Signal Processing 4 (1982) 79–84
North-Holland Publishing Company

SHORT COMMUNICATION

ALIASING ERRORS IN SHORT-TIME ANALYSIS

R. YARLAGADDA
School of Electrical Engineering, Oklahoma State University, Stillwater, OK 74078, USA

Jont B. ALLEN
Acoustics Research Department, Bell Laboratories, Murray Hill, NJ 07974, USA

Received 31 May 1980
Revised 5 January 1981

Abstract. This paper analyzes the error made in the synthesis of a digital signal from its short-time Fourier transform. The error description is given in terms of the digital Poisson summation formula. The results are then discussed in terms of computer simulations using rectangular, Hamming, Dolph–Chebyshev, and Kaiser windows.

Zusammenfassung. Dieser Beitrag analysiert den Fehler, der bei der Synthese digitaler Signale aus ihrer Kurzzeittransformierten auftritt. Der Fehler wird mit Hilfe der digitalen Poissonsummenformel beschrieben. Die Ergebnisse werden dann anhand einer Rechnersimulation diskutiert, bei welcher Rechteck-, Hamming-, Dolph–Tschebysheff- und Kaiser-fenster verwendet werden.

Résumé. Cet article est une étude de l'erreur liée à la synthèse d'un signal digital à partir de sa transformée évolutive. L'erreur est décrite au moyen de la formule de Poissen digitale. Les résultats de simulations sont présentés pour le cas des fenêtres de Hamming, Dolph–Tschebysheff, Kaiser et rectangulaire.

Keywords. Short-time analysis, aliasing errors, window functions, digital Poisson sum.

Introduction

The short-time Fourier transform of a signal $x(n)$, at time mR, is defined as

$$X_m(e^{j\omega_p}) = \sum_{l=mR-L+1}^{mR} x(l)w(mR-l)\,e^{-j\omega_p l} \quad (1a)$$

$$= \mathbf{F}[x(l)w(mR-l)] \quad (1b)$$

$$= \mathbf{F}[x_{mR}(l)], \quad p=0,1,\ldots,L-1, \quad (1c)$$

where $\omega_p = 2\pi p/L$ are the frequencies at which the short-time Fourier transform is evaluated, R is the decimation period (in samples) between adjacent samples of the short-time transform of the signal, $w(n)$ is a causal window function of length L and $\mathbf{F}[\cdot]$ denotes the discrete Fourier transform operation (DFT). It has been shown in [1] that any signal $x(n)$ can be recovered within small aliasing errors from its short-time transform $X_m(e^{j\omega_p})$, when the short-time transform is specified at or above the Nyquist decimation rate of the window [1–3], by 'overlap adding' the windowed signals $x_{mR}(n)$

$$x(n) = x(n)\left[\frac{R}{W(e^{j0})}\sum_{m=-\infty}^{\infty} w(mR-n)\right] \quad (2a)$$

$$= \frac{R}{W(e^{j0})}\sum_{m=-\infty}^{\infty} x_{mR}(n), \quad (2b)$$

where $W(e^{i0})$ corresponds to the zero frequency value of the window.

The derivation of eq. (2) makes use of the approximate relation

$$\frac{R}{W(e^{i0})}\sum_{m=-\infty}^{\infty} w(mR-n)\approx 1. \qquad (3)$$

This relation is exact if $w(n)$ is bandlimited to a frequency of $(L/2RT)$, where L is the length of the window in samples and T is the length of the time interval for which $w(t)$ is nonzero [1]. Since $w(n)$ is time limited, its spectrum cannot be bandlimited. Hence in order to obtain an equality, eq. (3) must be modified to include an error term

$$\frac{R}{W(e^{i0})}\sum_{m=-\infty}^{\infty} w(mR-n)=1+e_R(n), \qquad (4)$$

where $e_R(n)$ takes into consideration the aliasing errors due to the non-bandlimited window function. Note that the subscript R is introduced to denote that the error signal $e_R(n)$ depends on R, the decimation period.

