Audio Engineering Society

Convention Paper

Presented at the 119th Convention
2005 October 7–10 New York, New York USA

A New Class of Smooth Power Complementary Windows and their Application to Audio Signal Processing

Deepen Sinha[1], Anibal J. S. Ferreira[1,2]

[1] *ATC Labs, New Jersey, USA*

[2] *University of Porto, Portugal*

Correspondence should be addressed to D. Sinha (sinha@atc-labs.com)

ABSTRACT

In this paper we describe a new family of smooth power complementary windows which exhibit a very high level of localization in both time and frequency domain. This window family is parameterized by a "smoothness quotient". As the smoothness quotient increases the window becomes increasingly localized in time (most of the energy gets concentrated in the center half of the window) and frequency (far field rejection becomes increasing stronger to the order of 150 dB or higher). A closed form solution for such window function exists and the associated design procedure is described. The new class of windows is quite attractive for a number of applications as switching functions, equalization functions, or as windows for overlap-add and modulated filter banks. An extension to the family of smooth windows which exhibits improved near-field response in the frequency domain is also discussed. More information is available at http://www.atc-labs.com/technology/misc/windows.

1. INTRODUCTION

Windowing is a prevalent signal processing operation which has important implications to the performance of a range of signal processing algorithms. Several window functions have been proposed and utilized in signal processing. Some examples include Hanning Window, Hamming Window, Kaiser Window, Raised Cosine Window, etc. [1-4].

In many applications it is desirable to use a *symmetric* window function $w(n)$ (with a maximum value of 1.0) that satisfies an additional *power complementary condition*. To formalize these conditions, let's define a window function of size *2N*. The symmetry condition implies

$$w(n) = w(2N-n-1),\ n = 0,1,..,2N-1 \tag{1}$$

The *power complementary condition* is indicated by the constraint

$$|w(n)|^2 + |w(n+N)|^2 = 1.0 \ for \ n = 0,1,..,N-1 \tag{2}$$

The above condition is at times also referred to as Princen-Bradley condition [6-7]. For the reasons that will become obvious in the later sections we will refer to this condition as the power complementary condition.

In choosing a particular window function the frequency selectivity is an important consideration. The frequency selectivity is characterized by the stop-band energy of the frequency response of the window function. Furthermore, in many applications the time selectivity of the window is also an important consideration. In the framework of power complementary windows, time localization is quantified by the fraction of window energy concentrated in the center half of the window. Obviously the most time localized window is the rectangular window albeit at the cost of extremely poor frequency selectivity. It may also be noted that the popular Raised Cosine window [4, 6] is an example of a power complementary window that has relatively good frequency selectivity but poor time localizing. In general frequency selectivity and temporal localization are two conflicting requirements and it is difficult to design a window function that has good performance with respect to both of these performance criteria.

Here we describe a new family of smooth window functions satisfying the above constraints. The new class of windows is designed using the realization that the power complementary condition (2) is a time domain dual of an identical frequency domain condition used in the design of *wavelets with high degree of regularity* [2, 8, 16]. This leads to a systematic design procedure for a family of smooth windows parameterized by a *smoothness quotient*. The window family is unique in the sense that as the smoothness quotient is increased the window becomes better localized in both time and frequency domain. It may be noted that although we have used this family of windows in our work for a few years, such windows appear to be otherwise not well known in the signal processing literature.

2. THE CLASS OF SMOOTH WINDOWS

It is well known in the theory of signal approximation and splines [9] that the smoothness of a function increases as a function of the *highest order continuous differential* it possesses. The increased smoothness of a function results in improved frequency response for the function. We construct the proposed family of smooth window based on such differentiability criterion. In particular the window function $w(n)$ is constructed as samples of a continuous function $w(t)$ with compact support, in other words

$$\begin{aligned} & w(t) = 0 \ \forall t \notin [0,1] \\ & and \\ & w(n) = w(t)\big|_{t=\frac{i+0.5}{2N}}, \ i = 0,1,..,2N-1 \end{aligned} \tag{3}$$

The window is said to be smooth to a degree p (also called the *order* of the smooth window) if it possesses a continuous differential of order $p-1$. Here we describe a procedure for constructing a power complementary window with an arbitrary degree of smoothness, p. As will be seen below, higher degree of smoothness for a window function (satisfying the power complementary condition) results in improved temporal localization (in addition to improved far-field frequency response).

