Bark and ERB Bilinear Transforms

Julius O. Smith III
Center for Computer Research in Music and Acoustics*
(CCRMA), Stanford University
Stanford, CA 94305 USA

Jonathan S. Abel†
Human Factors Research Division
NASA-Ames Research Center
Moffet Field, CA 94035 USA

(Final draft accepted for publication in the IEEE Transactions on Speech and Audio Processing, November, 1999.)

Abstract

Use of a bilinear conformal map to achieve a frequency warping nearly identical to that of the Bark frequency scale is described. Because the map takes the unit circle to itself, its form is that of the transfer function of a first-order allpass filter. Since it is a first-order map, it preserves the model order of rational systems, making it a valuable frequency warping technique for use in audio filter design. A closed-form weighted-equation-error method is derived which computes the optimal mapping coefficient as a function of sampling rate, and the solution is shown to be generally indistinguishable from the optimal least-squares solution. The optimal Chebyshev mapping is also found to be essentially identical to the optimal least-squares solution. The expression `0.8517*sqrt(atan(0.06583*Fs))-0.1916` is shown to accurately approximate the optimal allpass coefficient as a function of sampling rate Fs in kHz for sampling rates greater than 1 kHz. A filter design example is included which illustrates improvements due to carrying out the design over a Bark scale. Corresponding results are also given and compared for approximating the related "equivalent rectangular bandwidth (ERB) scale" of Moore and Glasberg using a first-order allpass transformation. Due to the higher frequency resolution called for by the ERB scale, particularly at low frequencies, the first-order conformal map is less able to follow the desired mapping, and the error is two to three times greater than the Bark-scale case, depending on the sampling rate.

1 Contents

Contents

*`http://ccrma.stanford.edu/`

†Work supported in part by San Jose State University Cooperative Agreement NCC-2-327.

2 Introduction

With the increasing use of frequency-domain techniques in audio signal processing applications such as audio compression, there is increasing emphasis on psychoacoustic-based spectral measures [36, 2, 11, 12]. One of the classic approaches is to analyze and process signal spectra over the *Bark frequency scale* (also called "critical band rate") [41, 42, 39, 21, 7]. Based on the results of many psychoacoustic experiments, the Bark scale is defined so that the critical bands of human hearing have a width of one Bark. By representing spectral energy (in dB) over the Bark scale, a closer correspondence is obtained with spectral information processing in the ear.

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[1] in such a way that, for $0 < \rho < 1$, low frequencies are stretched and high frequencies are compressed, as in a transformation from frequency in Hertz to the Bark scale. Because the conformal map $\mathcal{A}_\rho(\zeta)$ is identical in form to a first-order allpass transfer function (having a pole at $\zeta = -1/\rho$), 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 warping, 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 for a sampling rate of 31 kHz. The fit is so good, in fact, that there is almost no difference between the optimal least-squares and optimal Chebyshev approximations, as the figure shows. The purpose of this paper is to spread awareness of this useful fact and to present new methods for computing the optimal warping parameter ρ as a function of sampling rate.

2.1 Auditory Filter Banks

Auditory frequency-scale warping is closely related to the topic of *auditory filter banks* which are non-uniform bandpass filter banks designed to imitate the frequency resolution of human hearing [30, 32]. Classical auditory filter banks include constant-Q filter banks such as the widely used third-octave filter bank. More recently, constant-Q filter banks for audio have been devised based on the *wavelet transform*, including the auditory wavelet filter bank [8]. Auditory filter banks have also been based more directly on psychoacoustic measurements, leading to approximations of the auditory filter frequency response in terms of a Gaussian function [27], a "rounded exponential" [29], and more recently the *gammatone* (or "Patterson-Holdsworth") filter bank [30, 32]. The *gammachirp* filter bank further adds a level-dependent asymmetric correction to the basic gammatone channel frequency response, thereby providing a yet more accurate approximation to the auditory frequency response [10, 9].

All auditory filter banks can be seen as defining some linear to warped frequency mapping, since the filter-bank output signals are non-uniformly distributed versus frequency. While this paper is concerned primarily with approximating the Bark frequency scale using a first-order conformal map, the same approach can be used to approximate the warping defined by any pre-existing auditory filter bank.

[1] Note that the image of the conformal map corresponds to the *domain* variable ζ of the allpass transformation, while the input of the map corresponds to the *range* variable z.

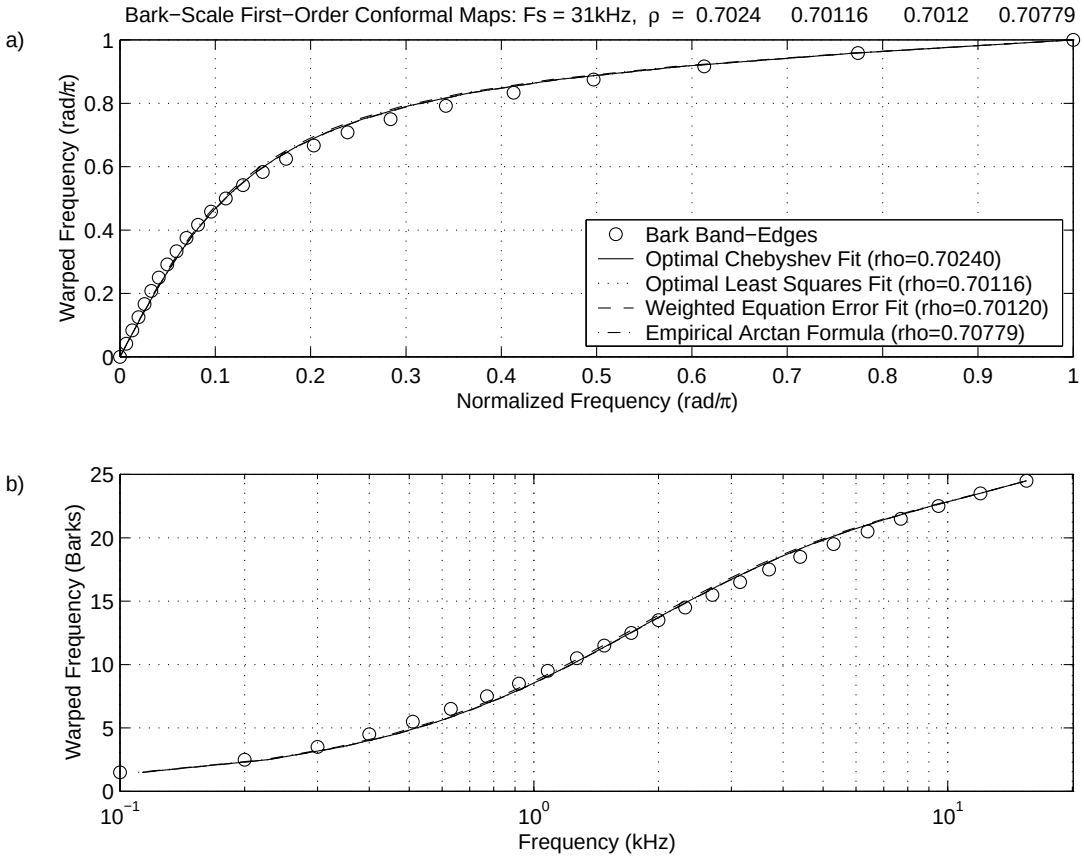

Figure 1: Bark and allpass frequency warpings at a sampling rate of 31 kHz (the highest possible without extrapolating the published Bark scale bandlimits). a) Bark frequency warping viewed as a conformal mapping of the interval $[0, \pi]$ to itself on the unit circle. b) Same mapping interpreted as an auditory frequency warping from Hz to Barks; the legend shown in plot a) also applies to plot b). The legend additionally displays the optimal allpass parameter ρ used for each map. The discrete band-edges which define the Bark scale are plotted as circles. The optimal Chebyshev (solid), least-squares (dashed), and weighted equation-error (dot-dashed) allpass parameters produce mappings which are nearly identical. Also plotted (dotted) is the mapping based on an allpass parameter given by an analytic expression in terms of the sampling rate, which will be described. It should be pointed out that the fit improves as the sampling rate is decreased.

As another application of the results of this paper, an *alternative* to the use of an auditory filter bank is a simpler *uniform* filter bank, such as an FFT, applied to a signal having a *warped* spectrum, where the warping is designed to approximate whatever auditory frequency axis is deemed most appropriate. It so happens that the earliest related work we are aware of was concerned with exactly this application, as we take up in the next subsection.

2.2 Prior Use of First-Order Conformal Maps as Frequency Warpings

In 1971, Oppenheim, Johnson, and Steiglitz proposed forming an FFT filter bank with non-uniformly spaced bins by taking the FFT of the outputs of a cascade chain of first-order allpass filters [25].

In 1980, "warped linear prediction" was proposed by Strube [37] for obtaining better formant models of speech: The frequency axis "seen" by LPC is made to approximate a Bark scale using the first-order allpass transformation. It was noted in [37] 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.

In 1983, the Bark bilinear transformation was also developed independently for audio digital filter design [34]. In that work, the frequency response fit was carried out over an approximate Bark 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 allpass transformation $\mathcal{A}_\rho$ determined an interpolated resampling of the desired filter frequency response $H(e^{j\omega})$ which converted its support to an approximate Bark 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 $H^*[\mathcal{A}_{-\rho}(z)]$.

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 [26, 4]. Higher order allpass transformations have been used to convert lowpass or highpass prototype filters into multiple bandpass/bandstop filters [23]. Allpass transformations of order greater than one appear not to have been used in frequency warping applications, since allpass transformations of order N map the unit circle to N traversals of the unit circle, and a one-to-one mapping of the unit circle to itself is desired.[2]

[2]In general, the unit circle is mapped once to itself by any allpass transformation for which the number of poles N_p minus the number of zeros N_z inside the unit circle is $|N_p - N_z| = 1$. Therefore, higher order allpass transfer functions can be used having N_p poles inside the unit circle, say, and $N_z = N_p \pm 1$ poles *outside* the unit circle. However, such a transformation cannot be used for audio digital filter design, our principle application, because it results in an *unstable* final filter $H^*[\mathcal{A}_{-\rho}(z)]$. It similarly cannot be used in any applications requiring time-domain implementation of the unstable allpass filter in place of a unit delay element.

More recently, in 1994 [15], an allpass coefficient of 0.62 was used to generate a frequency warping closely approximating the Bark scale for a sampling rate of 22 kHz. Experiments comparing the performance of warped LPC and "normal" LPC for speech coding and speech recognition applications showed that warped LPC required less than half the predictor model order for comparable performance.

Very recently, the first-order allpass transformation was used to implement audio-warped filters directly in the warped domain [13, 14]. In this application, a digital filter is designed over the warped frequency axis, and in its implementation, each delay element is replaced by a first-order allpass filter which implements the unwarping on the fly. Advantages of this scheme include (a) reducing the necessary filter order by a factor of 5 to 10 (more than compensating for the increased cost of implementing a delay element as a first-order allpass filter), (b) avoiding numerical failures which can occur (even in double-precision floating point) when attempting to unwarp very high-order filters (e.g., much larger than 30), and (c) providing a dynamic warping modulation control which tends to act as a frequency-scaling parameter associated with "acoustic size" and is therefore musically useful.

The critical feature of the first-order conformal map in the z plane 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" filter transformation such as this can so closely match the Bark frequency scale.

2.3 Paper Outline

In the next section, the Bark frequency scale is reviewed, followed by a section reviewing the bilinear transformation and its specialization to the first-order allpass transformation. Section IV is concerned with optimally choosing the allpass parameter: A weighted equation-error solution is derived which is shown to be essentially equal to the optimal least-squares solution. The optimal Chebyshev solution is compared and found to be insignificantly different from the least-squares solutions. A variation on the error criterion which is only concerned with mapped bandwidth error, as opposed to the absolute error in the mapping of Hz to Barks, is introduced and evaluated. A simple closed-form expression relating the sampling rate to the optimal warping parameter is presented. Section IV concludes with a filter-design example illustrating the benefits of working over a Bark frequency scale. Finally, Section V applies the methods of Section IV to approximating the ERB scale, and the results, while potentially useful, are found to be significantly less accurate than for the Bark scale. The paper concludes with a summary of findings.