The purpose of this paper is to quantify the magnitude of the error term $e_R(n)$. This may be done in terms of a discrete version of the Poisson summation formula. We apply this result in order to find explicit formulae for $\max_n(|e_R(n)|)$ and $\mathrm{RMS}(e_R(n))$. These results are then illustrated using rectangular, Hamming, Dolph–Chebyshev, and Kaiser windows as a function of R.

1. Numerical examples

For notational purposes, Fig. 1 identifies the basic definitions of the characteristic time T and the frequency length F for a Dolph–Chebyshev window. T denotes the length of the time interval for which $w(t)$ is nonzero, $W(f)$ denotes the Fourier transform of $w(t)$, and F denotes the smallest frequency for which $|W(f)|<\delta$ for $|f|>\frac{1}{2}F$. F is approximately equal to the main lobe width. It is clear that T and F can be defined for other windows similarly. Using $(1/F)$ as the sampling period (i.e., the Nyquist period), then $T/(1/F)=TF$ gives

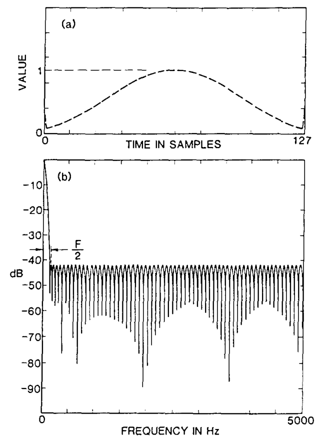

Fig. 1. Dolph–Chebyshev window with a −42 dB side-lobe level and its spectrum.

the number of sample periods in time T for the short-time analysis output. Therefore the ratio (L/TF) gives the Nyquist upper bound for the decimation period R identified earlier. The product TF (the time-bandwidth product of the window) will be useful, and is identified here by Q,

$$Q=TF. \qquad (5)$$

Using this notation, $R<L/Q$. For a Hamming window which has a −42 dB side-lobe level, Q is approximately 4.0 [1]. For a Dolph–Chebyshev window, Q is approximately [4, eq. (17)]

$$Q\cong\frac{2}{\pi}\ln\left[\frac{2}{\delta}\right]. \qquad (6)$$

For example, for a Chebyshev window with a −42 dB ($\delta=0.0079$) side-lobe level, Q is about 3.52. For the sake of comparison, windows with the same Q were used here. Thus, a Chebyshev window with $Q=4$, which has $\delta=0.00374$ (−48.5 dB) side-lobe level was chosen. For a is concerned. The Kaiser window has the smallest RMS error and is also flexible. Therefore, the Kaiser window is the most suitable window studied here, in terms of its RMS error. From this analysis it appears that one possible measure of window quality is $e_{\mathrm{RMS}}/e_{\max}$ evaluated at its cutoff $R = L/Q$.

Conclusions

In this paper we have first applied the digital Poisson sum formula to the problem of aliasing errors in short-time Fourier Synthesis (overlap add) and then directly compared various windows based on our error criterion. We believe that the formulae (eqs. (12), (13)) are useful in specifying the aliasing errors incurred during short-time Fourier synthesis, spectral analysis [6] and system identification by DFT [6].

Acknowledgement

The authors thank L. R. Rabiner for providing the computer programs for the generation of Chebyshev windows.

References

[1] Jont B. Allen, "Short-term spectral analysis, synthesis, and modification by Discrete Fourier Transform", *IEEE Trans. Acoust. Speech Signal Process*, Vol. ASSP-25, No. 3, June 1977, pp. 235–238.

[2] Jont B. Allen and L.R. Rabiner, "A unified approach to short-time Fourier analysis and synthesis", *Proc. IEEE*, Vol. 65, No. 11, Nov. 1977, pp. 1558–1564.

[3] R.W. Schafer and L.R. Rabiner, "Design and simulation of a speech analysis-synthesis system based on short-time Fourier analysis", *IEEE Trans. Audio Electroacoust.*, Vol. AU-21, No. 3, June 1973, pp. 165–174.

[4] O. Herrmann, L.R. Rabiner, and D.S.K. Chan, "Practical design rules for optimum finite impulse response lowpass digital filters", *Bell System Tech. J.*, Vol. 52, No. 6, July 1973, pp. 769–799.

