The Bark Bilinear Transform

Julius O. Smith, III
CCRMA
Stanford University
Stanford, CA 94305 USA
jos@ccrma.stanford.edu

Jonathan S. Abel
Abel Innovations
184 Tennyson Avenue
Palo Alto, CA 94301 USA
jsa@ccrma.stanford.edu

Abstract

Use of a bilinear conformal map to achieve a frequency warping nearly identical to the Bark scale is described. Because the map takes the unit circle to itself, its form is that of an allpass transfer function. Since it is a first-order map, it preserves the model order of rational systems. A direct-form expression for computing the optimal allpass coefficient as a function of sampling rate is developed, and a filter design example is presented.

1 Introduction

With the increasing use of frequency-domain techniques in audio compression, there is increasing emphasis on psychoacoustic based spectral measures [4]. One of the important spectral transformations for audio is warping to the Bark frequency scale. Based on the results of many psychoacoustic experiments, the Bark frequency scale is designed so that the critical bands of human hearing are uniform in size [1, 2].

The bilinear conformal map, defined by the substitution

$$z = \mathcal{A}_\rho(\zeta) \triangleq \frac{\zeta + \rho}{1 + \zeta\rho} \tag{1}$$

takes the unit circle in the z plane to the unit circle in the ζ plane in such a way that low frequencies are stretched and high frequencies are compressed, much like a transformation to the Bark frequency scale. Note that the conformal map $\mathcal{A}_\rho(\zeta)$ is identical in form to the transfer function of the first-order allpass filter with its pole at $\zeta = -1/\rho$; for this reason, we also call it the *first-order allpass transformation*, and ρ the *allpass coefficient*.

Since the allpass mapping possesses only a single degree of freedom, we have no reason to expect a particularly good match to the Bark frequency axis, even for an optimal choice of ρ. It turns out, however, that the match is *surprisingly good* over a wide range of sampling rates, as illustrated in Fig. 1. The purpose of this paper is to spread awareness of this important fact and present new methods for computing the optimal warping parameter ρ as a function of sampling rate.

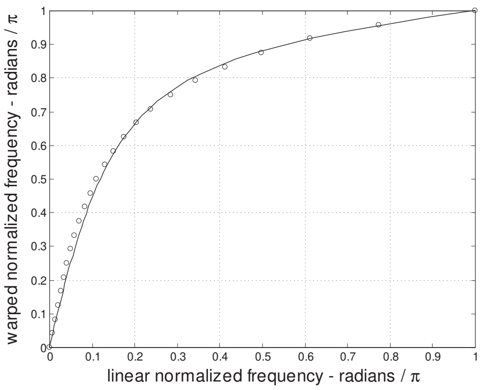

Figure 1: Bark and Allpass Warpings. Bark (discrete) and allpass (solid) frequency warpings are shown for a sampling rate of 31 kHz (the highest possible without extrapolating the published Bark scale bandlimits). The quality of the fit improves as the sampling rate is reduced. Published Bark band-edges are indicated by '∘', and the band-centers are plotted as '·'.

2 Prior Work

The earliest usage we've found applying the allpass transformation in audio signal processing is "warped linear prediction" [9, 10]. In warped LPC, the frequency axis "seen" by LPC is made to approximate a Bark scale using the first-order allpass transformation. It was noted in [9] that setting the allpass coefficient to 0.47 gave a "very good approximation to the subjective Bark scale based on the critical bands of the ear" at a 10 kHz sampling rate. It was concluded that low-order LPC was helped significantly by the frequency warping, because the first and second formants of speech become well separated on a Bark scale and therefore better resolved by a low-order predictor. However, higher order LPC fits could actually be made worse, e.g., due to splitting of the first formant as a result of four poles being used in the LPC fit instead of two.

Independently, the Bark bilinear transformation was developed for use in audio digital filter design [11, 12]. In that work, the frequency response fit was carried out over an approximate Bark frequency scale provided by the allpass transformation. The allpass coefficient ρ was optimized as a function of sampling rate using the method of bisection under a least-squares norm on the error between the allpass and Bark frequency warpings. The root mean square errors were found to range from $0.0034 f_s$ at $f_s = 6$ kHz to $0.0068 f_s$ at $f_s = 27$ kHz, where f_s denotes the sampling rate. The frequency warp dictated by the optimal $\mathcal{A}_\rho$ determined an interpolated resampling of the desired filter frequency response $H(e^{j\omega})$ which converted its support to an approximate Bark frequency scale $H(e^{j\omega}) = H[\mathcal{A}_\rho(e^{ja(\omega)})]$. Any filter design method could then be carried out to give an optimal match $H^*[e^{ja(\omega)}]$ over the warped, sampled frequency response. Many filter design methods were compared and evaluated with respect to their audio quality. Finally, the optimal warped filter $H^*(\zeta)$ was unwarped by applying the inverse allpass transformation $\mathcal{A}_{-\rho}$ to the warped filter transfer function using polynomial manipulations to obtain $H^*[\mathcal{A}_{-\rho}(z)] = H^*(\zeta)$.

