Windows for Transform Processing

Peter Kabal

Department of Electrical & Computer Engineering
McGill University
Montreal, Canada

December 2005

Abstract

This report examines the time windows used in processing of signals in a transformed domain. The goal of windowing is to create frames of data, each of which will be used to calculate a transformed sequence. The transform coefficients are then modified (filtering for instance for noise reduction) or coded (transform coding). The modified transform coefficients are then applied to an inverse transform and windowed again before creating an output signal using addition of the overlapped blocks. It is the analysis window (before the transform) and the synthesis window (after the inverse transform) that are examined in this report. The requirement for perfect reconstruction (when the transform coefficients are not modified) is developed. This gives a condition on the product of the analysis and synthesis windows. An argument is given to show that if additive noise is introduced in the transform domain, the windowing should be equally apportioned between these windows, i.e. the analysis and synthesis windows should be the same. The windowing requirements for systems implementing block-by-block filtering of the input signal in the transform domain are also examined.

Windows for Transform Processing

1 Introduction

This report examines the time windows used in processing of signals in a transformed domain. The transform can be, for instance, one which creates transform coefficients which have a frequency domain interpretation. These transform coefficients are then subject to spectral modification. The overall system uses analysis and synthesis windows and these will be the subject of this report.

2 Windows in Transform Processing

Consider a transform coding system using block-based linear transforms with overlapping windows. Let the block length be N and let the transform blocks advance by L samples. For block p, we can represent the N samples of the block as

$$x_p[m] = x[pL + m], \qquad 0 \leq m < N. \tag{1}$$

An analysis window $w_a[m]$ of length N is applied to the data before the transform. The windowed block of samples is

$$x_{pw}[m] = w_a[m]\, x_p[m], \qquad 0 \leq m < N. \tag{2}$$

We will represent the N transform coefficients for block p as $X_p[k]$. These operations are shown in the top part of Fig. 1. In this diagram, the transform is given as a DFT, but could in fact be any square transform.

The transform coefficients are then modified, for instance by coding (quantization). We can model the effect of coding as adding some noise to the transformed signal,

$$\hat{X}_p[k] = X_p[k] + \Xi_p[k], \qquad 0 \leq k < N. \tag{3}$$

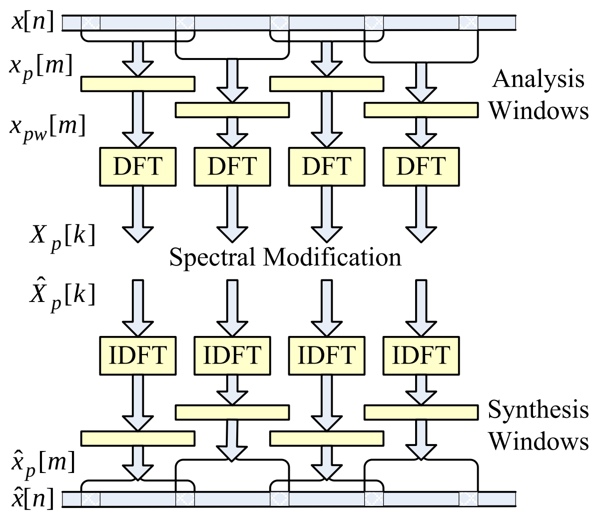

Fig. 1 Transform Analysis/Synthesis with Windowing.

An inverse transform is then applied. Assuming linearity,

$$\hat{x}_{pw}[m] = x_{pw}[m] + \xi_p[m], \qquad 0 \le m < N. \tag{4}$$

The output of the inverse transform is then windowed again with a synthesis window $w_s[n]$ to give

$$\hat{x}_p[m] = w_s[m]\, x_p w[m] + w_s[m]\, \xi_p[m]), \qquad 0 \le m < N. \tag{5}$$

The final signal is given by an overlap-add reconstruction

$$\begin{aligned}
\hat{x}[n] &= \sum_{p=-\infty}^{\infty} \hat{x}_p[n-pL] \\
&= \sum_{p=-\infty}^{\infty} w_s[n-pL]\, x_{pw}[n-pL] + \sum_{p=-\infty}^{\infty} w_s[n-pL]\, \xi_p[n-pL]) \\
&= \sum_{p=-\infty}^{\infty} w_s[n-pL]\, w_a[n-pL]\, x_p[n-pL] + \sum_{p=-\infty}^{\infty} w_s[n-pL]\, \xi_p[n-pL] \\
&= x[n] \sum_{p=-\infty}^{\infty} w_s[n-pL]\, w_a[n-pL] + \sum_{p=-\infty}^{\infty} w_s[n-pL] \xi_p[n-pL].
\end{aligned} \tag{6}$$

This analysis separates the signal part of the reconstructed signal and the noise part of the reconstructed signal. The reconstruction process is shown in the bottom part of Fig. 1.

The analysis window can be optimized for a particular application. Tapered analysis windows trade-off frequency resolution with sidelobe suppression. With modification of the spectrum, a tapered synthesis window which merges outputs from adjacent windows can reduce block-edge effects. First we will consider the effect of windowing on the signal components.