3 The Bark Frequency Scale

The Bark scale ranges from 1 to 24 Barks, corresponding to the first 24 critical bands of hearing [39]. 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]. These center-frequencies and bandwidths are to be interpreted as samplings of a continuous variation in the frequency response of the ear to a sinusoid or narrowband noise process. That is, critical-band-shaped masking patterns should be seen as forming around specific stimuli in the ear rather than

being associated with a specific fixed filter bank in the ear.

Note that since the Bark scale is defined only up to 15.5 kHz, the highest sampling rate for which the Bark scale is defined up to the Nyquist limit, without requiring extrapolation, is 31 kHz. The 25th Bark band certainly extends above 19 kHz (the sum of the 24th Bark band edge and the 23rd critical bandwidth), so that a sampling rate of 40 kHz is implicitly 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. While human hearing generally does not extend above 20 kHz, audio sampling rates as high as 48 kHz or higher are common in practice.

The Bark scale is defined above in terms of frequency in Hz versus Bark number. For computing optimal allpass transformations, it is preferable to optimize the allpass fit to the *inverse* of this map, i.e., Barks versus Hz, so that the mapping error will be measured in Barks rather than Hz.

4 The Bilinear Transform

The formula for a general first-order (bilinear) conformal mapping of functions of a complex variable is conveniently expressed by [3, page 75]

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

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 ($z_2 = \zeta_2 = -1$). Making these substitutions in Eq. (2) leaves us with transformations of the form

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

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

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

In this form, it is clear that ρ is real and that the inverse of $\mathcal{A}_\rho$ is $\mathcal{A}_{-\rho}$. Also, since $0 \le \{\omega, a(\omega)\} \le \pi$, and $a(\omega) \ge \omega$ for a Bark map, we have $\rho \in [0, 1)$ for a Bark map from the z plane to the ζ plane.

5 Optimal Bark Warping

Figure 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 least-squares fit to a Bark scale for a chosen sampling rate. As Fig. 1 shows, the error is so small that the solution is also very

close to the optimal Chebyshev fit. In fact, the L_2 optimal warping is within 0.04 Bark of the L_∞ optimal warping. Since the experimental uncertainty when measuring critical bands is on the order of a tenth of a Bark or more [20, 22, 31, 38], we consider the optimal Chebyshev and least-squares maps to be equivalent psychoacoustically.

5.1 Computing ρ

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

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

best approximates the Bark 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 sample, so that a frequency of half of the sampling rate corresponds to a value of π.)

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

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

where $\| \cdot \|$ represents the L_2 norm. (We use the superscript '$*$' to denote optimality in some sense.) Unfortunately, the frequency error

$$\epsilon_{\text{A}} \stackrel{\Delta}{=} a(\omega) - b(\omega) \tag{7}$$

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

$$\epsilon_{\text{C}} \stackrel{\Delta}{=} e^{ja(\omega)} - e^{jb(\omega)}, \tag{8}$$

has a norm which is more amenable to minimization. The first issue we address is how the minimizers of $\| \epsilon_{\text{A}} \|$ and $\| \epsilon_{\text{C}} \|$ are related.

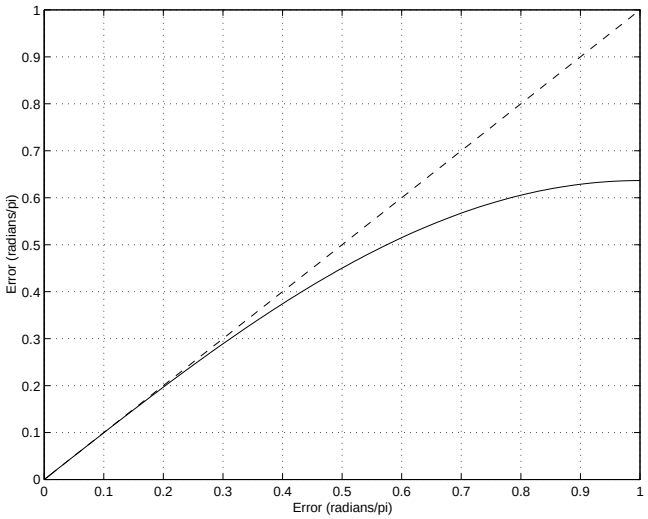

Figure 2: Frequency Map Errors

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 $|\epsilon_{\rm A}|$ is the arc length between the points ζ and β, whereas $|\epsilon_{\rm C}|$ is the chord length or distance:

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

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

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

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

The error $\epsilon_{\rm C}$ is also nonlinear in the parameter ρ, and to find its norm minimizer, an *equation error* is introduced, as is common practice in developing solutions to nonlinear system identification problems [17]. 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 Eq. (12) by the denominator $(1-z\rho)$, and substitute $\zeta = \beta + \epsilon_{\rm C}$ from Eq. (8), to get

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

Rearranging terms, we have

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

where $\epsilon_{\rm E}$ is an equation error defined by

$$\epsilon_{\rm E} \stackrel{\Delta}{=} (z\rho - 1)\epsilon_{\rm C}. \tag{15}$$

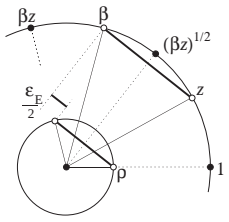

Figure 3: Geometric Interpretation of Equation Error

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 to the angle $[b(\omega) + \omega]/2 + \pi/2$.

This suggests defining a rotated equation error which is real valued. Multiply Eq. (14) by $(z\beta)^{-\frac{1}{2}}/2j$ to obtain

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

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 $\boldsymbol{d}$ and $\boldsymbol{s}$ the columns

$$\boldsymbol{d} = \begin{bmatrix}\sin\{[b(\omega_1)-\omega_1]/2\}\\ \vdots\\ \sin\{[b(\omega_K)-\omega_K]/2\}\end{bmatrix} \tag{18}$$

$$\boldsymbol{s} = \begin{bmatrix}\sin\{[b(\omega_1)+\omega_1]/2\}\\ \vdots\\ \sin\{[b(\omega_K)+\omega_K]/2\}\end{bmatrix}. \tag{19}$$

Then Eq. (16) becomes

$$\boldsymbol{d}-\boldsymbol{s}\rho=\boldsymbol{\epsilon}_{\mathrm{R}}, \tag{20}$$

where $\boldsymbol{\epsilon}_{\mathrm{R}}$ is a column of rotated equation errors. Eq. (20) is now in the form of a standard least squares problem. As is well known [17, 34], the solution ρ^* which minimizes the weighted sum of squared errors, $\boldsymbol{\epsilon}_{\mathrm{R}}^\top\boldsymbol{V}\boldsymbol{\epsilon}_{\mathrm{R}}$, the matrix $\boldsymbol{V}$ being an arbitrary positive-definite weighting, may be obtained by premultiplying both sides of Eq. (20) by $\boldsymbol{s}^\top\boldsymbol{V}$ and solving for ρ, noting that at $\rho=\rho^*$, $\boldsymbol{s}^\top\boldsymbol{V}\boldsymbol{\epsilon}_{\mathrm{R}}=0$ by the orthogonality principle. Doing this yields the optimal weighted least-squares conformal map parameter

$$\rho^*=\frac{\boldsymbol{s}^\top\boldsymbol{V}\boldsymbol{d}}{\boldsymbol{s}^\top\boldsymbol{V}\boldsymbol{s}}. \tag{21}$$

If the weighting matrix $\boldsymbol{V}$ is diagonal with kth diagonal element $v(\omega_k)>0$, then the weighted least-squares solution Eq. (21) reduces to

$$\begin{aligned}\rho^* &= \frac{\sum_{k=1}^K v(\omega_k)\sin\left[\frac{b(\omega_k)-\omega_k}{2}\right]\sin\left[\frac{b(\omega_k)+\omega_k}{2}\right]}{\sum_{k=1}^K v(\omega_k)\sin^2\left[\frac{b(\omega_k)+\omega_k}{2}\right]}\\ &= \frac{\sum_{k=1}^K v(\omega_k)\left\{\cos\left[b(\omega_k)\right]-\cos(\omega_k)\right\}}{\sum_{k=1}^K v(\omega_k)\left\{\cos\left[b(\omega_k)+\omega_k\right]-1\right\}},\end{aligned}$$

where we have used Equations Eq. (18) and Eq. (19), and the trigonometric identities

$$\begin{aligned}\cos(A+B)-\cos(A-B) &= -2\sin(A)\sin(B)\\ \cos(A)-1 &= -2\sin^2(A/2)\end{aligned}$$

to simplify the numerator and denominator, respectively.

It remains to choose a weighting matrix $\boldsymbol{V}$. Recall that we initially wanted to minimize the sum of squared chord-length errors $\boldsymbol{\epsilon}_{\mathrm{C}}^{\top}\boldsymbol{\epsilon}_{\mathrm{C}}$. The rotated equation error ϵ_{R} is proportional to the chord length error ϵ_{C}, viz. Eq. (17), and it is easily verified that when $\boldsymbol{V}$ is diagonal with kth diagonal element

$$v(\omega_k) = \frac{1}{1+\rho^2-2\rho\cos\omega_k}, \tag{22}$$

the chord length error and the weighted equation error coincide. Thus, with this diagonal weighting matrix, the solution Eq. (21) 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 $\boldsymbol{V} = \boldsymbol{I}$, where $\boldsymbol{I}$ denotes the identity matrix, and then computing ρ^* using the weighting Eq. (22) 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 [16].

5.2 Optimal Frequency Warpings

Optimal allpass coefficients ρ^* were computed for sampling rates of twice the Bark band-edge frequencies by means of four different optimization methods:

1. Minimize the peak arc-length error $\|\,\epsilon_{\mathrm{A}}\,\|_\infty$ at each sampling rate to obtain the optimal Chebyshev allpass parameter $\rho^*_\infty(f_s)$.

2. Minimize the sum of squared arc-length errors $\|\,\epsilon_{\mathrm{A}}\,\|_2^2$ to obtain the optimal least-squares allpass parameter $\rho^*_2(f_s)$.

3. Use the closed-form weighted equation-error solution Eq. (21) computed twice, first with $\boldsymbol{V} = \boldsymbol{I}$, and second with $\boldsymbol{V}$ set from Eq. (22) to obtain the optimal "weighted equation error" solution $\rho^*_{\mathrm{E}}(f_s)$.

4. Fit the function $\gamma_1\left[\frac{2}{\pi}\arctan(\gamma_2 f_s)\right]^{\frac{1}{2}}+\gamma_3$ to the optimal Chebyshev allpass parameter $\rho^*_\infty(f_s)$ via Chebyshev optimization with respect to $\gamma \triangleq \{\gamma_1,\gamma_2,\gamma_3\}$. We will refer to the resulting function as the "arctangent approximation" $\rho^*_\gamma(f_s)$ (or, less formally, the "Barktan formula"), and note that it is easily computed directly from the sampling rate.

In all cases, the error minimized was in units proportional to Barks. The discrete frequency grid in all cases was taken to be the Bark band-edges given in Section II. The resulting allpass coefficients are plotted as a function of sampling rate in Fig. 4.

The peak and rms frequency-mapping errors are plotted versus sampling rate in Fig. 5. Peak and rms errors in Barks[3] are plotted for all four cases (Chebyshev, least squares, weighted equation-error, and arctangent approximation). The conformal-map fit to the Bark scale is generally excellent in all cases. We see that the rms error is essentially identical in the first three cases, although the Chebyshev rms error is visibly larger below 10 kHz. Similarly, the peak error is essentially the same for least squares and weighted equation error, with the Chebyshev case being able to shave almost 0.1 Bark from the maximum error at high sampling rates. The arctangent formula shows up to a

[3]The normalized warped-frequency interval $\omega \in [0,\pi]$ was converted to Barks b by the affine transformation $b = (\omega/\pi)*(N_b-1)+0.5$, where N_b is the number of Bark bands in use. For example, $N_b = 25$ for a 31 kHz sampling rate.