The first-order allpass transformation has been used traditionally in digital filter design to scale the cut-off frequency of digital lowpass and highpass filters, preserving optimality in the Chebyshev sense [8]. Higher order allpass transformations have been used to convert lowpass or highpass prototype filters into multiple bandpass/bandstop filters [7].

The critical feature of the first-order conformal map is that it does not increase filter order; it is the most general order-preserving frequency-warping transformation for rational digital filters. In view of this constraint, it is remarkable indeed that a "free" transformation such as this can so closely match the Bark frequency scale.

3 The Bark Frequency Scale

The Bark frequency scale ranges from 1 to 24 Barks, corresponding to the first 24 critical bands of hearing [2]. The published Bark band edges are given in Hertz as [0, 100, 200, 300, 400, 510, 630, 770, 920, 1080, 1270, 1480, 1720, 2000, 2320, 2700, 3150, 3700, 4400, 5300, 6400, 7700, 9500, 12000, 15500]. The published band centers in Hertz are [50, 150, 250, 350, 450, 570, 700, 840, 1000, 1170, 1370, 1600, 1850, 2150, 2500, 2900, 3400, 4000, 4800, 5800, 7000, 8500, 10500, 13500]. Note that the highest sampling rate for which the Bark scale is defined is 31 kHz, although the 25th Bark band certainly extends above 19 kHz, so that a sampling rate of 40 kHz is well supported by the data. We have extrapolated the Bark band-edges in our work, appending the values $[20500, 27000]$ so that sampling rates up to 54 kHz are defined.

4 The Bilinear Transform

The formula for a general first-order (bilinear) conformal mapping of functions of a complex variable is conveniently expressed by [6, page 75]

$$\frac{(\zeta - \zeta_1)(\zeta_2 - \zeta_3)}{(\zeta_2 - \zeta_1)(\zeta - \zeta_3)} = \frac{(z - z_1)(z_2 - z_3)}{(z_2 - z_1)(z - z_3)}. \qquad (2)$$

It can be seen that choosing three specific points and their images determines the mapping for all z and ζ.

Bilinear transformations map circles and lines into circles and lines (lines being viewed as circles passing through the point at infinity). In digital audio, where both domains are "z planes," we normally want to map the unit circle to itself, with dc mapping to dc ($z_1 = \zeta_1 = 1$) and half the sampling rate mapping to half the sampling rate ($z_2 = \zeta_2 = -1$). Thus, we are left with transformations of the form

$$z = \mathcal{A}_\rho(\zeta) = \frac{\zeta + \rho}{1 + \zeta\rho}, \qquad \rho = \frac{\zeta_3 - z_3}{1 - z_3\zeta_3}. \qquad (3)$$

The constant ρ provides one remaining degree of freedom which can be used to map any particular frequency ω (corresponding to the point $e^{j\omega}$ on the unit circle) to a new location $a(\omega)$. All other frequencies will be *warped* accordingly. The allpass coefficient ρ can be written in terms of these frequencies as

$$\rho = \frac{\sin\{[a(\omega) - \omega]/2\}}{\sin\{[a(\omega) + \omega]/2\}}, \qquad (4)$$

In this form, it is clear that ρ is real and that the inverse of $\mathcal{A}_\rho$ is $\mathcal{A}_{-\rho}$. Also, since $0 < \{\omega, a(\omega)\} < \pi$, and $a(\omega) > \omega$ for a Bark map, we have $\rho \in (0, 1)$.

5 Optimal Warping

Fig. 1 illustrates the surprisingly good match between the allpass transformation $\mathcal{A}_\rho$ and a Bark frequency warping when the map parameter ρ is properly chosen. In the following, a simple direct-form expression is developed for the map parameter giving the best fit to a Bark scale.

5.1 Computing ρ

Our goal is to find the allpass coefficient ρ such that the frequency mapping

$$a(\omega) = \text{angle}\left\{\mathcal{A}_{-\rho}(e^{j\omega})\right\} \qquad (5)$$

best approximates the Bark frequency scale $b(\omega)$ for a given sampling rate f_s. (Note that the frequencies ω, $a(\omega)$, and $b(\omega)$ are all expressed in radians per second normalized by the sampling rate, so that a frequency of half of the sampling rate corresponds to a value of π.)