2.1 Perfect Reconstruction

At a given time point n, the input signal may appear as input to several transforms since the blocks overlap. The signal component at the output after analysis and synthesis is

$$\hat{x}[n] = x[n] \sum_{p=-\infty}^{\infty} w_a[n-pL]\, w_s[n-pL]. \tag{7}$$

In the absence of modification of the transform coefficients, this is the output of the analysis/synthesis system. The requirement for perfect reconstruction is

$$\sum_{p=-\infty}^{\infty} w_a[n-pL]\, w_s[n-pL] = 1, \qquad \text{all } n. \tag{8}$$

Denote the product of the analysis and synthesis windows as $w[n]$,

$$w[n] = w_a[n]\, w_s[n]. \tag{9}$$

The perfect reconstruction property can be written as a convolution,

$$\sum_{p=-\infty}^{\infty} \delta[n-pL] * w[n] = 1, \qquad \text{all } n. \tag{10}$$

In the frequency domain, this convolution corresponds to a product,

$$\sum_{l=-\infty}^{\infty} \delta(\omega - \frac{2\pi l}{L})\, W(\omega) = 2\pi \sum_{k=-\infty}^{\infty} \delta(\omega - 2\pi k), \tag{11}$$

or equivalently

$$W(\frac{2\pi l}{L}) = \begin{cases} 2\pi, & l = pL, \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

This is a requirement on the zero crossings of the frequency response of the combined window.[1]

The minimum length window satisfying the perfect reconstruction property is a rectangular

[1] The perfect reconstruction property is analogous to the Nyquist condition for no-intersymbol interference of pulses in data transmission.

window of length L. From Eq. (12), one can see that if its frequency response is multiplied by another frequency response, $P(\omega)$, the combined response, $W(\omega) = P(\omega)W_L(\omega)$ also has the perfect reconstruction property (to within a constant). This multiplication corresponds to a convolution in the time domain. This means we can convolve the rectangular window of length L with another time function and still end up with a window satisfying the perfect reconstruction property,

$$w[n] = p[n] * w_L[n]. \tag{13}$$

The only condition on $p[n]$ is that its area should be unity (i.e. $(P(\omega) = 1)$. If the sequence $p[n]$ is symmetric, the window $w[n]$ will be symmetric. As an example of this type of construction, in Appendix A, raised-cosine windows are shown to be expressible as the convolution of a rectangular window and another time function (a sine lobe).

2.2 Optimizing the Signal-to-Noise Ratio

Let us model the spectral modification process in the transform domain as generating additive noise. The noise component is different for each transform, so the noise output at time n is

$$\xi[n] = \sum_{p=-\infty}^{\infty} \xi_p[n-pL]\, w_s[n-pL], \tag{14}$$

where $\xi_p[m]$ is the noise in the mth sample in the pth block. It is not unreasonable to assume the noise components from different transforms are uncorrelated. We will also assume that the noise terms have the same variance and the variance does not depend on the time index m. The noise power at the output at time n is then

$$\begin{aligned} N_o[n] &= E\big[\hat{\xi}^2[n]\big] \\ &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} E\big[\xi_k[n-kL]\, \xi_l[n-lL]\big]\, w_s[n-kL]\, w_s[n-lL] \\ &= \sigma_\xi^2 \sum_{k=-\infty}^{\infty} w_s^2[n-kL]. \end{aligned} \tag{15}$$

The signal power is (assuming a zero mean signal)

$$\begin{aligned} S_o[n] &= E\big[\hat{x}^2[n]\big] \\ &= E\big[x^2[n]\big] \Big| \sum_{k=-\infty}^{\infty} w_a[n-kL]\, w_s[n-kL] \Big|^2 \\ &= \sigma_x^2[n] \Big| \sum_{k=-\infty}^{\infty} w_a[n-kL]\, w_s[n-kL] \Big|^2 . \end{aligned} \tag{16}$$

The signal-to-noise ratio (SNR) at time n is

$$\frac{S_o[n]}{N_o[n]} = \frac{\sigma_x^2[n]}{\sigma_\xi^2} \frac{\Big| \displaystyle\sum_{k=-\infty}^{\infty} w_a[n-kL]\, w_s[n-kL] \Big|^2}{\displaystyle\sum_{k=-\infty}^{\infty} w_s^2[n-kL]} . \tag{17}$$

For a given choice of $w_a[n]$, the SNR is maximized (Schwartz's inequality) by choosing $w_s[n] = w_a[n]$.[2] This is a matched-filter result. SNR maximization requires that the analysis and synthesis windows have the same shape.

In perfect reconstruction systems, the product of the two windows is often chosen to be a modified Hann window (see Appendix A). If the analysis and synthesis windows are made equal, they become sine windows. These are the minimum bias windows discussed in [3].