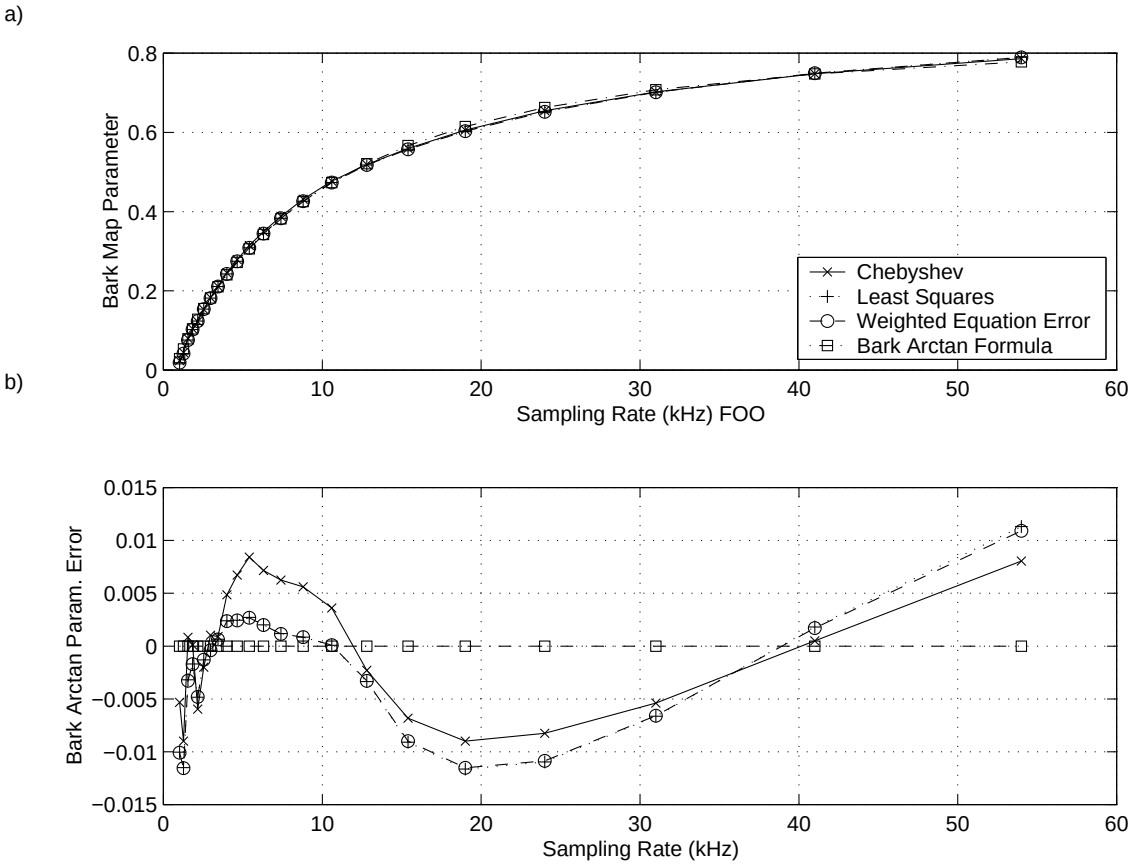

Figure 4: a) Optimal allpass coefficients ρ_∞^*, ρ_2^*, and ρ_{E}^*, plotted as a function of sampling rate f_s. Also shown is the arctangent approximation $\rho_\gamma^* = 1.0674\sqrt{(2/\pi)\arctan(0.06583 f_s)} - 0.1916$. b) Same as a) with the arctangent approximation subtracted out. Note the nearly identical behavior of optimal least-squares (plus signs) and weighted equation-error (circles).

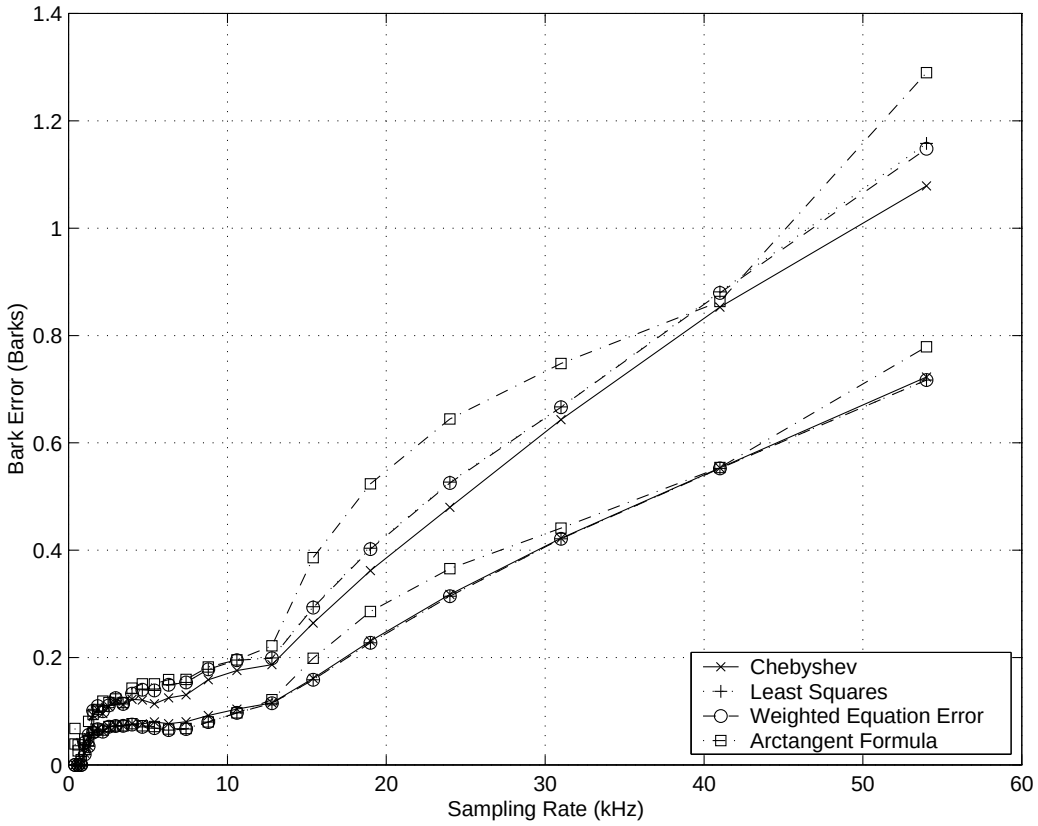

Figure 5: Root-mean-square and peak frequency-mapping errors versus sampling rate for Chebyshev, least squares, weighted equation-error, and arctangent optimal maps. The rms errors are nearly coincident along the lower line, while the peak errors a little more spread out well above the rms errors.

tenth of a Bark larger peak error at sampling rates 15–30 and 54 kHz, but otherwise it performs very well; at 41 kHz and below 12 kHz the arctangent approximation is essentially optimal in all senses considered.

At sampling rates up to the maximum non-extrapolated sampling rate of 31 kHz, the peak mapping errors are all much less than one Bark (0.64 Barks for the Chebyshev case and 0.67 Barks for the two least squares cases). The mapping errors in Barks can be seen to increase almost linearly with sampling rate. However, the irregular nature of the Bark-scale data results in a nonmonotonic relationship at lower sampling rates.

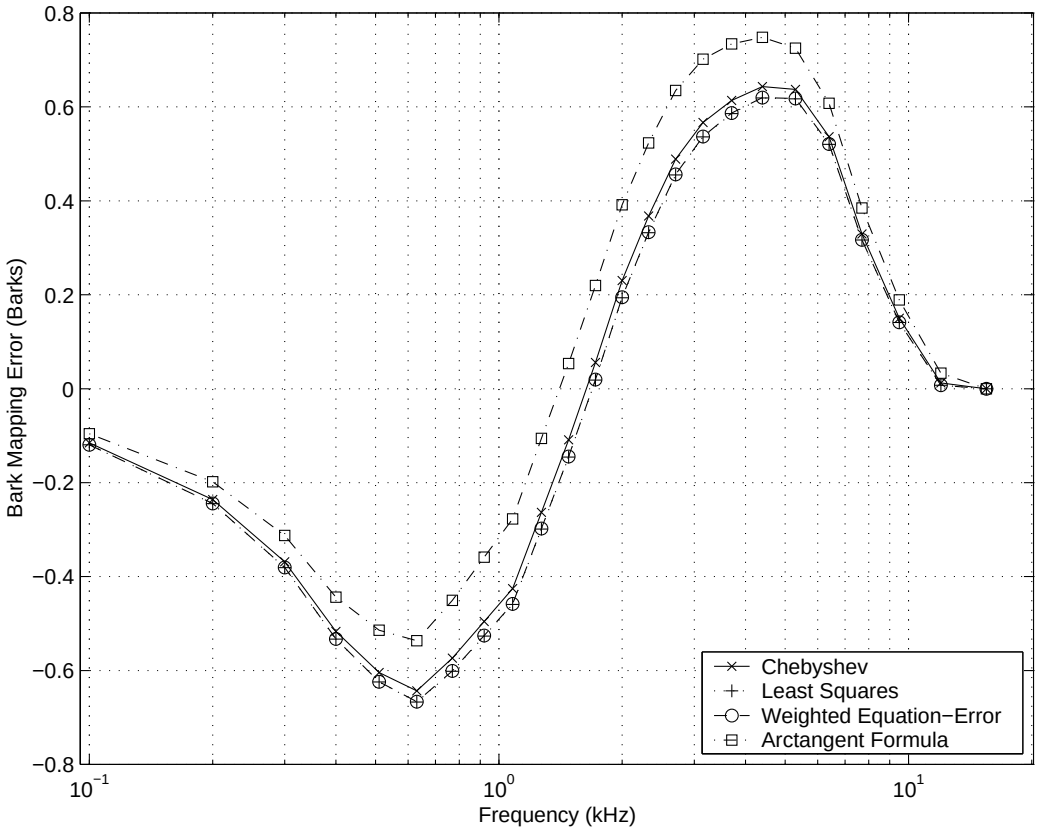

Figure 6: Frequency mapping errors versus frequency for a sampling rate of 31 kHz.

The specific frequency mapping errors versus frequency at the 31 kHz sampling rate (the same case shown in Fig. 1) are plotted in Fig. 6. Again, all four cases are overlaid, and again the least squares and weighted equation-error cases are essentially identical. By forcing equal and opposite peak errors, the Chebyshev case is able to lower the peak error from 0.67 to 0.64 Barks. A difference of 0.03 Barks is probably insignificant for most applications. The peak errors occur at 1.3 kHz and 8.8 kHz where the error is approximately 2/3 Bark. The arctangent formula peak error is 0.73 Barks at 8.8 kHz, but in return, its secondary error peak at 1.3 kHz is only 0.55 Barks. In some applications, such as when working with oversampled signals, higher accuracy at low frequencies at the expense of higher error at very high frequencies may be considered a desirable tradeoff.

We see that the mapping falls "behind" a bit as frequency increases from zero to 1.3 kHz, mapping linear frequencies slightly below the desired corresponding Bark values; then, the mapping

"catches up," reaching an error of 0 Barks near 3 kHz. Above 3 kHz, it gets "ahead" slightly, with frequencies in Hz being mapped a little too high, reaching the positive error peak at 8.8 kHz, after which it falls back down to zero error at $z = e^{j\pi}$. (Recall that dc and half the sampling-rate are always points of zero error by construction.)