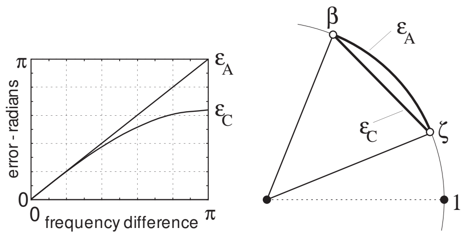

Figure 2: Frequency Map Errors.

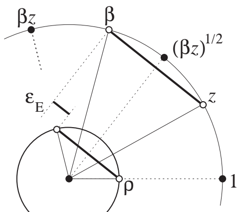

Figure 3: Equation Error Interpretation.

Using squared frequency errors to gauge the fit between $a(\omega)$ and its Bark warped counterpart, the desired parameter ρ^* may be written as

$$\rho^* = \operatorname{Arg}\left[\min_{\rho}\left\{\| a(\omega) - b(\omega) \|\right\}\right], \tag{6}$$

where $\|\cdot\|$ represents the L_2 norm. Unfortunately, the frequency error

$$\epsilon_{\mathrm{A}} \triangleq a(\omega) - b(\omega) \tag{7}$$

is nonlinear in ρ, and its norm is not easily minimized directly. It turns out, however, that a related error,

$$\epsilon_{\mathrm{C}} \triangleq e^{ja(\omega)} - e^{jb(\omega)}, \tag{8}$$

has a norm which is more amenable to minimization. The immediate question is how are the minimizers of $\| \epsilon_{\mathrm{A}} \|$ and $\| \epsilon_{\mathrm{C}} \|$ related.

Denote by ζ and β the complex representations of the frequencies $a(\omega)$ and $b(\omega)$ on the unit circle,

$$\zeta = e^{ja(\omega)}, \qquad \beta = e^{jb(\omega)}. \tag{9}$$

As seen in Fig. 2, the absolute frequency error $|\epsilon_{\mathrm{A}}|$ is the arc length between the points ζ and β, whereas $|\epsilon_{\mathrm{C}}|$ is the chord length or distance:

$$|\epsilon_{\mathrm{C}}| = 2\sin(|\epsilon_{\mathrm{A}}|/2). \tag{10}$$

The desired arc length error ϵ_{A} gives more weight to large errors than the chord length error ϵ_{C}; however, in the presence of small discrepancies between ζ and β, the absolute errors are very similar,

$$|\epsilon_{\mathrm{C}}| \approx |\epsilon_{\mathrm{A}}|, \quad \text{when } |\epsilon_{\mathrm{A}}| \ll 1. \tag{11}$$

Accordingly, essentially the same ρ^* results from minimizing $\| \epsilon_{\mathrm{A}} \|$ or $\| \epsilon_{\mathrm{C}} \|$ when the fit is uniformly good over frequency.

The error ϵ_{C} is also nonlinear in the parameter ρ, and to find its norm minimizer, an *equation error* is introduced, as is common practice in system identification [13]. The idea is to develop an error which is related in a simple way to the desired error, but has a norm which is more readily minimized. Consider mapping the frequency $z = e^{j\omega}$ via the allpass transformation $\mathcal{A}_{-\rho}(z)$,

$$\zeta = \frac{z - \rho}{1 - z\rho}. \tag{12}$$

Now, multiply (12) by the denominator $(1 - z\rho)$, and substitute $\zeta = \beta + \epsilon_{\mathrm{C}}$ from (8), to get

$$(\beta + \epsilon_{\mathrm{C}})(1 - z\rho) = z - \rho. \tag{13}$$

Rearranging terms, we have

$$(\beta - z) - (\beta z - 1)\rho = \epsilon_{\mathrm{E}}, \tag{14}$$

where ϵ_{E} is an equation error defined by

$$\epsilon_{\mathrm{E}} \triangleq (z\rho - 1)\epsilon_{\mathrm{C}}. \tag{15}$$

Referring to Fig. 3, note that the equation error is the difference between the unit circle chord $(\beta - z)$ and the ρ circle chord $(\beta z - 1)\rho$. The average of the input and Bark warped angles $[b(\omega) + \omega]/2$ bisects both these chords, and therefore the chords are parallel. The equation error may then be interpreted as the difference in chord lengths, rotated by the chord angle $[b(\omega) + \omega]/2 - \frac{\pi}{2}$. This suggests defining a rotated equation error which is real valued. Multiply (14) by $(z\beta)^{-\frac{1}{2}}/2j$ to obtain