3 Analysis/Synthesis Windows for Linear Filtering

One form of modification of the transform coefficients is multiplication term-by-term by another transform response. For the case of the DFT, this multiplication can be used to implement filtering using an overlap-add approach.

Following the formalism above, we set up the problem as follows. Let the block length be N and let the blocks advance by L samples. For block p, we can represent the N samples of the block as

$$x_p[m] = x[pL+m], \qquad 0 \le m < N. \tag{18}$$

An analysis window $w_a[m]$ of length N is applied to the data before filtering. The windowed block of samples is

$$x_{pw}[m] = w_a[m]\, x_p[m], \qquad 0 \le m < N. \tag{19}$$

[2] The version of Schwartz's inequality applicable here is $|\sum_k w_1[k]\, w_2[k]|^2 \le \sum_k |w_1[k]|^2 \sum_k |w_2[k]|^2$, with equality if and only if $w_1[k] = a w_2[k]$.

These values will be filtered block-by-block with a causal FIR filter $h[k]$ with N_h coefficients. The filtered values are

$$\begin{aligned} y_{fp}[m] &= x_{pw}[m] * h[m], \qquad 0 \leq m < N + N_h - 1 \\ &= \sum_{k=0}^{N_h-1} h[k]\, w_a[m-k]\, x_p[m-k]. \end{aligned} \tag{20}$$

Each such block is then windowed again with a synthesis window $w_s[m]$ to give

$$y_{fpw}[m] = w_s[m]\, y_{fp}[m], \qquad 0 \leq m < N + N_h - 1. \tag{21}$$

The final signal is given by an overlap-add reconstruction

$$\begin{aligned} y[n] &= \sum_{p=-\infty}^{\infty} y_{fpw}[n-pL] \\ &= \sum_{p=-\infty}^{\infty} w_s[n-pL]\, y_{fp}[n-pL] \\ &= \sum_{p=-\infty}^{\infty} w_s[n-pL] \sum_{k=0}^{N_h-1} h[k]\, w_a[n-pL-k]\, x_p[n-pL-k] \\ &= \sum_{k=0}^{N_h-1} h[k]\, x[n-k] \sum_{p=-\infty}^{\infty} w_a[n-pL-k]\, w_s[n-pL]. \end{aligned} \tag{22}$$

In order for this to be just a convolution of $h[\cdot]$ and $x[\cdot]$, the last sum must be constant as a function of n for k between 0 and $N_h - 1$,

$$\sum_{p=-\infty}^{\infty} w_a[n-pL-k]\, w_s[n-pL] = 1, \qquad \text{all } n,\ 0 \leq k < N_h - 1. \tag{23}$$

3.1 Window Requirements for Filtering

Consider rectangular analysis and synthesis windows as illustrated in Fig. 2

$$\begin{aligned} w_a[m] &= \begin{cases} 1 & 0 \leq m < L, \\ 0 & L \leq m < N. \end{cases} \\ w_s[m] &= 1 \qquad 0 \leq m < N. \end{aligned} \tag{24}$$

The analysis window pads the L input samples with $N_h - 1$ zeros to give a total block length of N. This is the zero-padding needed to implement linear convolution with an N-point DFT (overlap-add method). The zero-padding avoids the time-domain aliasing present in the circular convolution produced by DFT processing (see [4, 5] for discussions of overlap-add filtering using DFT's).

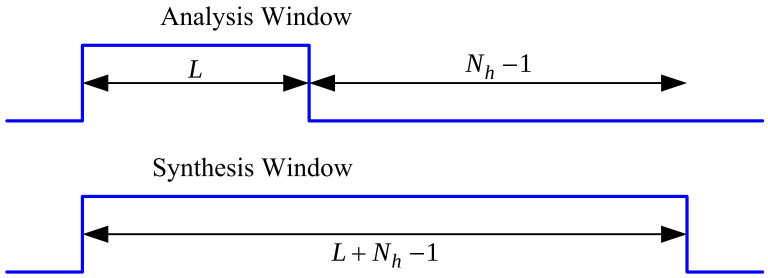

Fig. 2 Rectangular windows for linear filtering

Consider a tapered analysis window of length $L + M$, padded with with $N_h - 1$ zeros to give a total length $N = L + M + N_h - 1$. The synthesis window will be a rectangular window of length N. These are illustrated in Fig. 3. The analysis window by itself must satisfy the perfect reconstruction property. This setup allows for a tapered analysis window of total non-zero length $L + M$, with a block advance of L samples.

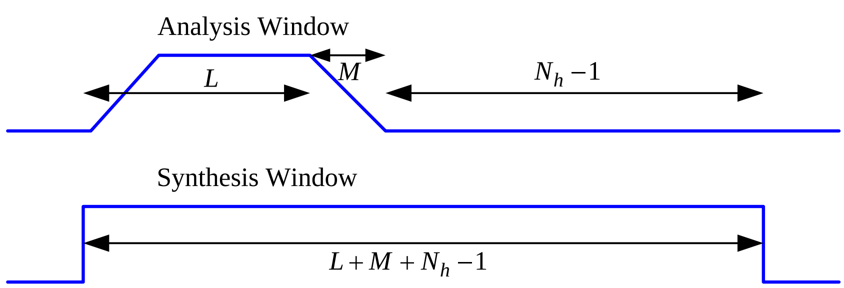

Fig. 3 Tapered and rectangular windows for linear filtering

4 Discussion

We have considered two scenarios. In the first, the transform coefficients are modified by adding noise to the coefficients. This models the effect of quantization of the coefficients. In the second scenario, the transform coefficients are subject to modification to implement filtering.