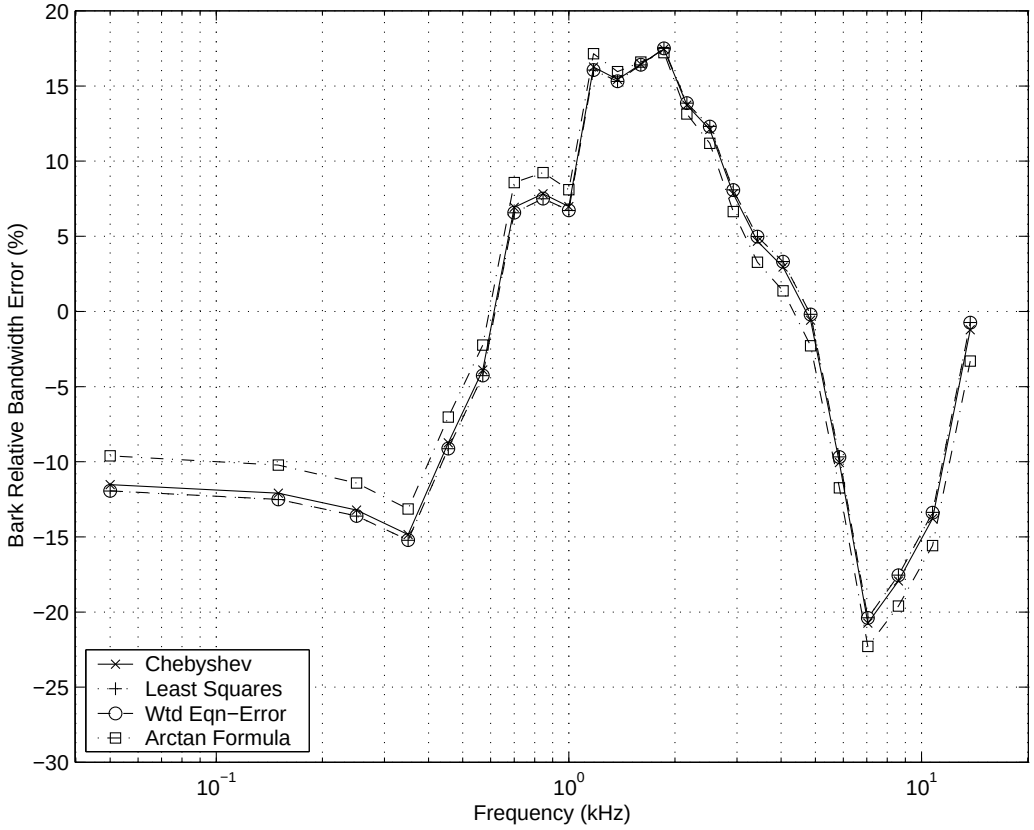

Figure 7: Relative bandwidth mapping error (RBME) for a 31 kHz sampling rate using the optimized allpass warpings of Fig. 4 at 31 kHz. The optimal Chebyshev, least squares, and weighted equation-error cases are almost indistinguishable.

5.3 Relative Bandwidth Mapping Error

The *slope* of the frequency versus warped-frequency curve can be interpreted as being proportional to critical bandwidth, since a unit interval (one Bark) on the warped-frequency axis is magnified by the slope to restore the band to its original size (one critical bandwidth). It is therefore interesting to look at the *relative slope error*, i.e., the error in the slope of the frequency mapping divided by the ideal Bark-map slope. We interpret this error measure as the *relative bandwidth-mapping error* (RBME). The RBME is plotted in Fig. 7 for a 31 kHz sampling rate. The worst case is 21% for the Chebyshev case and 20% for both least-squares cases. When the mapping coefficient is explicitly optimized to minimize RBME, the results of Fig. 8 are obtained: the Chebyshev peak error drops from 21% down to 18%, while the least-squares cases remain unchanged at 20% maximum RBME. A 3% change in RBME is comparable to the 0.03 Bark peak-error reduction seen in Fig. 6 when

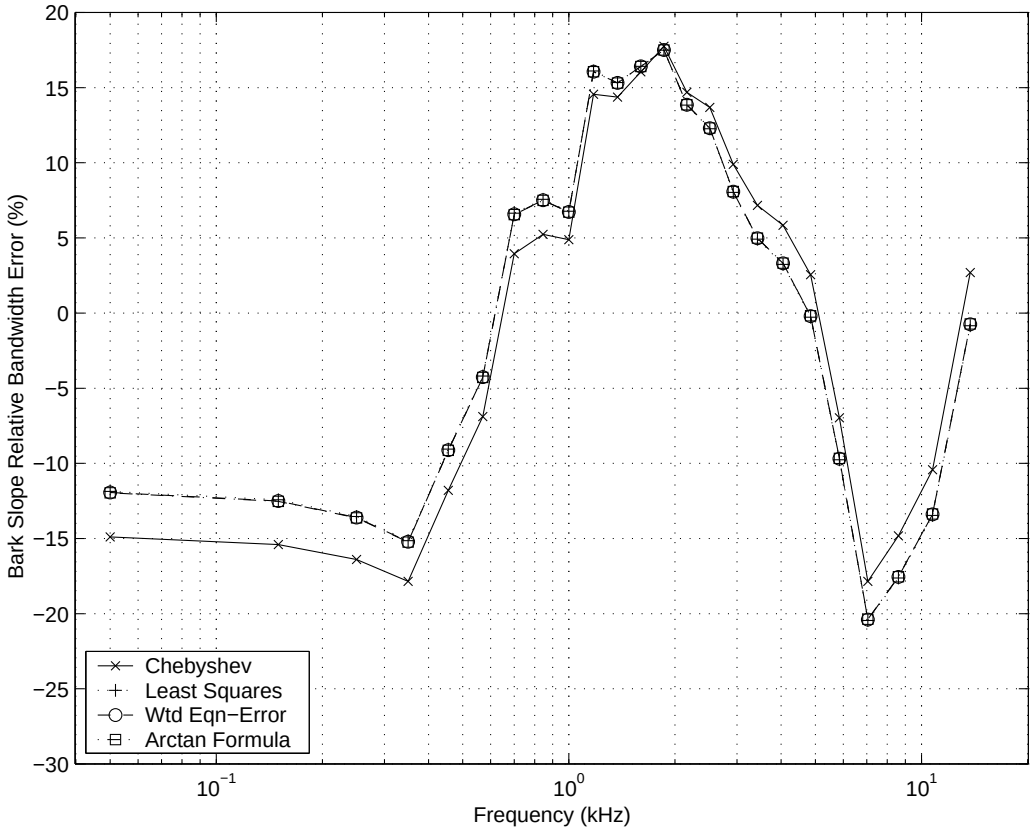

Figure 8: RBME for a 31 kHz sampling rate, with explicit minimization of RBME in the optimizations.

using the Chebyshev norm instead of the L_2 norm; again, such a small difference is not likely to be significant in most applications.

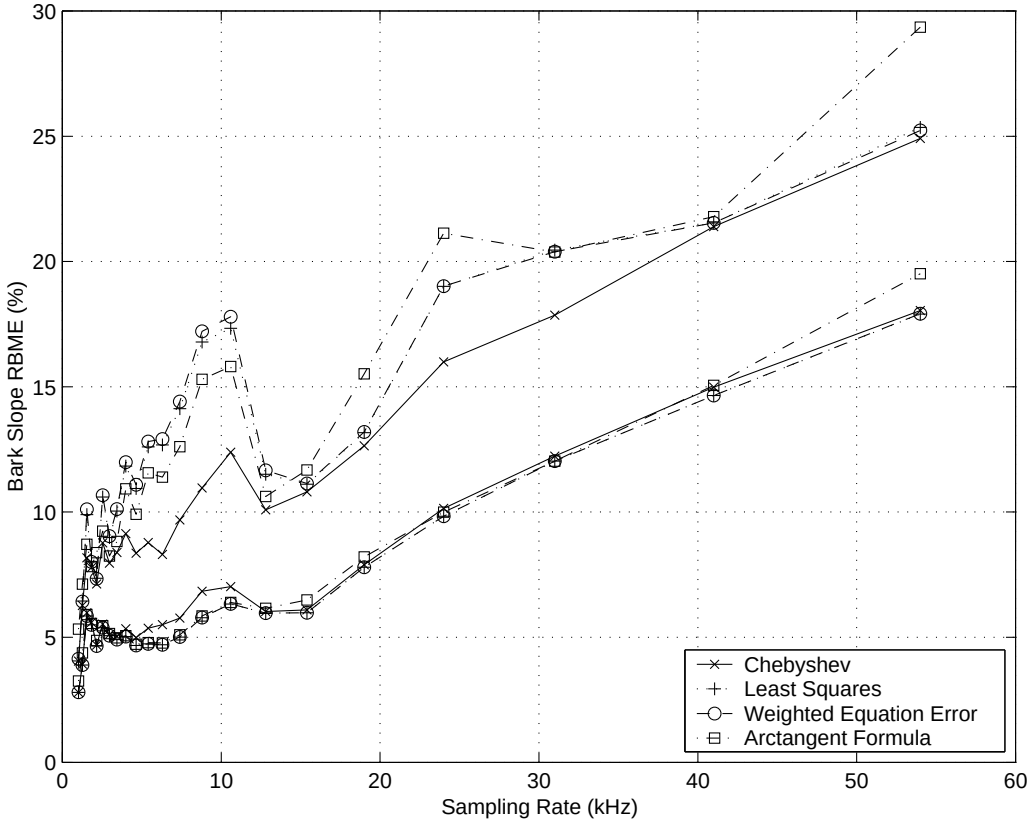

Figure 9: Root-mean-square and peak relative-bandwidth-mapping errors versus sampling rate for Chebyshev, least squares, weighted equation-error, and arctangent optimal maps, with explicit minimization of RBME used in all optimizations. The peak errors form a group lying well above the lower lying rms group.

Similar observations are obtained at other sampling rates, as shown in Fig. 9. Near a 10 kHz sampling rate, the Chebyshev RBME is reduced from 17% when minimizing absolute error in Barks (not shown in any figure) to around 12% by explicitly minimizing the RBME, and this is the sampling-rate range of maximum benefit. At 15.2, 19, 41, and 54 kHz sampling rates, the difference is on the order of only 1%. Other cases generally lie between these extremes. The arctangent formula generally falls between the Chebyshev and optimal least-squares cases, except at the highest (extrapolated) sampling rate 54kHz. The rms error is very similar in all four cases, although the Chebyshev case has a little larger rms error near a 10 kHz sampling rate, and the arctangent case gives a noticeably larger rms error at 54 kHz.

5.4 Error Significance

In one study, young normal listeners exhibited a standard deviation in their measured auditory bandwidths (based on notched-noise masking experiments) on the order of 10% of center frequency [20]. Therefore, a 20% peak error in mapped bandwidth (typical for sampling rates approaching 40

kHz) could be considered significant. However, the *range* of auditory-filter bandwidths measured in 93 young normal subjects at 2kHz [20] was 230 to 410 Hz, which is -26% to +32% relative to 310 Hz. In [38], 40 subjects were measured, yielding auditory-filter bandwidths between -33% and +65%, with a standard deviation of 18%. It may thus be concluded that a worst-case mapping error on the order of 20%, while probably detectable by "golden ears" listeners, lies well within the range of experimental deviations in the empirical measurement of auditory bandwidth.

As a worst-case example of how the 18% peak bandwidth-mapping error in Fig. 8 might correspond to an audible distortion, consider one critical band of noise centered at the frequency of maximum negative mapping error, scaled to be the same loudness as a single critical band of noise centered at the frequency of maximum positive error. The systematic nature of the mapping error results in a narrowing of the lower band and expansion of the upper band by about 1.7 dB. As a result, over the warped frequency axis, the upper band will be effectively *emphasized* over the lower band by about 3 dB.