$$\sin\{[b(\omega) - \omega]/2\} - \sin\{[b(\omega) + \omega]/2\}\rho = \epsilon_{\mathrm{R}}, \tag{16}$$

where the rotated equation error ϵ_{R} is defined by

$$\epsilon_{\mathrm{R}} \triangleq \frac{1}{2j}(z\beta)^{-\frac{1}{2}}\epsilon_{\mathrm{E}} = \frac{1}{2j}(z\beta)^{-\frac{1}{2}}(z\rho - 1)\epsilon_{\mathrm{C}}. \tag{17}$$

The rotated equation error ϵ_{R} is linear in the unknown ρ, and its norm minimizer is easily expressed in closed form. Denote by $\omega_k, k = 1, \ldots, K$ a set of frequencies corresponding to Bark frequencies $b(\omega_k)$, and by $\mathbf{d}$ and $\mathbf{s}$ the columns

$$\mathbf{d} = \begin{bmatrix} \sin\{[b(\omega_1) - \omega_1]/2\} \\ \vdots \\ \sin\{[b(\omega_K) - \omega_K]/2\} \end{bmatrix}, \; \mathbf{s} = \begin{bmatrix} \sin\{[b(\omega_1) + \omega_1]/2\} \\ \vdots \\ \sin\{[b(\omega_K) + \omega_K]/2\} \end{bmatrix}.$$

Then (16) becomes

$$\mathbf{d} - \mathbf{s}\rho = \epsilon_{\mathrm{R}}, \tag{18}$$

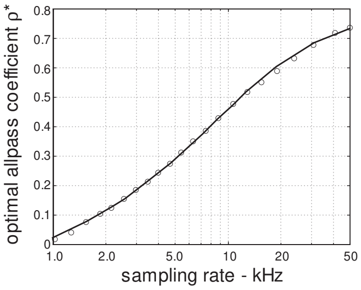

Figure 4: OPTIMAL ALLPASS COEFFICIENT.
The optimal allpass coefficient ρ^* is plotted as a function of sampling rate f_s using the marker '∘'. Also shown is the function $1.0211\left[\frac{2}{\pi}\tan^{-1}(0.076 f_s)\right]^{\frac{1}{2}} - 0.19877$, for f_s in kHz.

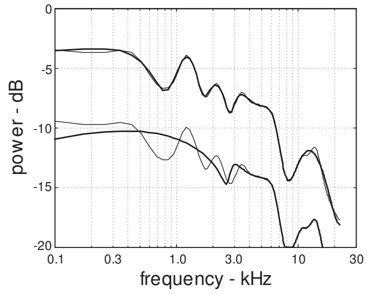

Figure 5: FILTER DESIGN EXAMPLE.
Measured and modeled magnitude transfer functions are shown for a 12th-order Prony model with (top) and without (bottom, offset 6 dB) Bark bilinear transform prewarping.

where ϵ_{R} is a column of rotated equation errors. The conformal map parameter

$$\rho^* = \frac{\mathbf{s}^\top \mathbf{V}\mathbf{d}}{\mathbf{s}^\top \mathbf{V}\mathbf{s}} \tag{19}$$

then minimizes the weighted sum of squared equation errors, $\epsilon_{\mathrm{R}}^\top \mathbf{V}\epsilon_{\mathrm{R}}$, for a positive definite weighting matrix $\mathbf{V}$. If the weighting $\mathbf{V}$ is diagonal with kth element $v(\omega_k)$, ρ^* may be written as a ratio of sums,

$$\rho^* = \frac{\sum_{k=1}^{K}\left[\cos b(\omega_k) - \cos\omega_k\right] v(\omega_k)}{\sum_{k=1}^{K}\left\{\cos\left[b(\omega_k) + \omega_k\right] - 1\right\} v(\omega_k)}. \tag{20}$$

What remains is to specify the weighting matrix $\mathbf{V}$. Recall that we want to minimize the sum of squared chord length errors $\epsilon_{\mathrm{C}}^\top \epsilon_{\mathrm{C}}$. The rotated equation error ϵ_{R} is proportional to the chord length error ϵ_{C}, viz. (17), and it is easily verified that when $\mathbf{V}$ is diagonal with kth diagonal element

$$v(\omega_k) = \frac{1}{1+\rho^2 - 2\rho\cos\omega_k}, \tag{21}$$

the chord length error and the weighted equation error coincide, and the solution (19) minimizes the chord length error norm.

