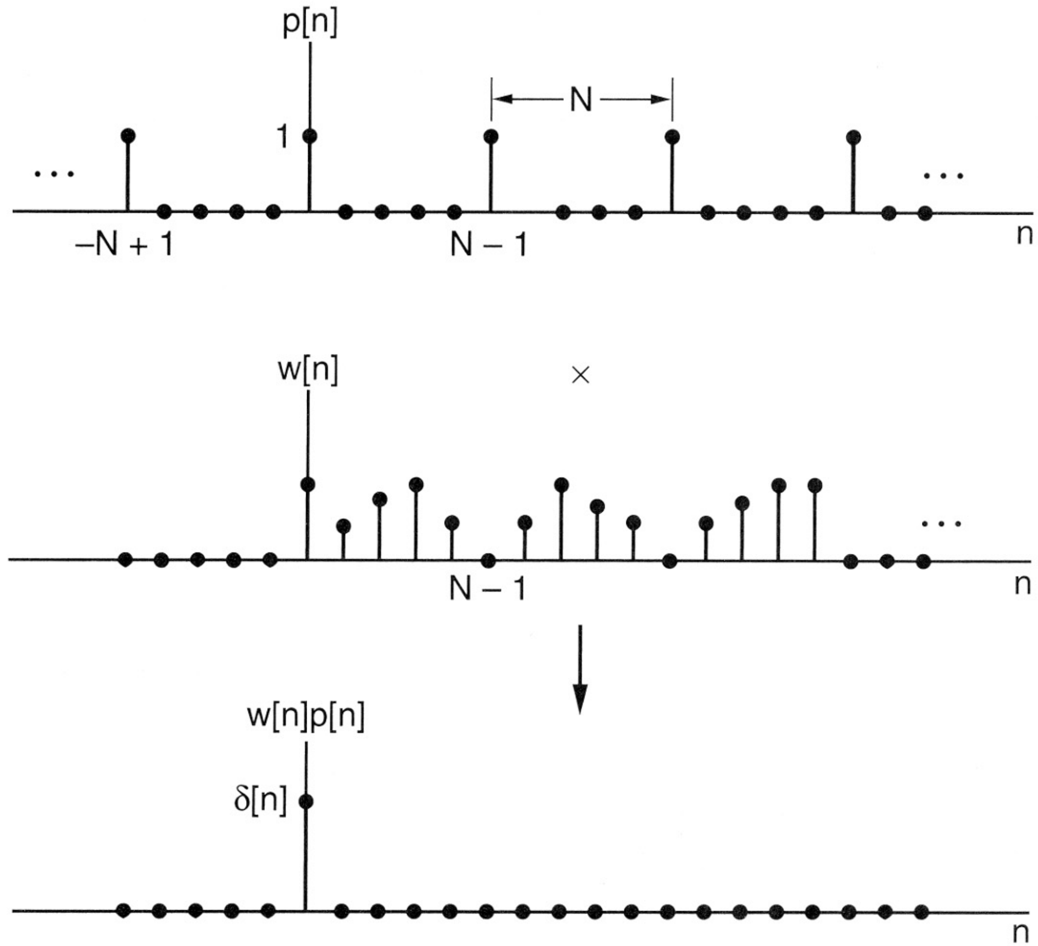


**Figure 7.11** Example of an analysis window and how it satisfies the FBS constraint. The analysis-window length is longer than the frequency sampling factor. The sequence $p[n] = \sum_{r=-\infty}^{\infty} \delta[n, n-r]$.

SOURCE: S.H. Nawab and T.F. Quatieri, "Short-Time Fourier Transform" [13]. ©1987, Pearson Education, Inc. Used by permission.

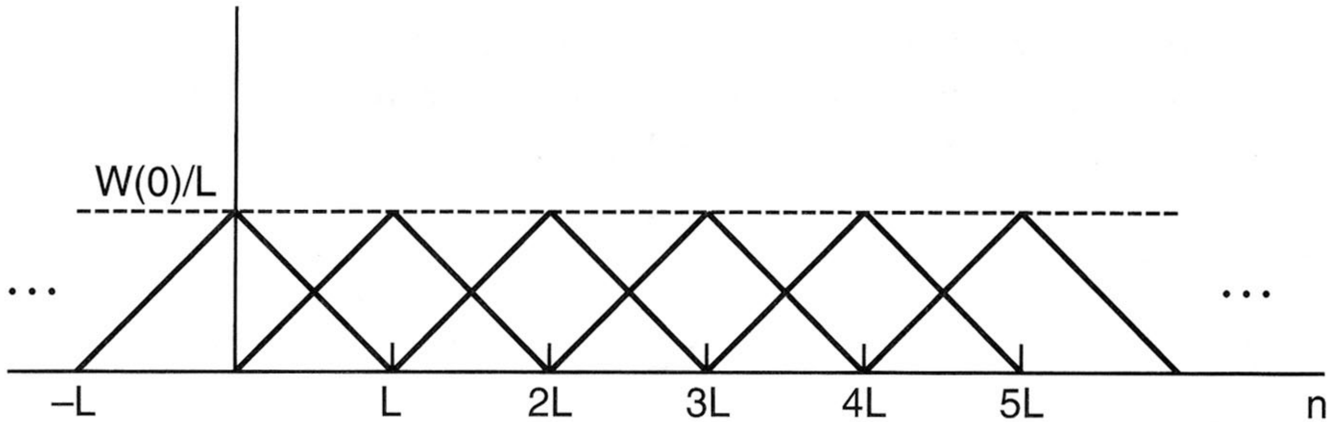


**Figure 7.12** The OLA constraint visualized in the time domain.

SOURCE: S.H. Nawab and T.F. Quatieri, "Short-Time Fourier Transform" [13]. ©1987, Pearson Education, Inc. Used by permission.