To describe the design procedure we first define a function $P_0(\xi)$ which may be considered as a complex periodic extension of $w(t)$ with conjugate symmetry; i.e.,

$$\begin{aligned} & P_0(-\xi) = P_0^*(\xi) \\ & P_0(\xi + 2\pi) = P_0(\xi) \\ & and \\ & w(t) = |P_0(\xi)|\big|_{\xi=\pi(2t-1)} \ \forall t \in [0,1] \end{aligned} \tag{4}$$

The power symmetry condition in (2) puts the following constraint on $P_0(\xi)$; i.e.,

$$|P_0(\xi)|^2 + |P_0(\xi+\pi)|^2 = 1 \tag{5}$$

And smoothness quotient, p, translates into the condition that $P_0(\xi)$ has a zero order p at $\xi = \pi$. Readers familiar with the theory of wavelets and filter banks will readily recognize that the above formulation makes this family of windows the time domain dual of the well known orthogonal wavelet filter bank used in 2 band QMF analysis that satisfy a *regularity or flatness* constraint [2, 8]. A solution for $P_0(\xi)$ that satisfies (4) can be found using a famous result due to Daubechies [8]. In particular we reproduce Proposition 4.5 from this reference [8] as follows:

Proposition [Daubechies, 1988]: Any Trigonometric polynomial solution of equation (4) above is of the form

$$P_0(\xi) = \left[\frac{1}{2}\left(1 + e^{j\xi}\right)\right]^p Q(e^{j\xi}) \tag{6}$$

where $p \geq 1$ is the order of zeros it possesses at $\xi = \pi$ and Q is a polynomial such that

$$\left|Q(e^{j\xi})\right|^2 = \sum_{k-0}^{p-1} \binom{N-1+k}{k} \sin^{2k} \frac{1}{2}\xi + \left[\sin^{2N} \frac{1}{2}\xi\right] R(\frac{1}{2}\cos\xi) \tag{7}$$

where R is an odd polynomial. In the above $\binom{n}{k}$ implies the Choice function [4]. The polynomial R is not arbitrary, and for our purpose (as will be explained below) may be set to 0.

A solution for $P_0(\xi)$ may be found by using *spectral factorization* technique on the polynomial, $Q(e^{j\xi})$. For this we substitute

$$\sin\frac{1}{2}\xi = \frac{z^{\frac{1}{2}} - z^{-\frac{1}{2}}}{2} \tag{8}$$

where z is a free complex variable. We then have

$$\left|Q(z)\right|^2 = \sum_{k-0}^{p-1} \binom{N-1+k}{k} \frac{\left[z - 2 + z^{-1}\right]^k}{4^k} \tag{9}$$

The procedure for constructing the smooth power complementary window of an order p is as follows:

Step 1: For the chosen p form the polynomial $\left|Q(z)\right|^2$ as in (7) above.

Step 2: Find all the roots of $\left|Q(z)\right|^2$. It may be noted that all the roots will occur in complex conjugate pairs which are also mirror imaged across the unit circle, in other words if $z = \rho e^{j\phi}$ is a root then so will be $z = \rho e^{-j\phi}, \frac{1}{\rho} e^{j\phi}, and \frac{1}{\rho} e^{-j\phi}$ (this results from the fact that $\left|Q(z)\right|^2$ is a zero phase, real polynomial).

Step 3: For each set of roots choose one pair which is on the same side of the unit circle. Form a polynomial $\overline{Q}(z)$ by combining all the chosen roots. We have $\left|\overline{Q}(z)\right| = \left|Q(z)\right|$ (this is spectral factorization).

Step 4: Combine $\overline{Q}(z)$ with the p^{th} order zero in (6) to form one possible polynomial, $P_0(z)$. The taps of this $P_0(z)$ form a FIR filter $p(n)$.

Step 5: $P_0(z)$ can be found by performing a Fourier analysis of the sequence $p(n)$. The window coefficients are then computed by sampling the Fourier transform of $p(n)$. as per equations (3) and (4).