[5] J.F. Kaiser, "Nonrecursive digital filter design using the I_0-sinh window function", *Proc. 1974 IEEE Int. Symp. on Circuits and Syst.*, April 22–25, 1974, pp. 20–23.

[6] L.R. Rabiner and Jont B. Allen, "Short-time Fourier analysis techniques for FIR system identification and power spectrum estimation," *IEEE Trans. Acoust. Speech Signal Process*, Vol. ASSP-27, April 1979, pp. 182–192.

[7] J.B. Allen and L.R. Rabiner, "Unbiased spectral estimation and system identification using short-time analysis methods", *Bell System Tech. J.*, Vol. 58, No. 8, Oct. 1979, pp. 1743–1763.

[8] A. Papoulis, *The Fourier Integral and Its Applications*, McGraw-Hill, New York, 1962, p. 47.

[9] A.V. Oppenheim and R.W. Schafer, *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1975, p. 97 and p. 125.

[10] L.R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1975, p. 19.

Kaiser window, assuming that F is defined by the first spectral zero of the window, Q is approximately [5, eq. (2)],

$$Q \cong 2\sqrt{1+(\alpha/\pi)^2},$$

where α is the Kaiser window design parameter. Thus, if $Q=4$, $\alpha = 5.44$.

For numerical reasons, only a finite sum for eq. (4) was considered. That is,

$$\tilde{e}_R(n) = \frac{R}{W(e^{j0})} \sum_{m=1}^{M_1} w(mR+n) - 1. \qquad (7)$$

M_1 was chosen to be large enough to assure that $\tilde{e}_R(n)$ was equal to $e_R(n)$ over the midrange of n, namely $M_1 R \gg L$.

The computed result of eq. (7) is plotted (on a dB scale) in Figs. 2a and3a respectively for Chebyshev and Hamming windows for $1 \le n \le 1024$ with $Q=4$, $L=128$, $R=32$, and $M_1 = 28$. From the figures, one can see that except at the beginning and at the end of the interval [6, 7], the error function is periodic and this periodic part is identical to $e_R(n)$ of eq. (4). For our analysis, consider

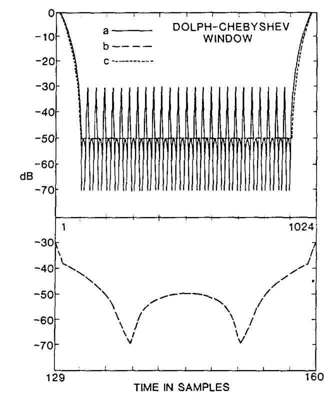

Fig. 2. For a Dolph–Chebyshev window with a −48.5 dB side-lobe level: a. error function $\tilde{e}_R(n)$, b. error function $e_{RP}(n)$, c. windowed average error $e_{RA}(k)$.

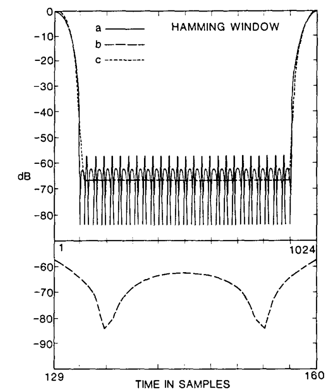

Fig. 3. For a Hamming window: a. error function $\tilde{e}_R(n)$, b. error function $e_{RP}(n)$, c. windowed average error $e_{RA}(k)$.

a period in the periodic portion of $\tilde{e}_R(n)$ in (7), say $\tilde{e}_R(n)$, $n = L+1, L+2, \ldots, L+R$. Here onwards, one such period is identified as $e_{RP}(n)$. The errors in that central periodic portion are small, while the errors in the end portion are large. Plots of $e_R(n) = e_{RP}(n)$ are given in Figs. 2b and 3b for Chebyshev and Hamming windows ($Q=4$) respectively. Dashed curves labeled 2c and 3c display plots of a one period average of $\tilde{e}_R(n)$, namely

$$e_{RA}(k) = \frac{1}{R} \sum_{k-R/2}^{k-R/2} \tilde{e}_R(n).$$

Analytical bounds on $e_R(n) = e_{RP}(n)$, in terms of the RMS error and peak error are derived in Section 3.