5.5 Arctangent Approximations for $\rho^*(f_s)$

This subsection provides further details on the arctangent approximation for the optimal allpass coefficient as a function of sampling rate. Compared with other spline or polynomial approximations, the arctangent form

$$\rho_\gamma(f_s) \triangleq \max\left\{0, \gamma_1\left[\frac{2}{\pi}\arctan(\gamma_2 f_s)\right]^{\frac{1}{2}} + \gamma_3\right\} \tag{23}$$

was found to provide a more parsimonious expression at a given accuracy level. The idea was that the arctangent function provided a mapping from the interval $[0,\infty)$, the domain of f_s, to the interval $[0,1)$, the range of $\rho(f_s)$. The additive component γ_3 allowed $\rho_\gamma(f_s)$ to be zero at smaller sampling rates, where the Bark scale is linear with frequency. As an additional benefit, the arctangent expression was easily inverted to give sampling rate f_s in terms of the allpass coefficient ρ_γ:

$$f_s = \frac{1}{\gamma_2}\tan\left[\frac{\pi}{2}\left(\frac{\rho_\gamma - \gamma_3}{\gamma_1}\right)^2\right]. \tag{24}$$

To obtain the optimal arctangent form $\rho^*_\gamma(f_s)$, the expression for $\rho_\gamma(f_s)$ in Eq. (23) was optimized with respect to its free parameters $\gamma = \{\gamma_1, \gamma_2, \gamma_3\}$ to match the optimal Chebyshev allpass coefficient as a function of sampling rate:

$$\rho^*_\gamma(f_s) \triangleq \mathrm{Arg}\left[\min_\gamma\left\{\| \rho^*_\infty(f_s) - \rho_\gamma(f_s) \|_\infty\right\}\right]. \tag{25}$$

For a Bark warping, the optimized arctangent formula was found to be

$$\rho^*_\gamma(f_s) = 1.0674\left[\frac{2}{\pi}\arctan(0.06583 f_s)\right]^{\frac{1}{2}} - 0.1916, \tag{26}$$

where f_s is expressed in units of kHz. This formula is plotted along with the various optimal ρ^* curves in Fig. 4a, and the approximation error is shown in Fig. 4b. It is extremely accurate below 15 kHz and near 40 kHz, and adds generally less than 0.1 Bark to the peak error at other sampling rates. The rms error versus sampling rate is very close to optimal at all sampling rates, as Fig. 5 also shows.

When the optimality criterion is chosen to minimize relative bandwidth mapping error (relative map *slope* error), the arctangent formula optimization yields

$$\rho_\gamma^*(f_s) = 1.0480 \left[\frac{2}{\pi}\arctan(0.07212 f_s)\right]^{\frac{1}{2}} - 0.1957. \tag{27}$$

The performance of this formula is shown in Fig. 9. It tends to follow the performance of the optimal least squares map parameter even though the peak parameter error was minimized relative to the optimal Chebyshev map. At 54 kHz there is an additional 3% bandwidth error due to the arctangent approximation, and near 10 kHz the additional error is about 4%; at other sampling rates, the performance of the RBME arctangent approximation is better, and like Eq. (26), it is extremely accurate at 41 kHz.

5.6 Filter Design Example

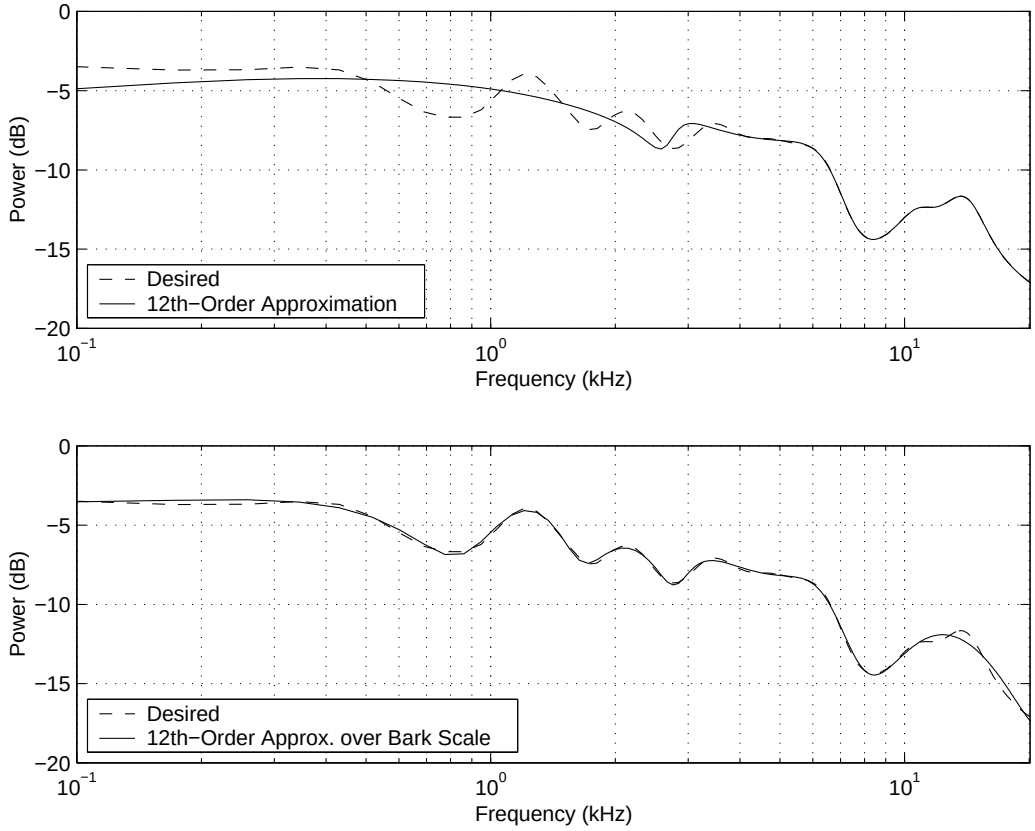

Figure 10: Filter Design Example: Overlay of measured and modeled magnitude transfer functions, where the model is a 12th-order filter designed by Prony's method. a) Results without prewarping of the frequency axis. b) Results using the Bark bilinear transform prewarping.

We conclude discussion of the Bark bilinear transform with the filter design example of Fig. 10. A 12th-order pole-zero filter was fit using Prony's method [18] to the equalization function plotted in the figure as a dashed line. Prony's method was applied normally over a uniformly sampled

linear frequency grid in the example of Fig. 10a, and over an approximate Bark-scale axis in the example of Fig. 10b. The procedure in the Bark-scale case was as follows [34]:[4]

1. The optimal allpass coefficient $\rho_\gamma^*(f_s)$ was found using Eq. (26).
2. The desired frequency response $H(e^{j\omega})$ defined on a linear frequency axis ω was warped to an approximate Bark scale $a(\omega)$ using the Bark bilinear transform Eq. (1), $\tilde{H}(e^{j\omega}) \triangleq H[\mathcal{A}_\rho(e^{ja(\omega)})]$.
3. A parametric ARMA model $\tilde{H}^*(\zeta)$ was fit to the desired Bark-warped frequency response $\tilde{H}(e^{j\omega})$ over the unit circle $\zeta = e^{j\omega}$.
4. Finally, the inverse Bark bilinear transform $\zeta \leftarrow \mathcal{A}_{-\rho}(z)$ was used to "unwarp" the modeled system to a linear frequency axis.

Referring to Fig. 10, it is clear that the warped solution provides a better overall fit than the direct solution which sacrifices accuracy below 4 kHz to achieve a tighter fit above 10 kHz. In some part, the spacing of spectral features is responsible for the success of the Bark-warped model in this particular example. However, we generally recommend using the Bark bilinear transform to design audio filters, since doing so weights the error norm (for norms other than Chebyshev types) in a way which gives equal importance to matching features having equal Bark bandwidths. Even in the case of Chebyshev optimization, auditory warping appears to improve the *numerical conditioning* of the filter design problem; this applies also to optimization under the Hankel norm which includes an optimal Chebyshev design internally as an intermediate step. Further filter-design examples, including more on the Hankel-norm case, may be found in [34].

6 Equivalent Rectangular Bandwidth

Moore and Glasberg [19] have revised Zwicker's loudness model to better explain (1) how equal-loudness contours change as a function of level, (2) why loudness remains constant as the bandwidth of a fixed-intensity sound increases up to the critical bandwidth, and (3) the loudness of partially masked sounds. The modification that is relevant here is the replacement of the Bark scale by the *equivalent rectangular bandwidth* (ERB) scale. The ERB of the auditory filter is assumed to be closely related to the critical bandwidth, but it is measured using the *notched-noise* method [27, 28, 31, 22, 5] rather than on classical masking experiments involving a narrowband masker and probe tone [41, 42, 39]. As a result, the ERB is said not to be affected by the detection of beats or intermodulation products between the signal and masker. Since this scale is defined analytically, it is also more smoothly behaved than the Bark scale data.

At moderate sound levels, the ERB in Hz is defined by [19]

$$\mathrm{ERB}(f) = 0.108f + 24.7 \tag{28}$$

where f is center-frequency in Hz, normally in the range 100 Hz to 10kHz. The ERB is generally narrower than the classical critical bandwidth (CB), being about 11% of center frequency at high frequencies, and leveling off to about 25 Hz at low frequencies. The classical CB, on the other hand,

[4] Matlab functions bark2lin.m and lin2bark.m for transforming between linear and bark-warped frequency representations are available on the internet at http://ccrma.stanford.edu/~jos/bbt/bbt.html.

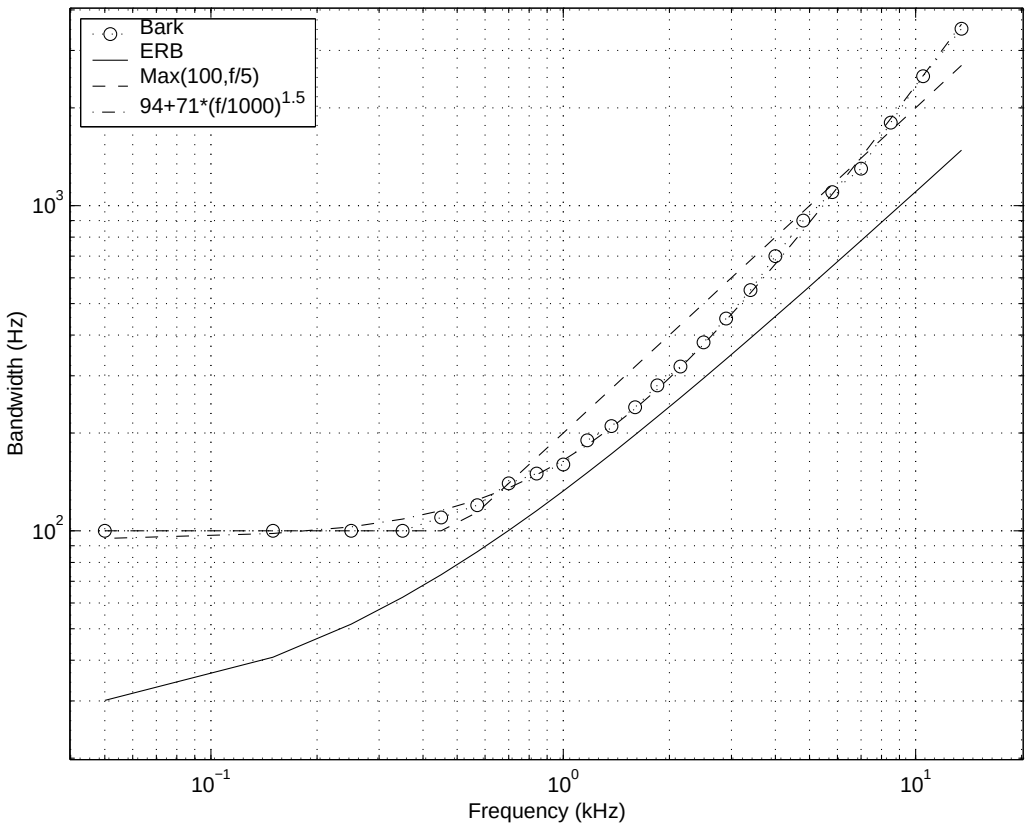

Figure 11: Bark critical bandwidth and equivalent rectangular bandwidth as a function of frequency. Also plotted is the classical rule of thumb that a critical band is 100 Hz wide for center frequencies below 500 Hz, and 20% of the center frequency above 500 Hz. Also plotted is the emprically determined formula, CB bandwidth in Hz $\approx 94 + 71f^{3/2}$, with f in kHz [37]. The ERBs are computed from Eq. (28), and the Bark CB bandwidths were computed by differencing the band-edge frequencies listed in Section 3, plotting each difference over its corresponding band center (also listed in Section 3).

is approximately 20% of center frequency, leveling off to 100 Hz below 500 Hz. An overlay of ERB and CB bandwidths is shown in Fig. 11. Also shown is the approximate classical CB bandwidth, as well as a more accurate analytical expression for Bark bandwidth vs. Hz [1]. Finally, note that the frequency interval [400 Hz, 6.5 kHz] corresponds to good agreement between the psychophysical ERB and the directly physical audio filter bandwidths defined in terms of *place* along the basilar membrane [6, p. 2601].