4.1 Transform Coding

This scenario applies to coding of transform coefficients. Perhaps the most familiar example is in audio coding, see for example [6]. This coder uses MDCT's as a transform and uses sine lobe windows for analysis and synthesis. Quantization of the signal occurs in the MDCT domain. The analysis in Section 2.2 assumes that the variance of the noise introduced is the same for each frame. This is generally not true since the quantization step sizes are signal level dependent. That means that the noise in an energetic frame is usually larger than the noise in a less energetic frame. In the region of overlap between frames, the noise from the more energetic frame will dominate. Nonetheless, in regions of fairly constant quantization noise, the analysis suggests that equi-partition of the windowing is appropriate.

An optimization of the windows for controlling the sidelobes in a transform coder, while maintaining perfect reconstruction, was carried out in [7].

4.2 Spectral Modification for Signal Enhancement

A common approach for speech enhancement is to use spectral subtraction or Wiener filtering. In such systems, the transform domain coefficients are scaled based on an estimate of the signal-to-noise ratio — coefficients with poor SNR are attenuated more than coefficients with good SNR. See Quatieri [8] for an overview of spectral subtraction techniques.

The scaling of the transform domain coefficients can be recast as a zero-phase filter with frequency dependent attenuation. The filter is time-varying, changing from block-to-block. The analysis given in Section 3.1 for filtering is still relevant.

An example of speech enhancement is the noise suppressor designed for the IS-127 EVRC (Enhanced Variable Rate Codec) Coder, a standard for CDMA cellular systems (see [9] for an overview, and [10] for details). This system uses a raised-cosine analysis window) and a rectangular synthesis window. In terms of the parameters in Fig. 3, $L = 80$, $M = 24$, and $N_h = 25$, for a total block length of $N = 128$. The analysis window provides for zero-padding. The filter gain is constant in subbands. However, the length of the filter is 128 samples, so that some time-domain aliasing will be inevitable.

The problem of time-domain aliasing has been considered in [11]. That article points out that the magnitude-only modification of the spectrum corresponds to a non-causal filter. For instance if the filter response is lowpass in nature, the inverse transform (IDFT) of the filter response has large samples near the beginning of the filter array (corresponding to positive time indices) and near the end of the filter array (corresponding to negative time indices). The authors suggest adding a linear phase (corresponding to half the filter length) to shift the response to be centred in the filter coefficient array. The problem of long filter responses can be handled by estimating the filter response on a coarser frequency grid. The frequency response of the short filter is then interpolated to the denser frequency grid.

5 Summary

This report has developed the perfect reconstruction property for analysis and synthesis windows used in transform processing. For transform coding, it was shown that the conventional use of equal analysis and synthesis windows is justified from a SNR maximization argument. Means to construct windows satisfying the perfect reconstruction property have also been given.

Appendix A - Raised-Cosine Windows

A.1 Continuous-time Raised-Cosine Windows

Consider a raised-cosine time-window defined in continuous time,

$$w(t) = \begin{cases} \frac{1}{2} + \frac{1}{2}\sin\big(\dfrac{\pi}{U}(t + \dfrac{W}{2})\big) & -\frac{W+U}{2} \leq t < -\frac{W-U}{2}, \\ 1 & -\frac{W-U}{2} < t \leq \frac{W+U}{2}, \\ \frac{1}{2} - \frac{1}{2}\sin\big(\dfrac{\pi}{U}(t - \dfrac{W}{2})\big) & \frac{W-U}{2} \leq t \leq \frac{W+U}{2}, \\ 0 & \text{elsewhere.} \end{cases} \tag{25}$$

This window is of length $W + U$ and centred at the origin. For a fixed W, changing U allows the characteristics to change from a rectangular window ($U = 0$) to a Hann window ($U = W$). The window has a rising portion (length U), a constant part (length $W - U$), and a falling section (length U). When the constant part vanishes ($U = W$), the pulse is a *full* raised-cosine or $\cos^2(\cdot)$ function. This raised-cosine time-window is the time domain analog to the raised-cosine spectra used to mitigate intersymbol interference in data transmission [1, 2].

The raised-cosine window satisfies the perfect reconstruction property with the time advance W,

$$\sum_{k=-\infty}^{\infty} w(t - kW) = 1. \tag{26}$$

The continuous-time window has a continuous time-derivative, implying that the frequency response falls off asymptotically as $1/|f|^3$ [1, 2].[3] A discrete-time window can be formed by sampling the continuous-time window. We cannot talk about derivatives of a discrete-time function, but the continuity properties of the underlying continuous-time function will help ensure a “smoothness” of the discrete-time window. Also, the asymptotic fall-off of the frequency response will influence the discrete-time frequency response fall-off. The smoothness / frequency response properties will be most evident for long discrete-time windows.