Windows designed using the procedure for $p = 1,2,3$ are shown in Figure 1. The magnitude response of these windows is included in Figure 2(a), (b), (c) respectively. It is readily apparent that as p is increased the windows become more localized in time and at the same time the far field frequency responses improve substantially.

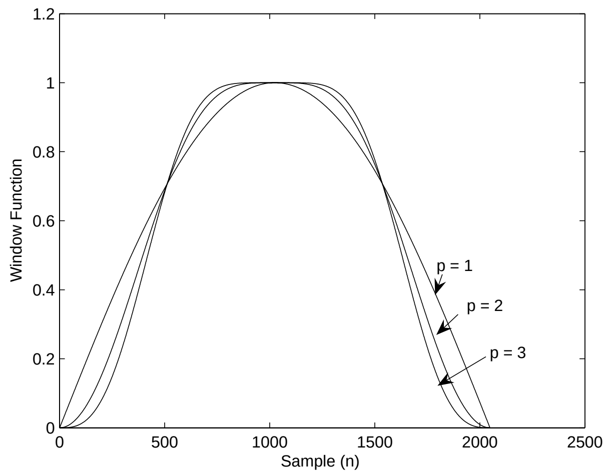

Figure 1. Example of smooth windows of various order. Three windows of size 2048 samples and order respectively of *p=1*, *p=2*, and, *p=3* are shown above.

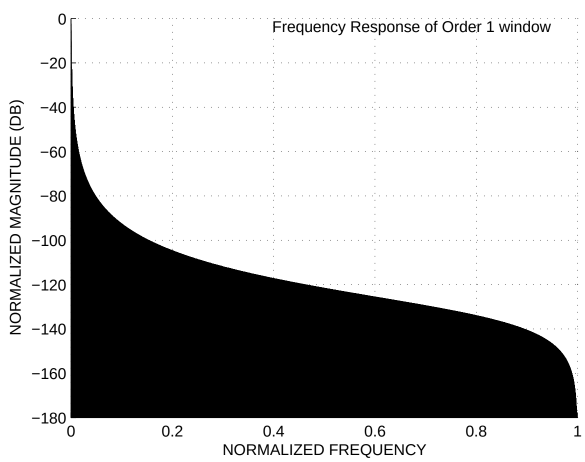

Figure 2(a). Frequency Response of an order 1 window (*p = 1*). This window is identical to the Raised Cosine (Sine) window (see section 3.3)

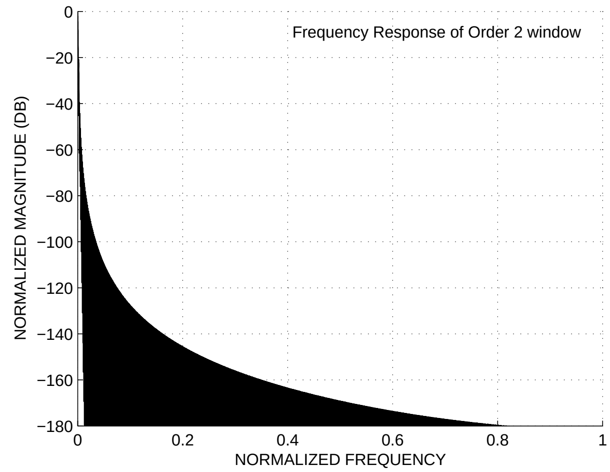

Figure 2 (b). Frequency Response of an order 2 window.

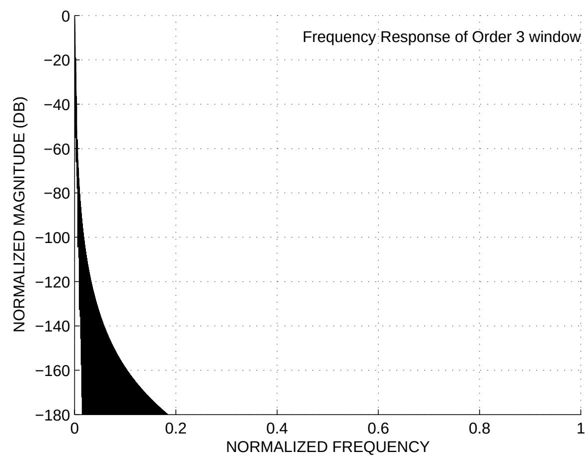

Figure 2 (c). Frequency Response of an order 3 window.