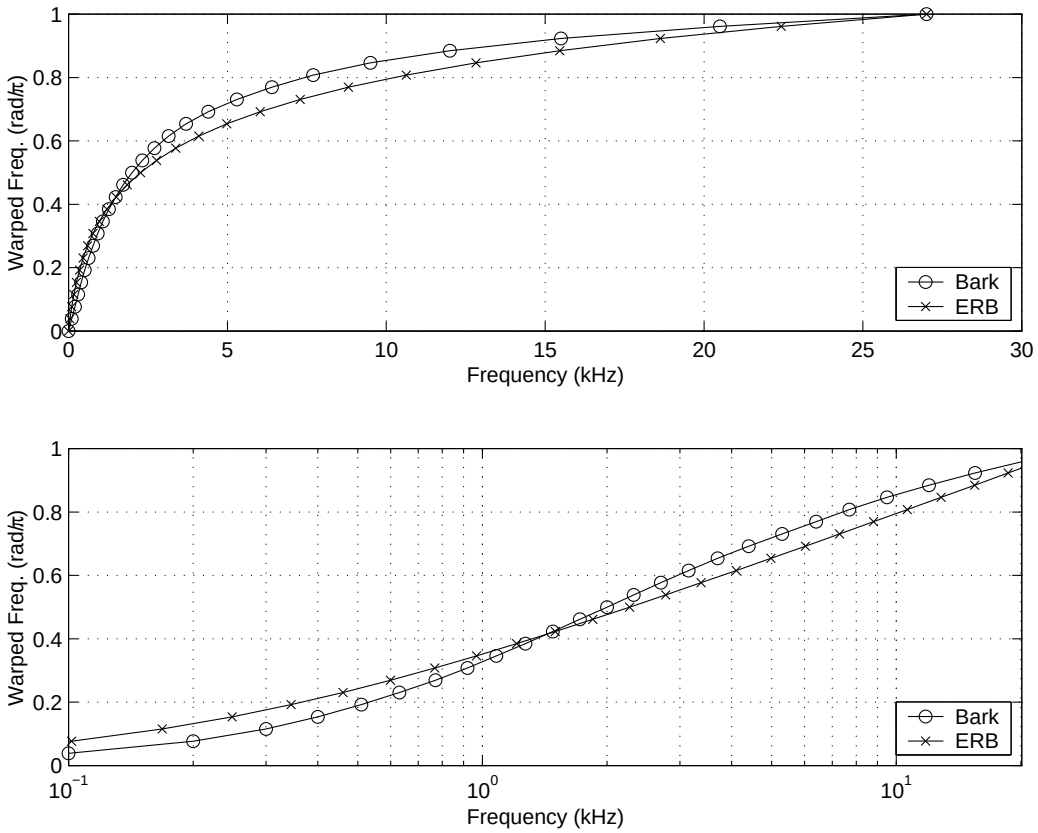

Figure 12: Bark and ERB frequency warpings for a sampling rate of 31 kHz. a) Linear input frequency scale. b) Log input frequency scale. Note that sampling is uniform across the vertical axis (corresponding to the desired audio frequency scale). As a result, the plotted samples align horizontally rather than vertically.

The *ERB scale* is defined as the number of ERBs below each frequency [19]:

$$\mathrm{ERBS}(f) = 21.4 \log_{10}(0.00437 f + 1) \tag{29}$$

for f in Hz. An overlay of the normalized Bark and ERB frequency warpings is shown in Fig. 12. The ERB warping is determined by scaling the inverse of Eq. (29), evaluated along a uniform frequency grid from zero to the number of ERBs at half the sampling rate, so that dc maps to zero and half the sampling rate maps to π.

Proceeding in the same manner as for the Bark-scale case, allpass coefficients giving a best approximation to the ERB-scale warping were computed for sampling rates near twice the Bark band edge frequencies (chosen to facilitate comparison between the ERB and Bark cases). The

resulting optimal map coefficients are shown in Fig. 13. The allpass parameter increases with increasing sampling rate, as in the Bark-scale case, but it covers a significantly narrower range, as a comparison with Fig. 4 shows. Also, the Chebyshev solution is now systematically larger than the least-squares solutions, and the least-squares and weighted equation-error cases are no longer essentially identical. The fact that the arctangent formula is optimized for the Chebyshev case is much more evident in the error plot of Fig. 13b than it was in Fig. 4b for the Bark warping parameter.

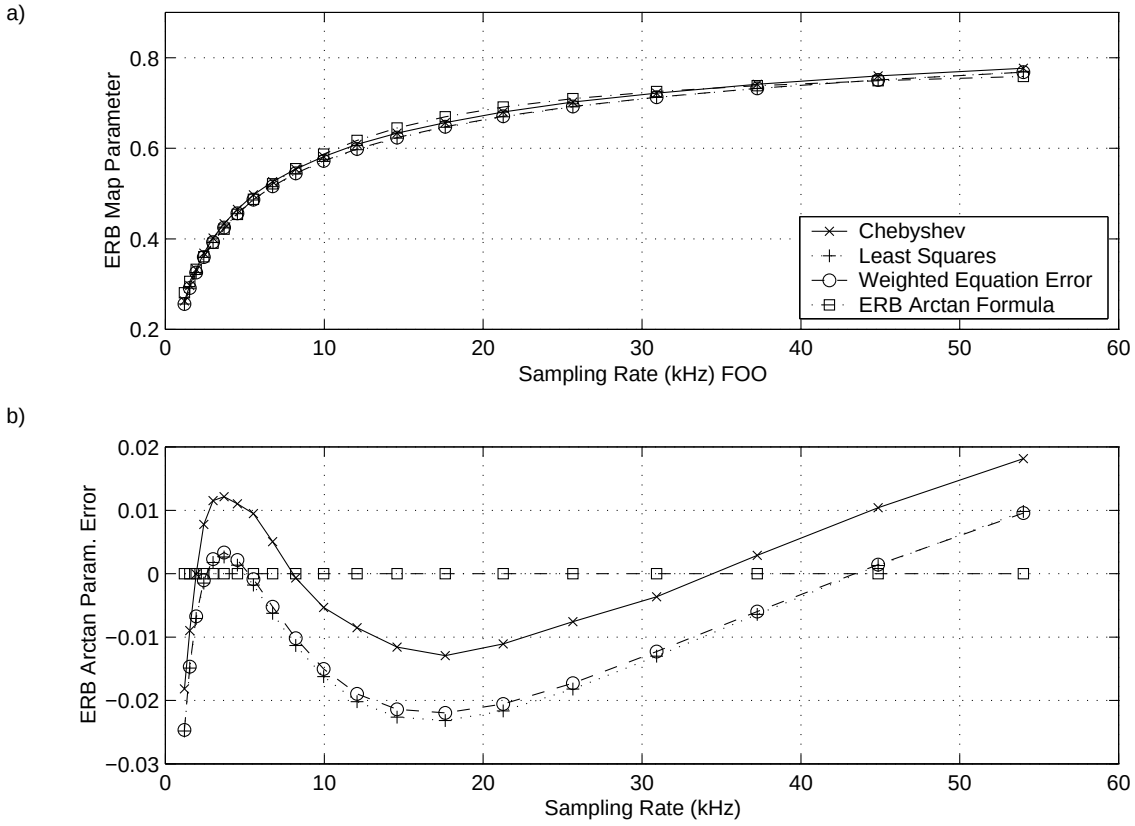

Figure 13: a) Optimal allpass coefficients ρ^* for the ERB case, plotted as a function of sampling rate f_s. Also shown is the arctangent approximation. b) Same as a) with the arctangent formula subtracted out.

The peak and rms mapping errors are plotted versus sampling rate in Fig. 14. Compare these results for the ERB scale with those for the Bark scale in Fig. 5. The ERB map errors are plotted in Barks to facilitate comparison. The rms error of the conformal map fit to the ERB scale increases nearly linearly with log-sampling-rate. The ERB-scale error increases very smoothly with frequency while the Bark-scale error is non-monotonic (see Fig. 5). The smoother behavior of the ERB errors appears due in part to the fact that the ERB scale is defined analytically while the Bark scale is defined more directly in terms of experimental data: The Bark-scale fit is so good as to be within experimental deviation, while the ERB-scale fit has a much larger systematic error component.

The peak error in Fig. 14 also grows close to linearly on a log-frequency scale and is similarly

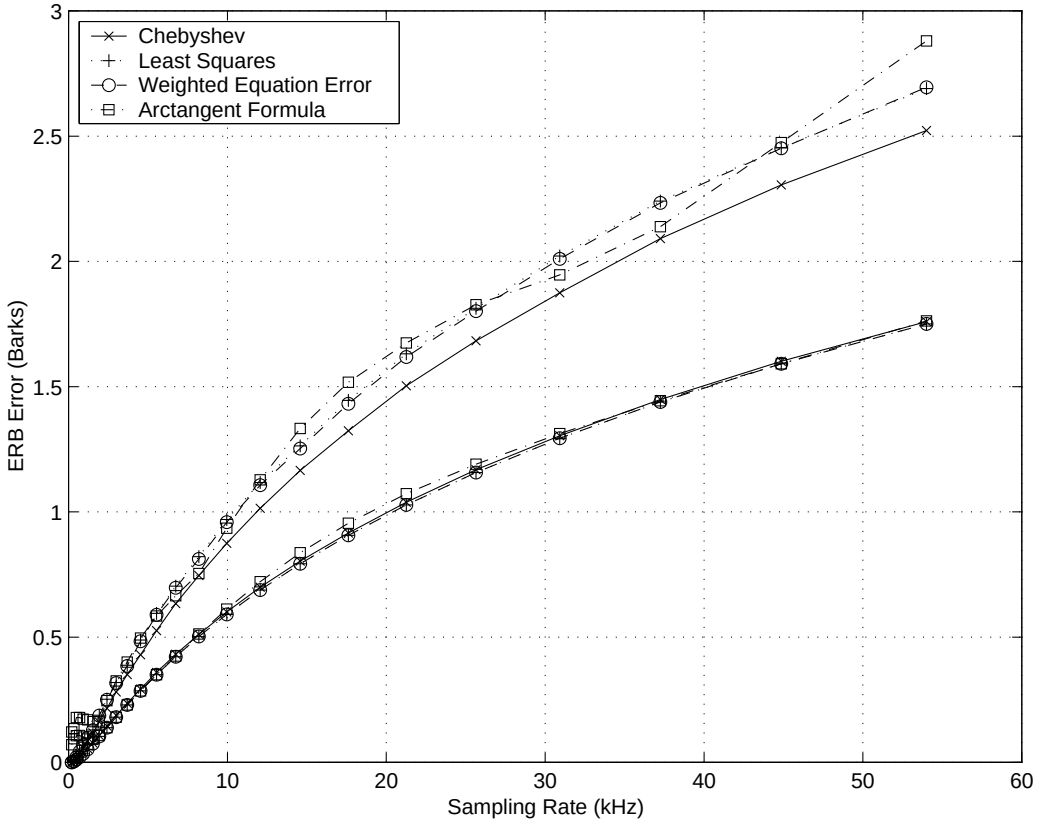

Figure 14: Root-mean-square and peak frequency-mapping errors (conformal map minus ERB) versus sampling rate for Chebyshev, least squares, weighted equation-error, and arctangent optimal maps. The rms errors are nearly coincident along the lower line, while the peak errors form an upper group well above the rms errors.

two to three times the Bark-scale errors of Fig. 5.