The error $e_R(n)$ for the case of a Hamming window (Fig. 3) is characteristic of the general shape of the error function for many of the well-known window functions. As discussed in Section 3, the rectangular window has a peculiarity that the period error $e_{RP}(n)$ is identically zero for certain values of R. However since the side lobes of the spectrum of a rectangular window drop off at 12 dB/oct, the error curves for a rectangular window do not compare favorably with the error curves of the other windows for arbitrary values of R, as will be shown analytically.

2. Digital Poisson summation formula

If $w(t)$ is any arbitrary function, with $W(f)$ its Fourier transform, then according to the classical Poisson formula [8]

$$\sum_{m=-\infty}^{\infty} w(t+mT)=\frac{1}{T}\sum_{m=-\infty}^{\infty} W\left(\frac{m}{T}\right) e^{j2\pi(m/T)t}$$

The corresponding digital equivalent formula is given as follows. Let $w(n)$ be any sequence with

$$W(z)=\sum_{l=-\infty}^{\infty} w(l)z^{-l}, \tag{8}$$

its z transform. Then the digital Poisson formula is [9, eqs. (3-15), (3-16)],

$$\sum_{m=-\infty}^{\infty} w(n+mR) =\frac{1}{R}\sum_{p=0}^{R-1} W(e^{j(2\pi/R)p})\, e^{j(2\pi/R)pn}. \tag{9}$$

Eq. (9) differs from the classical "analogue" Poisson formula in that the righthand sum is finite and is over the z transform of $w(n)$ evaluated on the unit circle at R equispaced points.

3. Analysis of overlap-add errors

In this section, the error, $e_R(n)$, in eq. (4) is quantified using the results of the last section. From eq. (9) and eq. (4), and letting $\omega'_p=2\pi p/R$, we have

$$e_R(n)=\frac{1}{W(e^{j0})}\sum_{p=0}^{R-1} W(e^{j\omega'_p})\, e^{jn\omega'_p}-1 =\sum_{p=1}^{R-1}\left[\frac{W(e^{j\omega'_p})}{W(e^{j0})}\right] e^{jn\omega'_p}, \tag{10}$$

since the $p=0$ term is identically 1. This implies that the error can be computed from the samples of the z transform of the window sequence. It is clear that when $R=1$ the error is exactly zero. For $R>1$, eq. (10) gives a simple procedure for the computation of $e_R(n)$ using an inverse DFT. This can be seen by noting that the DFT coefficients of $e_R(n)$ can be determined from eq. (10) and are given by

$$E_R(p)=\mathrm{DFT}(e_R(n)) =\begin{cases}0, & p=0\\ \dfrac{R}{W(e^{j0})}W(e^{j\omega'_p}), & 1\le p\le R-1.\end{cases} \tag{11}$$

Two useful bounds on $e_R(n)$ can be obtained. First, using the triangle inequality on eq. (10), we have

$$|e_R(n)|\le\frac{1}{W(e^{j0})}\sum_{p=1}^{R-1}|W(e^{j\omega'_p})|. \tag{12}$$

We define $\hat{e}_{\max}$ as the right side of eq. (12). Second, the root mean square error (RMS(e)) per sample can be obtained using Parseval's theorem [9] and eq. (11), where the RMS error per sample is defined as

$$e_{\mathrm{RMS}}=\left(\frac{1}{R}\sum_{n=0}^{R-1} e_R^2(n)\right)^{1/2} =\frac{1}{W(e^{j0})}\left(\sum_{p=1}^{R-1}|W(e^{j\omega'_p})|^2\right)^{1/2}. \tag{13}$$

Interestingly, if

$$\left|\frac{W(e^{j\omega'_p})}{W(e^{j0})}\right|\le\delta \quad \text{for } p=1,\ldots,R-1, \tag{14}$$

then

$$e_{\mathrm{RMS}}\le\sqrt{R-1}\,\delta=\hat{e}_{\mathrm{RMS}}. \tag{15}$$

Note that eq. (14) is realistic when R satisfies the Nyquist decimation criterion discussed earlier and δ bounds the out-of-band-peak spectral error. For larger values of R, eq. (14) and, therefore, eq. (15) are not valid. Eq. (15) is an excellent bound relating the RMS error per sample, window decimation period R, and the side-lobe attenuation of the window δ, when the Nyquist condition $R<L/Q$ has been satisfied.