A couple of comments about the above procedure are in order. Firstly, it may be noted that although spectral factorization may yield several different choices for $\overline{Q}(z)$ all these result in identical magnitude response for $P_0(z)$ and hence the same window function. Furthermore, the choice of $R = 0$ in equation (7) is justified because inclusion of this will increase the order of $p(n)$ thus adversely affecting the frequency response of the window without adding to the temporal localization of the window.

3. FIGURE OF MERIT AND COMPARISON TO OTHER POWER COMPLEMENTARY WINDOWS

3.1. Figures of Merit for Windows

Several different criteria have traditionally been used for evaluating the performance of signal processing windows [4]. Some of these are particularly meaningful for the power complementary windows. Below we list a few of the most pertinent performance criteria.

Time Domain Localization and Aliasing

Time domain localization can be measured with the help of a criterion based on its energy concentration in the center half. A converse temporal spread criterion (*TDS*) based on the energy concentration outside the center band is defined as below.

$$TDS = \frac{1 - \sum_{n=N/2}^{3N/2-1} w^2(n)}{\sum_{n} w^2(n)} \qquad (10)$$

A lower value for *TDS (*measured in dB) is desirable. As described in the next section a prominent area of application for power complementary windows is in the area of perfect reconstruction Modulated Lapped Transforms which employ 50% overlap between consecutive analysis frames. These filter banks utilize the so-called Time Domain Alias mechanism [6]. Estimation of un-cancelled TDAC energy in the absence of the overlap-add alias cancellation operation is an important consideration. This is directly tied to the *TDS* criterion defined above.

Stop-Band Energy

The Frequency selectivity of Window may be measured with the help of its stop band energy. A stop band edge frequency f_c may be chosen as $f_c = k.(\pi / N)$ (where *k* is an integer).

$$SBE \;=\; \int_{f \geq f_c}^{\pi} \left| W(f) \right|^2 df \qquad (11)$$
$$(unweighted)$$

where $W(f)$ is the frequency response of the window. This value is also expressed in dB and a higher negative value indicates better frequency selectivity. For smaller values of k this measure is dominated by the so-called *Near Field Response* (i.e., energy of the first few side-lobes). For higher values of k this measure represents the *Far Field Response* (or Far Field Rejection) of the windows. In the above definition *SBE* is defined in an un-*weighted* form. A *weighted* definition of Stop Band Energy is as follows

$$SBE \;=\; \int_{f \geq f_c}^{\pi} \alpha(f) \left| W(f) \right|^2 df \qquad (12)$$
$$(weighted)$$

The weight function $\alpha(f)$ for example may be used to weight the frequency bands farther away from the center frequency more strongly.

Main Lobe Spectral Response

Spectral interval between Peak Gain and the -3.0dB and/or -6.0dB response level measures the width of the main lobe. The main lobe width, for example, affects the *coding gain* [17] of a transform using a particular window function.

Relative Strength of First and Second Side Lobes

The strength of first and second side lobes in relation to the spectral peak is also an interesting performance metric for many applications.

Scalloping Loss

Scalloping Loss (*SL*) is defined as below is sometimes a useful measure in evaluating the performance of a window function. It is defined as the ratio of the gain of a tone located at the offset of half frequency bin (for frequency bins in *DFT* analysis) with respect to the gain of the tone located at the center of the bin.

$$SL = \frac{\left|\sum_n w(n) e^{j\frac{\pi}{2N}}\right|}{\sum_n w(n)} \tag{13}$$

3.2. Some Other Power Complementary Windows

Several power complementary windows have been reported in literature. A commonly used window is the *Sine* (or *Raised Cosine*) window defined as [3]:

$$w_k = \sin\left[\frac{\pi}{2N}\left(k + \frac{1}{2}\right)\right] \tag{14}$$

It may be noted that this window is identical to a smooth window with a degree $p = 1$.(i.e., it possesses no continuous differential). Another such window is the one used in Vorbis audio coding scheme [15]. The Vorbis window is defined as:

$$w(n) = \sin\left[\frac{\pi}{2}\sin^2\left(\frac{\pi}{2N}\left(k + \frac{1}{2}\right)\right)\right] \tag{15}$$

In addition the AC-3 audio codec [10](as well as AAC audio codec [11]) uses a so-called "Dolby Window" which is also referred to as Kaiser-Bessel Derived Window (KBD) and is based on the integration of a Kaiser-Bessel function.. Here we use the Dolby window with $\alpha = 3.0$ for comparisons.