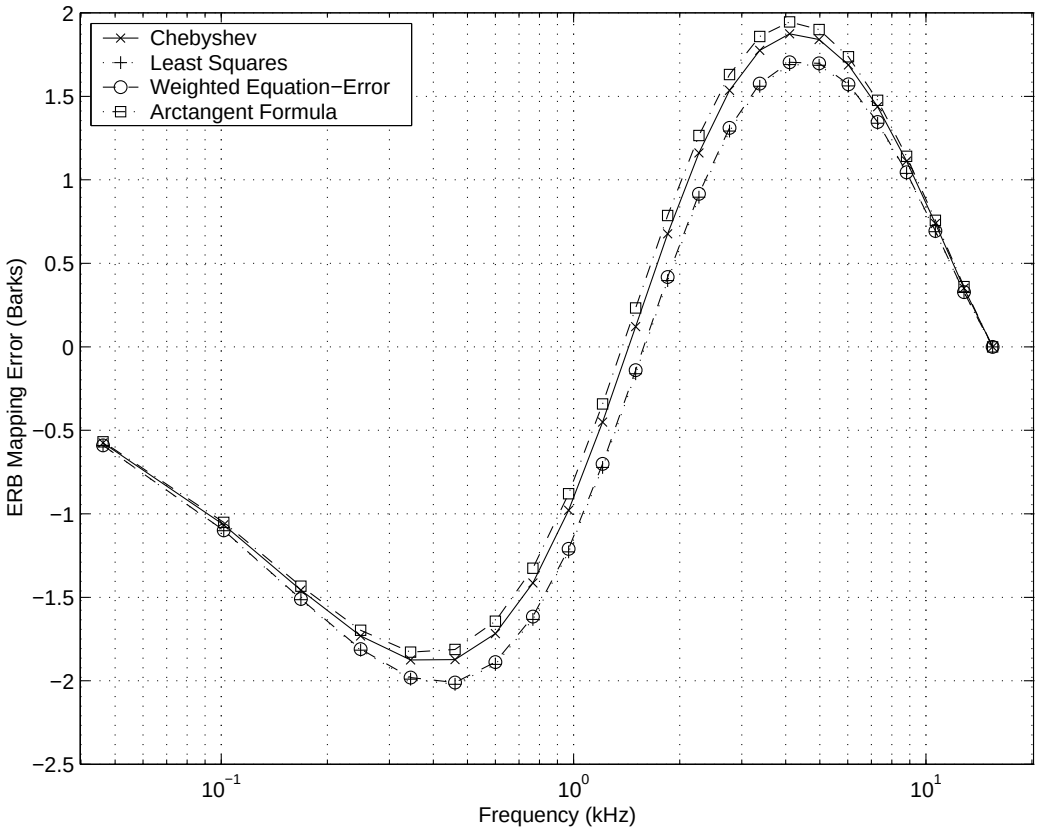

Figure 15: ERB frequency mapping errors versus frequency for the sampling rate 31 kHz.

The frequency mapping errors are plotted versus frequency in Fig. 15 for a sampling rate of 31 kHz. Unlike the Bark-scale case in Fig. 6, there is now a visible difference between the weighted equation-error and optimal least-squares mappings for the ERB scale. The figure shows also that the peak error when warping to an ERB scale is about three times larger than the peak error when warping to the Bark scale, growing from 0.64 Barks to 1.9 Barks. The locations of the peak errors are also at lower frequencies (moving from 1.3 and 8.8 kHz in the Bark-scale case to 0.7 and 8.2 kHz in the ERB-scale case).

6.1 Relative Bandwidth Mapping Error

The optimal relative bandwidth-mapping error (RBME) for the ERB case is plotted in Fig. 16 for a 31 kHz sampling rate. The peak error has grown from close to 20% for the Bark-scale case to more than 60% for the ERB case. Thus, frequency intervals are mapped to the ERB scale with up to three times as much relative error (60%) as when mapping to the Bark scale (20%). The continued narrowing of the auditory filter bandwidth as frequency decreases on the ERB scale results in the conformal map not being able to supply sufficient stretching of the low-frequency axis. The Bark scale case, on the other hand, is much better provided at low frequencies by the first-order conformal map.

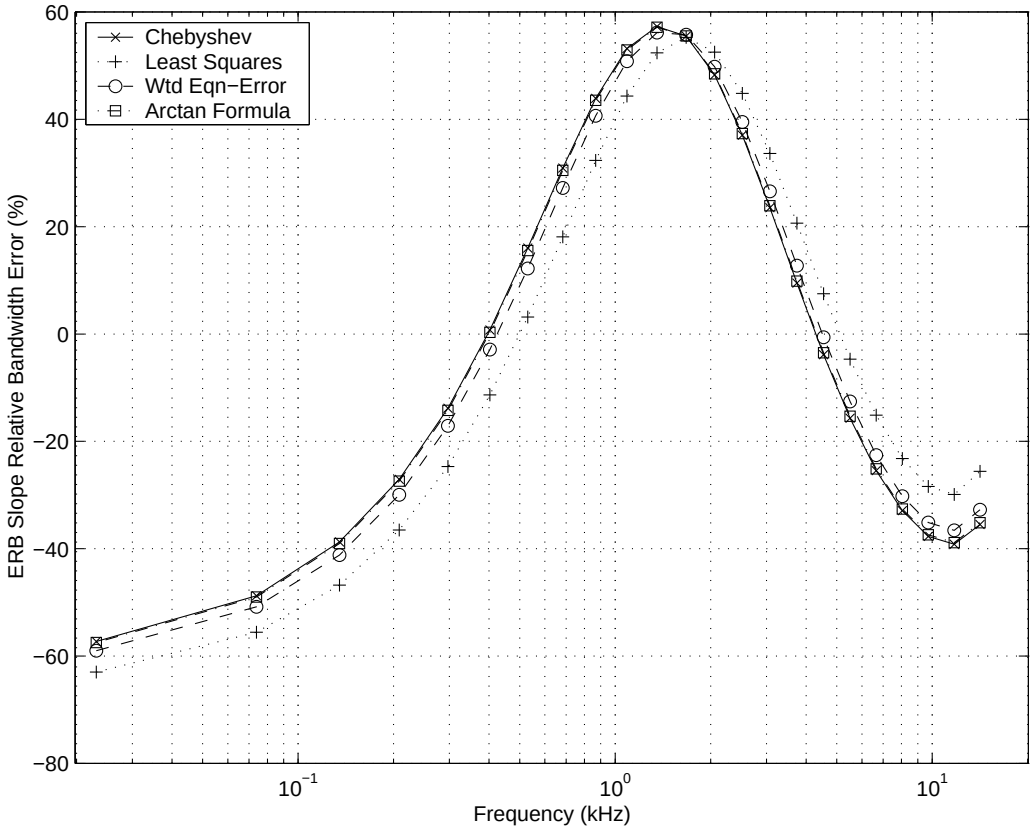

Figure 16: ERB RBME for $f_s = 31$ kHz, with explicit minimization of RBME.

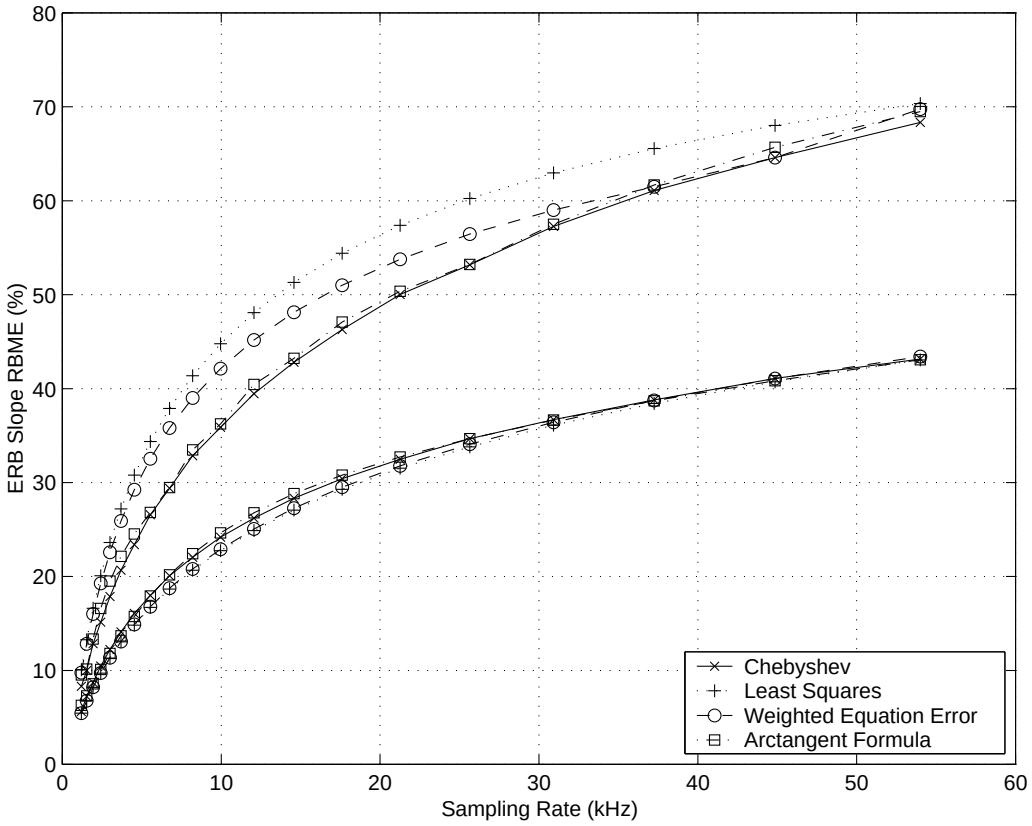

Figure 17: RMS and peak relative-bandwidth-mapping errors versus sampling rate for Chebyshev, least squares, weighted equation-error, and arctangent optimal maps, with explicit minimization of RBME used in all optimizations. The peak errors form a group lying well above the lower lying rms group.

Figure 17 shows the rms and peak ERB RBME as a function of sampling rate. Near a 10 kHz sampling rate, for example, the Chebyshev ERB RBME is increased from 12% in the Bark-scale case to around 37%, again a tripling of the peak error. We can also see in Fig. 17 that the arctangent formula gives a very good approximation to the optimal Chebyshev solution at all sampling rates. The optimal least-squares and weighted equation-error solutions are quite different, with the weighted equation-error solution moving from being close to the least-squares solution at low sampling rates, to being close to the Chebyshev solution at the higher sampling rates. The rms error is very similar in all four cases, as it was in the Bark-scale case, although the Chebyshev and arctangent formula solutions show noticeable increase in the rms error at low sampling rates where they also show a reduction in peak error by 5% or so.

6.2 Arctangent Approximations for $\rho^*(f_s)$, ERB Case

For an approximation to the optimal Chebyshev ERB frequency mapping, the arctangent formula becomes

$$\rho_\gamma^*(f_s) = 0.7446 \left[\frac{2}{\pi}\arctan(0.1418 f_s)\right]^{\frac{1}{2}} + 0.03237, \tag{30}$$

where f_s is in kHz. This formula is plotted along with the various optimal ρ^* curves in Fig. 13a, and the approximation error is shown in Fig. 13b. The performance of the arctangent approximation can be seen in Fig. 14.

When the optimality criterion is chosen to minimize relative bandwidth mapping error in the ERB case, the arctangent formula optimization yields

$$\rho_\gamma^*(f_s) = 0.7164 \left[\frac{2}{\pi}\arctan(0.09669 f_s)\right]^{\frac{1}{2}} + 0.08667. \tag{31}$$

The performance of this formula is shown in Fig. 17. It follows the optimal Chebyshev map parameter very well.

7 Directions for Improvements

Audio conformal maps can be adjusted by using a more general error weighting versus frequency. For example, the weighting can be set to zero above some frequency limit along the unit circle. A more general weighting can also be used to obtain improved accuracy in specific desired frequency ranges. Again, these refinements would seem to be of interest primarily for the ERB-scale and other mappings, since the Bark-scale warping is excellent already. The diagonal weighting matrix $\boldsymbol{V}$ in the weighted equation error solution Eq. (21) can be multiplied by any desired application-dependent weighting.