[3] See [2] for a discussion of the relationship between the number of continuous derivatives and the asymptotic decay rate. In that citation, the decay rate is in the time-domain and the derivatives are in the frequency-domain. The results carry over to our problem with an interchange of the time and frequency responses.

A.2 Direct Construction

We want an N sample discrete-time window which satisfies the perfect reconstruction property for a time advance of L samples,

$$\sum_{k=-\infty}^{\infty} w[n-kL] = 1. \tag{27}$$

A discrete-time window can be obtained by sampling an appropriate version of the continuous-time raised-cosine window. Define $M = N - L$. There will be M samples taken from the rise, $L - M$ samples taken from the constant portion, and M samples taken from the fall. Denote the sample times as t_n, $0 \leq n < N$, with the sample interval being Δt. The first sampling time, t_0 will be a distance $\tau_l \Delta t$ from the beginning of the continuous-time window and the last sampling time will be a distance $\tau_u \Delta t$ from the end of the continuous-time window. If t_n is in the rise region $(0 \leq n < M)$, then the perfect reconstruction property requires that

$$w(t_n) + w(t_{n+L}) = 1. \tag{28}$$

This condition is satisfied by setting $W = L\Delta t$ and having the fall region extend from $t_L - \tau_l \Delta t$ to $t_{L+M-1} + \tau_u \Delta t$. The parameters of the underlying continuous-time window are

$$W = L\Delta t, \qquad U = (M - 1 + \tau_l + \tau_u)\Delta t. \tag{29}$$

The sample times are

$$t_n = (n - \frac{N-1}{2} + \frac{\tau_u - \tau_l}{2})\Delta t. \tag{30}$$

The sampled window will be symmetric if $\tau_u = \tau_l$.

A.2.1 Discrete-Time Window I

For a discrete-time window of Type I, we require that the window be symmetric with the sampling pattern starting and ending half a sample from the ends of each segment of the continuous-time window. Then

$$\tau_l = \tau_u = \frac{1}{2}. \tag{31}$$

For convenience, we normalize Δt to be unity. Then the continuous-time window and the sampling points are described by

$$W = L, \qquad U = M, \qquad t_n = n - \frac{N-1}{2}. \tag{32}$$

The resulting discrete-time window is

$$w_I[n] = \begin{cases} \frac{1}{2} + \frac{1}{2}\sin\bigl(\frac{\pi}{M}(n - \frac{M-1}{2})\bigr) & 0 \le n < M, \\ 1 & M \le n < L, \\ \frac{1}{2} - \frac{1}{2}\sin\bigl(\frac{\pi}{M}(n - L - \frac{M-1}{2})\bigr) & L \le n < L+M, \\ 0 & \text{elsewhere.} \end{cases} \tag{33}$$

This window has a rise (M samples), a flat top ($L-M$ samples), and fall (M samples). For $M = 0$, we get a rectangular window; when $M = L$, we get the modified Hann window (see [3] for more details). This construction is no longer applicable if $M > L$.

An example of a Type I raised-cosine window is shown in Fig. 4.

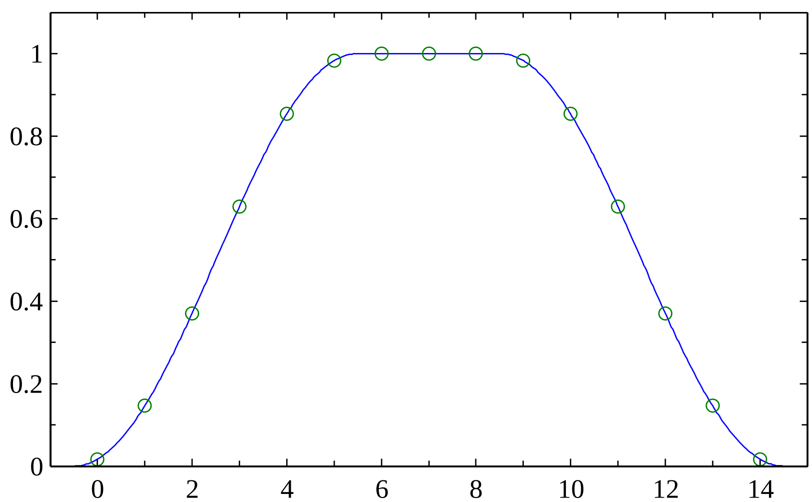

Fig. 4 Raised-cosine window (Type I) for $N = 15$, $L = 9$ (circles). The line is the shifted underlying continuous-time window.