Furthermore, many researchers have proposed the utilization of optimized window functions, which are designed by solving a constraint optimization problem that minimizes a cost function (e.g., one based on Time Domain Aliasing and Stop Band Energy). The optimization may be performed using a non linear optimization technique which is part of MATLAB computing platform. One such optimized window reported in [14] ("AJF Optimum Window") is used in the comparisons in next section.

3.3. Smooth Windows Compared to Other Windows

Table I summarizes several figures of merit for various power complementary window functions. The following conclusions may be drawn

- It is obvious that the higher order smooth windows provide substantially better temporal localization (lower *TDS*) and Frequency Selectivity (Lower *SBE*, particularly in the Far Field Response) in comparison to the Raised Cosine (Sine) window at the cost of a slight increase in the main lobe width

- As the order of the smooth window is increased there is a very substantial improvement in Far Field Response ($k \geq 5$) and temporal localization (*TDS*). This gain comes at the cost of somewhat lower Near Field rejection.

- Many performance metrics of 2nd order smooth window match closely with the Dolby Window and Vorbis Window. The Far Field rejection however is substantially better. This is not apparent from the data in Table I, but is readily apparent in Figure 3(b), where the frequency response of this window is compared to the Dolby window.

- The 3rd order smooth window represents a good compromise with very high temporal localization and far field rejection with only a small penalty in terms of first few side lobe heights.

A direct comparison of 2nd and 3rd order smooth windows is presented in Figures 3 and 4 respectively. The windows shapes are shown in Figure 3(a) (2nd order with Dolby window) and Figure 4(a) (3rd order with Dolby window). The corresponding full spectrum plot (highlighting the far field rejection) is shown in Figure 3(b) and Figure 4(b). A closer comparative look at the Near Field responses of the 2nd and 3rd order smooth windows is shown in Figure 3(c) and Figure 4(c) respectively.