**FBS Method**

$$y[n] = \left[\frac{1}{Nw[0]}\right] \underbrace{\sum_{k=0}^{N-1} X(n, k)\, e^{j\frac{2\pi}{N}kn}}$$

Adding Frequency Components For Each n

FBS Constraint: $\sum_{k=0}^{N-1} W(\omega - \frac{2\pi}{N}k) = Nw[0]$

For $N_w < N \rightarrow y[n] = x[n]$

**OLA Method**

$$y[n] = \left[\frac{L}{W(0)}\right] \underbrace{\sum_{p=-\infty}^{\infty} x[n]\, w[pL-n]}$$

Adding Time Components For Each n

OLA Constraint: $\sum_{p=-\infty}^{\infty} w[pL-n] = \frac{W(0)}{L}$

For $\omega_c < \frac{2\pi}{L} \rightarrow y[n] = x[n]$

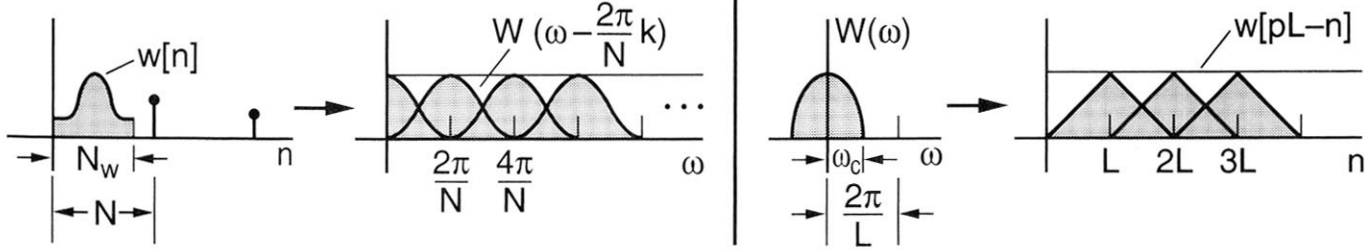


**Figure 7.13** Duality between the FBS and OLA constraints. Relaxation of constraints to allow zero crossings in time and in frequency is not shown.

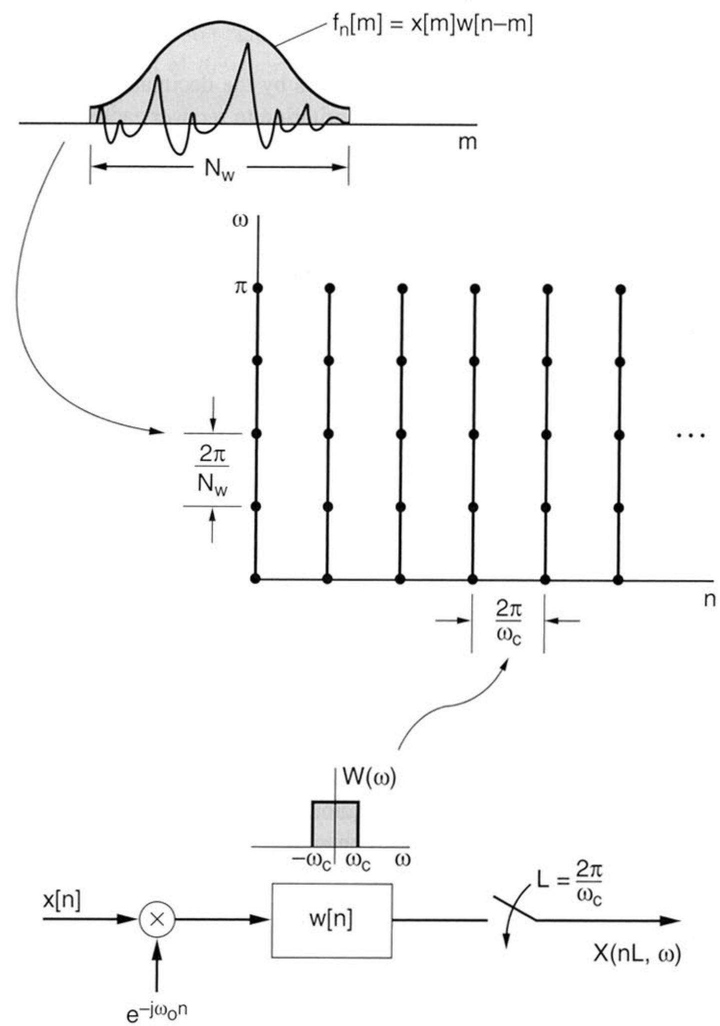


**Figure 7.14** Time-frequency sampling constraints from the perspective of classical time- and frequency-domain aliasing. The time sampling must satisfy the Nyquist criterion to avoid aliasing in frequency (but the OLA constraint allows relaxing the finite filter bandwidth constraint), while the frequency sampling must be fine enough to avoid aliasing in time (but the FBS constraint allows relaxing the finite window duration condition).