A.2.2 Discrete-Time Window II

A second version of the discrete-time window can be created with a sampling pattern which starts and ends one sample from the ends of each segment of the continuous-time window,

$$\tau_l = \tau_u = 1. \tag{34}$$

With Δt set to unity, the continuous-time window and the sampling points are given by

$$W = L, \qquad U = M + 1, \qquad t_n = n - \frac{N-1}{2}. \tag{35}$$

The rise is sampled starting one sample from its initial (zero-valued) start and ending one sample from its end. The end of the rise is also the start point of the constant portion. The rise which is of length U will span $M+1$ sampling intervals. The constant part will be sampled with the first sample its starting point and last sample at its ending point. Thus the constant part which is of length $W-U$ will span $L-1$ sampling intervals.

The resulting discrete-time window is

$$w_{II}[n] = \begin{cases} \frac{1}{2} + \frac{1}{2}\sin\Big(\frac{\pi}{M+1}(n - \frac{M-1}{2})\Big) & 0 \le n < M, \\ 1 & M \le n < L, \\ \frac{1}{2} - \frac{1}{2}\sin\Big(\frac{\pi}{M+1}(n - L - \frac{M-1}{2})\Big) & L \le n < L+M, \\ 0 & \text{elsewhere.} \end{cases} \tag{36}$$

This window has a rise (M samples), a flat top ($L-M$ samples), and fall (M samples). For $M = 0$, we get a rectangular window.

An example of a Type II raised-cosine window is shown in Fig. 5. For the same discrete-time window parameters, the underlying continuous-time window has a shorter flat portion and longer rise and fall portions than for the Type I window.

A.3 Construction via Convolution

The continuous-time window can be created as a convolution of a unit-area cosine pulse (length U) with a rectangular window (length W),

$$w(t) = p(t) * w_W(t), \tag{37}$$

where $p(t)$ is a cosine pulse,

$$p(t) = \frac{\pi}{2U}\cos\Big(\frac{\pi t}{U}\Big), \qquad -\frac{U}{2} \le t \le \frac{U}{2}. \tag{38}$$

The frequency response of the cosine pulse (continuous function, but discontinuous first derivative) falls off as $1/|f|^2$ and the frequency response of the rectangular pulse (discontinuous function) falls off as $1/|f|$, giving an overall asymptotic decay of $1/|f|^3$. In the time-domain, this implies that the continuous-time raised-cosine window has a continuous first derivative.

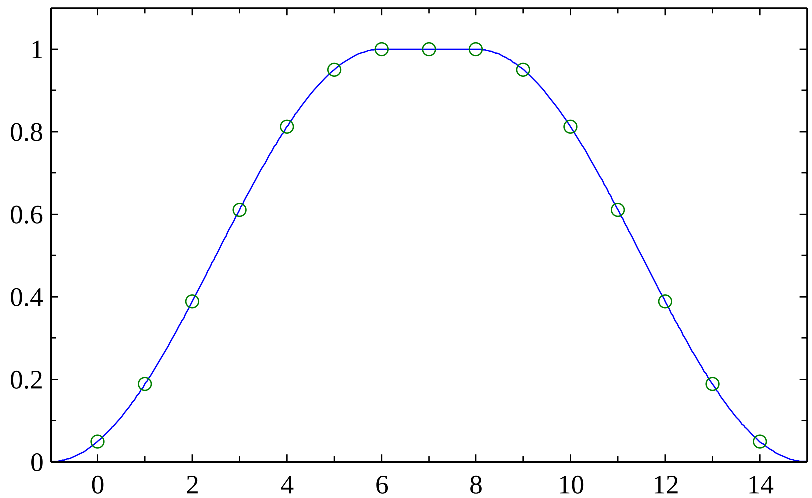

Fig. 5 Raised-cosine window (Type II) for $N = 15$, $L = 9$ (circles). The line is the shifted underlying continuous-time window.

Forming the window as a convolution does not limit the length, U, of the cosine pulse relative to the length of the rectangular window. The perfect reconstruction property holds for any length of the cosine pulse. The overall window will have a flat top for $U < 2W$. For $U = 2W$, the flat top part vanishes: the overall pulse is a $\cos^2(\cdot)$ pulse. For $U > 2W$, the first and last parts of the window (length W) will follow the same function as for a longer window. The middle portion (length $U - 2W$) will be different, but will have continuous derivatives at the transition points. In the limit as U is much larger than $2W$ (the rectangular window becoming closer to an impulse), the overall window will approach that of the cosine pulse.