Note that the desired weighting depends on the unknown map parameter ρ. To overcome this difficulty, we suggest first estimating ρ^* using $\mathbf{V} = \mathbf{I}$, and then computing ρ^* using the weighting (21) based on the unweighted solution. This is analogous to the Steiglitz-McBride algorithm for converting an equation error minimizer to the more desired "output error" minimizer using an iteratively computed weight function [14].

5.2 Solution Behavior

Allpass coefficients were computed for sampling rates of twice the Bark band edge frequencies by iteratively minimizing $\|\epsilon_{\mathrm{A}}\|$, and by using (19) with $\mathbf{V}$ derived from the $\mathbf{V} = \mathbf{I}$ solution. The resulting optimal coefficients are shown in Fig. 4. The allpass parameter increases with increasing sampling rate, reflecting the widening critical bandwidth with increasing Bark number.

It was noted that the weighted equation error norm minimizer (19) agreed with the arc length error norm minimizer to within a few parts in 10^{-4} for all sampling rates tried, thus validating the approximation (11) and the weighting (21). In addition, the conformal map fit to the Bark scale was excellent at all sample rates, never exceeding $0.012 f_s$ in absolute frequency error, and $0.005 f_s$ in root mean square error.

Finally, a simple analytic expression for the optimal allpass coefficient was developed:

$$\rho^*(f_s) \approx 1.0211\left[\frac{2}{\pi}\tan^{-1}(0.076 f_s)\right]^{\frac{1}{2}} - 0.19877, \tag{22}$$

where f_s is measured in kHz. This expression agrees with the optimal allpass coefficient ρ^* to within 0.01 for sampling rates in the range $f_s \in [1.0, 54]$ kHz.

6 Filter Design Example

We conclude with the filter design example shown in Fig. 5. A 12th-order ARMA model was fit to the equalization function shown using the Prony method, both directly and with prewarping of the data via $\mathcal{A}_\rho$ followed by modeling and "unwarping" of the modeled poles and zeros. It is clear that the warped solution provides an overall better fit than the direct solution which sacrifices accuracy below 4 kHz to achieve a better fit above 10 kHz.

References

[1] E. Zwicker and B. Scharf, "A Model of Loudness Summation," *Psych. Rev.*, vol. 72, pp. 3–26, 1965.

[2] E. Zwicker, *Psychoacoustics*, Springer-Verlag: Berlin, 1990.

[3] M. Karjalainen, "A New Auditory Model for the Evaluation of Sound Quality of Audio Systems," *Proc. IEEE ICASSP-85*, Tampa, Florida, 1985.

[4] T. Sporer and K. Brandenburg, "Constraints of Filter Banks Used for Perceptual Measurement," *J. Audio Eng. Soc.*, vol. 43, no. 3, pp. 107–116, March 1995.

[5] Z. Nehari, *Conformal Mapping*, Dover, 1952.

[6] R. V. Churchill, *Complex Variables and Applications*, McGraw-Hill, 1960.

[7] J. A. Moorer, "The Manifold Joys of Conformal Mapping: Applications to Digital Filtering in the Studio," *J. Audio Eng. Soc.*, vol. 31, pp. 826–841, Nov. 1983.

[8] T. W. Parks and C. S. Burrus, "Digital Filter Design," J. Wiley: New York NY, June 1987.

[9] H. W. Strube, "Linear prediction on a warped frequency scale," *J. Acoust. Soc. Am.*, vol. 68, no. 4, pp. 1071–1076, Oct. 1980.

[10] U. K. Laine, M. Karjalainen, T. Altosaar, "Warped Linear Prediction (WLP) in Speech and Audio Processing," *Proc. IEEE ICASSP-94*, Adelaide, Australia, 1994.

[11] J. O. Smith. *Techniques for Digital Filter Design and System Identification with Application to the Violin,* Ph.D. Dissertation, Elec. Eng. Dept., Stanford University, June 1983.

[12] J. O. Smith. "Spectral Pre-Processing for Audio Digital Filter Design," Proc. Int. Computer Music Conf. (ICMC), Rochester NY, 1983. (Substantially contained within [11].)

[13] L. Ljung and T. L. Soderstrom, *Theory and Practice of Recursive Identification*, MIT Press: Cambridge MA, 1983.

[14] P. Stoica and T. Soderstrom, "The Steiglitz-McBride Algorithm Revisited—Convergence Analysis and Accuracy Aspects," *IEEE Trans. Auto. Cont.*, vol. AC-26, no. 3, pp. 712–717, June 1981.