The window bandwidth is less than $\frac{2\pi}{L}$. We can relax this constraint if the shifted window transforms pass through zero at the frequency origin; in particular, if the window transform equals zero at $\omega = \frac{2\pi}{L}k$, the constraint holds. This is analogous to relaxing the FBS constraint that the window length N_w is less than the frequency sampling factor N by letting the analysis window pass through zero at $n = kN$.

7.3.4 Time-Frequency Sampling

We now give a different qualitative discussion of the above time-frequency sampling concepts for the OLA and FBS constraints from the perspective of classical time- and frequency-domain aliasing [20]. This discussion also serves to further summarize the sampling issues for these methods, and gives motivation for our earlier statement that sufficient but not necessary conditions for invertibility of the discrete STFT are that the analysis window is non-zero over its finite length N_w, the temporal decimation factor $L \leq N_w$, and the frequency sampling interval $\frac{2\pi}{N} \leq \frac{2\pi}{N_w}$.

Consider a short-time segment $f_n[m] = w[m]x[n-m]$ and its Fourier transform $X(n,\omega)$ with the analysis window of duration N_w. From the Fourier transform view, recovering time sequence $f_n[m]$ (with respect to the variable m) by an inverse DFT of the discrete STFT $X(n,k)$ requires a frequency sampling interval of $\frac{2\pi}{N_w}$ or finer to avoid time-domain aliasing of the time segment $f_n[m]$. Consider now a time decimation factor L. From the filtering view of the STFT, recovering the time sequence $X(n,k)$ (with respect to the time variable n) requires that this time sampling interval L meets the Nyquist criterion based on the bandwidth of the "filter" $w[n]$. This implies that we sample $X(n,k)$ at a time interval $L \leq \frac{2\pi}{\omega_c}$, where ω_c is the filter bandwidth (i.e., $W(\omega)$ is zero outside $[-\omega_c, \omega_c]$ within the interval $[-\pi, \pi]$), to avoid frequency-domain aliasing of the time sequence $X(n,\omega)$. This time-frequency sampling is illustrated in Figure 7.14. Selecting a Hamming window, typically used in speech analysis, and defining the bandwidth with respect to the 3 dB attenuation points, we find that the above sampling requirements, over all N filters in our filter bank, imply four times the number of samples in the original representation of the sequence $x[n]$ [20].

The above time-frequency sampling constraints, derived from simple aliasing considerations, are consistent with the OLA constraint (filter bandwidth $\omega_c \leq \frac{2\pi}{L}$, L being the time decimation factor) and the FBS constraint (window duration $N_w \leq N$, $\frac{2\pi}{N}$ being the frequency sampling interval). These window length and bandwidth constraints can, however, be relaxed, as we have seen in the OLA and FBA constraints, by allowing zeros in the window or its transform at the appropriate time or frequency points, respectively. This implies that we can avoid the four-fold increase in sampling requirements in the above example with a Hamming window analysis. We return to this issue later in Chapter 12 in our discussion of time-frequency analysis with application to speech coding.

Integrating our discussion of aliasing with the OLA and FBS methods, we summarize the following time-frequency constraint considerations from the perspective of each method.

OLA Method

1. For a window length N_w and with a DFT length chosen to give sufficient frequency sampling, i.e., a frequency sampling interval less than $\frac{2\pi}{N_w}$, then each short-time segment can be recovered because there is no time-domain aliasing.