For discrete-time windows which satisfy the perfect reconstruction property, we can write the windows as the convolution of a rectangular window of length L with another function $p[n]$ (see Section 2.1),

$$w[n] = p[n] * w_L[n]. \tag{39}$$

To get perfect reconstruction, the area of $p[n]$ should be unity. The general form of the pulse is given by,

$$p_I[n] = \begin{cases} \sin^2\bigl(\frac{\pi}{2U}\tau_l\bigr) & n = 0, \\ \sin\bigl(\frac{\pi}{2U}\bigr)\,\sin\bigl(\frac{\pi}{U}(n - \frac{1}{2} + \tau_l)\bigr) & 1 \le n < M, \\ \sin^2\bigl(\frac{\pi}{2U}\tau_u\bigr) & n = M, \\ 0 & \text{elsewhere.} \end{cases} \tag{40}$$

A.3.1 Discrete-Time Window I

For the Type I raised-cosine window ($U = M$, $\tau_l = \tau_u = 1/2$),

$$p_I[n] = \begin{cases} \sin^2(\dfrac{\pi}{4M}) & n = 0, \\ \sin(\dfrac{\pi}{2M})\,\sin(\dfrac{\pi n}{M}) & 1 \le n < M, \\ \sin^2(\dfrac{\pi}{4M}) & n = M, \\ 0 & \text{elsewhere.} \end{cases} \tag{41}$$

This formula can be seen to be a scaled sine lobe (which has zero-valued end points), plus scaled impulses at $n = 0$ and $n = M$. This pulse is plotted in Fig. 6. Note that the end points lie above the sine lobe.

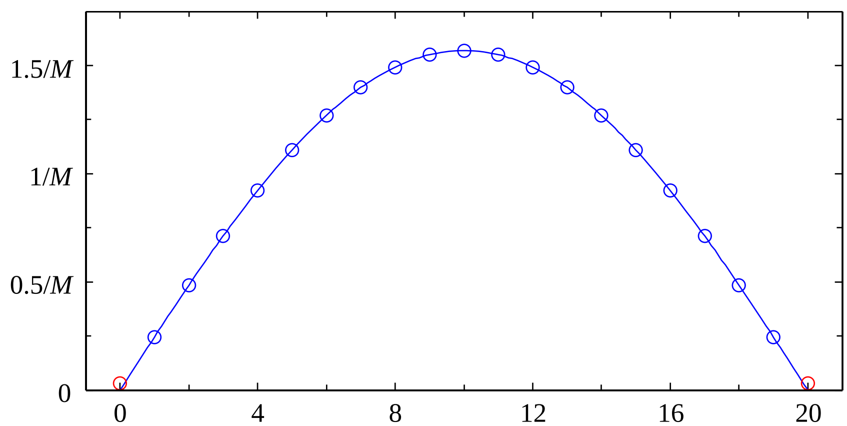

Fig. 6 Sine pulse for raised-cosine window (Type I) for $M = 20$ (circles). The end points lie above zero. The line is the underlying sine function.

Using the convolution construction of the raised-cosine window allows more generality — the requirement that the windows advance M be at most L is not required. If $M > L$, a given window stretches beyond its immediate neighbours and the overlap-add reconstruction involves three terms or more.

The discrete-time Fourier transform of the pulse is

$$\begin{aligned} P_I(\omega) = e^{-j\omega\frac{M}{2}} \big[&(1-\cos(\frac{\omega_0}{2}))\cos(\omega\frac{M}{2}) \\ &+ \frac{1}{2}\sin(\frac{\omega_0}{2})\big(\cos(\frac{\omega-\omega_0}{2})\,\text{Dsinc}(\frac{\omega-\omega_0}{2},M) + \cos(\frac{\omega+\omega_0}{2})\,\text{Dsinc}(\frac{\omega+\omega_0}{2},M)\big)\big], \end{aligned} \tag{42}$$

where the function Dsinc is defined as[4]

$$\text{Dsinc}(\omega, N) = \frac{\sin(\omega N/2)}{\sin(\omega/2)}. \tag{43}$$

and $\omega_0 = \pi/M$. This response takes on unity gain at $\omega = 0$. The term $\cos(\omega M/2)$ has zero crossings at $\omega = (2k+1)\pi/M$, i.e. at odd multiples of π/M. The $\text{Dsinc}(\omega \pm \omega_0, M)$ terms have zero crossings at odd multiples of π/M, except at $\pm\pi/M$.

The final frequency response of the rectangular window of length L is

$$W_L(\omega) = e^{-j\omega\frac{L-1}{2}}\, \text{Dsinc}(\omega, L). \tag{44}$$

This response has zero crossings at $\omega = \pi k/L$ The frequency response of the raised-cosine window is then the product of this response and the pulse response.

$$\begin{aligned} W_I(\omega) = e^{-j\omega\frac{N-1}{2}}\, \text{Dsinc}(\omega, L)\big[&(1 - \cos(\frac{\omega_0}{2}))\cos(\omega\frac{M}{2}) \\ &+ \frac{1}{2}\sin(\frac{\omega_0}{2})\big(\cos(\frac{\omega-\omega_0}{2})\,\text{Dsinc}(\frac{\omega-\omega_0}{2}, M) + \cos(\frac{\omega+\omega_0}{2})\,\text{Dsinc}(\frac{\omega+\omega_0}{2}, M)\big)\big], \end{aligned} \tag{45}$$

A special case occurs when $L = M$. The zero crossings due to the rectangular window and those due to the cosine pulse interlace, creating a frequency response with zero crossings at double the rate. This case is the modified Hann window discussed in [3].