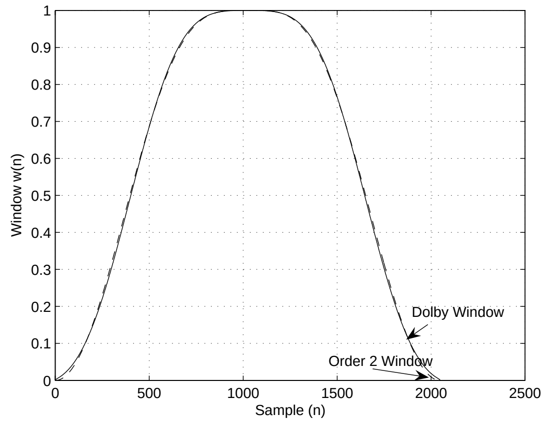

Figure 3(a). The 2nd order smooth window compared to the Dolby window

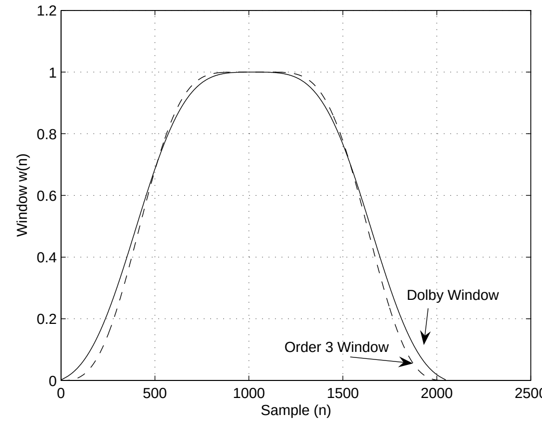

Figure 4(a). Shape of the 3rd order smooth window compared to the Dolby window

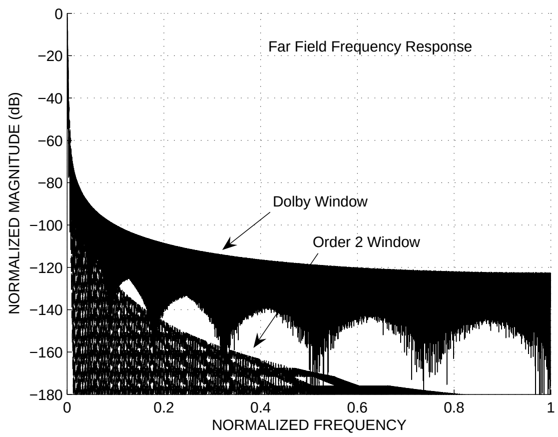

Figure 3(b). Spectrum of the 2nd order smooth window compared to the Dolby window

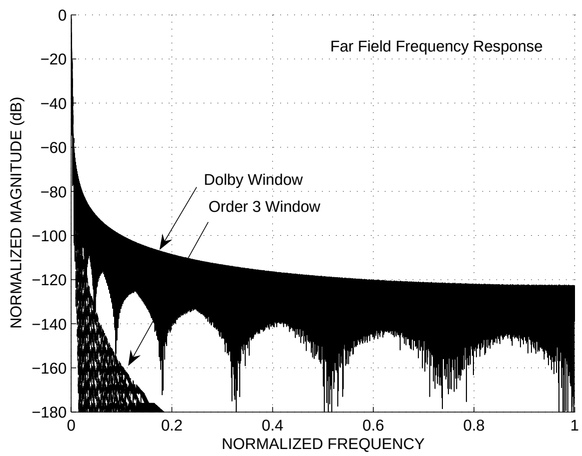

Figure 4(b). Spectrum of the 3rd order smooth window compared to the Dolby window

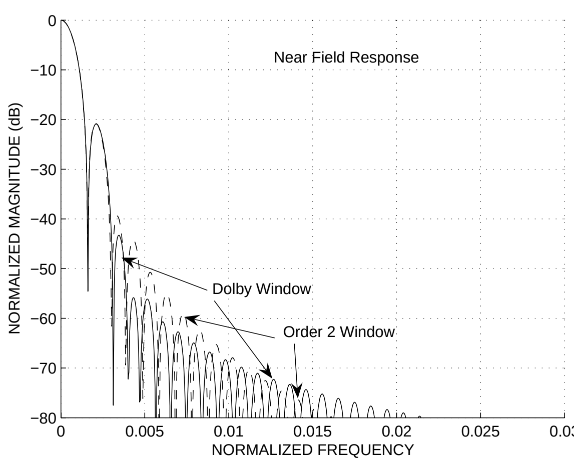

Figure 3(c). Near Field Spectrum of the 2nd order smooth window compared to the Dolby

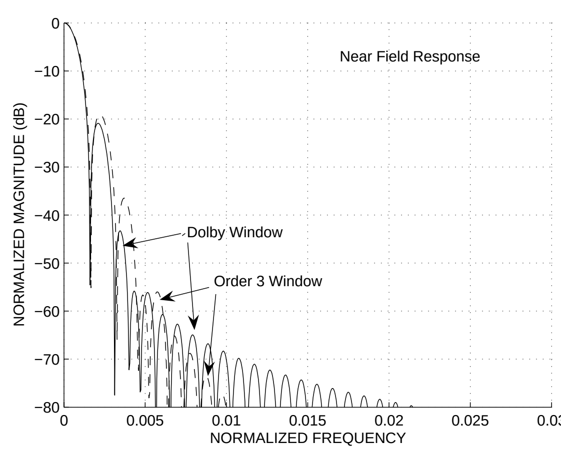

Figure 4(c). Near Field Spectrum of the 3rd order smooth window compared to the Dolby window