As another variation, an auditory frequency scale could be defined based on the cochlear frequency-to-place function [6]. In this case, a close relationship still exists between equal-place increments along the basilar membrane and equal bandwidth increments in the defined audio filterbank. Preliminary comparisons [6, Fig. 9] indicate that the first-order conformal map errors for this case are qualitatively between the ERB and Bark-scale cases. The first-order conformal map works best when the auditory filter bandwidths level off to a minimum width at low frequencies, as they do in the Bark-scale case below 500 Hz. Thus, the question of the "audio fidelity" of the first-order conformal map is directly tied to the question of what is really the best frequency resolution to provide at low frequencies in the auditory filterbank.

8 Conclusions

The first-order "allpass" conformal map which maps the unit circle to itself was configured to approximate frequency warpings from a linear frequency scale to either a Bark scale or an ERB frequency scale for a wide variety of sampling rates. The accuracy of this warping is extremely good for the Bark-scale case, and fair also for the ERB case; the first-order conformal map shows significantly more error in the ERB case (about three times that of the Bark-scale case) due to its narrower resolution bandwidths at low frequencies.

A closed-form expression was derived for the allpass coefficient which minimizes the norm of the weighted equation error between samples of the allpass warping and the desired Bark or ERB warpings. The weighting function was designed to give estimates as close as possible to the optimal least-squares estimate, and comparisons showed this to be well achieved, especially in the Bark-scale case.

A simple, closed-form, invertible expression which comes very close to the optimal Chebyshev allpass coefficient vs. sampling rate was given in Eq. (26) for the Bark-scale case and in Eq. (30) for the ERB-scale case.

Three optimal conformal maps were defined based on Chebyshev, least squares, and weighted equation-error approximation, and all three mappings were found to be psychoacoustically identical, for most practical purposes, in the Bark-scale case. When using optimal maps, the peak relative bandwidth mapping error is about 20% in the Bark-scale case and 60% in the ERB-scale case.

We conclude that the first-order conformal map is a highly useful tool for audio digital filter design and related applications in digital audio signal processing which may benefit from an order-invariant mapping of the unit circle from a linear frequency scale to an approximate auditory frequency scale.

Matlab code for plots, optimizations, and the filter design example presented here may be obtained at
http://ccrma.stanford.edu/~jos/bbt/bbt.html.

References

[1] J. S. Abel, "Expressions relating frequency, critical-band rate, and critical bandwidth," 1997, submitted for publication.

[2] M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagin, H. Fuchs, M. Dietz, J. Herre, G. Davidson, and Y. Oikawa, "ISO / IEC MPEG-2 advanced audio coding," *Audio Engineering Society Convention*, vol. Preprint 4382, Nov. 1996, 36 pages. See also ISO/IEC International Standard IS 13818-7 entitled "MPEG-2 Advanced Audio Coding," April, 1997.

[3] R. V. Churchill, *Complex Variables and Applications*, New York: McGraw-Hill, 1960.

[4] A. G. Constantinides, "Spectral transformations for digital filters," *Proceedings of the IEE*, vol. 117, pp. 1585–1590, Aug. 1970.

[5] B. R. Glasberg and B. C. J. Moore, "Derivation of auditory filter shapes from notched-noise data," *Hearing Research*, vol. 47, pp. 103–138, 1990.

[6] D. D. Greenwood, "A cochlear frequency-position function for several species—29 years later," *Journal of the Acoustical Society of America*, vol. 87, pp. 2592–605, June 1990.

[7] W. Hartmann, *Signals, Sound, and Sensation*, New York: AIP Press, 1997, 647 pp., 221 illustrations, hardcover.

[8] T. Irino and H. Kawahara, "Signal reconstruction from modified auditory wavelet transform," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3549–3554, 1993.

[9] T. Irino and R. D. Patterson, "A time-domain, level-dependent auditory filter: The gammachirp," *Journal of the Acoustical Society of America*, vol. 101, pp. 412–419, 1997.

[10] T. Irino and M. Unoki, "A time-varying analysis/synthesis auditory filterbank using the gammachirp," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Seattle*, (New York), IEEE Press, 1998.

[11] ISE/IEC JTC 1/SC 29/WG 11, *ISO/IEC 11172-3: Information Technology - Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s - Part 3: Audio*, Motion Picture Experts Group, 1993.

[12] M. Karjalainen, "A new auditory model for the evaluation of sound quality of audio systems," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Tampa, Florida*, (New York), pp. 608–611, IEEE Press, 1985.

[13] M. Karjalainen and J. O. Smith, "Body modeling techniques for string instrument synthesis," in *Proceedings of the 1996 International Computer Music Conference, Hong Kong*, pp. 232–239, Computer Music Association, Aug. 1996.

[14] M. Karjalainen, A. Harma, U. K. Laine, and J. Huopaniemi, "Warped filters and their audio applications," in *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY*, (New York), IEEE Press, Oct. 1997, Session 11, paper 2, 4 pages.

[15] U. K. Laine, M. Karjalainen, and T. Altosaar, "Warped linear prediction (WLP) in speech and audio processing," *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Adelaide, Australia*, pp. III:349–352, 1994.

[16] L. Ljung and T. L.Soderstrom, "The Steiglitz-McBride algorithm revisited—convergence analysis and accuracy aspects," *IEEE Transactions on Automatic Control*, vol. 26, pp. 712–717, June 1981, See also the function `stmcb()` in the Matlab Signal Processing Toolbox.

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

[18] J. D. Markel and A. H. Gray, *Linear Prediction of Speech*, New York: Springer Verlag, 1976.

[19] B. C. J. Moore and B. R. Glasberg, "A revision of Zwicker's loudness model," *Acta Acustica*, vol. 82, pp. 335–345, 1996.

[20] B. C. J. Moore, "Distribution of auditory-filter bandwidths at 2 kHz in young normal listeners," *Journal of the Acoustical Society of America*, vol. 81, pp. 1633–1635, May 1987.

[21] B. C. J. Moore, *An Introduction to the Psychology of Hearing*, New York: Academic Press, 1997.

[22] B. C. J. Moore, R. W. Peters, and B. R. Glasberg, "Auditory filter shapes at low center frequencies," *Journal of the Acoustical Society of America*, vol. 88, pp. 132–140, July 1990.

[23] J. A. Moorer, "The manifold joys of conformal mapping: Applications to digital filtering in the studio," *Journal of the Audio Engineering Society*, vol. 31, pp. 826–841, Nov 1983.

[24] Z. Nehari, *Conformal Mapping*, New York: Dover, 1952.

[25] A. V. Oppenheim, D. H. Johnson, and K. Steiglitz, "Computation of spectra with unequal resolution using the fast Fourier transform," *Proceedings of the IEEE*, vol. 59, pp. 299–301, 1971.

[26] T. W. Parks and C. S. Burrus, *Digital Filter Design*, New York: John Wiley and Sons, Inc., June 1987, contains FORTRAN software listings.

[27] R. D. Patterson, "Auditory filter shapes derived with noise stimuli," *Journal of the Acoustical Society of America*, vol. 76, pp. 640–654, Mar. 1982.

[28] R. D. Patterson and G. B. Henning, "Stimulus variability and auditory filter shape," *Journal of the Acoustical Society of America*, vol. 62, pp. 649–663, Sept. 1977.

[29] R. D. Patterson, I. Nimmo-Smith, D. L. Weber, and R. Milroy, "The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold," *Journal of the Acoustical Society of America*, vol. 72, pp. 1788–1803, Dec. 1982.

[30] R. D. Patterson, M. Allerhand, and C. Giguere, "Time-domain modelling of peripheral auditory processing: A modular architecture and software platform," *Journal of the Acoustical Society of America*, vol. 98, pp. 1890–1894, 1995.

[31] M. J. Shailer, B. C. J. Moore, B. R. Glasberg, and N. Watson, "Auditory filter shapes at 8 and 10 kHz," *Journal of the Acoustical Society of America*, vol. 88, pp. 141–148, July 1990.

[32] M. Slaney, "An efficient implementation of the Patterson-Holdsworth auditory filter bank," Tech. Rep. 35, Apple Computer, Inc., 1993, `http://-www.slaney.org/malcolm/apple/tr35/PattersonsEar.pdf`.

[33] J. O. Smith, "Spectral pre-processing for audio digital filter design," in *Proceedings of the 1983 International Computer Music Conference, Eastman School of Music*, Computer Music Association, 1983, Essentially fully contained in [34].

[34] J. O. Smith, *Techniques for Digital Filter Design and System Identification with Application to the Violin*, PhD thesis, Elec. Engineering Dept., Stanford University (CCRMA), June 1983, CCRMA Technical Report STAN-M-14, `http://-ccrma.stanford.edu/STANM/stanms/stanm14/`.

[35] J. O. Smith and J. S. Abel, "The Bark bilinear transform," in *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY*, (New York), IEEE Press, Oct. 1995, Session 8, paper 6, 4 pages. `http://-ccrma.stanford.edu/~jos/gz/bbtmh.tgz`.

[36] T. Sporer and K. Brandenburg, "Constraints of filter banks used for perceptual measurement," *Journal of the Audio Engineering Society*, vol. 43, pp. 107–116, Mar. 1995.

[37] H. W. Strube, "Linear prediction on a warped frequency scale," *Journal of the Acoustical Society of America*, vol. 68, no. 4, pp. 1071–1076, 1980.

[38] B. A. Wright, "Auditory filter asymmetry at 2000 Hz in 80 normal-hearing ears," *Journal of the Acoustical Society of America*, vol. 100, pp. 1717–1721, Sept. 1996.

[39] E. Zwicker and H. Fastl, *Psychoacoustics, Facts and Models*, Berlin: Springer Verlag, 1990, see also later 1999 edition.

[40] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models*, Berlin: Springer Verlag, 1999, second updated edition, 80pp., CD-ROM/softcover.

[41] E. Zwicker and B. Scharf, "A model of loudness summation," *Psych. Rev.*, vol. 72, pp. 3–26, 1965.

[42] E. Zwicker and E. Terhardt, "Analytical expressions for critical band rate and critical bandwidth as a function of frequency," *Journal of the Acoustical Society of America*, vol. 68, pp. 1523–1525, 1980.

Julius O. Smith is an Associate Professor of Music and (by courtesy) Electrical Engineering at the Center for Computer Research in Music and Acoustics (CCRMA), Department of Music, Stanford University (http://ccrma.stanford.edu/CCRMA/Overview/Overview.html). His activities include teaching audio signal processing courses, advising graduate students, and pursuing research in signal processing techniques applied to music, acoustics, and audio. From 1986 to 1991, he was a software engineer at NeXT Computer, Inc., responsible for signal processing software pertaining to music and audio. From 1982 to 1986 he was with the Adaptive Systems Department at Systems Control Technology, Palo Alto, CA, where he worked in the areas of adaptive filtering and spectral estimation. He received the M.S. and Ph.D. degrees in E.E. from Stanford University, Stanford, CA, in 1978 and 1983, respectively. His Ph.D. research involved the application of digital filter design and system identification techniques to the modeling and synthesis of the violin, clarinet, reverberant spaces, and other musical systems. From 1975 to 1977 he worked in the Signal Processing Department at ESL, Sunnyvale, CA, on systems for digital communications. He received the B.S.E.E. degree from Rice University, Houston, TX, in 1975. For more information, see `http://ccrma.stanford.edu/~jos/`.

Jonathan S. Abel is a researcher with the San Jose State University Foundation studying spatial hearing on a grant from the Human Factors Research Division of NASA Ames Research Center. He also owns Abel Innovations, an engineering consulting firm specializing in audio signal processing. He was chief scientist of Crystal River Engineering, Inc. where he developed efficient methods for synthesizing spatial audio cues and measuring head-related transfer functions. Prior to joining Crystal River, Dr. Abel was vice president of Tetra Systems Incorporated, a lecturer at Yale University, and a consultant to Northwest Digital Research, Systems Control Technology, Saxpy Computer, and Apple Computer, among others. He holds a PhD and MS from Stanford University, and an SB from MIT, all in electrical engineering.