A.3.2 Discrete-Time Window II

For the Type II discrete-time window ($U = M+1$, $\tau_l = \tau_u = 1$),

$$p_{II}[n] = \sin(\frac{\pi}{2(M+1)})\sin(\frac{\pi}{M+1}\frac{2n+1}{2}), \qquad 0 \le n \le M. \tag{46}$$

This pulse is plotted in Fig. 7.

The discrete-time Fourier transform of the raised-cosine window is the product of the DTFT for each of $w_L[n]$ and $p_{II}[n]$,

$$W_{II}(\omega) = \frac{1}{2}e^{-j\omega\frac{N-1}{2}}\, \text{Dsinc}(\omega, L)\big[\text{Dsinc}(\frac{\omega-\omega_0}{2}, M+1) + \text{Dsinc}(\frac{\omega+\omega_0}{2}, M+1)\big], \tag{47}$$

where $\omega_0 = \pi/(M+1)$. For this case, interlacing of the zero crossings occurs when $L = M+1$. This is the standard Hann window (see [3]).

[4] For N odd, the Dsinc function can be expressed in terms of the Dirichlet kernel, $\text{Dsinc}(\omega, N) = D_{(N-1)/2}(\omega)$, where $D_n(\omega)$ is the Dirichlet kernel.

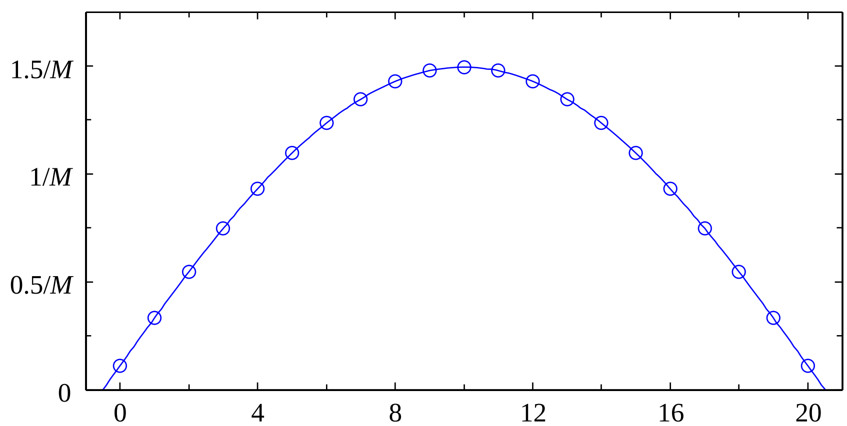

Fig. 7 Sine pulse for raised-cosine window (Type II) for $M = 20$ (circles). The line is the underlying sine function.

References

[1] N. Sheikholeslami Alagha and P. Kabal "Generalized Raised-Cosine Filters", *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 989–997, July 1999

[2] N. C. Beaulieu and M. O. Damen, "Parametric Construction of Nyquist-I Pulses", *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2134–2142, Dec. 2004.

[3] P. Kabal, "Time Windows for Linear Prediction of Speech" Technical Report, Electrical & Computer Engineering, McGill University, Oct. 2003 (on-line at `http://WWW.TSP.ECE.McGill.CA/MMSP/Documents/`).

[4] A. V. Oppenheim, R. W. Schafer, with J. R. Buck, *Discrete-Time Signal Processing*, 2nd ed., Prentice-Hall, 1999.

[5] J. G. Proakis and D. F. Manolakis, *Digital Signal Processing: Principles, Algorithms and Applications*, 3rd edition, Prentice-Hall, 1996.

[6] M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, M. Dietz, J. Herre, G. Davidson, and Y. Oikawa "ISO/IEC MPEG-2 Advanced Audio Coding" *J. Audio Eng. Soc.*, vol. 45, no. 10, p. 789–812, Oct. 1997.

[7] H. Najafzadeh-Azghandi, *Perceptual Coding of Narrowband Audio Signals*, Ph.D. Thesis, Electrical & Computer Engineering, McGill University, April 2000 (on-line at `http://WWW.TSP.ECE.McGill.CA/MMSP/Theses/`).

[8] T. F. Quatieri, *Discrete-Time Speech Signal Processing* Prentice Hall PTR, 2002.

[9] T. V. Ramabadran, J. P. Ashley, and M. J. McLaughlin, "Background Noise Suppression for Speech Enhancement and Coding", *Proc. IEEE Speech Coding Workshop* (Pocono Manor, PA), pp. 43–44, Sept. 1997.

[10] Telecommunications Industry Association, *Enhanced Variable Rate Codec (EVRC)*, Interim Standard TIA/EIA-IS-127, Dec. 1992 (on-line at `http://www.3gpp2.org` as document C.S0014-0, 1992-12).

[11] H. Gustafsson, S. E. Nordholm, and I. Claesson, "Spectral Subtraction Using Reduced Delay Convolution and Adaptive Averaging", *IEEE Trans. Speech, Audio Processing*, vol. 9, no. 8, pp. 799–807, Nov. 2001.