Table I: Figure of Merits for the Known and New Family of Windows

Performance Metric	Sine Window (Raised Cosine)	Dolby Window (alpha = 3.0)	Vorbis Window	AJF Optimum Window	Higher Order Smooth Windows		
					p=2	*P=3*	*P=3*
Main Lobe Width (-6.0 dB, normalized frequency)	0.018	0.02	0.02	0.019	0.02	0.02	0.02
TSD (dB)	-15	-18	-18	-17	-18	-20	-21
SBE(dB) (Un-Weighted) $f_c = 2.(\pi / N)$	-22	-23	-22	-23	-22	-21	-18
SBE(dB) (Un-Weighted) $f_c = 3.(\pi / N)$	-25	-38	-37	-32	-37	-36	-31
SBE(dB) (Un-Weighted) $f_c = 4.(\pi / N)$	-27	-43	-42	-34	-40	-44	-38
SBE(dB) (Un-Weighted) $f_c = 5.(\pi / N)$	-29	-44	-44	-36	-44	-53	-57
SL (dB)	-1.1	-0.85	-0.8	-0.9	-0.85	-0.8	-0.8
First Side Lobe(dB)	-23	-21	-20	-22	-21	-19	-18
Third Side Lobe(dB)	-36	-56	-47	-44	-44	-57	-56
Fifth Side Lobe(dB)	-43	-61	-56	-48	-55	-68	-69

4. OPTIMALITY OF THE PROPOSED WINDOW FAMILY

The proposed family of smooth windows is optimal in the sense of being *maximally flat*. The term maximally flat is often used in the context of filter/filter bank design [2] and refers to the closeness of the filter magnitude response to an ideal rectangular function. The time domain duality of the smooth windows indicates similar characteristics for these windows in time domain. In other words it can be shown that for a fixed cost a smooth window of an appropriate order is closest to the rectangular window in time domain. The cost in this case is in the form of main lobe width and near field frequency selectivity. The maximal flatness of the window function comes from requiring that the window posses a continuous differential (hence a vanishing differential) at the edges. The flatness at the edges leads to flatness at the center of the window due to the power complementary condition.

Also as noted earlier, another very interesting aspect of the smooth window family is that as the order (p) is increased the window approaches the rectangular window in time domain at the same time its far field frequency selectivity improves substantially.

5. APPLICATIONS OF THE SMOOTH POWER COMPLEMENATRY WINDOW

5.1. Modified Discrete Cosine Transformation (MDCT)

The lapped MDCT transform is quite popular in audio compression algorithms [10-13]. In MDCT a real sequence of length $2N$, $x(n)$, is transformed into a real sequence of length N, $f(k)$, as

$$f(k) = \sum_{n} w(n) \cdot x(n) \cdot \cos\left[\frac{\pi}{N} \left(k + \frac{1}{2} \right) \left(n + \frac{1}{2} + \frac{N}{2} \right) \right] \tag{16}$$

The MDCT filter bank utilizes the symmetry of cosine basis function to achieve perfect reconstruction. This mechanism is often referred to as Time Domain Alias Cancellation (TDAC). The frequency selectivity of the windows plays an important role in the coding gain of the transform as well as the cleanliness of harmonic reconstruction in a signal rich in harmonics. Also lower *TDS* is important to control the temporal spread of quantization noise and un-cancelled time domain alias terms. A smooth window of order $p = 2$ or 3 was found to improve the performance of audio codecs significantly.

5.2. Energy Preserving Cross Fade Function

Use of a cross fade or switching function is prevalent in audio processing, music synthesis and other audio applications [5]. A power complementary cross fade has the advantage that it preserves the instantaneous energy when making the transition. Use of higher order smooth windows (rather than a raised cosine function) allows for a more rapid transition that is still smooth and low in frequency smearing. This is the direct result of improved time frequency localization of the smooth windows.

5.3. Smooth Inter-band Coupling in Multi Band Audio Processing

The smooth windows also find application as a transition function (for multi-band adaptive processing (inter-band transition). Since windowing in frequency domain is equivalent to convolution in time the superior far field response of the higher order smooth windows results in lower temporal ringing due to frequency domain processing (particularly when high frequency resolution analysis is used.

5.4. Spectrum Analysis and Overlap-Add Processing with Low Effective Overlap

DFT analysis with overlapping blocks is frequently used [3, 5]. The advantage of using overlapping blocks is that discontinuities at the boundary are avoided and better frequency selectivity is achieved. However, because of the overlap, the effect of processing in a single block is spread over adjacent blocks. By using the higher order smooth windows one is able to localize the effects better in time while preserving the advantages of overlap operation.