For a rectangular window, $e_R(n)$ is exactly zero for some values of R. This can be seen by noting that $W(e^{j\omega_p})=0$ for $\omega_p=p2\pi/L$, $p\ne 0$, where L corresponds to length of the window [10]. When $\omega'_p=p2\pi/R$, $p=1,\ldots,R-1$, coincide with these frequencies, the error is zero. This happens when L is divisible by R. An illustrated numerical example of this is given in the next section.

4. Computer simulations

In this section, computer simulations of some of the results in the last few sections are given. Let

$$e_{\max}=\max|e_{Rp}(n)|,$$

$$\hat{e}_{\max}=\frac{1}{W(e^{j0})}\sum_{p=1}^{R-1}|W(e^{j\omega'_p})|.$$

From our theory leading to eq. (12), we have

$$\hat{e}_{\max}\ge e_{\max}.$$

Fig. 4 gives the results for a 128 point Dolph-Chebyshev window having a 42 dB side-lobe attenuation with R varying from 4 to 128. Figs. 4a and 4b show respectively the theoretic $\hat{e}_{\max}$ and measurement $e_{\max}$ (eq. (12)) for the peak error. Figs. 4c and 4d respectively give the measured (true) and theoretic RMS error (right-hand side of eq. (13)). The spectral values required by the theory are computed using a 1024 point

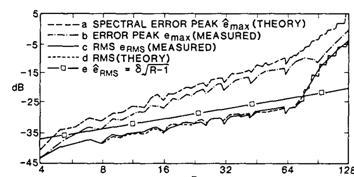

Fig. 4. For a Dolph–Chebyshev window with a −42 dB side-lobe level: a. Plot of $\hat{e}_{\max}$, b. Plot of $e_{\max}$, c. Plot of e_{RMS} from time values, d. Plot of $\hat{e}_{\mathrm{RMS}}$ from spectral values, e. Plot of $\hat{e}_{\mathrm{RMS}}=\sqrt{R-1}\delta$, $\delta=0.0079$. Squares are used on this curve to distinguish this curve from others.

fast Fourier transform (FFT). The deviations in the two curves Figs. 4c, d is due to the fact that the $|W(e^{j\omega'_p})|$ cannot be computed exactly for all p and R values required from a single 1024 point FFT. Fig. 4e gives the plot of $(\sqrt{R-1})\delta$, illustrating the utility of eq. (15) for the Chebyshev window.

Figs. 5a, b, c and d give respectively the RMS errors e_{RMS} (eq. (13)) for four different 128 point windows, a Dolph-Chebyshev, a Hamming window, a rectangular window, and a Kaiser window with R from 4 to 128. Fig. 5e is $(\sqrt{R-1})\delta$, with

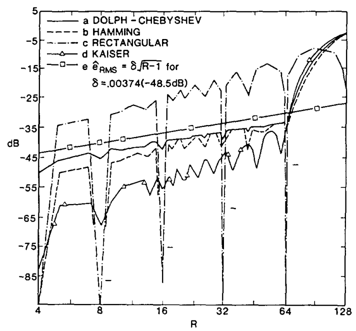

Fig. 5. Comparison of e_{RMS} for: a. Dolph–Chebyshev window, b. Hamming window, c. Rectangular window, d. Kaiser window, e. Plot of $\hat{e}_{\mathrm{RMS}}=\sqrt{R-1}\delta$, $\delta=0.00374$ (−48.5 dB). These results are computed using the time domain method for each integer value of R, $4\le R\le 128$. Squares and triangles are used on some curves to distinguish these curves from others.

$\delta=0.00374$ (−48.5 dB). The rectangular window has zero RMS error for R a power of 2. The Hamming, Chebyshev, and Kaiser windows all have time-band width products Q of 4. For the case of $Q=4$, the Chebyshev window has a slightly larger RMS error (and much larger peak error, Figs. 2 and 3) when compared to the Hamming window. However, the Chebyshev window is more flexible as far as the range of side-lobe attenuation