Various audio samples illustrating the advantage offered by higher order smooth windows in relation to the above applications are available at http://www.atc-labs.com/technology/misc/windows.

6. EXTENSION TO THE SMOOTH WINDOWS FAMILY

As noted above the improved time and frequency (far field) selectivity of the smooth window family is at times associated with a slight (or significant for higher orders) increase in the first few side lobes of the frequency spectrum. In some applications improvement in temporal and far field rejection beyond a certain point is not of particular use but rather a better near field response is desirable. In these applications a mechanism that allows for the trade-off of some of the localization gains for an improved near field response is desirable. This can be accomplished using a set of windows derived from a smooth window of an appropriate order. Details of such windows derived from the smooth window family will be presented during our convention presentation and future publications.

7. CONCLUSIONS

We have introduced a family of smooth power complementary windows and described a constructive procedure for generating a smooth window with any desired degree of smoothness. Depending upon the application, and its associated time and frequency selectivity requirement, a window of desired order can be generated and employed. The smooth windows have utility in a variety of signal processing applications such as audio coding, processing, analysis, equalization, music synthesis, etc. Further details on the this class of windows as well as future extensions may be found at http://www.atc-labs.com/technology/misc/windows.

8. ACKNOWLEDGEMENTS

The authors would like to acknowledge very useful discussions with Mr. James D. ("JJ") Johnston on the topic of window characteristics and design.

9. REFERENCES

[1] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*, Prentice-Hall, 1975.

[2] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice Hall, 1993.

[3] Henrique S. Malvar, *Signal Processing with Lapped Transforms*, Artech House, 1992.

[4] Alexander D. Poularikas, *The Handbook of Formulas and Tables for Signal Processing*, CRC Press, 1999.

[5] Vijay K. Madisetti and Douglas B. Williams, *The Digital Signal Processing Handbook*, CRC Press, 1998.

[6] John P. Princen and Alan B. Bradley, "Analysis/synthesis filter bank design based on time domain aliasing cancellation," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-34 (5), 1153-1161 October, 1986.

[7] J. P. Princen, A. W. Johnson, and A. B. Bradley, "Subband/Transform Coding Using Filter Bank Designs Based on Time Domain Alias Cancellation," in *IEEE Int. Conf. on Acoust., Speech and Signal Proc.*, 1987, pp. 2161-2164.

[8] I. Daubechies, "Orthonormal bases of compactly supported wavelets," Commun. Pure Appl. Math., vol. 41, pp. 909--996, 1988.

[9] Milton Abramowitz and Irene A. Stegun, *Handbook of Mathematical Functions*, (1964) Dover Publications, New York. ISBN 486-61272-4

[10] Mark Davis, "The AC-3 Multichannel Coder," *95th Convention of the Audio Engineering Society*, October 1993, Preprint n. 3774.

[11] Marina Bosi et al., "ISO/IEC MPEG-2 Advanced Audio Coding," *101st Convention of the Audio Engineering Society*, November 1996, Preprint n. 4382.

[12] Anibal J. S. Ferreira, *Spectral Coding and Post-Processing of High Quality Audio*, Ph.D. thesis, Faculdade de Engenharia da Universidade do Porto-Portugal, 1998, http://telecom.inescn.pt/doc/phd_en.html.

[13] J. D. Johnston, D. Sinha, S. Dorward, and S. R Quackenbush, "AT&T Perceptual Audio Coding (PAC)," *in AES Collected Papers on Digital Audio Bit-Rate Reduction*, N. Gilchrist and C. Grewin, Eds. 1996, pp. 73-82.

[14] Aníbal J. S. Ferreira, *Convolutional Effects in Transform Coding with TDAC: An Optimal Window*, IEEE Transactions on Speech and Audio Processing, vol. 4, n° 2, pp. 104-114, March 1996.

[15] Ogg Vorbis Documentation at http://www.xiph.org/ogg/vorbis/docs.html.

[16] D. Sinha, *Low bit rate transparent audio compression using adapted wavelets.* Ph.D. thesis, University of Minnesota, 1993.

[17] N.S. Jayant and P. Noll, *Digital Coding of Waveforms,* Prentice Hall, 1984.